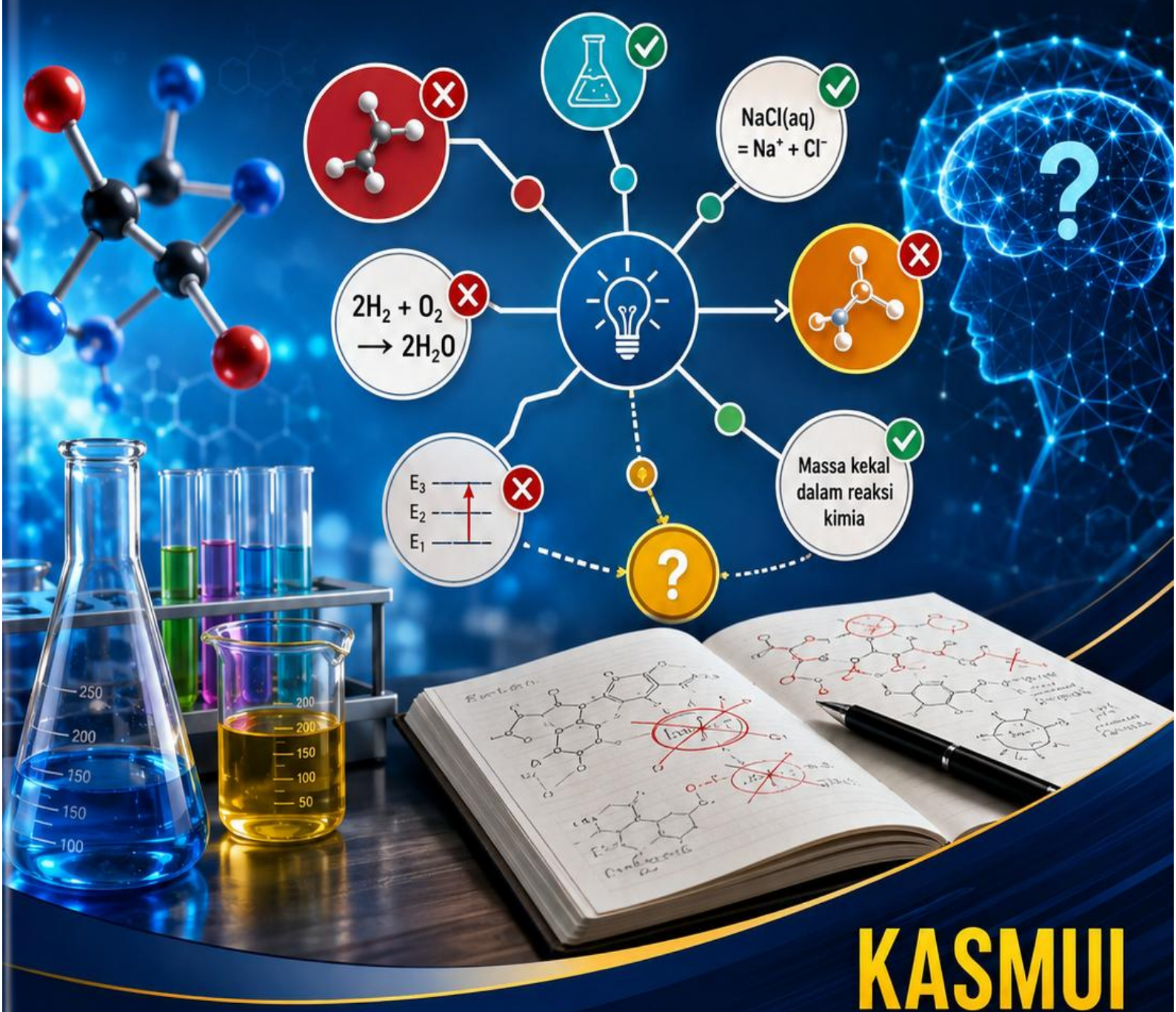


MISKONSEPSI KONSEP KIMIA SMA

BAGIAN 2: DIAGNOSIS DAN PEMETAAN MISKONSEPSI KIMIA



KASMUI

MISKONSEPSI KONSEP KIMIA SMA

BAGIAN II

DIAGNOSIS DAN PEMETAAN MISKONSEPSI KIMIA

KASMUI

Jaten Press

MISKONSEPSI KONSEP KIMIA SMA

BAGIAN 2: DIAGNOSIS DAN PEMETAAN MISKONSEPSI KIMIA

Penulis KASMUI

Desain Sampul KASMUI / Tim Artistik Penerbit

Penata Letak KASMUI / Tim Layout

Penerbit Jaten & Siblings Patemon, Gunungpati, Semarang

Cetakan dan Edisi Cetakan I, Edisi Tahun 2026

Ukuran Buku 21,59 × 27,94 cm

Jumlah Halaman 1068

ISBN [Diisi setelah ISBN resmi diterbitkan]

Hak Cipta © 2026 pada Penulis

Hak Cipta Dilindungi Undang-Undang.

Dilarang mengutip, memperbanyak, atau menerjemahkan sebagian atau seluruh isi buku ini tanpa izin tertulis dari Penerbit, kecuali untuk kepentingan tinjauan pustaka, resensi, atau kutipan ilmiah yang wajar.

Perpustakaan Nasional RI: Katalog Dalam Terbitan (KDT)

Kasmui

Miskonsepsi Konsep Kimia SMA. Bagian 2: Diagnosis dan Pemetaan Miskonsepsi Kimia / Kasmui. — Cetakan I, Edisi Tahun 2026. — Semarang: Jaten & Siblings, 2026.

1. Kimia—Pembelajaran.
2. Miskonsepsi Kimia.
3. Asesmen Diagnostik.
4. Pendidikan Kimia—Sekolah Menengah Atas.
5. Pemetaan Konsep.

I. Judul.

DDC:

Dicetak oleh Jaten Press.

KATA PENGANTAR

Bismillāhirrahmānirrahīm

Puji syukur ke hadirat Allah Swt. atas limpahan rahmat, taufik, hidayah, dan inayah-Nya sehingga buku **Miskonsepsi Konsep Kimia SMA: Bagian II—Diagnosis dan Pemetaan Miskonsepsi Kimia** ini dapat disusun dan diselesaikan. Selawat dan salam semoga senantiasa tercurah kepada Nabi Muhammad saw., keluarga, sahabat, dan seluruh umat beliau yang dengan sungguh-sungguh mengikuti petunjuknya hingga akhir zaman.

Buku ini merupakan kelanjutan dari pembahasan mengenai miskonsepsi dalam pembelajaran kimia pada jenjang Sekolah Menengah Atas. Apabila bagian sebelumnya lebih banyak memberikan landasan konseptual tentang hakikat konsep, konsepsi, miskonsepsi, perubahan konseptual, sumber-sumber kesalahan pemahaman, serta karakteristik konsep kimia, maka bagian kedua ini diarahkan pada persoalan yang lebih operasional, metodologis, dan aplikatif, yaitu bagaimana miskonsepsi kimia didiagnosis, dipetakan, dianalisis, ditafsirkan, dan dilaporkan secara bertanggung jawab.

Penyusunan buku ini dilatarbelakangi oleh kenyataan bahwa kesalahan jawaban peserta didik tidak selalu memiliki penyebab yang sama. Jawaban yang salah dapat muncul karena peserta didik belum mengetahui konsep, lupa terhadap informasi tertentu, melakukan kesalahan perhitungan, keliru membaca soal, tidak mampu menghubungkan berbagai representasi kimia, memiliki pengetahuan yang belum lengkap, atau telah membangun suatu konsepsi alternatif yang diyakini benar. Oleh karena itu, guru tidak cukup hanya membedakan jawaban benar dan salah. Guru perlu memperoleh informasi yang lebih mendalam mengenai alasan, tingkat keyakinan, pola berpikir, model mental, dan hubungan antarkonsep yang mendasari respons peserta didik.

Dalam konteks tersebut, diagnosis miskonsepsi harus dipahami sebagai proses sistematis untuk mengumpulkan, menganalisis, dan menafsirkan bukti mengenai struktur pemahaman peserta didik. Diagnosis bukan sekadar pemberian tes awal, bukan pula kegiatan untuk memberikan label negatif kepada peserta didik. Diagnosis merupakan bagian integral dari pembelajaran yang bertujuan membantu guru menentukan tindakan pedagogis yang paling tepat. Penanganan terhadap peserta didik yang belum memahami konsep tentu berbeda dari penanganan terhadap peserta didik yang telah memiliki miskonsepsi kuat dan disertai tingkat keyakinan tinggi.

Kompleksitas diagnosis semakin meningkat karena ilmu kimia dipelajari melalui berbagai tingkat representasi. Peserta didik harus menghubungkan fenomena yang dapat diamati secara makroskopik dengan model partikel pada tingkat submikroskopik, simbol, rumus, persamaan reaksi, grafik, tabel, perhitungan matematis, dan berbagai bentuk representasi lainnya. Seseorang dapat menjawab soal numerik dengan benar, tetapi belum tentu

memahami proses partikulat yang terjadi. Sebaliknya, peserta didik mungkin mampu menjelaskan fenomena secara verbal, tetapi kesulitan menerjemahkannya ke dalam persamaan kimia atau representasi matematis. Kondisi tersebut menuntut penggunaan instrumen diagnosis yang beragam dan tidak bergantung pada satu jenis tes saja.

Buku ini membahas prinsip-prinsip dasar diagnosis miskonsepsi, tujuan dan waktu pelaksanaannya, perbedaan antara kesalahan biasa dan miskonsepsi, keterkaitan antara jawaban, alasan, serta tingkat keyakinan, dan penyusunan profil konseptual peserta didik. Pembahasan juga mencakup validitas, reliabilitas, kualitas inferensi, potensi bias, etika penggunaan hasil diagnosis, serta pemanfaatannya untuk memperbaiki proses pembelajaran kimia.

Berbagai instrumen diagnostik dijelaskan secara sistematis, mulai dari tes pilihan ganda konvensional, tes pilihan ganda beralasan, *two-tier diagnostic test*, *three-tier diagnostic test*, *four-tier diagnostic test*, dan *five-tier diagnostic test*. Selain itu, dibahas pula penggunaan *Certainty of Response Index*, wawancara klinis, wawancara berbasis tugas, pertanyaan terbuka, *Predict–Observe–Explain*, *think-aloud protocol*, tes menggambar, peta konsep, analisis kesalahan penyelesaian masalah, analisis percakapan kelas, dan pemanfaatan data aktivitas digital.

Setiap instrumen tidak hanya dijelaskan dari segi bentuk dan prosedur penggunaannya, tetapi juga ditinjau berdasarkan landasan konseptual, perkembangan historis, konstruk yang diukur, desain stimulus, sistem penskoran, pengodean respons, interpretasi kategori pemahaman, validitas, reliabilitas, kelebihan, keterbatasan, serta kemungkinan bias. Dengan cara tersebut, pembaca diharapkan tidak sekadar mengetahui nama-nama instrumen, tetapi mampu memilih, mengembangkan, menggunakan, dan mengevaluasinya sesuai dengan tujuan diagnosis.

Perhatian khusus juga diberikan pada pengembangan distraktor berbasis miskonsepsi. Distraktor yang baik tidak semata-mata berupa jawaban salah, melainkan harus merepresentasikan pola pemikiran atau konsepsi alternatif yang benar-benar mungkin dimiliki peserta didik. Proses pengembangan instrumen karena itu perlu didukung oleh kajian literatur, wawancara, respons terbuka, analisis buku ajar, observasi kelas, validasi ahli, wawancara kognitif, uji coba empiris, serta analisis butir yang memadai.

Selain diagnosis melalui instrumen, buku ini menguraikan metodologi penelitian miskonsepsi kimia secara luas. Pembahasan mencakup paradigma penelitian, penelitian deskriptif, studi kasus, survei, eksperimen, kuasi-eksperimen, *mixed methods*, penelitian pengembangan instrumen, *design-based research*, penelitian tindakan kelas, dan *systematic literature review*. Pembaca juga diperkenalkan pada prinsip pemilihan subjek dan sampel, teknik pengumpulan data, analisis kuantitatif, analisis kualitatif, triangulasi, analisis pola respons, serta pemetaan jaringan konsep.

Pemetaan miskonsepsi dalam buku ini tidak hanya dipandang sebagai penyajian daftar kesalahan peserta didik. Pemetaan diarahkan untuk mengungkap hubungan antarkonsep, tingkat sentralitas suatu konsep, pola keterhubungan proposisi, konsistensi respons, distribusi miskonsepsi dalam kelas, serta perubahan struktur pengetahuan sebelum dan setelah pembelajaran. Dengan pemetaan yang tepat, guru dan peneliti dapat mengenali miskonsepsi yang bersifat lokal maupun yang menjadi pusat dari berbagai kesalahan konseptual lainnya.

Aspek etika memperoleh perhatian penting dalam keseluruhan pembahasan. Hasil diagnosis tidak boleh digunakan untuk memperlakukan, memberi cap, atau mengurangi martabat peserta didik. Istilah miskonsepsi harus digunakan secara hati-hati sebagai kategori analitis terhadap suatu respons atau pola pemahaman, bukan sebagai identitas tetap seseorang. Pengumpulan, penyimpanan, analisis, dan pelaporan data juga harus memperhatikan persetujuan partisipan, kerahasiaan, perlindungan data pribadi, keadilan, transparansi, serta penggunaan bahasa yang tidak menstigmatisasi.

Buku ini diharapkan dapat digunakan oleh guru kimia, mahasiswa pendidikan kimia, calon guru, dosen, pengembang instrumen, peneliti pendidikan sains, pengelola laboratorium pembelajaran, serta pihak-pihak yang memiliki perhatian terhadap peningkatan mutu pembelajaran kimia. Bagi guru, buku ini dapat menjadi panduan dalam merancang asesmen diagnostik dan menindaklanjuti hasilnya. Bagi mahasiswa dan peneliti, buku ini dapat menjadi referensi dalam menentukan desain penelitian, mengembangkan instrumen, menganalisis data, dan menyajikan temuan secara ilmiah. Bagi pengembang kurikulum dan pengambil kebijakan, buku ini diharapkan memberikan perspektif bahwa keberhasilan pembelajaran tidak cukup diukur melalui capaian nilai, tetapi juga melalui kualitas struktur pemahaman yang dibangun peserta didik.

Penulis menyadari bahwa pembahasan mengenai diagnosis dan pemetaan miskonsepsi merupakan bidang yang sangat luas dan terus berkembang. Perkembangan teknologi digital, analitik pembelajaran, kecerdasan buatan, pemodelan jaringan, dan asesmen berbasis komputer membuka peluang baru untuk memperoleh gambaran pemahaman peserta didik secara lebih cepat dan mendalam. Namun, teknologi tetap harus ditempatkan sebagai alat bantu. Keputusan diagnostik yang baik tetap memerlukan kejelasan konstruk, kualitas instrumen, ketepatan interpretasi, pertimbangan konteks, serta kebijaksanaan profesional guru dan peneliti.

Dalam penyusunan buku ini, penulis berusaha mengintegrasikan teori, metodologi, hasil penelitian, contoh penerapan, prinsip pengukuran, pertimbangan etika, serta kebutuhan nyata pembelajaran kimia di sekolah. Meskipun demikian, penulis menyadari bahwa buku ini masih mungkin memiliki keterbatasan, baik dalam keluasan cakupan, kedalaman pembahasan, ketepatan penggunaan istilah, penyajian contoh, maupun kemutakhiran referensi. Oleh karena itu, kritik, koreksi, dan saran yang konstruktif dari pembaca sangat diharapkan untuk penyempurnaan buku ini pada edisi berikutnya.

Penulis menyampaikan penghargaan dan terima kasih kepada para guru, dosen, mahasiswa, peneliti, penulis buku, pengembang instrumen, dan seluruh pemerhati pendidikan kimia yang karya, pengalaman, pemikiran, dan hasil penelitiannya telah memberikan inspirasi bagi penyusunan buku ini. Terima kasih juga disampaikan kepada semua pihak yang telah memberikan dukungan moral, akademik, teknis, dan kelembagaan selama proses penulisan dan penerbitan.

Akhirnya, penulis berharap buku ini dapat memberikan manfaat nyata bagi peningkatan kualitas pendidikan kimia. Semoga buku ini membantu para pendidik beralih dari pembelajaran yang hanya berorientasi pada penyampaian materi dan pencapaian nilai menuju pembelajaran yang lebih peka terhadap cara berpikir peserta didik. Diagnosis yang tepat diharapkan menghasilkan keputusan pembelajaran yang tepat, sedangkan pemetaan yang baik diharapkan menjadi dasar bagi penyusunan strategi perubahan konseptual yang efektif, manusiawi, dan berkelanjutan.

Semoga ikhtiar sederhana ini menjadi bagian dari pengembangan ilmu pengetahuan dan memberikan kontribusi bagi terwujudnya pembelajaran kimia yang lebih bermakna, kritis, ilmiah, dan mencerdaskan.

Semarang, Juli 2026

Penulis,



KASMUI

DAFTAR ISI

KATA PENGANTAR.....	4
DAFTAR ISI.....	8
BAGIAN II	41
DIAGNOSIS DAN PEMETAAN MISKONSEPSI KIMIA	41
BAB 5.....	42
PRINSIP DIAGNOSIS MISKONSEPSI	42
Pengantar Bab	42
5.1 Hakikat Asesmen Diagnostik.....	43
5.1.1 Asesmen diagnostik sebagai proses pengumpulan bukti	43
5.1.2 Perbedaan pengukuran, penilaian, asesmen, dan evaluasi	44
5.1.3 Asesmen diagnostik sebagai proses inferensi	45
5.1.4 Kedudukan diagnosis dalam siklus pembelajaran kimia	46
5.1.5 Diagnosis sebagai upaya membuat pemikiran peserta didik terlihat	48
5.1.6 Objek diagnosis: konsep, relasi, model, dan prosedur	49
5.1.7 Diagnosis terhadap pengetahuan awal dan prasyarat	50
5.1.8 Hubungan asesmen diagnostik dan asesmen formatif.....	51
5.1.9 Perbedaan asesmen diagnostik dan asesmen sumatif	53
5.1.10 Spektrum metode diagnosis miskonsepsi	54
5.1.11 Tes tertulis sebagai alat diagnosis awal.....	55
5.1.12 Wawancara diagnostik dan probing	56
5.1.13 Observasi penalaran dan tindakan eksperimen	57
5.1.14 Peta konsep sebagai representasi struktur pengetahuan	59
5.1.15 Gambar, diagram, dan model partikel sebagai bukti pemahaman	60
5.1.16 Predict-Observe-Explain dalam diagnosis kimia	61
5.1.17 Think-aloud untuk menelusuri proses penalaran	62
5.1.18 Analisis pekerjaan peserta didik dan jejak penyelesaian	63
5.1.19 Pemanfaatan data digital dan respons waktu nyata	65
5.1.20 Triangulasi metode dan sumber bukti.....	66
5.1.21 Keterbatasan asesmen diagnostik.....	67

5.1.22 Peran guru sebagai penafsir bukti	68
5.1.23 Implementasi diagnosis pada kelas yang heterogen	69
5.1.24 Sintesis hakikat asesmen diagnostik	70
5.1.25 Implikasi praktis bagi guru kimia	72
5.2 Tujuan Diagnosis Miskonsepsi	72
5.2.1 Mengidentifikasi konsepsi awal sebelum pembelajaran.....	72
5.2.2 Menemukan miskonsepsi yang tersembunyi di balik jawaban benar	73
5.2.3 Membedakan ketidaktahuan, pengetahuan parsial, dan miskonsepsi	74
5.2.4 Memetakan konsep prasyarat yang belum kokoh	75
5.2.5 Menentukan prioritas pembelajaran dan remediasi	77
5.2.6 Menyesuaikan representasi dan strategi pengajaran	78
5.2.7 Mencegah penguatan miskonsepsi melalui pembelajaran lanjutan.....	79
5.2.8 Menentukan kebutuhan diferensiasi pembelajaran	80
5.2.9 Memberikan umpan balik yang spesifik dan dapat ditindaklanjuti.....	82
5.2.10 Mendukung regulasi diri dan metakognisi peserta didik	83
5.2.11 Mengevaluasi efektivitas penjelasan, analogi, dan model.....	84
5.2.12 Mengevaluasi kualitas buku teks dan sumber belajar.....	85
5.2.13 Mengidentifikasi miskonsepsi kolektif pada tingkat kelas	87
5.2.14 Mendeteksi miskonsepsi yang khas secara individual.....	88
5.2.15 Menilai perubahan konseptual dan kestabilannya.....	89
5.2.16 Menguji transfer pemahaman ke konteks baru	90
5.2.17 Mengidentifikasi hubungan antara konsep dan prosedur.....	91
5.2.18 Mengurangi keputusan pembelajaran yang hanya berbasis skor	93
5.2.19 Mendukung komunikasi guru, peserta didik, dan orang tua	94
5.2.20 Menyediakan data bagi pengembangan kurikulum	95
5.2.21 Mendukung penelitian tindakan kelas	96
5.2.22 Menjaga fokus diagnosis pada perbaikan, bukan pelabelan	97
5.2.23 Indikator ketercapaian tujuan diagnosis	98
5.2.24 Sintesis tujuan diagnosis miskonsepsi.....	100
5.2.25 Implikasi praktis bagi guru kimia	101
5.3 Diagnosis sebelum, selama, dan setelah Pembelajaran	101
5.3.1 Prinsip kontinuitas diagnosis	101

5.3.2	Diagnosis sebelum pembelajaran sebagai pemetaan kesiapan	102
5.3.3	Menentukan konsep prasyarat yang harus diperiksa	103
5.3.4	Menyusun asesmen awal yang singkat tetapi bermakna	105
5.3.5	Menggunakan fenomena pemantik untuk elisitasi gagasan awal	106
5.3.6	Diagnosis selama pembelajaran melalui pertanyaan berjenjang	107
5.3.7	Diagnosis melalui diskusi kelas dan argumentasi	108
5.3.8	Diagnosis melalui kerja kelompok dan interaksi sosial	109
5.3.9	Diagnosis melalui praktikum dan demonstrasi	111
5.3.10	Diagnosis melalui representasi yang berubah-ubah.....	112
5.3.11	Exit ticket dan minute paper sebagai diagnosis cepat.....	113
5.3.12	Penggunaan kuis formatif berulang	114
5.3.13	Diagnosis setelah pembelajaran untuk memeriksa perubahan konsep	116
5.3.14	Membedakan peningkatan skor dan perubahan konseptual.....	117
5.3.15	Diagnosis tertunda untuk menguji retensi.....	118
5.3.16	Diagnosis transfer pada konteks baru	119
5.3.17	Membandingkan bukti sebelum dan sesudah pembelajaran	121
5.3.18	Menentukan tindak lanjut pascadiagnosis	122
5.3.19	Siklus diagnosis-remediasi-rediagnosis.....	123
5.3.20	Pengelolaan waktu dalam diagnosis berkelanjutan	124
5.3.21	Dokumentasi longitudinal perkembangan konsep	125
5.3.22	Contoh alur diagnosis pada satu unit kimia	126
5.3.23	Analisis kritis terhadap model pretest-posttest sederhana	128
5.3.24	Sintesis diagnosis sepanjang pembelajaran	129
5.3.25	Implikasi praktis bagi guru kimia	130
5.4	Membedakan Kesalahan dan Miskonsepsi	130
5.4.1	Mengapa setiap jawaban salah bukan miskonsepsi	130
5.4.2	Kesalahan karena tidak tahu	131
5.4.3	Kesalahan karena lupa atau gagal memanggil pengetahuan	133
5.4.4	Kesalahan membaca dan memahami bahasa soal	134
5.4.5	Kesalahan aritmetika dan manipulasi aljabar	135
5.4.6	Kesalahan simbolik dan notasi kimia	136
5.4.7	Kesalahan prosedural dalam algoritma penyelesaian	137

5.4.8 Kesalahan akibat beban kognitif	139
5.4.9 Kesalahan akibat ketidaktelitian	140
5.4.10 Kesalahan karena strategi tebakan	141
5.4.11 Kesalahan kontekstual dan ketidakstabilan respons	142
5.4.12 Miskonsepsi sebagai pola penalaran yang sistematis	143
5.4.13 Konsistensi lintas butir sebagai indikator	145
5.4.14 Keyakinan tinggi sebagai bukti tambahan	146
5.4.15 Resistensi terhadap koreksi sebagai indikator	147
5.4.16 Peran wawancara dalam diferensiasi kesalahan	148
5.4.17 Peran tugas paralel dan isomorfik	150
5.4.18 Analisis jejak kerja dan langkah penyelesaian	151
5.4.19 Matriks klasifikasi kesalahan dan miskonsepsi	152
5.4.20 Risiko false positive dalam pelabelan miskonsepsi	153
5.4.21 Risiko false negative dalam diagnosis	155
5.4.22 Contoh diferensiasi pada stoikiometri	156
5.4.23 Contoh diferensiasi pada kesetimbangan kimia	157
5.4.24 Sintesis pembedaan kesalahan dan miskonsepsi	158
5.4.25 Implikasi praktis bagi guru kimia	159
5.5 Hubungan Jawaban, Alasan, dan Keyakinan	159
5.5.1 Jawaban sebagai bukti yang belum lengkap	160
5.5.2 Alasan sebagai jendela struktur pengetahuan	161
5.5.3 Keyakinan sebagai dimensi metakognitif	162
5.5.4 Empat pola dasar jawaban dan alasan	163
5.5.5 Jawaban benar dengan alasan benar	164
5.5.6 Jawaban benar dengan alasan salah	166
5.5.7 Jawaban salah dengan alasan benar sebagian	167
5.5.8 Jawaban salah dengan alasan salah	168
5.5.9 Makna keyakinan tinggi dan rendah	169
5.5.10 Certainty of Response Index dan penggunaannya	170
5.5.11 Tes dua tingkat untuk diagnosis konseptual	171
5.5.12 Tes tiga tingkat dan penambahan keyakinan	173
5.5.13 Tes empat tingkat dan pemisahan keyakinan jawaban-alasan	174

5.5.14 Alasan terbuka dan alasan pilihan ganda	175
5.5.15 Kualitas distraktor sebagai representasi konsepsi alternatif.....	176
5.5.16 Kalibrasi keyakinan dan meta-ignorance	177
5.5.17 Ketidaksesuaian antara performa dan keyakinan.....	179
5.5.18 Analisis pola respons per butir	180
5.5.19 Analisis pola respons per peserta didik.....	181
5.5.20 Batas penggunaan skor keyakinan	182
5.5.21 Wawancara konfirmasi terhadap klasifikasi tier	184
5.5.22 Contoh matriks jawaban-alasan-keyakinan	185
5.5.23 Implikasi bagi umpan balik dan remediasi.....	186
5.5.24 Sintesis hubungan jawaban, alasan, dan keyakinan	187
5.5.25 Implikasi praktis bagi guru kimia	188
5.6 Prinsip Validitas Diagnosis	189
5.6.1 Makna validitas dalam diagnosis miskonsepsi	189
5.6.2 Validitas melekat pada interpretasi dan penggunaan hasil	190
5.6.3 Argumentasi interpretasi dan penggunaan skor.....	191
5.6.4 Representasi konstruk miskonsepsi	192
5.6.5 Batas domain konsep yang didiagnosis.....	193
5.6.6 Analisis konsep sebagai dasar kisi-kisi	195
5.6.7 Bukti validitas berbasis isi.....	196
5.6.8 Bukti validitas berbasis proses respons.....	197
5.6.9 Bukti validitas berbasis struktur internal	198
5.6.10 Bukti validitas berbasis hubungan dengan variabel lain	199
5.6.11 Bukti validitas berbasis konsekuensi penggunaan	200
5.6.12 Validitas distraktor konseptual	201
5.6.13 Validitas alasan dan tingkat keyakinan	203
5.6.14 Validitas wawancara diagnostik	204
5.6.15 Validitas tugas representasional.....	205
5.6.16 Validitas lintas konteks dan transfer	206
5.6.17 Validitas lintas kelompok peserta didik.....	208
5.6.18 Ancaman underrepresentation konstruk	209
5.6.19 Ancaman construct-irrelevant variance	210

5.6.20 Bahasa, keterbacaan, dan validitas	211
5.6.21 Bias budaya dan pengalaman kontekstual	212
5.6.22 Expert judgment dan indeks validitas isi	213
5.6.23 Uji coba kognitif dan revisi butir	214
5.6.24 Triangulasi sebagai penguat argumentasi validitas	216
5.6.25 Dokumentasi bukti validitas	217
5.6.26 Sintesis prinsip validitas diagnosis	218
5.6.27 Implikasi praktis bagi guru kimia	219
5.7 Prinsip Reliabilitas Instrumen	220
5.7.1 Makna reliabilitas dalam diagnosis konseptual	220
5.7.2 Reliabilitas sebagai konsistensi bukti	221
5.7.3 Sumber kesalahan pengukuran	222
5.7.4 Konsistensi internal dan keterbatasannya	223
5.7.5 Koefisien alpha dan asumsi dasarnya	224
5.7.6 Alternatif koefisien omega dan reliabilitas komposit	226
5.7.7 Reliabilitas skor dikotomis	227
5.7.8 Reliabilitas skor politomus dan tier	228
5.7.9 Stabilitas waktu melalui test-retest	229
5.7.10 Kesetaraan bentuk paralel	230
5.7.11 Reliabilitas antarpemilai	231
5.7.12 Reliabilitas pengodean alasan terbuka	232
5.7.13 Reliabilitas wawancara diagnostik	234
5.7.14 Reliabilitas observasi dan rubrik	235
5.7.15 Reliabilitas pada sampel kecil	236
5.7.16 Pengaruh panjang tes terhadap reliabilitas	237
5.7.17 Pengaruh heterogenitas peserta didik	239
5.7.18 Analisis butir klasik	240
5.7.19 Model Rasch dan fungsi informasi	241
5.7.20 Person reliability dan item reliability	242
5.7.21 Differential item functioning	243
5.7.22 Standard error pada skor individual	245
5.7.23 Keputusan kategori dan konsistensi klasifikasi	246

5.7.24 Menyeimbangkan reliabilitas dan kedalaman diagnosis.....	247
5.7.25 Pelaporan reliabilitas secara bertanggung jawab	248
5.7.26 Sintesis prinsip reliabilitas instrumen	249
5.7.27 Implikasi praktis bagi guru kimia	250
5.8 Diagnosis Individual dan Diagnosis Kelas	251
5.8.1 Perbedaan unit analisis individual dan kelas	251
5.8.2 Tujuan diagnosis individual.....	252
5.8.3 Tujuan diagnosis kelas.....	253
5.8.4 Profil respons individual	254
5.8.5 Pola dominan pada tingkat kelas.....	256
5.8.6 Prevalensi miskonsepsi dan maknanya	257
5.8.7 Distribusi keyakinan dalam kelas	258
5.8.8 Kelompok konseptual dan diferensiasi	259
5.8.9 Diagnosis peserta didik dengan respons tidak konsisten	260
5.8.10 Diagnosis peserta didik berkemampuan tinggi.....	262
5.8.11 Diagnosis peserta didik yang mengalami hambatan belajar.....	263
5.8.12 Membaca variasi di balik rerata kelas	264
5.8.13 Analisis butir untuk pemetaan kelas	265
5.8.14 Heatmap konsepsi dan matriks peserta didik-konsep.....	266
5.8.15 Penggunaan peta konsep kelas.....	267
5.8.16 Wawancara sampel untuk konfirmasi kelas.....	269
5.8.17 Pemilihan kasus tipikal dan kasus menyimpang.....	270
5.8.18 Tindak lanjut individual	271
5.8.19 Tindak lanjut kelompok kecil	272
5.8.20 Tindak lanjut pembelajaran klasikal.....	273
5.8.21 Pengelolaan data diagnosis pada kelas besar	275
5.8.22 Komunikasi hasil tanpa mempermalukan peserta didik	276
5.8.23 Monitoring perkembangan individual dan kelas	277
5.8.24 Contoh dashboard diagnosis sederhana	278
5.8.25 Sintesis diagnosis individual dan kelas.....	280
5.8.26 Implikasi praktis bagi guru kimia	281
5.9 Diagnosis Berbasis Representasi.....	281

5.9.1 Representasi sebagai bagian dari pengetahuan kimia	281
5.9.2 Tingkat makroskopik dalam diagnosis	283
5.9.3 Tingkat submikroskopik dalam diagnosis	284
5.9.4 Tingkat simbolik dalam diagnosis.....	285
5.9.5 Representasi matematis dan kuantitatif	286
5.9.6 Representasi grafis dan diagram	287
5.9.7 Representasi verbal dan bahasa ilmiah.....	288
5.9.8 Koordinasi antarrepresentasi	290
5.9.9 Translasi dari fenomena ke model partikel	291
5.9.10 Translasi dari model partikel ke persamaan	292
5.9.11 Translasi dari persamaan ke makna kuantitatif	293
5.9.12 Konsistensi representasional sebagai bukti pemahaman	294
5.9.13 Miskonsepsi akibat pencampuran tingkat representasi	296
5.9.14 Gambar partikel sebagai tugas diagnostik	297
5.9.15 Grafik sebagai tugas diagnostik	298
5.9.16 Model molekul dan keterbatasannya.....	299
5.9.17 Animasi dan simulasi dalam diagnosis.....	301
5.9.18 Tugas multiple external representations.....	302
5.9.19 Representational competence dan representational fluency	303
5.9.20 Rubrik penilaian representasi	304
5.9.21 Validitas representasi diagnostik.....	305
5.9.22 Aksesibilitas dan beban visual	307
5.9.23 Contoh diagnosis representasional pada reaksi kimia	308
5.9.24 Contoh diagnosis representasional pada kesetimbangan	309
5.9.25 Sintesis diagnosis berbasis representasi	310
5.9.26 Implikasi praktis bagi guru kimia	311
5.10 Etika dalam Diagnosis Peserta Didik.....	312
5.10.1 Diagnosis sebagai praktik pendidikan yang berorientasi kemaslahatan	312
5.10.2 Prinsip tidak merugikan peserta didik.....	313
5.10.3 Menghindari pelabelan permanen.....	314
5.10.4 Menghindari bahasa defisit dan stigmatisasi.....	315
5.10.5 Keadilan dan kesempatan yang setara	317

5.10.6 Persetujuan, pemberitahuan, dan transparansi	318
5.10.7 Kerahasiaan hasil diagnosis	319
5.10.8 Pelindungan data pribadi peserta didik	320
5.10.9 Data anak sebagai data yang memerlukan perlindungan khusus	322
5.10.10 Minimalisasi data dan pembatasan tujuan	323
5.10.11 Keamanan penyimpanan data diagnosis	324
5.10.12 Akses terhadap data dan akuntabilitas	325
5.10.13 Etika penggunaan platform digital	326
5.10.14 Etika penggunaan kecerdasan artifisial	327
5.10.15 Bias algoritmik dan keputusan otomatis	329
5.10.16 Hak peserta didik untuk memperoleh penjelasan	330
5.10.17 Komunikasi hasil kepada orang tua atau wali	331
5.10.18 Penggunaan hasil untuk remediasi, bukan hukuman	332
5.10.19 Etika publikasi dan penelitian kelas	333
5.10.20 Anonimisasi dan pseudonimisasi data	335
5.10.21 Keadilan bagi peserta didik berkebutuhan khusus	336
5.10.22 Sensitivitas bahasa dan budaya	337
5.10.23 Konflik peran guru sebagai pengajar dan penilai	338
5.10.24 Prosedur penanganan temuan sensitif	339
5.10.25 Audit etika diagnosis	341
5.10.26 Sintesis etika diagnosis peserta didik	342
5.10.27 Implikasi praktis bagi guru kimia	343
5.11 Interpretasi Hasil Diagnosis	343
5.11.1 Interpretasi sebagai proses argumentatif	344
5.11.2 Memulai interpretasi dari tujuan diagnosis	345
5.11.3 Membaca skor total secara hati-hati	346
5.11.4 Membaca pola per konsep dan subkonsep	347
5.11.5 Membaca pola jawaban-alasan-keyakinan	348
5.11.6 Membedakan frekuensi dan tingkat keparahan	349
5.11.7 Menafsirkan konsistensi lintas butir	351
5.11.8 Menafsirkan konsistensi lintas representasi	352
5.11.9 Menafsirkan perubahan dari waktu ke waktu	353

5.11.10 Menafsirkan hasil pretest dan posttest	354
5.11.11 Menafsirkan retensi dan transfer	355
5.11.12 Menafsirkan respons tidak lengkap	356
5.11.13 Menafsirkan data wawancara	358
5.11.14 Menafsirkan peta konsep dan gambar	359
5.11.15 Menafsirkan data observasi.....	360
5.11.16 Menggabungkan data kuantitatif dan kualitatif	361
5.11.17 Menentukan ambang klasifikasi	363
5.11.18 Ketidakpastian dan interval keputusan	364
5.11.19 Menghindari overinterpretation	365
5.11.20 Menghindari causal fallacy	366
5.11.21 Menghubungkan interpretasi dengan tindak lanjut.....	367
5.11.22 Menyusun laporan diagnosis yang komunikatif	369
5.11.23 Visualisasi hasil yang tidak menyesatkan.....	370
5.11.24 Contoh interpretasi satu butir diagnostik	371
5.11.25 Contoh interpretasi satu kelas	372
5.11.26 Sintesis interpretasi hasil diagnosis.....	373
5.11.27 Implikasi praktis bagi guru kimia	374
5.12 Penyusunan Profil Konseptual Peserta Didik	375
5.12.1 Pengertian profil konseptual	375
5.12.2 Tujuan penyusunan profil konseptual	376
5.12.3 Unit konsep dan subkonsep dalam profil.....	377
5.12.4 Dimensi ketepatan jawaban	378
5.12.5 Dimensi kualitas alasan.....	380
5.12.6 Dimensi keyakinan	381
5.12.7 Dimensi konsistensi.....	382
5.12.8 Dimensi representasi	383
5.12.9 Dimensi transfer	384
5.12.10 Dimensi retensi	385
5.12.11 Dimensi sumber miskonsepsi	386
5.12.12 Dimensi konsep prasyarat.....	388
5.12.13 Kategori paham, parsial, tidak tahu, dan miskonsepsi	389

5.12.14 Profil konseptual berbasis matriks	390
5.12.15 Profil konseptual berbasis narasi	391
5.12.16 Profil konseptual berbasis peta jaringan	392
5.12.17 Profil konseptual longitudinal	394
5.12.18 Profil individual dan profil kelompok	395
5.12.19 Pemutakhiran profil setelah remediasi	396
5.12.20 Keterhubungan profil dengan diferensiasi pembelajaran	397
5.12.21 Keterhubungan profil dengan umpan balik	398
5.12.22 Keterhubungan profil dengan perencanaan unit berikutnya	400
5.12.23 Validasi profil melalui triangulasi	401
5.12.24 Reliabilitas pengodean profil	402
5.12.25 Etika penyimpanan dan penggunaan profil	403
5.12.26 Contoh profil konseptual topik asam-basa	404
5.12.27 Contoh profil konseptual topik ikatan kimia	405
5.12.28 Template operasional profil konseptual	407
5.12.29 Analisis kritis terhadap reduksi peserta didik menjadi kategori	408
5.12.30 Sintesis penyusunan profil konseptual	409
5.12.31 Implikasi praktis bagi guru kimia	410
5.13 Sintesis Bab	411
Daftar Pustaka Bab 5	412
BAB 6	417
INSTRUMEN DIAGNOSIS MISKONSEPSI	417
Pengantar Bab	417
6.1 Tes Pilihan Ganda Konvensional	419
6.1.1 Landasan Konseptual dan Kedudukan Instrumen	419
6.1.2 Perkembangan Historis dan Rasional Penggunaan	420
6.1.3 Konstruk, Klaim, dan Bukti Diagnosis	421
6.1.4 Struktur, Komponen, dan Alur Respons	422
6.1.5 Desain Stimulus, Tugas, atau Pertanyaan	423
6.1.6 Prosedur Pengembangan Secara Sistematis	424
6.1.7 Pelaksanaan dan Pengumpulan Data	425
6.1.8 Penskoran, Pengodean, dan Pengolahan Data	426

6.1.9 Interpretasi dan Penetapan Kategori	427
6.1.10 Validitas dan Kualitas Inferensi.....	428
6.1.11 Reliabilitas, Keajegan, dan Analisis Empiris	429
6.1.12 Contoh Penerapan dalam Pembelajaran Kimia	430
6.1.13 Kelebihan, Keterbatasan, dan Potensi Bias	431
6.1.14 Integrasi Digital, Etika, dan Tata Kelola Data.....	432
6.1.15 Sintesis dan Pedoman Implementasi	433
6.2 Tes Pilihan Ganda Beralasan	434
6.2.1 Landasan Konseptual dan Kedudukan Instrumen	434
6.2.2 Perkembangan Historis dan Rasional Penggunaan.....	435
6.2.3 Konstruk, Klaim, dan Bukti Diagnosis.....	436
6.2.4 Struktur, Komponen, dan Alur Respons	437
6.2.5 Desain Stimulus, Tugas, atau Pertanyaan.....	438
6.2.6 Prosedur Pengembangan Secara Sistematis	439
6.2.7 Pelaksanaan dan Pengumpulan Data	440
6.2.8 Penskoran, Pengodean, dan Pengolahan Data	441
6.2.9 Interpretasi dan Penetapan Kategori	442
6.2.10 Validitas dan Kualitas Inferensi.....	443
6.2.11 Reliabilitas, Keajegan, dan Analisis Empiris	444
6.2.12 Contoh Penerapan dalam Pembelajaran Kimia	445
6.2.13 Kelebihan, Keterbatasan, dan Potensi Bias	446
6.2.14 Integrasi Digital, Etika, dan Tata Kelola Data.....	447
6.2.15 Sintesis dan Pedoman Implementasi	448
6.3 Two-Tier Diagnostic Test	449
6.3.1 Landasan Konseptual dan Kedudukan Instrumen	449
6.3.2 Perkembangan Historis dan Rasional Penggunaan.....	450
6.3.3 Konstruk, Klaim, dan Bukti Diagnosis.....	451
6.3.4 Struktur, Komponen, dan Alur Respons	452
6.3.5 Desain Stimulus, Tugas, atau Pertanyaan.....	453
6.3.6 Prosedur Pengembangan Secara Sistematis	454
6.3.7 Pelaksanaan dan Pengumpulan Data	455
6.3.8 Penskoran, Pengodean, dan Pengolahan Data	456

6.3.9 Interpretasi dan Penetapan Kategori	457
6.3.10 Validitas dan Kualitas Inferensi.....	458
6.3.11 Reliabilitas, Keajegan, dan Analisis Empiris	459
6.3.12 Contoh Penerapan dalam Pembelajaran Kimia	460
6.3.13 Kelebihan, Keterbatasan, dan Potensi Bias	461
6.3.14 Integrasi Digital, Etika, dan Tata Kelola Data.....	462
6.3.15 Sintesis dan Pedoman Implementasi	463
6.4 Three-Tier Diagnostic Test.....	464
6.4.1 Landasan Konseptual dan Kedudukan Instrumen	464
6.4.2 Perkembangan Historis dan Rasional Penggunaan.....	465
6.4.3 Konstruk, Klaim, dan Bukti Diagnosis.....	466
6.4.4 Struktur, Komponen, dan Alur Respons	467
6.4.5 Desain Stimulus, Tugas, atau Pertanyaan.....	468
6.4.6 Prosedur Pengembangan Secara Sistematis	469
6.4.7 Pelaksanaan dan Pengumpulan Data	470
6.4.8 Penskoran, Pengodean, dan Pengolahan Data	471
6.4.9 Interpretasi dan Penetapan Kategori	472
6.4.10 Validitas dan Kualitas Inferensi.....	473
6.4.11 Reliabilitas, Keajegan, dan Analisis Empiris	474
6.4.12 Contoh Penerapan dalam Pembelajaran Kimia	475
6.4.13 Kelebihan, Keterbatasan, dan Potensi Bias	476
6.4.14 Integrasi Digital, Etika, dan Tata Kelola Data.....	477
6.4.15 Sintesis dan Pedoman Implementasi	478
6.5 Four-Tier Diagnostic Test.....	479
6.5.1 Landasan Konseptual dan Kedudukan Instrumen	479
6.5.2 Perkembangan Historis dan Rasional Penggunaan.....	480
6.5.3 Konstruk, Klaim, dan Bukti Diagnosis.....	481
6.5.4 Struktur, Komponen, dan Alur Respons	482
6.5.5 Desain Stimulus, Tugas, atau Pertanyaan.....	483
6.5.6 Prosedur Pengembangan Secara Sistematis	484
6.5.7 Pelaksanaan dan Pengumpulan Data	485
6.5.8 Penskoran, Pengodean, dan Pengolahan Data	486

6.5.9 Interpretasi dan Penetapan Kategori	487
6.5.10 Validitas dan Kualitas Inferensi.....	488
6.5.11 Reliabilitas, Keajegan, dan Analisis Empiris	489
6.5.12 Contoh Penerapan dalam Pembelajaran Kimia	490
6.5.13 Kelebihan, Keterbatasan, dan Potensi Bias	491
6.5.14 Integrasi Digital, Etika, dan Tata Kelola Data.....	492
6.5.15 Sintesis dan Pedoman Implementasi	493
6.6 Five-Tier Diagnostic Test.....	494
6.6.1 Landasan Konseptual dan Kedudukan Instrumen	494
6.6.2 Perkembangan Historis dan Rasional Penggunaan.....	495
6.6.3 Konstruk, Klaim, dan Bukti Diagnosis.....	496
6.6.4 Struktur, Komponen, dan Alur Respons	497
6.6.5 Desain Stimulus, Tugas, atau Pertanyaan.....	498
6.6.6 Prosedur Pengembangan Secara Sistematis	499
6.6.7 Pelaksanaan dan Pengumpulan Data	500
6.6.8 Penskoran, Pengodean, dan Pengolahan Data	501
6.6.9 Interpretasi dan Penetapan Kategori	502
6.6.10 Validitas dan Kualitas Inferensi.....	503
6.6.11 Reliabilitas, Keajegan, dan Analisis Empiris	504
6.6.12 Contoh Penerapan dalam Pembelajaran Kimia	505
6.6.13 Kelebihan, Keterbatasan, dan Potensi Bias	506
6.6.14 Integrasi Digital, Etika, dan Tata Kelola Data.....	507
6.6.15 Sintesis dan Pedoman Implementasi	508
6.7 Certainty of Response Index	509
6.7.1 Landasan Konseptual dan Kedudukan Instrumen	509
6.7.2 Perkembangan Historis dan Rasional Penggunaan.....	510
6.7.3 Konstruk, Klaim, dan Bukti Diagnosis.....	511
6.7.4 Struktur, Komponen, dan Alur Respons	512
6.7.5 Desain Stimulus, Tugas, atau Pertanyaan.....	513
6.7.6 Prosedur Pengembangan Secara Sistematis	514
6.7.7 Pelaksanaan dan Pengumpulan Data	515
6.7.8 Penskoran, Pengodean, dan Pengolahan Data	516

6.7.9 Interpretasi dan Penetapan Kategori	517
6.7.10 Validitas dan Kualitas Inferensi.....	518
6.7.11 Reliabilitas, Keajegan, dan Analisis Empiris	519
6.7.12 Contoh Penerapan dalam Pembelajaran Kimia	520
6.7.13 Kelebihan, Keterbatasan, dan Potensi Bias	521
6.7.14 Integrasi Digital, Etika, dan Tata Kelola Data.....	522
6.7.15 Sintesis dan Pedoman Implementasi	523
6.8 Wawancara Klinis	524
6.8.1 Landasan Konseptual dan Kedudukan Instrumen	524
6.8.2 Perkembangan Historis dan Rasional Penggunaan.....	525
6.8.3 Konstruk, Klaim, dan Bukti Diagnosis.....	526
6.8.4 Struktur, Komponen, dan Alur Respons	527
6.8.5 Desain Stimulus, Tugas, atau Pertanyaan.....	528
6.8.6 Prosedur Pengembangan Secara Sistematis	529
6.8.7 Pelaksanaan dan Pengumpulan Data	530
6.8.8 Penskoran, Pengodean, dan Pengolahan Data	531
6.8.9 Interpretasi dan Penetapan Kategori	532
6.8.10 Validitas dan Kualitas Inferensi.....	533
6.8.11 Reliabilitas, Keajegan, dan Analisis Empiris	534
6.8.12 Contoh Penerapan dalam Pembelajaran Kimia	535
6.8.13 Kelebihan, Keterbatasan, dan Potensi Bias	536
6.8.14 Integrasi Digital, Etika, dan Tata Kelola Data.....	537
6.8.15 Sintesis dan Pedoman Implementasi	538
6.9 Wawancara Berbasis Tugas	539
6.9.1 Landasan Konseptual dan Kedudukan Instrumen	539
6.9.2 Perkembangan Historis dan Rasional Penggunaan.....	540
6.9.3 Konstruk, Klaim, dan Bukti Diagnosis.....	541
6.9.4 Struktur, Komponen, dan Alur Respons	542
6.9.5 Desain Stimulus, Tugas, atau Pertanyaan.....	543
6.9.6 Prosedur Pengembangan Secara Sistematis	544
6.9.7 Pelaksanaan dan Pengumpulan Data	545
6.9.8 Penskoran, Pengodean, dan Pengolahan Data	546

6.9.9 Interpretasi dan Penetapan Kategori	547
6.9.10 Validitas dan Kualitas Inferensi.....	548
6.9.11 Reliabilitas, Keajegan, dan Analisis Empiris	549
6.9.12 Contoh Penerapan dalam Pembelajaran Kimia	550
6.9.13 Kelebihan, Keterbatasan, dan Potensi Bias	551
6.9.14 Integrasi Digital, Etika, dan Tata Kelola Data.....	552
6.9.15 Sintesis dan Pedoman Implementasi	553
6.10 Peta Konsep	554
6.10.1 Landasan Konseptual dan Kedudukan Instrumen	554
6.10.2 Perkembangan Historis dan Rasional Penggunaan.....	555
6.10.3 Konstruk, Klaim, dan Bukti Diagnosis.....	556
6.10.4 Struktur, Komponen, dan Alur Respons	557
6.10.5 Desain Stimulus, Tugas, atau Pertanyaan.....	558
6.10.6 Prosedur Pengembangan Secara Sistematis	559
6.10.7 Pelaksanaan dan Pengumpulan Data	560
6.10.8 Penskoran, Pengodean, dan Pengolahan Data	561
6.10.9 Interpretasi dan Penetapan Kategori	562
6.10.10 Validitas dan Kualitas Inferensi.....	563
6.10.11 Reliabilitas, Keajegan, dan Analisis Empiris	564
6.10.12 Contoh Penerapan dalam Pembelajaran Kimia	565
6.10.13 Kelebihan, Keterbatasan, dan Potensi Bias	566
6.10.14 Integrasi Digital, Etika, dan Tata Kelola Data.....	567
6.10.15 Sintesis dan Pedoman Implementasi	568
6.11 Open-Ended Diagnostic Test	569
6.11.1 Landasan Konseptual dan Kedudukan Instrumen	569
6.11.2 Perkembangan Historis dan Rasional Penggunaan.....	570
6.11.3 Konstruk, Klaim, dan Bukti Diagnosis.....	571
6.11.4 Struktur, Komponen, dan Alur Respons	572
6.11.5 Desain Stimulus, Tugas, atau Pertanyaan.....	573
6.11.6 Prosedur Pengembangan Secara Sistematis	574
6.11.7 Pelaksanaan dan Pengumpulan Data	575
6.11.8 Penskoran, Pengodean, dan Pengolahan Data	576

6.11.9 Interpretasi dan Penetapan Kategori	577
6.11.10 Validitas dan Kualitas Inferensi.....	578
6.11.11 Reliabilitas, Keajegan, dan Analisis Empiris	579
6.11.12 Contoh Penerapan dalam Pembelajaran Kimia	580
6.11.13 Kelebihan, Keterbatasan, dan Potensi Bias	581
6.11.14 Integrasi Digital, Etika, dan Tata Kelola Data.....	582
6.11.15 Sintesis dan Pedoman Implementasi	583
6.12 Predict–Observe–Explain	584
6.12.1 Landasan Konseptual dan Kedudukan Instrumen	584
6.12.2 Perkembangan Historis dan Rasional Penggunaan.....	585
6.12.3 Konstruk, Klaim, dan Bukti Diagnosis.....	586
6.12.4 Struktur, Komponen, dan Alur Respons	587
6.12.5 Desain Stimulus, Tugas, atau Pertanyaan.....	588
6.12.6 Prosedur Pengembangan Secara Sistematis	589
6.12.7 Pelaksanaan dan Pengumpulan Data	590
6.12.8 Penskoran, Pengodean, dan Pengolahan Data	591
6.12.9 Interpretasi dan Penetapan Kategori	592
6.12.10 Validitas dan Kualitas Inferensi.....	593
6.12.11 Reliabilitas, Keajegan, dan Analisis Empiris	594
6.12.12 Contoh Penerapan dalam Pembelajaran Kimia	595
6.12.13 Kelebihan, Keterbatasan, dan Potensi Bias	596
6.12.14 Integrasi Digital, Etika, dan Tata Kelola Data.....	597
6.12.15 Sintesis dan Pedoman Implementasi	598
6.13 Drawing Test dan Representasi Gambar	599
6.13.1 Landasan Konseptual dan Kedudukan Instrumen	599
6.13.2 Perkembangan Historis dan Rasional Penggunaan.....	600
6.13.3 Konstruk, Klaim, dan Bukti Diagnosis.....	601
6.13.4 Struktur, Komponen, dan Alur Respons	602
6.13.5 Desain Stimulus, Tugas, atau Pertanyaan.....	603
6.13.6 Prosedur Pengembangan Secara Sistematis	604
6.13.7 Pelaksanaan dan Pengumpulan Data	605
6.13.8 Penskoran, Pengodean, dan Pengolahan Data	606

6.13.9 Interpretasi dan Penetapan Kategori	607
6.13.10 Validitas dan Kualitas Inferensi.....	608
6.13.11 Reliabilitas, Keajegan, dan Analisis Empiris	609
6.13.12 Contoh Penerapan dalam Pembelajaran Kimia	610
6.13.13 Kelebihan, Keterbatasan, dan Potensi Bias	611
6.13.14 Integrasi Digital, Etika, dan Tata Kelola Data.....	612
6.13.15 Sintesis dan Pedoman Implementasi	613
6.14 Analisis Kesalahan Penyelesaian Masalah	614
6.14.1 Landasan Konseptual dan Kedudukan Instrumen	614
6.14.2 Perkembangan Historis dan Rasional Penggunaan.....	615
6.14.3 Konstruk, Klaim, dan Bukti Diagnosis.....	616
6.14.4 Struktur, Komponen, dan Alur Respons	617
6.14.5 Desain Stimulus, Tugas, atau Pertanyaan.....	618
6.14.6 Prosedur Pengembangan Secara Sistematis	619
6.14.7 Pelaksanaan dan Pengumpulan Data	620
6.14.8 Penskoran, Pengodean, dan Pengolahan Data	621
6.14.9 Interpretasi dan Penetapan Kategori	622
6.14.10 Validitas dan Kualitas Inferensi.....	623
6.14.11 Reliabilitas, Keajegan, dan Analisis Empiris	624
6.14.12 Contoh Penerapan dalam Pembelajaran Kimia	625
6.14.13 Kelebihan, Keterbatasan, dan Potensi Bias	626
6.14.14 Integrasi Digital, Etika, dan Tata Kelola Data.....	627
6.14.15 Sintesis dan Pedoman Implementasi	628
6.15 Think-Aloud Protocol.....	629
6.15.1 Landasan Konseptual dan Kedudukan Instrumen	629
6.15.2 Perkembangan Historis dan Rasional Penggunaan.....	630
6.15.3 Konstruk, Klaim, dan Bukti Diagnosis.....	631
6.15.4 Struktur, Komponen, dan Alur Respons	632
6.15.5 Desain Stimulus, Tugas, atau Pertanyaan.....	633
6.15.6 Prosedur Pengembangan Secara Sistematis	634
6.15.7 Pelaksanaan dan Pengumpulan Data	635
6.15.8 Penskoran, Pengodean, dan Pengolahan Data	636

6.15.9 Interpretasi dan Penetapan Kategori	637
6.15.10 Validitas dan Kualitas Inferensi.....	638
6.15.11 Reliabilitas, Keajegan, dan Analisis Empiris	639
6.15.12 Contoh Penerapan dalam Pembelajaran Kimia	640
6.15.13 Kelebihan, Keterbatasan, dan Potensi Bias	641
6.15.14 Integrasi Digital, Etika, dan Tata Kelola Data.....	642
6.15.15 Sintesis dan Pedoman Implementasi	643
6.16 Analisis Percakapan Kelas	644
6.16.1 Landasan Konseptual dan Kedudukan Instrumen	644
6.16.2 Perkembangan Historis dan Rasional Penggunaan.....	645
6.16.3 Konstruk, Klaim, dan Bukti Diagnosis.....	646
6.16.4 Struktur, Komponen, dan Alur Respons	647
6.16.5 Desain Stimulus, Tugas, atau Pertanyaan.....	648
6.16.6 Prosedur Pengembangan Secara Sistematis	649
6.16.7 Pelaksanaan dan Pengumpulan Data	650
6.16.8 Penskoran, Pengodean, dan Pengolahan Data	651
6.16.9 Interpretasi dan Penetapan Kategori	652
6.16.10 Validitas dan Kualitas Inferensi.....	653
6.16.11 Reliabilitas, Keajegan, dan Analisis Empiris	654
6.16.12 Contoh Penerapan dalam Pembelajaran Kimia	655
6.16.13 Kelebihan, Keterbatasan, dan Potensi Bias	656
6.16.14 Integrasi Digital, Etika, dan Tata Kelola Data.....	657
6.16.15 Sintesis dan Pedoman Implementasi	658
6.17 Penggunaan Log Aktivitas Digital	659
6.17.1 Landasan Konseptual dan Kedudukan Instrumen	659
6.17.2 Perkembangan Historis dan Rasional Penggunaan.....	660
6.17.3 Konstruk, Klaim, dan Bukti Diagnosis.....	661
6.17.4 Struktur, Komponen, dan Alur Respons	662
6.17.5 Desain Stimulus, Tugas, atau Pertanyaan.....	663
6.17.6 Prosedur Pengembangan Secara Sistematis	664
6.17.7 Pelaksanaan dan Pengumpulan Data	665
6.17.8 Penskoran, Pengodean, dan Pengolahan Data	666

6.17.9 Interpretasi dan Penetapan Kategori	667
6.17.10 Validitas dan Kualitas Inferensi.....	668
6.17.11 Reliabilitas, Keajegan, dan Analisis Empiris	669
6.17.12 Contoh Penerapan dalam Pembelajaran Kimia	670
6.17.13 Kelebihan, Keterbatasan, dan Potensi Bias	671
6.17.14 Integrasi Digital, Etika, dan Tata Kelola Data.....	672
6.17.15 Sintesis dan Pedoman Implementasi	673
6.18 Pengembangan Distraktor Berbasis Miskonsepsi	674
6.18.1 Landasan Konseptual dan Kedudukan Instrumen	674
6.18.2 Perkembangan Historis dan Rasional Penggunaan.....	675
6.18.3 Konstruk, Klaim, dan Bukti Diagnosis.....	676
6.18.4 Struktur, Komponen, dan Alur Respons	677
6.18.5 Desain Stimulus, Tugas, atau Pertanyaan.....	678
6.18.6 Prosedur Pengembangan Secara Sistematis	679
6.18.7 Pelaksanaan dan Pengumpulan Data	680
6.18.8 Penskoran, Pengodean, dan Pengolahan Data	681
6.18.9 Interpretasi dan Penetapan Kategori	682
6.18.10 Validitas dan Kualitas Inferensi.....	683
6.18.11 Reliabilitas, Keajegan, dan Analisis Empiris	684
6.18.12 Contoh Penerapan dalam Pembelajaran Kimia	685
6.18.13 Kelebihan, Keterbatasan, dan Potensi Bias	686
6.18.14 Integrasi Digital, Etika, dan Tata Kelola Data.....	687
6.18.15 Sintesis dan Pedoman Implementasi	688
6.19 Validasi Isi oleh Ahli	689
6.19.1 Landasan Konseptual dan Kedudukan Instrumen	689
6.19.2 Perkembangan Historis dan Rasional Penggunaan.....	690
6.19.3 Konstruk, Klaim, dan Bukti Diagnosis.....	691
6.19.4 Struktur, Komponen, dan Alur Respons	692
6.19.5 Desain Stimulus, Tugas, atau Pertanyaan.....	693
6.19.6 Prosedur Pengembangan Secara Sistematis	694
6.19.7 Pelaksanaan dan Pengumpulan Data	695
6.19.8 Penskoran, Pengodean, dan Pengolahan Data	696

6.19.9 Interpretasi dan Penetapan Kategori	697
6.19.10 Validitas dan Kualitas Inferensi.....	698
6.19.11 Reliabilitas, Keajegan, dan Analisis Empiris	699
6.19.12 Contoh Penerapan dalam Pembelajaran Kimia	700
6.19.13 Kelebihan, Keterbatasan, dan Potensi Bias	701
6.19.14 Integrasi Digital, Etika, dan Tata Kelola Data.....	702
6.19.15 Sintesis dan Pedoman Implementasi	703
6.20 Uji Coba dan Analisis Butir	704
6.20.1 Landasan Konseptual dan Kedudukan Instrumen	704
6.20.2 Perkembangan Historis dan Rasional Penggunaan.....	705
6.20.3 Konstruk, Klaim, dan Bukti Diagnosis.....	706
6.20.4 Struktur, Komponen, dan Alur Respons	707
6.20.5 Desain Stimulus, Tugas, atau Pertanyaan.....	708
6.20.6 Prosedur Pengembangan Secara Sistematis	709
6.20.7 Pelaksanaan dan Pengumpulan Data	710
6.20.8 Penskoran, Pengodean, dan Pengolahan Data	711
6.20.9 Interpretasi dan Penetapan Kategori	712
6.20.10 Validitas dan Kualitas Inferensi.....	713
6.20.11 Reliabilitas, Keajegan, dan Analisis Empiris	714
6.20.12 Contoh Penerapan dalam Pembelajaran Kimia	715
6.20.13 Kelebihan, Keterbatasan, dan Potensi Bias	716
6.20.14 Integrasi Digital, Etika, dan Tata Kelola Data.....	717
6.20.15 Sintesis dan Pedoman Implementasi	718
6.21 Penetapan Kategori Pemahaman Konsep	719
6.21.1 Landasan Konseptual dan Kedudukan Instrumen	719
6.21.2 Perkembangan Historis dan Rasional Penggunaan.....	720
6.21.3 Konstruk, Klaim, dan Bukti Diagnosis.....	721
6.21.4 Struktur, Komponen, dan Alur Respons	722
6.21.5 Desain Stimulus, Tugas, atau Pertanyaan.....	723
6.21.6 Prosedur Pengembangan Secara Sistematis	724
6.21.7 Pelaksanaan dan Pengumpulan Data	725
6.21.8 Penskoran, Pengodean, dan Pengolahan Data	726

6.21.9 Interpretasi dan Penetapan Kategori	727
6.21.10 Validitas dan Kualitas Inferensi.....	728
6.21.11 Reliabilitas, Keajegan, dan Analisis Empiris	729
6.21.12 Contoh Penerapan dalam Pembelajaran Kimia	730
6.21.13 Kelebihan, Keterbatasan, dan Potensi Bias	731
6.21.14 Integrasi Digital, Etika, dan Tata Kelola Data.....	732
6.21.15 Sintesis dan Pedoman Implementasi	733
6.22 Pelaporan Hasil Diagnosis	734
6.22.1 Landasan Konseptual dan Kedudukan Instrumen	734
6.22.2 Perkembangan Historis dan Rasional Penggunaan.....	735
6.22.3 Konstruksi, Klaim, dan Bukti Diagnosis.....	736
6.22.4 Struktur, Komponen, dan Alur Respons	737
6.22.5 Desain Stimulus, Tugas, atau Pertanyaan.....	738
6.22.6 Prosedur Pengembangan Secara Sistematis	739
6.22.7 Pelaksanaan dan Pengumpulan Data	740
6.22.8 Penskoran, Pengodean, dan Pengolahan Data	741
6.22.9 Interpretasi dan Penetapan Kategori	742
6.22.10 Validitas dan Kualitas Inferensi.....	743
6.22.11 Reliabilitas, Keajegan, dan Analisis Empiris	744
6.22.12 Contoh Penerapan dalam Pembelajaran Kimia	745
6.22.13 Kelebihan, Keterbatasan, dan Potensi Bias	746
6.22.14 Integrasi Digital, Etika, dan Tata Kelola Data.....	747
6.22.15 Sintesis dan Pedoman Implementasi	748
Daftar Pustaka Bab 6	749
BAB 7	755
METODOLOGI PENELITIAN MISKONSEPSI KIMIA	755
Pengantar Bab	755
7.1 Paradigma Penelitian Miskonsepsi	756
7.1.1 Hakikat Paradigma dalam Penelitian Miskonsepsi	756
7.1.2 Ontologi Konsep, Konsepsi, dan Miskonsepsi.....	757
7.1.3 Epistemologi Bukti tentang Pemahaman Peserta Didik	758
7.1.4 Postpositivisme dan Inferensi Berbasis Indikator	759

7.1.5 Konstruktivisme dalam Kajian Konsepsi Alternatif	761
7.1.6 Interpretivisme dan Makna Kontekstual	762
7.1.7 Pragmatisme sebagai Dasar Pemilihan Metode	763
7.1.8 Realisme Kritis dan Mekanisme Pembentukan Miskonsepsi	764
7.1.9 Paradigma Transformasional dan Keadilan Pendidikan	765
7.1.10 Posisi serta Peran Peneliti	766
7.1.11 Unit Analisis dan Tingkat Realitas	767
7.1.12 Bahasa Klaim: Indikasi, Pola, dan Penjelasan	768
7.1.13 Keselarasan Paradigma, Pertanyaan, dan Desain	769
7.1.14 Refleksivitas serta Audit Asumsi	770
7.1.15 Sintesis Kerangka Paradigmatik Penelitian Miskonsepsi	771
7.2 Penelitian Deskriptif	773
7.2.1 Tujuan dan Ruang Lingkup Studi Deskriptif	773
7.2.2 Desain Potong Lintang	774
7.2.3 Deskripsi Prevalensi Miskonsepsi	775
7.2.4 Penyusunan Kategori Konsepsi	776
7.2.5 Analisis Profil Individu dan Kelas	777
7.2.6 Analisis Tingkat Butir	778
7.2.7 Deskripsi Berdasarkan Tingkat Keyakinan	779
7.2.8 Perbandingan Kelompok Secara Deskriptif	780
7.2.9 Deskripsi Lintas Representasi Kimia	781
7.2.10 Deskripsi Miskonsepsi Berdasarkan Konteks	782
7.2.11 Studi Deskriptif Berulang	783
7.2.12 Penggunaan Data Sekunder dan Arsip Asesmen	784
7.2.13 Pengendalian Bias Deskriptif	785
7.2.14 Pelaporan Temuan Deskriptif	787
7.2.15 Sintesis Rancangan Penelitian Deskriptif	788
7.3 Penelitian Studi Kasus	789
7.3.1 Hakikat Kasus dan Batas Sistem	789
7.3.2 Pemilihan Kasus Tunggal atau Multikasus	790
7.3.3 Kasus Tipikal, Ekstrem, dan Kritis	791
7.3.4 Perumusan Proposisi Awal	792

7.3.5 Unit Analisis Terbenam	793
7.3.6 Kronologi Perkembangan Miskonsepsi	794
7.3.7 Wawancara Mendalam dalam Studi Kasus	795
7.3.8 Observasi Episode Pembelajaran	796
7.3.9 Analisis Artefak dan Dokumen	797
7.3.10 Penyusunan Basis Data Kasus	798
7.3.11 Triangulasi dalam Studi Kasus	799
7.3.12 Analisis Dalam Kasus	800
7.3.13 Analisis Lintas Kasus	801
7.3.14 Generalisasi Analitik dan Transferabilitas	802
7.3.15 Sintesis Protokol Studi Kasus Miskonsepsi	803
7.4 Penelitian Survei	804
7.4.1 Perbedaan Survei dan Tes Diagnostik	805
7.4.2 Perumusan Tujuan Survei	806
7.4.3 Kerangka Sampel Survei	807
7.4.4 Pemilihan Mode Pengumpulan	808
7.4.5 Desain Kuesioner dan Tes Terintegrasi	809
7.4.6 Urutan, Tampilan, dan Beban Respons	810
7.4.7 Uji Coba Kognitif Kuesioner	811
7.4.8 Strategi Kontak dan Pengingat	812
7.4.9 Pengendalian Nonrespons	813
7.4.10 Survei Daring dan CHERRIES	814
7.4.11 Pembobotan Data Survei	815
7.4.12 Analisis Kesetaraan Mode	816
7.4.13 Survei Lintas Wilayah dan Budaya	817
7.4.14 Pelaporan Survei yang Transparan	818
7.4.15 Sintesis Protokol Survei Miskonsepsi	819
7.5 Penelitian Eksperimen dan Kuasi-Eksperimen	820
7.5.1 Logika Kausal dalam Remediasi Miskonsepsi	820
7.5.2 Eksperimen Acak Individual	821
7.5.3 Eksperimen Acak Berkelompok	822
7.5.4 Desain Pretest-Posttest dengan Kelompok Kontrol	823

7.5.5 Kuasi-Eksperimen Kelompok Tidak Ekuivalen	824
7.5.6 Desain Deret Waktu Terputus	825
7.5.7 Pemilihan Intervensi dan Pembanding	826
7.5.8 Pengukuran Perubahan Konsepsi	827
7.5.9 Pengukuran Transfer dan Retensi	828
7.5.10 Fidelitas Implementasi	829
7.5.11 Ancaman Validitas Internal	830
7.5.12 Ancaman Validitas Eksternal	831
7.5.13 Analisis Ukuran Efek dan Ketidakpastian	832
7.5.14 Pelaporan CONSORT dan TREND	833
7.5.15 Sintesis Desain Eksperimental Miskonsepsi	834
7.6 Penelitian Mixed Methods	835
7.6.1 Rasional Penggunaan Mixed Methods	836
7.6.2 Desain Konvergen	837
7.6.3 Desain Eksplanatori Sekuensial	838
7.6.4 Desain Eksploratori Sekuensial	839
7.6.5 Desain Embedded dalam Eksperimen	840
7.6.6 Desain Multitahap	841
7.6.7 Sampling dalam Mixed Methods	842
7.6.8 Connecting, Building, Merging, dan Embedding	843
7.6.9 Joint Display untuk Data Miskonsepsi	844
7.6.10 Analisis Divergensi dan Kasus Negatif	845
7.6.11 Meta-Inferensi	846
7.6.12 Validitas dan Legitimasi Integrasi	847
7.6.13 Manajemen Tim Peneliti Multimetode	848
7.6.14 Pelaporan Mixed Methods	849
7.6.15 Sintesis Desain Mixed Methods Miskonsepsi	850
7.7 Penelitian Pengembangan Instrumen	851
7.7.1 Definisi Konstruk dan Tujuan Penggunaan	851
7.7.2 Analisis Domain Konsep Kimia	852
7.7.3 Penyusunan Blueprint Instrumen	853
7.7.4 Eksplorasi Miskonsepsi untuk Distraktor	854

7.7.5 Pengembangan Butir Satu dan Dua Tingkat.....	855
7.7.6 Pengembangan Butir Tiga dan Empat Tingkat	856
7.7.7 Instrumen Respons Terbuka dan Gambar	857
7.7.8 Validasi Isi oleh Pakar.....	858
7.7.9 Wawancara Kognitif dan Proses Respons	859
7.7.10 Uji Coba Awal dan Revisi.....	860
7.7.11 Analisis Teori Tes Klasik.....	861
7.7.12 Analisis Rasch dan IRT	862
7.7.13 Reliabilitas dan Ketepatan Pengukuran	863
7.7.14 Argumen Validitas dan Fairness.....	864
7.7.15 Sintesis Siklus Pengembangan Instrumen	865
7.8 Penelitian Design-Based Research.....	866
7.8.1 Karakteristik Design-Based Research	867
7.8.2 Analisis Masalah Praktik	868
7.8.3 Perumusan Teori dan Dugaan Desain.....	869
7.8.4 Co-Design dengan Guru dan Peserta Didik	870
7.8.5 Pengembangan Prototipe Awal	871
7.8.6 Siklus Implementasi dan Revisi.....	872
7.8.7 Pengumpulan Data Multisumber	873
7.8.8 Analisis Proses dan Hasil	874
7.8.9 Dokumentasi Keputusan Desain.....	875
7.8.10 Fidelitas dan Adaptasi Lokal	876
7.8.11 Pembentukan Prinsip Desain.....	877
7.8.12 Pengembangan Teori Lokal	878
7.8.13 Skalabilitas dan Transfer	879
7.8.14 Pelaporan DBR yang Transparan.....	880
7.8.15 Sintesis Protokol DBR Miskonsepsi.....	881
7.9 Penelitian Tindakan Kelas	882
7.9.1 Hakikat PTK dalam Penanganan Miskonsepsi.....	882
7.9.2 Identifikasi Masalah Konseptual Kelas	883
7.9.3 Perencanaan Tindakan.....	884
7.9.4 Penetapan Indikator Keberhasilan	885

7.9.5 Siklus Diagnosis Awal	886
7.9.6 Pelaksanaan Intervensi	887
7.9.7 Observasi Proses Perubahan Konsep	888
7.9.8 Refleksi Berbasis Bukti	889
7.9.9 Revisi Antar-Siklus	890
7.9.10 Kolaborasi Guru dan Pengamat	891
7.9.11 Penggunaan Data Kuantitatif dalam PTK	892
7.9.12 Penggunaan Data Kualitatif dalam PTK	893
7.9.13 Etika Peran Ganda Guru-Peneliti	894
7.9.14 Pelaporan dan Diseminasi PTK	895
7.9.15 Sintesis Protokol PTK Miskonsepsi	896
7.10 Systematic Literature Review	897
7.10.1 Perumusan Pertanyaan Review	898
7.10.2 Penyusunan dan Registrasi Protokol	899
7.10.3 Pemilihan Basis Data	900
7.10.4 Pengembangan Strategi Pencarian	901
7.10.5 Pencarian Literatur Abu-Abu	902
7.10.6 Manajemen Rekaman dan Duplikasi	903
7.10.7 Kriteria Inklusi dan Eksklusi	904
7.10.8 Proses Screening	905
7.10.9 Ekstraksi Data	906
7.10.10 Penilaian Kualitas dan Risiko Bias	907
7.10.11 Sintesis Naratif	908
7.10.12 Meta-Analisis Bila Memenuhi Syarat	909
7.10.13 Analisis Heterogenitas	910
7.10.14 Pelaporan PRISMA 2020	911
7.10.15 Sintesis SLR Miskonsepsi Kimia	912
7.11 Pemilihan Subjek dan Sampel	913
7.11.1 Populasi Target dan Terjangkau	913
7.11.2 Unit Analisis dan Unit Sampling	914
7.11.3 Probability Sampling	915
7.11.4 Simple Random dan Systematic Sampling	916

7.11.5 Stratified Sampling	917
7.11.6 Cluster dan Multistage Sampling	918
7.11.7 Purposive Sampling	919
7.11.8 Maximum Variation dan Criterion Sampling	920
7.11.9 Sampling Kasus Negatif dan Ekstrem	921
7.11.10 Ukuran Sampel Kuantitatif	922
7.11.11 Ukuran Sampel Kualitatif	923
7.11.12 Sampling untuk Mixed Methods	924
7.11.13 Rekrutmen dan Retensi Partisipan	925
7.11.14 Analisis Bias Seleksi	926
7.11.15 Sintesis Strategi Pemilihan Sampel	927
7.12 Teknik Pengumpulan Data	928
7.12.1 Pemilihan Teknik Berdasarkan Konstruk	928
7.12.2 Tes Pilihan Ganda Diagnostik	929
7.12.3 Tes Dua, Tiga, dan Empat Tingkat	930
7.12.4 Pertanyaan Terbuka	931
7.12.5 Wawancara Klinis dan Semi-Terstruktur	932
7.12.6 Think-Aloud Protocol	933
7.12.7 Tugas Menggambar dan Model Partikel	934
7.12.8 Peta Konsep	935
7.12.9 Observasi Kelas	936
7.12.10 Diskusi Kelompok Terarah	937
7.12.11 Dokumentasi Artefak Belajar	938
7.12.12 Data Digital dan Log Aktivitas	940
7.12.13 Prosedur Pilot Pengumpulan Data	941
7.12.14 Manajemen dan Keamanan Data	942
7.12.15 Sintesis Strategi Pengumpulan Data	943
7.13 Analisis Data Kuantitatif	944
7.13.1 Persiapan dan Pembersihan Data	944
7.13.2 Statistik Deskriptif	945
7.13.3 Estimasi Prevalensi dan Interval Kepercayaan	946
7.13.4 Analisis Tabulasi Silang	947

7.13.5 Uji Perbedaan Kelompok	948
7.13.6 Korelasi dan Regresi	949
7.13.7 Analisis Data Pretest-Posttest	950
7.13.8 Ukuran Efek	951
7.13.9 Analisis Data Berjenjang	952
7.13.10 Penanganan Data Hilang	953
7.13.11 Reliabilitas Skor	954
7.13.12 Analisis Butir Klasik	955
7.13.13 Rasch dan Item Response Theory	956
7.13.14 Kontrol Kesalahan Multipel dan Sensitivitas	957
7.13.15 Sintesis Analisis Kuantitatif Miskonsepsi	958
7.14 Analisis Data Kualitatif	959
7.14.1 Orientasi Analisis Kualitatif	959
7.14.2 Transkripsi dan Familiarisasi	960
7.14.3 Penentuan Unit Makna	961
7.14.4 Coding Awal	962
7.14.5 Pengembangan Codebook	963
7.14.6 Analisis Tematik Refleksif	964
7.14.7 Analisis Isi Kualitatif	965
7.14.8 Analisis Fenomenografis Konsepsi	966
7.14.9 Analisis Model Mental	967
7.14.10 Analisis Gambar dan Representasi	968
7.14.11 Analisis Interaksi Kelas	969
7.14.12 Kasus Negatif dan Disconfirming Evidence	970
7.14.13 Kredibilitas serta Audit Trail	971
7.14.14 Penyajian Kutipan dan Interpretasi	973
7.14.15 Sintesis Analisis Kualitatif Miskonsepsi	974
7.15 Triangulasi Data	975
7.15.1 Pengertian dan Tujuan Triangulasi	975
7.15.2 Triangulasi Sumber	976
7.15.3 Triangulasi Metode	977
7.15.4 Triangulasi Peneliti	978

7.15.5 Triangulasi Teori	979
7.15.6 Triangulasi Waktu	980
7.15.7 Penyusunan Matriks Konvergensi	981
7.15.8 Analisis Konvergensi	982
7.15.9 Analisis Komplementaritas	983
7.15.10 Analisis Divergensi	984
7.15.11 Protokol Penyelesaian Ketidaksesuaian	985
7.15.12 Triangulasi dalam Mixed Methods	986
7.15.13 Triangulasi dan Validitas Konstruk	987
7.15.14 Pelaporan Hasil Triangulasi	988
7.15.15 Sintesis Protokol Triangulasi Miskonsepsi	989
7.16 Analisis Pola Respons	990
7.16.1 Rasional Analisis Pola Respons	990
7.16.2 Matriks Individu × Butir	991
7.16.3 Analisis Distraktor	992
7.16.4 Kombinasi Jawaban dan Alasan	993
7.16.5 Integrasi Tingkat Keyakinan	994
7.16.6 Indeks Konsistensi Lintas Butir	995
7.16.7 Pola Lintas Representasi	996
7.16.8 Cluster Analysis	997
7.16.9 Latent Class Analysis	998
7.16.10 Latent Transition Analysis	999
7.16.11 Model Respons Nominal	1000
7.16.12 Analisis Waktu Respons	1001
7.16.13 Validasi Profil dengan Wawancara	1002
7.16.14 Visualisasi Pola Respons	1003
7.16.15 Sintesis Analisis Pola Respons	1004
7.17 Pemetaan Jaringan Konsep	1005
7.17.1 Hakikat Jaringan Konsep	1006
7.17.2 Peta Konsep Terbuka dan Terarah	1007
7.17.3 Pemilihan Konsep Kunci	1008
7.17.4 Proposisi dan Kata Penghubung	1009

7.17.5 Peta Ahli sebagai Pembanding	1010
7.17.6 Penskoran Struktural.....	1011
7.17.7 Penskoran Proposisional	1012
7.17.8 Reliabilitas Pengodean Peta.....	1013
7.17.9 Konversi Peta menjadi Graf	1014
7.17.10 Analisis Derajat dan Sentralitas	1015
7.17.11 Deteksi Proposisi Miskonsepsi.....	1016
7.17.12 Perbandingan Peta Sebelum-Sesudah.....	1017
7.17.13 Pemetaan Jaringan Kelas	1018
7.17.14 Visualisasi dan Interpretasi.....	1019
7.17.15 Sintesis Pemetaan Jaringan Konsep	1020
7.18 Etika Penelitian.....	1021
7.18.1 Prinsip Dasar Etika Penelitian Pendidikan.....	1021
7.18.2 Persetujuan Etik Institusional	1022
7.18.3 Informed Consent dan Assent	1023
7.18.4 Penelitian dengan Peserta Didik Minor.....	1024
7.18.5 Relasi Kuasa Guru-Peneliti	1025
7.18.6 Risiko Stigmatisasi Miskonsepsi	1026
7.18.7 Privasi dan Kerahasiaan	1027
7.18.8 Pelindungan Data Pribadi	1028
7.18.9 Rekaman Audio, Video, dan Gambar	1029
7.18.10 Hak Menolak dan Menarik Diri	1030
7.18.11 Insentif dan Kompensasi.....	1031
7.18.12 Umpan Balik Hasil Diagnosis	1032
7.18.13 Etika Berbagi Data	1033
7.18.14 Etika Publikasi dan Kepengarangan.....	1034
7.18.15 Sintesis Protokol Etika Penelitian Miskonsepsi	1035
7.19 Penyajian Temuan Miskonsepsi	1036
7.19.1 Prinsip Penyajian yang Bertanggung Jawab	1037
7.19.2 Penyajian Tabel Prevalensi.....	1038
7.19.3 Grafik dan Visualisasi	1039
7.19.4 Penyajian Kategori Konsepsi	1040

7.19.5 Penyajian Data Keyakinan.....	1041
7.19.6 Penyajian Pola Respons	1042
7.19.7 Penyajian Kutipan Kualitatif	1043
7.19.8 Penyajian Gambar dan Peta Konsep.....	1044
7.19.9 Joint Display Mixed Methods	1045
7.19.10 Penyajian Hasil Eksperimen	1046
7.19.11 Penyajian SLR dan Meta-Analisis.....	1047
7.19.12 Bahasa Ketidakpastian	1048
7.19.13 Bahasa Nonstigmatisasi	1049
7.19.14 Pelaporan Sesuai Pedoman	1050
7.19.15 Sintesis Struktur Pelaporan Temuan	1051
7.20 Keterbatasan Penelitian Miskonsepsi	1052
7.20.1 Perbedaan Keterbatasan dan Delimitasi	1052
7.20.2 Keterbatasan Konstruk.....	1053
7.20.3 Keterbatasan Instrumen	1054
7.20.4 Keterbatasan Sampel	1055
7.20.5 Keterbatasan Desain Deskriptif	1056
7.20.6 Keterbatasan Eksperimen dan Kuasi-Eksperimen	1057
7.20.7 Keterbatasan Data Kualitatif.....	1058
7.20.8 Keterbatasan Mixed Methods	1059
7.20.9 Keterbatasan SLR.....	1060
7.20.10 Keterbatasan Analisis Statistik.....	1061
7.20.11 Keterbatasan Interpretasi Keyakinan	1062
7.20.12 Keterbatasan Lintas Budaya dan Bahasa.....	1063
7.20.13 Keterbatasan Ekologis dan Implementasi	1064
7.20.14 Strategi Mitigasi dan Analisis Sensitivitas	1065
7.20.15 Sintesis Pelaporan Keterbatasan	1066
7.21 Sintesis Metodologi Penelitian Miskonsepsi Kimia	1067
Daftar Pustaka Bab 7	1068
GLOSARIUM	1074

BAGIAN II

DIAGNOSIS DAN PEMETAAN MISKONSEPSI KIMIA

BAB 5

PRINSIP DIAGNOSIS MISKONSEPSI

Pengantar Bab

Diagnosis miskonsepsi merupakan salah satu tahap paling menentukan dalam pembelajaran kimia yang responsif terhadap cara berpikir peserta didik. Guru tidak cukup mengetahui bahwa suatu jawaban salah; guru perlu memahami hubungan konsep, alasan, representasi, dan tingkat keyakinan yang menghasilkan jawaban tersebut. Perbedaan ini penting karena tindakan pembelajaran yang sesuai untuk ketidaktahuan, kesalahan hitung, pengetahuan parsial, dan miskonsepsi tidaklah sama.

Kimia memiliki karakter pengetahuan yang membuat diagnosis menjadi kompleks. Fenomena makroskopik harus dijelaskan melalui model submikroskopik, dinyatakan dengan simbol, dan sering diolah secara matematis. Peserta didik dapat berhasil pada satu tingkat representasi tetapi gagal pada tingkat lain. Mereka dapat menyetarakan persamaan tanpa memahami perubahan partikel, menghitung pH tanpa membedakan kekuatan dan konsentrasi asam, atau menghafal asas Le Chatelier tanpa memahami kuosien reaksi.

Asesmen diagnostik karena itu tidak boleh disamakan dengan tes awal yang sekadar menghasilkan nilai. Diagnosis adalah proses sistematis untuk mengumpulkan bukti, menilai kualitas bukti, mempertimbangkan penjelasan alternatif, dan menyusun inferensi yang dapat digunakan untuk memperbaiki pembelajaran. Standar pengukuran modern menegaskan bahwa validitas berkaitan dengan interpretasi dan penggunaan hasil, sedangkan reliabilitas berkaitan dengan konsistensi bukti serta keputusan yang dibuat dari bukti tersebut (AERA et al., 2014; Kane, 2013).

Dalam praktik kelas, diagnosis dapat berlangsung sebelum, selama, dan setelah pembelajaran. Sebelum pembelajaran, diagnosis memetakan kesiapan dan konsep prasyarat. Selama pembelajaran, diagnosis membantu guru menyesuaikan penjelasan, pertanyaan, representasi, dan aktivitas. Setelah pembelajaran, diagnosis menilai perubahan konseptual, transfer, dan retensi. Dengan demikian, diagnosis merupakan bagian dari siklus pembelajaran, bukan kegiatan tambahan yang terpisah.

Bab ini membahas dua belas prinsip pokok. Pembahasan dimulai dari hakikat dan tujuan asesmen diagnostik, kemudian menguraikan waktu pelaksanaan diagnosis, perbedaan kesalahan dan miskonsepsi, serta hubungan jawaban, alasan, dan keyakinan. Bagian berikutnya mengkaji validitas dan reliabilitas, perbedaan diagnosis individual dan kelas, diagnosis berbasis representasi, etika, interpretasi hasil, serta penyusunan profil konseptual peserta didik.

Seluruh pembahasan menempatkan diagnosis sebagai sarana perbaikan. Hasil diagnosis tidak boleh digunakan untuk memberi cap permanen, mempermalukan peserta didik, atau membatasi kesempatan belajar. Data harus dikelola secara proporsional, aman, dan transparan. Kesimpulan harus selalu menyebutkan dasar bukti dan tingkat ketidakpastiannya, serta membuka kemungkinan revisi ketika bukti baru tersedia.

Pendekatan tersebut sejalan dengan panduan pembelajaran dan asesmen yang menempatkan asesmen dalam satu siklus dengan perencanaan dan perbaikan pembelajaran. Diagnosis yang baik akhirnya tidak diukur dari banyaknya kategori yang dihasilkan, melainkan dari ketepatan tindakan yang mengikuti hasilnya dan dari peningkatan kualitas pemahaman peserta didik (Anggraena et al., 2025; Black & Wiliam, 2009).

5.1 Hakikat Asesmen Diagnostik

Subbab ini mendefinisikan fungsi, objek, dan cara kerja asesmen diagnostik sebagai proses pengumpulan bukti tentang struktur pemahaman. Uraian disusun dari prinsip teoretis menuju prosedur kelas, contoh kimia, risiko interpretasi, dan tindak lanjut. Setiap bagian menekankan bahwa kesimpulan diagnostik harus didasarkan pada pola bukti, bukan pada satu respons yang terpisah.

5.1.1 Asesmen diagnostik sebagai proses pengumpulan bukti

Pembahasan tentang asesmen diagnostik sebagai proses pengumpulan bukti berangkat dari prinsip bahwa data asesmen selalu merupakan jejak tidak langsung dari proses berpikir. Satu pilihan jawaban hanya menunjukkan produk akhir, sedangkan struktur pengetahuan yang melahirkannya dapat berbeda. Diagnosis yang kuat karena itu menggabungkan koherensi alasan, konsistensi lintas representasi, dan bukti kontekstual sehingga keputusan pedagogis tidak bergantung pada satu indikator saja (Taber, 2002a; Black & Wiliam, 2009).

Secara operasional, guru dapat menggunakan exit ticket untuk menelusuri koherensi alasan. Data tersebut perlu dibaca bersama konsistensi lintas representasi, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mencampur kesalahan matematika dengan kesalahan konsep. Untuk mengurangnya, guru dapat menggunakan pertanyaan tindak lanjut dan mengulang diagnosis setelah remediasi. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai asesmen diagnostik sebagai proses pengumpulan bukti, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan asesmen diagnostik sebagai proses pengumpulan bukti melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan asesmen diagnostik sebagai proses pengumpulan bukti;

- Bandingkan bukti asesmen diagnostik sebagai proses pengumpulan bukti pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek asesmen diagnostik sebagai proses pengumpulan bukti dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian asesmen diagnostik sebagai proses pengumpulan bukti, topik stoikiometri dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa koefisien reaksi diperlakukan sebagai perbandingan massa langsung. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta translasi persamaan dari partikel ke mol dan massa. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa koefisien terutama menyatakan rasio jumlah entitas atau mol. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mendefinisikan fungsi, objek, dan cara kerja asesmen diagnostik sebagai proses pengumpulan bukti tentang struktur pemahaman, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, asesmen diagnostik sebagai proses pengumpulan bukti harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.1.2 Perbedaan pengukuran, penilaian, asesmen, dan evaluasi

Dalam pembelajaran kimia, perbedaan pengukuran, penilaian, asesmen, dan evaluasi memiliki arti strategis karena banyak jawaban dapat diperoleh melalui hafalan prosedur atau pengenalan pola. Peserta didik mungkin tampak berhasil tetapi tidak mampu menerjemahkan makna simbol, model, atau besaran yang digunakan. Prinsip diagnosis menuntut bukti tentang cara konsep dipakai, bukan hanya apakah hasil akhirnya sesuai kunci (Black & Wiliam, 2009; Anggraena et al., 2025).

Secara operasional, guru dapat menggunakan tes pilihan ganda dua tingkat untuk menelusuri tingkat keyakinan. Data tersebut perlu dibaca bersama kemampuan transfer, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menggunakan ambang kategori tanpa bukti. Untuk mengurangnya, guru dapat menyajikan tugas paralel dan mencatat derajat keyakinan. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai perbedaan pengukuran, penilaian, asesmen, dan evaluasi, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;

- Pilih tugas yang merepresentasikan perbedaan pengukuran, penilaian, asesmen, dan evaluasi melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan perbedaan pengukuran, penilaian, asesmen, dan evaluasi;
- Bandingkan bukti perbedaan pengukuran, penilaian, asesmen, dan evaluasi pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek perbedaan pengukuran, penilaian, asesmen, dan evaluasi dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian perbedaan pengukuran, penilaian, asesmen, dan evaluasi, topik asam-basa dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa asam kuat dianggap selalu berkonsentrasi tinggi. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta pemisahan konsep derajat ionisasi, K_a , konsentrasi, dan pH. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa kekuatan asam dan konsentrasi adalah dimensi berbeda. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mendefinisikan fungsi, objek, dan cara kerja asesmen diagnostik sebagai proses pengumpulan bukti tentang struktur pemahaman, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, perbedaan pengukuran, penilaian, asesmen, dan evaluasi harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.1.3 Asesmen diagnostik sebagai proses inferensi

Aspek asesmen diagnostik sebagai proses inferensi menegaskan bahwa diagnosis adalah kegiatan ilmiah berskala kelas: guru merumuskan pertanyaan, mengumpulkan bukti, mempertimbangkan penjelasan alternatif, lalu menentukan kesimpulan yang paling masuk akal. Inferensi harus proporsional terhadap kualitas data. Semakin spesifik label yang diberikan, semakin kuat pula bukti yang diperlukan (Anggraena et al., 2025; Treagust, 1988).

Secara operasional, guru dapat menggunakan alasan terbuka untuk menelusuri konsistensi lintas butir. Data tersebut perlu dibaca bersama stabilitas setelah jeda waktu, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah memperlakukan konsepsi sebagai sifat permanen. Untuk mengurangnya, guru dapat memeriksa konsep prasyarat dan menelusuri langkah penyelesaian. Prosedur tersebut

membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai asesmen diagnostik sebagai proses inferensi, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan asesmen diagnostik sebagai proses inferensi melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan asesmen diagnostik sebagai proses inferensi;
- Bandingkan bukti asesmen diagnostik sebagai proses inferensi pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek asesmen diagnostik sebagai proses inferensi dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian asesmen diagnostik sebagai proses inferensi, topik kinetika dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa reaksi eksoterm dianggap selalu cepat. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta perbandingan profil energi dan faktor kinetik. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa termodinamika dan kinetika menjawab pertanyaan berbeda. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mendefinisikan fungsi, objek, dan cara kerja asesmen diagnostik sebagai proses pengumpulan bukti tentang struktur pemahaman, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, asesmen diagnostik sebagai proses inferensi harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.1.4 Kedudukan diagnosis dalam siklus pembelajaran kimia

Penerapan kedudukan diagnosis dalam siklus pembelajaran kimia harus menghubungkan teori pengukuran dengan pengetahuan pedagogis kimia. Instrumen yang teknis tetapi tidak memahami struktur konsep akan menghasilkan kategori dangkal; sebaliknya, pengetahuan konten tanpa prosedur pengumpulan bukti dapat menghasilkan kesimpulan intuitif yang sulit diuji. Keduanya perlu dipadukan dalam satu alur yang transparan (Treagust, 1988; Taber, 2002a).

Secara operasional, guru dapat menggunakan peta konsep untuk menelusuri konsistensi lintas representasi. Data tersebut perlu dibaca bersama urutan langkah penyelesaian, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mengabaikan ketidakpastian pengukuran. Untuk menguranginya, guru dapat meminta translasi antarrepresentasi dan menggunakan triangulasi bukti. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai kedudukan diagnosis dalam siklus pembelajaran kimia, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan kedudukan diagnosis dalam siklus pembelajaran kimia melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan kedudukan diagnosis dalam siklus pembelajaran kimia;
- Bandingkan bukti kedudukan diagnosis dalam siklus pembelajaran kimia pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek kedudukan diagnosis dalam siklus pembelajaran kimia dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian kedudukan diagnosis dalam siklus pembelajaran kimia, topik elektrokimia dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa elektron dianggap bergerak melalui jembatan garam. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta penelusuran jalur elektron dan ion pada sel galvanik. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa elektron mengalir pada rangkaian luar, sedangkan ion menjaga kenetralan muatan. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mendefinisikan fungsi, objek, dan cara kerja asesmen diagnostik sebagai proses pengumpulan bukti tentang struktur pemahaman, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Kerangka operasional

Aspek	Bukti yang dicari	Risiko interpretasi	Tindak lanjut
Target konsep	Penjelasan tentang elektrokimia	Target terlalu luas	Batasi subkonsep dan prasyarat
Respons	Konsistensi lintas representasi	Mengabaikan ketidakpastian pengukuran	Meminta translasi antarrepresentasi
Penalaran	Urutan langkah penyelesaian	Alasan diperlakukan sekadar tambahan	Menggunakan triangulasi bukti

Aspek	Bukti yang dicari	Risiko interpretasi	Tindak lanjut
Konfirmasi	Peta konsep atau tugas isomorfik	Kesimpulan dibuat dari satu konteks	Gunakan bukti kedua dan rediagnosis

Implikasinya, kedudukan diagnosis dalam siklus pembelajaran kimia harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.1.5 Diagnosis sebagai upaya membuat pemikiran peserta didik terlihat

Diagnosis sebagai upaya membuat pemikiran peserta didik terlihat tidak berdiri sebagai prosedur administratif. Ia menjadi bagian dari keputusan tentang apa yang perlu dijelaskan ulang, representasi apa yang perlu dibandingkan, dan siapa yang memerlukan dukungan berbeda. Oleh sebab itu, bukti diagnosis harus langsung terhubung dengan tindakan pembelajaran, bukan berhenti pada daftar kesalahan (Taber, 2002a; Black & Wiliam, 2009).

Secara operasional, guru dapat menggunakan analisis respons tertulis untuk menelusuri kemampuan transfer. Data tersebut perlu dibaca bersama kualitas model mental yang dieksternalisasi, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menyimpulkan miskonsepsi hanya dari satu jawaban. Untuk mengurangnya, guru dapat mengulang diagnosis setelah remediasi dan menyusun tindak lanjut berbeda untuk kategori respons yang berbeda. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai diagnosis sebagai upaya membuat pemikiran peserta didik terlihat, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan diagnosis sebagai upaya membuat pemikiran peserta didik terlihat melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan diagnosis sebagai upaya membuat pemikiran peserta didik terlihat;
- Bandingkan bukti diagnosis sebagai upaya membuat pemikiran peserta didik terlihat pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek diagnosis sebagai upaya membuat pemikiran peserta didik terlihat dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian diagnosis sebagai upaya membuat pemikiran peserta didik terlihat, topik praktikum titrasi dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa titik akhir dianggap selalu identik dengan titik

ekuivalen. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta analisis indikator, kurva titrasi, dan ketidakpastian. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa titik akhir adalah observasi operasional yang mendekati titik ekuivalen. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mendefinisikan fungsi, objek, dan cara kerja asesmen diagnostik sebagai proses pengumpulan bukti tentang struktur pemahaman, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, diagnosis sebagai upaya membuat pemikiran peserta didik terlihat harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.1.6 Objek diagnosis: konsep, relasi, model, dan prosedur

Objek diagnosis: konsep, relasi, model, dan prosedur perlu ditempatkan dalam kerangka diagnosis yang mendefinisikan fungsi, objek, dan cara kerja asesmen diagnostik sebagai proses pengumpulan bukti tentang struktur pemahaman. Fokusnya bukan sekadar menghasilkan skor, melainkan memperoleh bukti yang cukup untuk menjelaskan bagaimana peserta didik membangun, menghubungkan, dan menggunakan konsep. Karena konsepsi tidak dapat diamati secara langsung, guru harus menarik inferensi dari respons verbal, tulisan, gambar, tindakan, dan perubahan strategi yang tampak selama penyelesaian tugas (Black & Wiliam, 2009; Anggraena et al., 2025).

Secara operasional, guru dapat menggunakan diskusi argumentatif untuk menelusuri stabilitas setelah jeda waktu. Data tersebut perlu dibaca bersama respons terhadap probing, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mencampur kesalahan matematika dengan kesalahan konsep. Untuk mengurangnya, guru dapat mencatat derajat keyakinan dan membandingkan jawaban dengan alasan. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai objek diagnosis: konsep, relasi, model, dan prosedur, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan objek diagnosis: konsep, relasi, model, dan prosedur melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan objek diagnosis: konsep, relasi, model, dan prosedur;

- Bandingkan bukti objek diagnosis: konsep, relasi, model, dan prosedur pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek objek diagnosis: konsep, relasi, model, dan prosedur dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian objek diagnosis: konsep, relasi, model, dan prosedur, topik sistem periodik dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa jari-jari atom dianggap selalu bertambah bersama nomor atom. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta prediksi tren disertai penjelasan muatan inti efektif dan kulit elektron. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa tren periodik perlu dijelaskan melalui struktur dan interaksi, bukan hafalan arah panah. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mendefinisikan fungsi, objek, dan cara kerja asesmen diagnostik sebagai proses pengumpulan bukti tentang struktur pemahaman, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, objek diagnosis: konsep, relasi, model, dan prosedur harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.1.7 Diagnosis terhadap pengetahuan awal dan prasyarat

Pembahasan tentang diagnosis terhadap pengetahuan awal dan prasyarat berangkat dari prinsip bahwa data asesmen selalu merupakan jejak tidak langsung dari proses berpikir. Satu pilihan jawaban hanya menunjukkan produk akhir, sedangkan struktur pengetahuan yang melahirkannya dapat berbeda. Diagnosis yang kuat karena itu menggabungkan urutan langkah penyelesaian, ketepatan pilihan, dan bukti kontekstual sehingga keputusan pedagogis tidak bergantung pada satu indikator saja (Anggraena et al., 2025; Treagust, 1988).

Secara operasional, guru dapat menggunakan analisis grafik untuk menelusuri urutan langkah penyelesaian. Data tersebut perlu dibaca bersama ketepatan pilihan, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menggunakan ambang kategori tanpa bukti. Untuk menguranginya, guru dapat menelusuri langkah penyelesaian dan menggunakan pertanyaan tindak lanjut. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai diagnosis terhadap pengetahuan awal dan prasyarat, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan diagnosis terhadap pengetahuan awal dan prasyarat melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan diagnosis terhadap pengetahuan awal dan prasyarat;
- Bandingkan bukti diagnosis terhadap pengetahuan awal dan prasyarat pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek diagnosis terhadap pengetahuan awal dan prasyarat dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian diagnosis terhadap pengetahuan awal dan prasyarat, topik stoikiometri dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa koefisien reaksi diperlakukan sebagai perbandingan massa langsung. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta translasi persamaan dari partikel ke mol dan massa. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa koefisien terutama menyatakan rasio jumlah entitas atau mol. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mendefinisikan fungsi, objek, dan cara kerja asesmen diagnostik sebagai proses pengumpulan bukti tentang struktur pemahaman, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, diagnosis terhadap pengetahuan awal dan prasyarat harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.1.8 Hubungan asesmen diagnostik dan asesmen formatif

Dalam pembelajaran kimia, hubungan asesmen diagnostik dan asesmen formatif memiliki arti strategis karena banyak jawaban dapat diperoleh melalui hafalan prosedur atau pengenalan pola. Peserta didik mungkin tampak berhasil tetapi tidak mampu menerjemahkan makna simbol, model, atau besaran yang digunakan. Prinsip diagnosis menuntut bukti tentang cara konsep dipakai, bukan hanya apakah hasil akhirnya sesuai kunci (Treagust, 1988; Taber, 2002a).

Secara operasional, guru dapat menggunakan tes empat tingkat untuk menelusuri kualitas model mental yang dieksternalisasi. Data tersebut perlu dibaca bersama koherensi alasan, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama

pada tahap ini adalah memperlakukan konsepsi sebagai sifat permanen. Untuk menguranginya, guru dapat menggunakan triangulasi bukti dan menyajikan tugas paralel. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai hubungan asesmen diagnostik dan asesmen formatif, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan hubungan asesmen diagnostik dan asesmen formatif melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan hubungan asesmen diagnostik dan asesmen formatif;
- Bandingkan bukti hubungan asesmen diagnostik dan asesmen formatif pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek hubungan asesmen diagnostik dan asesmen formatif dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian hubungan asesmen diagnostik dan asesmen formatif, topik asam-basa dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa asam kuat dianggap selalu berkonsentrasi tinggi. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta pemisahan konsep derajat ionisasi, K_a , konsentrasi, dan pH. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa kekuatan asam dan konsentrasi adalah dimensi berbeda. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mendefinisikan fungsi, objek, dan cara kerja asesmen diagnostik sebagai proses pengumpulan bukti tentang struktur pemahaman, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Kerangka operasional

Aspek	Bukti yang dicari	Risiko interpretasi	Tindak lanjut
Target konsep	Penjelasan tentang asam-basa	Target terlalu luas	Batasi subkonsep dan prasyarat
Respons	Kualitas model mental yang dieksternalisasi	Memperlakukan konsepsi sebagai sifat permanen	Menggunakan triangulasi bukti
Penalaran	Koherensi alasan	Alasan diperlakukan sekadar tambahan	Menyajikan tugas paralel
Konfirmasi	Tes empat tingkat atau tugas isomorfik	Kesimpulan dibuat dari satu konteks	Gunakan bukti kedua dan rediagnosis

Implikasinya, hubungan asesmen diagnostik dan asesmen formatif harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.1.9 Perbedaan asesmen diagnostik dan asesmen sumatif

Aspek perbedaan asesmen diagnostik dan asesmen sumatif menegaskan bahwa diagnosis adalah kegiatan ilmiah berskala kelas: guru merumuskan pertanyaan, mengumpulkan bukti, mempertimbangkan penjelasan alternatif, lalu menentukan kesimpulan yang paling masuk akal. Inferensi harus proporsional terhadap kualitas data. Semakin spesifik label yang diberikan, semakin kuat pula bukti yang diperlukan (Taber, 2002a; Black & Wiliam, 2009).

Secara operasional, guru dapat menggunakan think-aloud untuk menelusuri respons terhadap probing. Data tersebut perlu dibaca bersama tingkat keyakinan, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mengabaikan ketidakpastian pengukuran. Untuk mengurangnya, guru dapat menyusun tindak lanjut berbeda untuk kategori respons yang berbeda dan memeriksa konsep prasyarat. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai perbedaan asesmen diagnostik dan asesmen sumatif, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan perbedaan asesmen diagnostik dan asesmen sumatif melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan perbedaan asesmen diagnostik dan asesmen sumatif;
- Bandingkan bukti perbedaan asesmen diagnostik dan asesmen sumatif pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek perbedaan asesmen diagnostik dan asesmen sumatif dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian perbedaan asesmen diagnostik dan asesmen sumatif, topik kinetika dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa reaksi eksoterm dianggap selalu cepat. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta perbandingan profil energi dan faktor kinetik. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa termodinamika dan kinetika menjawab pertanyaan berbeda. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mendefinisikan fungsi, objek, dan cara kerja asesmen diagnostik sebagai proses pengumpulan bukti tentang struktur pemahaman, bukan diperlakukan sebagai

skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, perbedaan asesmen diagnostik dan asesmen sumatif harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.1.10 Spektrum metode diagnosis miskonsepsi

Penerapan spektrum metode diagnosis miskonsepsi harus menghubungkan teori pengukuran dengan pengetahuan pedagogis kimia. Instrumen yang teknis tetapi tidak memahami struktur konsep akan menghasilkan kategori dangkal; sebaliknya, pengetahuan konten tanpa prosedur pengumpulan bukti dapat menghasilkan kesimpulan intuitif yang sulit diuji. Keduanya perlu dipadukan dalam satu alur yang transparan (Black & Wiliam, 2009; Anggraena et al., 2025).

Secara operasional, guru dapat menggunakan predict-observe-explain untuk menelusuri ketepatan pilihan. Data tersebut perlu dibaca bersama konsistensi lintas butir, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menyimpulkan miskonsepsi hanya dari satu jawaban. Untuk menguranginya, guru dapat membandingkan jawaban dengan alasan dan meminta translasi antarrepresentasi. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai spektrum metode diagnosis miskonsepsi, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan spektrum metode diagnosis miskonsepsi melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan spektrum metode diagnosis miskonsepsi;
- Bandingkan bukti spektrum metode diagnosis miskonsepsi pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek spektrum metode diagnosis miskonsepsi dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian spektrum metode diagnosis miskonsepsi, topik elektrokimia dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa elektron dianggap bergerak melalui jembatan garam. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta penelusuran jalur elektron dan ion pada sel galvanik. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa

elektron mengalir pada rangkaian luar, sedangkan ion menjaga kenetralan muatan. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mendefinisikan fungsi, objek, dan cara kerja asesmen diagnostik sebagai proses pengumpulan bukti tentang struktur pemahaman, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, spektrum metode diagnosis miskonsepsi harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.1.11 Tes tertulis sebagai alat diagnosis awal

Tes tertulis sebagai alat diagnosis awal tidak berdiri sebagai prosedur administratif. Ia menjadi bagian dari keputusan tentang apa yang perlu dijelaskan ulang, representasi apa yang perlu dibandingkan, dan siapa yang memerlukan dukungan berbeda. Oleh sebab itu, bukti diagnosis harus langsung terhubung dengan tindakan pembelajaran, bukan berhenti pada daftar kesalahan (Anggraena et al., 2025; Treagust, 1988).

Secara operasional, guru dapat menggunakan tugas isomorfik untuk menelusuri koherensi alasan. Data tersebut perlu dibaca bersama konsistensi lintas representasi, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mencampur kesalahan matematika dengan kesalahan konsep. Untuk mengurangnya, guru dapat menggunakan pertanyaan tindak lanjut dan mengulang diagnosis setelah remediasi. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai tes tertulis sebagai alat diagnosis awal, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan tes tertulis sebagai alat diagnosis awal melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan tes tertulis sebagai alat diagnosis awal;
- Bandingkan bukti tes tertulis sebagai alat diagnosis awal pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek tes tertulis sebagai alat diagnosis awal dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian tes tertulis sebagai alat diagnosis awal, topik praktikum titrasi dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa titik akhir dianggap selalu identik dengan titik ekuivalen. Diagnosis tidak cukup

dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta analisis indikator, kurva titrasi, dan ketidakpastian. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa titik akhir adalah observasi operasional yang mendekati titik ekuivalen. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mendefinisikan fungsi, objek, dan cara kerja asesmen diagnostik sebagai proses pengumpulan bukti tentang struktur pemahaman, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, tes tertulis sebagai alat diagnosis awal harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.1.12 Wawancara diagnostik dan probing

Wawancara diagnostik dan probing perlu ditempatkan dalam kerangka diagnosis yang mendefinisikan fungsi, objek, dan cara kerja asesmen diagnostik sebagai proses pengumpulan bukti tentang struktur pemahaman. Fokusnya bukan sekadar menghasilkan skor, melainkan memperoleh bukti yang cukup untuk menjelaskan bagaimana peserta didik membangun, menghubungkan, dan menggunakan konsep. Karena konsepsi tidak dapat diamati secara langsung, guru harus menarik inferensi dari respons verbal, tulisan, gambar, tindakan, dan perubahan strategi yang tampak selama penyelesaian tugas (Treagust, 1988; Taber, 2002a).

Secara operasional, guru dapat menggunakan kuis formatif digital untuk menelusuri tingkat keyakinan. Data tersebut perlu dibaca bersama kemampuan transfer, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menggunakan ambang kategori tanpa bukti. Untuk mengurangnya, guru dapat menyajikan tugas paralel dan mencatat derajat keyakinan. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai wawancara diagnostik dan probing, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan wawancara diagnostik dan probing melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan wawancara diagnostik dan probing;
- Bandingkan bukti wawancara diagnostik dan probing pada lebih dari satu konteks, format tugas, atau bentuk representasi;

- Hubungkan kategori hasil pada aspek wawancara diagnostik dan probing dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian wawancara diagnostik dan probing, topik sistem periodik dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa jari-jari atom dianggap selalu bertambah bersama nomor atom. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta prediksi tren disertai penjelasan muatan inti efektif dan kulit elektron. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa tren periodik perlu dijelaskan melalui struktur dan interaksi, bukan hafalan arah panah. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mendefinisikan fungsi, objek, dan cara kerja asesmen diagnostik sebagai proses pengumpulan bukti tentang struktur pemahaman, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Kerangka operasional

Aspek	Bukti yang dicari	Risiko interpretasi	Tindak lanjut
Target konsep	Penjelasan tentang sistem periodik	Target terlalu luas	Batasi subkonsep dan prasyarat
Respons	Tingkat keyakinan	Menggunakan ambang kategori tanpa bukti	Menyajikan tugas paralel
Penalaran	Kemampuan transfer	Alasan diperlakukan sekadar tambahan	Mencatat derajat keyakinan
Konfirmasi	Kuis formatif digital atau tugas isomorfik	Kesimpulan dibuat dari satu konteks	Gunakan bukti kedua dan rediagnosis

Implikasinya, wawancara diagnostik dan probing harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.1.13 Observasi penalaran dan tindakan eksperimen

Pembahasan tentang observasi penalaran dan tindakan eksperimen berangkat dari prinsip bahwa data asesmen selalu merupakan jejak tidak langsung dari proses berpikir. Satu pilihan jawaban hanya menunjukkan produk akhir, sedangkan struktur pengetahuan yang melahirkannya dapat berbeda. Diagnosis yang kuat karena itu menggabungkan konsistensi

lintas butir, stabilitas setelah jeda waktu, dan bukti kontekstual sehingga keputusan pedagogis tidak bergantung pada satu indikator saja (Taber, 2002a; Black & Wiliam, 2009).

Secara operasional, guru dapat menggunakan tes tiga tingkat untuk menelusuri konsistensi lintas butir. Data tersebut perlu dibaca bersama stabilitas setelah jeda waktu, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah memperlakukan konsepsi sebagai sifat permanen. Untuk menguranginya, guru dapat memeriksa konsep prasyarat dan menelusuri langkah penyelesaian. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai observasi penalaran dan tindakan eksperimen, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan observasi penalaran dan tindakan eksperimen melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan observasi penalaran dan tindakan eksperimen;
- Bandingkan bukti observasi penalaran dan tindakan eksperimen pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek observasi penalaran dan tindakan eksperimen dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian observasi penalaran dan tindakan eksperimen, topik stoikiometri dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa koefisien reaksi diperlakukan sebagai perbandingan massa langsung. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta translasi persamaan dari partikel ke mol dan massa. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa koefisien terutama menyatakan rasio jumlah entitas atau mol. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mendefinisikan fungsi, objek, dan cara kerja asesmen diagnostik sebagai proses pengumpulan bukti tentang struktur pemahaman, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, observasi penalaran dan tindakan eksperimen harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.1.14 Peta konsep sebagai representasi struktur pengetahuan

Dalam pembelajaran kimia, peta konsep sebagai representasi struktur pengetahuan memiliki arti strategis karena banyak jawaban dapat diperoleh melalui hafalan prosedur atau pengenalan pola. Peserta didik mungkin tampak berhasil tetapi tidak mampu menerjemahkan makna simbol, model, atau besaran yang digunakan. Prinsip diagnosis menuntut bukti tentang cara konsep dipakai, bukan hanya apakah hasil akhirnya sesuai kunci (Black & Wiliam, 2009; Anggraena et al., 2025).

Secara operasional, guru dapat menggunakan wawancara semi-terstruktur untuk menelusuri konsistensi lintas representasi. Data tersebut perlu dibaca bersama urutan langkah penyelesaian, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mengabaikan ketidakpastian pengukuran. Untuk mengurangnya, guru dapat meminta translasi antarrepresentasi dan menggunakan triangulasi bukti. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai peta konsep sebagai representasi struktur pengetahuan, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan peta konsep sebagai representasi struktur pengetahuan melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan peta konsep sebagai representasi struktur pengetahuan;
- Bandingkan bukti peta konsep sebagai representasi struktur pengetahuan pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek peta konsep sebagai representasi struktur pengetahuan dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian peta konsep sebagai representasi struktur pengetahuan, topik asam-basa dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa asam kuat dianggap selalu berkonsentrasi tinggi. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta pemisahan konsep derajat ionisasi, K_a , konsentrasi, dan pH. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa kekuatan asam dan konsentrasi adalah dimensi berbeda. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mendefinisikan fungsi, objek, dan cara kerja asesmen diagnostik sebagai proses pengumpulan bukti tentang struktur pemahaman, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, peta konsep sebagai representasi struktur pengetahuan harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian

klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.1.15 Gambar, diagram, dan model partikel sebagai bukti pemahaman

Aspek gambar, diagram, dan model partikel sebagai bukti pemahaman menegaskan bahwa diagnosis adalah kegiatan ilmiah berskala kelas: guru merumuskan pertanyaan, mengumpulkan bukti, mempertimbangkan penjelasan alternatif, lalu menentukan kesimpulan yang paling masuk akal. Inferensi harus proporsional terhadap kualitas data. Semakin spesifik label yang diberikan, semakin kuat pula bukti yang diperlukan (Anggraena et al., 2025; Treagust, 1988).

Secara operasional, guru dapat menggunakan gambar partikel untuk menelusuri kemampuan transfer. Data tersebut perlu dibaca bersama kualitas model mental yang dieksternalisasi, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menyimpulkan miskonsepsi hanya dari satu jawaban. Untuk mengurangnya, guru dapat mengulang diagnosis setelah remediasi dan menyusun tindak lanjut berbeda untuk kategori respons yang berbeda. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai gambar, diagram, dan model partikel sebagai bukti pemahaman, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan gambar, diagram, dan model partikel sebagai bukti pemahaman melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan gambar, diagram, dan model partikel sebagai bukti pemahaman;
- Bandingkan bukti gambar, diagram, dan model partikel sebagai bukti pemahaman pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek gambar, diagram, dan model partikel sebagai bukti pemahaman dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian gambar, diagram, dan model partikel sebagai bukti pemahaman, topik kinetika dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa reaksi eksoterm dianggap selalu cepat. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta perbandingan profil energi dan faktor kinetik. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa termodinamika dan kinetika menjawab pertanyaan berbeda. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mendefinisikan fungsi, objek, dan cara kerja asesmen diagnostik sebagai proses pengumpulan bukti tentang struktur pemahaman, bukan diperlakukan sebagai

skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, gambar, diagram, dan model partikel sebagai bukti pemahaman harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.1.16 Predict-Observe-Explain dalam diagnosis kimia

Penerapan predict-observe-explain dalam diagnosis kimia harus menghubungkan teori pengukuran dengan pengetahuan pedagogis kimia. Instrumen yang teknis tetapi tidak memahami struktur konsep akan menghasilkan kategori dangkal; sebaliknya, pengetahuan konten tanpa prosedur pengumpulan bukti dapat menghasilkan kesimpulan intuitif yang sulit diuji. Keduanya perlu dipadukan dalam satu alur yang transparan (Treagust, 1988; Taber, 2002a).

Secara operasional, guru dapat menggunakan observasi praktikum untuk menelusuri stabilitas setelah jeda waktu. Data tersebut perlu dibaca bersama respons terhadap probing, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mencampur kesalahan matematika dengan kesalahan konsep. Untuk menguranginya, guru dapat mencatat derajat keyakinan dan membandingkan jawaban dengan alasan. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai predict-observe-explain dalam diagnosis kimia, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan predict-observe-explain dalam diagnosis kimia melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan predict-observe-explain dalam diagnosis kimia;
- Bandingkan bukti predict-observe-explain dalam diagnosis kimia pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek predict-observe-explain dalam diagnosis kimia dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian predict-observe-explain dalam diagnosis kimia, topik elektrokimia dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa elektron dianggap bergerak melalui jembatan garam. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta penelusuran jalur

elektron dan ion pada sel galvanik. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa elektron mengalir pada rangkaian luar, sedangkan ion menjaga kenetralan muatan. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mendefinisikan fungsi, objek, dan cara kerja asesmen diagnostik sebagai proses pengumpulan bukti tentang struktur pemahaman, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Kerangka operasional

Aspek	Bukti yang dicari	Risiko interpretasi	Tindak lanjut
Target konsep	Penjelasan tentang elektrokimia	Target terlalu luas	Batasi subkonsep dan prasyarat
Respons	Stabilitas setelah jeda waktu	Mencampur kesalahan matematika dengan kesalahan konsep	Mencatat derajat keyakinan
Penalaran	Respons terhadap probing	Alasan diperlakukan sekadar tambahan	Membandingkan jawaban dengan alasan
Konfirmasi	Observasi praktikum atau tugas isomorfik	Kesimpulan dibuat dari satu konteks	Gunakan bukti kedua dan rediagnosis

Implikasinya, predict-observe-explain dalam diagnosis kimia harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.1.17 Think-aloud untuk menelusuri proses penalaran

Think-aloud untuk menelusuri proses penalaran tidak berdiri sebagai prosedur administratif. Ia menjadi bagian dari keputusan tentang apa yang perlu dijelaskan ulang, representasi apa yang perlu dibandingkan, dan siapa yang memerlukan dukungan berbeda. Oleh sebab itu, bukti diagnosis harus langsung terhubung dengan tindakan pembelajaran, bukan berhenti pada daftar kesalahan (Taber, 2002a; Black & Wiliam, 2009).

Secara operasional, guru dapat menggunakan exit ticket untuk menelusuri urutan langkah penyelesaian. Data tersebut perlu dibaca bersama ketepatan pilihan, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menggunakan ambang kategori tanpa bukti. Untuk menguranginya, guru dapat menelusuri langkah penyelesaian dan menggunakan pertanyaan tindak lanjut. Prosedur tersebut

membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai think-aloud untuk menelusuri proses penalaran, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan think-aloud untuk menelusuri proses penalaran melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan think-aloud untuk menelusuri proses penalaran;
- Bandingkan bukti think-aloud untuk menelusuri proses penalaran pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek think-aloud untuk menelusuri proses penalaran dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian think-aloud untuk menelusuri proses penalaran, topik praktikum titrasi dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa titik akhir dianggap selalu identik dengan titik ekuivalen. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta analisis indikator, kurva titrasi, dan ketidakpastian. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa titik akhir adalah observasi operasional yang mendekati titik ekuivalen. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mendefinisikan fungsi, objek, dan cara kerja asesmen diagnostik sebagai proses pengumpulan bukti tentang struktur pemahaman, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, think-aloud untuk menelusuri proses penalaran harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.1.18 Analisis pekerjaan peserta didik dan jejak penyelesaian

Analisis pekerjaan peserta didik dan jejak penyelesaian perlu ditempatkan dalam kerangka diagnosis yang mendefinisikan fungsi, objek, dan cara kerja asesmen diagnostik sebagai proses pengumpulan bukti tentang struktur pemahaman. Fokusnya bukan sekadar menghasilkan skor, melainkan memperoleh bukti yang cukup untuk menjelaskan bagaimana peserta didik membangun, menghubungkan, dan menggunakan konsep. Karena konsepsi tidak dapat diamati secara langsung, guru harus menarik inferensi dari respons verbal, tulisan, gambar,

tindakan, dan perubahan strategi yang tampak selama penyelesaian tugas (Black & Wiliam, 2009; Anggraena et al., 2025).

Secara operasional, guru dapat menggunakan tes pilihan ganda dua tingkat untuk menelusuri kualitas model mental yang dieksternalisasi. Data tersebut perlu dibaca bersama koherensi alasan, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah memperlakukan konsepsi sebagai sifat permanen. Untuk menguranginya, guru dapat menggunakan triangulasi bukti dan menyajikan tugas paralel. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai analisis pekerjaan peserta didik dan jejak penyelesaian, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan analisis pekerjaan peserta didik dan jejak penyelesaian melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan analisis pekerjaan peserta didik dan jejak penyelesaian;
- Bandingkan bukti analisis pekerjaan peserta didik dan jejak penyelesaian pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek analisis pekerjaan peserta didik dan jejak penyelesaian dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian analisis pekerjaan peserta didik dan jejak penyelesaian, topik sistem periodik dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa jari-jari atom dianggap selalu bertambah bersama nomor atom. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta prediksi tren disertai penjelasan muatan inti efektif dan kulit elektron. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa tren periodik perlu dijelaskan melalui struktur dan interaksi, bukan hafalan arah panah. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mendefinisikan fungsi, objek, dan cara kerja asesmen diagnostik sebagai proses pengumpulan bukti tentang struktur pemahaman, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, analisis pekerjaan peserta didik dan jejak penyelesaian harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.1.19 Pemanfaatan data digital dan respons waktu nyata

Pembahasan tentang pemanfaatan data digital dan respons waktu nyata berangkat dari prinsip bahwa data asesmen selalu merupakan jejak tidak langsung dari proses berpikir. Satu pilihan jawaban hanya menunjukkan produk akhir, sedangkan struktur pengetahuan yang melahirkannya dapat berbeda. Diagnosis yang kuat karena itu menggabungkan respons terhadap probing, tingkat keyakinan, dan bukti kontekstual sehingga keputusan pedagogis tidak bergantung pada satu indikator saja (Anggraena et al., 2025; Treagust, 1988).

Secara operasional, guru dapat menggunakan alasan terbuka untuk menelusuri respons terhadap probing. Data tersebut perlu dibaca bersama tingkat keyakinan, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mengabaikan ketidakpastian pengukuran. Untuk mengurangnya, guru dapat menyusun tindak lanjut berbeda untuk kategori respons yang berbeda dan memeriksa konsep prasyarat. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai pemanfaatan data digital dan respons waktu nyata, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan pemanfaatan data digital dan respons waktu nyata melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan pemanfaatan data digital dan respons waktu nyata;
- Bandingkan bukti pemanfaatan data digital dan respons waktu nyata pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek pemanfaatan data digital dan respons waktu nyata dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian pemanfaatan data digital dan respons waktu nyata, topik stoikiometri dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa koefisien reaksi diperlakukan sebagai perbandingan massa langsung. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta translasi persamaan dari partikel ke mol dan massa. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa koefisien terutama menyatakan rasio jumlah entitas atau mol. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mendefinisikan fungsi, objek, dan cara kerja asesmen diagnostik sebagai proses pengumpulan bukti tentang struktur pemahaman, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, pemanfaatan data digital dan respons waktu nyata harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian

klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.1.20 Triangulasi metode dan sumber bukti

Dalam pembelajaran kimia, triangulasi metode dan sumber bukti memiliki arti strategis karena banyak jawaban dapat diperoleh melalui hafalan prosedur atau pengenalan pola. Peserta didik mungkin tampak berhasil tetapi tidak mampu menerjemahkan makna simbol, model, atau besaran yang digunakan. Prinsip diagnosis menuntut bukti tentang cara konsep dipakai, bukan hanya apakah hasil akhirnya sesuai kunci (Treagust, 1988; Taber, 2002a).

Secara operasional, guru dapat menggunakan peta konsep untuk menelusuri ketepatan pilihan. Data tersebut perlu dibaca bersama konsistensi lintas butir, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menyimpulkan miskonsepsi hanya dari satu jawaban. Untuk mengurangnya, guru dapat membandingkan jawaban dengan alasan dan meminta translasi antarrepresentasi. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai triangulasi metode dan sumber bukti, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan triangulasi metode dan sumber bukti melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan triangulasi metode dan sumber bukti;
- Bandingkan bukti triangulasi metode dan sumber bukti pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek triangulasi metode dan sumber bukti dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian triangulasi metode dan sumber bukti, topik asam-basa dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa asam kuat dianggap selalu berkonsentrasi tinggi. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta pemisahan konsep derajat ionisasi, K_a , konsentrasi, dan pH. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa kekuatan asam dan konsentrasi adalah dimensi berbeda. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mendefinisikan fungsi, objek, dan cara kerja asesmen diagnostik sebagai proses pengumpulan bukti tentang struktur pemahaman, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban

setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Kerangka operasional

Aspek	Bukti yang dicari	Risiko interpretasi	Tindak lanjut
Target konsep	Penjelasan tentang asam-basa	Target terlalu luas	Batasi subkonsep dan prasyarat
Respons	Ketepatan pilihan	Menyimpulkan miskonsepsi hanya dari satu jawaban	Membandingkan jawaban dengan alasan
Penalaran	Konsistensi lintas butir	Alasan diperlakukan sekadar tambahan	Meminta translasi antarrepresentasi
Konfirmasi	Peta konsep atau tugas isomorfik	Kesimpulan dibuat dari satu konteks	Gunakan bukti kedua dan rediagnosis

Implikasinya, triangulasi metode dan sumber bukti harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.1.21 Keterbatasan asesmen diagnostik

Aspek keterbatasan asesmen diagnostik menegaskan bahwa diagnosis adalah kegiatan ilmiah berskala kelas: guru merumuskan pertanyaan, mengumpulkan bukti, mempertimbangkan penjelasan alternatif, lalu menentukan kesimpulan yang paling masuk akal. Inferensi harus proporsional terhadap kualitas data. Semakin spesifik label yang diberikan, semakin kuat pula bukti yang diperlukan (Taber, 2002a; Black & Wiliam, 2009).

Secara operasional, guru dapat menggunakan analisis respons tertulis untuk menelusuri koherensi alasan. Data tersebut perlu dibaca bersama konsistensi lintas representasi, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mencampur kesalahan matematika dengan kesalahan konsep. Untuk mengurangnya, guru dapat menggunakan pertanyaan tindak lanjut dan mengulang diagnosis setelah remediasi. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai keterbatasan asesmen diagnostik, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;

- Pilih tugas yang merepresentasikan keterbatasan asesmen diagnostik melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan keterbatasan asesmen diagnostik;
- Bandingkan bukti keterbatasan asesmen diagnostik pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek keterbatasan asesmen diagnostik dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian keterbatasan asesmen diagnostik, topik kinetika dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa reaksi eksoterm dianggap selalu cepat. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta perbandingan profil energi dan faktor kinetik. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa termodinamika dan kinetika menjawab pertanyaan berbeda. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mendefinisikan fungsi, objek, dan cara kerja asesmen diagnostik sebagai proses pengumpulan bukti tentang struktur pemahaman, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, keterbatasan asesmen diagnostik harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.1.22 Peran guru sebagai penafsir bukti

Penerapan peran guru sebagai penafsir bukti harus menghubungkan teori pengukuran dengan pengetahuan pedagogis kimia. Instrumen yang teknis tetapi tidak memahami struktur konsep akan menghasilkan kategori dangkal; sebaliknya, pengetahuan konten tanpa prosedur pengumpulan bukti dapat menghasilkan kesimpulan intuitif yang sulit diuji. Keduanya perlu dipadukan dalam satu alur yang transparan (Black & Wiliam, 2009; Anggraena et al., 2025).

Secara operasional, guru dapat menggunakan diskusi argumentatif untuk menelusuri tingkat keyakinan. Data tersebut perlu dibaca bersama kemampuan transfer, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menggunakan ambang kategori tanpa bukti. Untuk mengurangnya, guru dapat menyajikan tugas paralel dan mencatat derajat keyakinan. Prosedur tersebut membantu memisahkan

pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai peran guru sebagai penafsir bukti, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan peran guru sebagai penafsir bukti melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan peran guru sebagai penafsir bukti;
- Bandingkan bukti peran guru sebagai penafsir bukti pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek peran guru sebagai penafsir bukti dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian peran guru sebagai penafsir bukti, topik elektrokimia dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa elektron dianggap bergerak melalui jembatan garam. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta penelusuran jalur elektron dan ion pada sel galvanik. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa elektron mengalir pada rangkaian luar, sedangkan ion menjaga kenetralan muatan. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mendefinisikan fungsi, objek, dan cara kerja asesmen diagnostik sebagai proses pengumpulan bukti tentang struktur pemahaman, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, peran guru sebagai penafsir bukti harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.1.23 Implementasi diagnosis pada kelas yang heterogen

Implementasi diagnosis pada kelas yang heterogen tidak berdiri sebagai prosedur administratif. Ia menjadi bagian dari keputusan tentang apa yang perlu dijelaskan ulang, representasi apa yang perlu dibandingkan, dan siapa yang memerlukan dukungan berbeda. Oleh sebab itu, bukti diagnosis harus langsung terhubung dengan tindakan pembelajaran, bukan berhenti pada daftar kesalahan (Anggraena et al., 2025; Treagust, 1988).

Secara operasional, guru dapat menggunakan analisis grafik untuk menelusuri konsistensi lintas butir. Data tersebut perlu dibaca bersama stabilitas setelah jeda waktu, karena pola yang

sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah memperlakukan konsepsi sebagai sifat permanen. Untuk mengurangnya, guru dapat memeriksa konsep prasyarat dan menelusuri langkah penyelesaian. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai implementasi diagnosis pada kelas yang heterogen, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan implementasi diagnosis pada kelas yang heterogen melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan implementasi diagnosis pada kelas yang heterogen;
- Bandingkan bukti implementasi diagnosis pada kelas yang heterogen pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek implementasi diagnosis pada kelas yang heterogen dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian implementasi diagnosis pada kelas yang heterogen, topik praktikum titrasi dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa titik akhir dianggap selalu identik dengan titik ekuivalen. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta analisis indikator, kurva titrasi, dan ketidakpastian. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa titik akhir adalah observasi operasional yang mendekati titik ekuivalen. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mendefinisikan fungsi, objek, dan cara kerja asesmen diagnostik sebagai proses pengumpulan bukti tentang struktur pemahaman, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, implementasi diagnosis pada kelas yang heterogen harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.1.24 Sintesis hakikat asesmen diagnostik

Sintesis hakikat asesmen diagnostik perlu ditempatkan dalam kerangka diagnosis yang mendefinisikan fungsi, objek, dan cara kerja asesmen diagnostik sebagai proses pengumpulan bukti tentang struktur pemahaman. Fokusnya bukan sekadar menghasilkan skor, melainkan memperoleh bukti yang cukup untuk menjelaskan bagaimana peserta didik membangun,

menghubungkan, dan menggunakan konsep. Karena konsepsi tidak dapat diamati secara langsung, guru harus menarik inferensi dari respons verbal, tulisan, gambar, tindakan, dan perubahan strategi yang tampak selama penyelesaian tugas (Treagust, 1988; Taber, 2002a).

Secara operasional, guru dapat menggunakan tes empat tingkat untuk menelusuri konsistensi lintas representasi. Data tersebut perlu dibaca bersama urutan langkah penyelesaian, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mengabaikan ketidakpastian pengukuran. Untuk menguranginya, guru dapat meminta translasi antarrepresentasi dan menggunakan triangulasi bukti. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai sintesis hakikat asesmen diagnostik, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan sintesis hakikat asesmen diagnostik melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan sintesis hakikat asesmen diagnostik;
- Bandingkan bukti sintesis hakikat asesmen diagnostik pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek sintesis hakikat asesmen diagnostik dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian sintesis hakikat asesmen diagnostik, topik sistem periodik dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa jari-jari atom dianggap selalu bertambah bersama nomor atom. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta prediksi tren disertai penjelasan muatan inti efektif dan kulit elektron. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa tren periodik perlu dijelaskan melalui struktur dan interaksi, bukan hafalan arah panah. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mendefinisikan fungsi, objek, dan cara kerja asesmen diagnostik sebagai proses pengumpulan bukti tentang struktur pemahaman, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, sintesis hakikat asesmen diagnostik harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.1.25 Implikasi praktis bagi guru kimia

Implikasi praktis subbab 5.1 dapat dirangkum sebagai alur: tentukan konsep target, pilih bukti yang relevan, kumpulkan respons dalam lebih dari satu bentuk, analisis pola, nyatakan derajat kepastian, susun tindak lanjut, dan lakukan rediagnosis. Alur ini menjaga agar hakikat asesmen diagnostik tidak berubah menjadi administrasi skor, tetapi berfungsi sebagai mekanisme perbaikan pembelajaran yang berulang.

1. Identifikasi konsep inti dan konsep prasyarat sebelum menyusun instrumen.
2. Gunakan pertanyaan yang meminta alasan, representasi, atau langkah kerja.
3. Bandingkan respons pada beberapa konteks dan waktu.
4. Pisahkan bukti faktual dari interpretasi guru.
5. Tentukan tindak lanjut yang berbeda untuk kategori kebutuhan yang berbeda.
6. Simpan data secukupnya dan lindungi kerahasiaan peserta didik.
7. Periksa kembali perubahan melalui tugas transfer dan retensi.

5.2 Tujuan Diagnosis Miskonsepsi

Subbab ini menempatkan diagnosis sebagai dasar keputusan pedagogis yang spesifik, proporsional, dan dapat ditindaklanjuti. Uraian disusun dari prinsip teoretis menuju prosedur kelas, contoh kimia, risiko interpretasi, dan tindak lanjut. Setiap bagian menekankan bahwa kesimpulan diagnostik harus didasarkan pada pola bukti, bukan pada satu respons yang terpisah.

5.2.1 Mengidentifikasi konsepsi awal sebelum pembelajaran

Pembahasan tentang mengidentifikasi konsepsi awal sebelum pembelajaran berangkat dari prinsip bahwa data asesmen selalu merupakan jejak tidak langsung dari proses berpikir. Satu pilihan jawaban hanya menunjukkan produk akhir, sedangkan struktur pengetahuan yang melahirkannya dapat berbeda. Diagnosis yang kuat karena itu menggabungkan koherensi alasan, konsistensi lintas representasi, dan bukti kontekstual sehingga keputusan pedagogis tidak bergantung pada satu indikator saja (Black & Wiliam, 2009; Taber, 2002a).

Secara operasional, guru dapat menggunakan analisis grafik untuk menelusuri koherensi alasan. Data tersebut perlu dibaca bersama konsistensi lintas representasi, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mencampur kesalahan matematika dengan kesalahan konsep. Untuk mengurangnya, guru dapat menggunakan pertanyaan tindak lanjut dan mengulang diagnosis setelah remediasi. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai mengidentifikasi konsepsi awal sebelum pembelajaran, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;

- Pilih tugas yang merepresentasikan mengidentifikasi konsepsi awal sebelum pembelajaran melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan mengidentifikasi konsepsi awal sebelum pembelajaran;
- Bandingkan bukti mengidentifikasi konsepsi awal sebelum pembelajaran pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek mengidentifikasi konsepsi awal sebelum pembelajaran dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian mengidentifikasi konsepsi awal sebelum pembelajaran, topik pereaksi pembatas dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa pereaksi dengan massa terkecil selalu dianggap pembatas. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta analisis rasio mol terhadap koefisien. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa status pembatas ditentukan oleh hubungan jumlah zat dan stoikiometri. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menempatkan diagnosis sebagai dasar keputusan pedagogis yang spesifik, proporsional, dan dapat ditindaklanjuti, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, mengidentifikasi konsepsi awal sebelum pembelajaran harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.2.2 Menemukan miskonsepsi yang tersembunyi di balik jawaban benar

Dalam pembelajaran kimia, menemukan miskonsepsi yang tersembunyi di balik jawaban benar memiliki arti strategis karena banyak jawaban dapat diperoleh melalui hafalan prosedur atau pengenalan pola. Peserta didik mungkin tampak berhasil tetapi tidak mampu menerjemahkan makna simbol, model, atau besaran yang digunakan. Prinsip diagnosis menuntut bukti tentang cara konsep dipakai, bukan hanya apakah hasil akhirnya sesuai kunci (Taber, 2002a; Shiddiqi et al., 2024).

Secara operasional, guru dapat menggunakan tes empat tingkat untuk menelusuri tingkat keyakinan. Data tersebut perlu dibaca bersama kemampuan transfer, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menggunakan ambang kategori tanpa bukti. Untuk menguranginya, guru dapat menyajikan tugas paralel dan mencatat derajat keyakinan. Prosedur tersebut membantu memisahkan

pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai menemukan miskonsepsi yang tersembunyi di balik jawaban benar, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan menemukan miskonsepsi yang tersembunyi di balik jawaban benar melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan menemukan miskonsepsi yang tersembunyi di balik jawaban benar;
- Bandingkan bukti menemukan miskonsepsi yang tersembunyi di balik jawaban benar pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek menemukan miskonsepsi yang tersembunyi di balik jawaban benar dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian menemukan miskonsepsi yang tersembunyi di balik jawaban benar, topik larutan penyangga dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa penyangga dianggap mempertahankan pH secara mutlak. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta analisis perubahan komposisi setelah penambahan asam atau basa. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa penyangga membatasi perubahan dalam kapasitas tertentu, bukan meniadakannya. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menempatkan diagnosis sebagai dasar keputusan pedagogis yang spesifik, proporsional, dan dapat ditindaklanjuti, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, menemukan miskonsepsi yang tersembunyi di balik jawaban benar harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.2.3 Membedakan ketidaktahuan, pengetahuan parsial, dan miskonsepsi

Aspek membedakan ketidaktahuan, pengetahuan parsial, dan miskonsepsi menegaskan bahwa diagnosis adalah kegiatan ilmiah berskala kelas: guru merumuskan pertanyaan, mengumpulkan bukti, mempertimbangkan penjelasan alternatif, lalu menentukan kesimpulan yang paling masuk akal. Inferensi harus proporsional terhadap kualitas data. Semakin spesifik label yang diberikan, semakin kuat pula bukti yang diperlukan (Shiddiqi et al., 2024; Treagust, 1988).

Secara operasional, guru dapat menggunakan think-aloud untuk menelusuri konsistensi lintas butir. Data tersebut perlu dibaca bersama stabilitas setelah jeda waktu, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah memperlakukan konsepsi sebagai sifat permanen. Untuk mengurangnya, guru dapat memeriksa konsep prasyarat dan menelusuri langkah penyelesaian. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai membedakan ketidaktahuan, pengetahuan parsial, dan miskonsepsi, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan membedakan ketidaktahuan, pengetahuan parsial, dan miskonsepsi melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan membedakan ketidaktahuan, pengetahuan parsial, dan miskonsepsi;
- Bandingkan bukti membedakan ketidaktahuan, pengetahuan parsial, dan miskonsepsi pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek membedakan ketidaktahuan, pengetahuan parsial, dan miskonsepsi dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian membedakan ketidaktahuan, pengetahuan parsial, dan miskonsepsi, topik katalis dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa katalis dianggap menambah jumlah produk kesetimbangan. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta analisis jalur energi dan laju maju-balik. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa katalis mempercepat tercapainya kesetimbangan tanpa mengubah K. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menempatkan diagnosis sebagai dasar keputusan pedagogis yang spesifik, proporsional, dan dapat ditindaklanjuti, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, membedakan ketidaktahuan, pengetahuan parsial, dan miskonsepsi harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.2.4 Memetakan konsep prasyarat yang belum kokoh

Penerapan memetakan konsep prasyarat yang belum kokoh harus menghubungkan teori pengukuran dengan pengetahuan pedagogis kimia. Instrumen yang teknis tetapi tidak

memahami struktur konsep akan menghasilkan kategori dangkal; sebaliknya, pengetahuan konten tanpa prosedur pengumpulan bukti dapat menghasilkan kesimpulan intuitif yang sulit diuji. Keduanya perlu dipadukan dalam satu alur yang transparan (Treagust, 1988; Black & Wiliam, 2009).

Secara operasional, guru dapat menggunakan predict-observe-explain untuk menelusuri konsistensi lintas representasi. Data tersebut perlu dibaca bersama urutan langkah penyelesaian, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mengabaikan ketidakpastian pengukuran. Untuk mengurangnya, guru dapat meminta translasi antarrepresentasi dan menggunakan triangulasi bukti. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai memetakan konsep prasyarat yang belum kokoh, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan memetakan konsep prasyarat yang belum kokoh melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan memetakan konsep prasyarat yang belum kokoh;
- Bandingkan bukti memetakan konsep prasyarat yang belum kokoh pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek memetakan konsep prasyarat yang belum kokoh dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian memetakan konsep prasyarat yang belum kokoh, topik elektrolisis dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa anoda selalu dianggap bermuatan negatif. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta pembedaan definisi anoda-katoda berdasarkan proses oksidasi-reduksi. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa tanda elektroda bergantung jenis sel, tetapi anoda tetap lokasi oksidasi. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menempatkan diagnosis sebagai dasar keputusan pedagogis yang spesifik, proporsional, dan dapat ditindaklanjuti, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Kerangka operasional

Aspek	Bukti yang dicari	Risiko interpretasi	Tindak lanjut
Target konsep	Penjelasan tentang elektrolisis	Target terlalu luas	Batasi subkonsep dan prasyarat

Aspek	Bukti yang dicari	Risiko interpretasi	Tindak lanjut
Respons	Konsistensi lintas representasi	Mengabaikan ketidakpastian pengukuran	Meminta translasi antarrepresentasi
Penalaran	Urutan langkah penyelesaian	Alasan diperlakukan sekadar tambahan	Menggunakan triangulasi bukti
Konfirmasi	Predict-observe-explain atau tugas isomorfik	Kesimpulan dibuat dari satu konteks	Gunakan bukti kedua dan rediagnosis

Implikasinya, memetakan konsep prasyarat yang belum kokoh harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.2.5 Menentukan prioritas pembelajaran dan remediasi

Menentukan prioritas pembelajaran dan remediasi tidak berdiri sebagai prosedur administratif. Ia menjadi bagian dari keputusan tentang apa yang perlu dijelaskan ulang, representasi apa yang perlu dibandingkan, dan siapa yang memerlukan dukungan berbeda. Oleh sebab itu, bukti diagnosis harus langsung terhubung dengan tindakan pembelajaran, bukan berhenti pada daftar kesalahan (Black & Wiliam, 2009; Taber, 2002a).

Secara operasional, guru dapat menggunakan tugas isomorfik untuk menelusuri kemampuan transfer. Data tersebut perlu dibaca bersama kualitas model mental yang dieksternalisasi, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menyimpulkan miskonsepsi hanya dari satu jawaban. Untuk mengurangnya, guru dapat mengulang diagnosis setelah remediasi dan menyusun tindak lanjut berbeda untuk kategori respons yang berbeda. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai menentukan prioritas pembelajaran dan remediasi, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan menentukan prioritas pembelajaran dan remediasi melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan menentukan prioritas pembelajaran dan remediasi;

- Bandingkan bukti menentukan prioritas pembelajaran dan remediasi pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek menentukan prioritas pembelajaran dan remediasi dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian menentukan prioritas pembelajaran dan remediasi, topik pengukuran laboratorium dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa angka lebih banyak selalu dianggap lebih akurat. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta perbedaan akurasi, presisi, resolusi, dan angka penting. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa kualitas data ditentukan oleh desain, kalibrasi, dan ketidakpastian. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menempatkan diagnosis sebagai dasar keputusan pedagogis yang spesifik, proporsional, dan dapat ditindaklanjuti, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, menentukan prioritas pembelajaran dan remediasi harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.2.6 Menyesuaikan representasi dan strategi pengajaran

Menyesuaikan representasi dan strategi pengajaran perlu ditempatkan dalam kerangka diagnosis yang menempatkan diagnosis sebagai dasar keputusan pedagogis yang spesifik, proporsional, dan dapat ditindaklanjuti. Fokusnya bukan sekadar menghasilkan skor, melainkan memperoleh bukti yang cukup untuk menjelaskan bagaimana peserta didik membangun, menghubungkan, dan menggunakan konsep. Karena konsepsi tidak dapat diamati secara langsung, guru harus menarik inferensi dari respons verbal, tulisan, gambar, tindakan, dan perubahan strategi yang tampak selama penyelesaian tugas (Taber, 2002a; Shiddiqi et al., 2024).

Secara operasional, guru dapat menggunakan kuis formatif digital untuk menelusuri stabilitas setelah jeda waktu. Data tersebut perlu dibaca bersama respons terhadap probing, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mencampur kesalahan matematika dengan kesalahan konsep. Untuk mengurangnya, guru dapat mencatat derajat keyakinan dan membandingkan jawaban dengan alasan. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai menyesuaikan representasi dan strategi pengajaran, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan menyesuaikan representasi dan strategi pengajaran melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan menyesuaikan representasi dan strategi pengajaran;
- Bandingkan bukti menyesuaikan representasi dan strategi pengajaran pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek menyesuaikan representasi dan strategi pengajaran dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian menyesuaikan representasi dan strategi pengajaran, topik ikatan ion dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa NaCl dipandang sebagai molekul diskrit. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta gambar partikel padatan NaCl dan penjelasan hubungan struktur-kisi dengan titik leleh. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa senyawa ionik membentuk kisi dan rumusnya menyatakan rasio ion. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menempatkan diagnosis sebagai dasar keputusan pedagogis yang spesifik, proporsional, dan dapat ditindaklanjuti, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, menyesuaikan representasi dan strategi pengajaran harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.2.7 Mencegah penguatan miskonsepsi melalui pembelajaran lanjutan

Pembahasan tentang mencegah penguatan miskonsepsi melalui pembelajaran lanjutan berangkat dari prinsip bahwa data asesmen selalu merupakan jejak tidak langsung dari proses berpikir. Satu pilihan jawaban hanya menunjukkan produk akhir, sedangkan struktur pengetahuan yang melahirkannya dapat berbeda. Diagnosis yang kuat karena itu menggabungkan urutan langkah penyelesaian, ketepatan pilihan, dan bukti kontekstual sehingga keputusan pedagogis tidak bergantung pada satu indikator saja (Shiddiqi et al., 2024; Treagust, 1988).

Secara operasional, guru dapat menggunakan tes tiga tingkat untuk menelusuri urutan langkah penyelesaian. Data tersebut perlu dibaca bersama ketepatan pilihan, karena pola

yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menggunakan ambang kategori tanpa bukti. Untuk mengurangnya, guru dapat menelusuri langkah penyelesaian dan menggunakan pertanyaan tindak lanjut. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai mencegah penguatan miskonsepsi melalui pembelajaran lanjutan, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan mencegah penguatan miskonsepsi melalui pembelajaran lanjutan melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan mencegah penguatan miskonsepsi melalui pembelajaran lanjutan;
- Bandingkan bukti mencegah penguatan miskonsepsi melalui pembelajaran lanjutan pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek mencegah penguatan miskonsepsi melalui pembelajaran lanjutan dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian mencegah penguatan miskonsepsi melalui pembelajaran lanjutan, topik pereaksi pembatas dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa pereaksi dengan massa terkecil selalu dianggap pembatas. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta analisis rasio mol terhadap koefisien. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa status pembatas ditentukan oleh hubungan jumlah zat dan stoikiometri. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menempatkan diagnosis sebagai dasar keputusan pedagogis yang spesifik, proporsional, dan dapat ditindaklanjuti, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, mencegah penguatan miskonsepsi melalui pembelajaran lanjutan harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.2.8 Menentukan kebutuhan diferensiasi pembelajaran

Dalam pembelajaran kimia, menentukan kebutuhan diferensiasi pembelajaran memiliki arti strategis karena banyak jawaban dapat diperoleh melalui hafalan prosedur atau pengenalan pola. Peserta didik mungkin tampak berhasil tetapi tidak mampu menerjemahkan makna simbol, model, atau besaran yang digunakan. Prinsip diagnosis menuntut bukti tentang cara

konsep dipakai, bukan hanya apakah hasil akhirnya sesuai kunci (Treagust, 1988; Black & Wiliam, 2009).

Secara operasional, guru dapat menggunakan wawancara semi-terstruktur untuk menelusuri kualitas model mental yang dieksternalisasi. Data tersebut perlu dibaca bersama koherensi alasan, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah memperlakukan konsepsi sebagai sifat permanen. Untuk menguranginya, guru dapat menggunakan triangulasi bukti dan menyajikan tugas paralel. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai menentukan kebutuhan diferensiasi pembelajaran, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan menentukan kebutuhan diferensiasi pembelajaran melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan menentukan kebutuhan diferensiasi pembelajaran;
- Bandingkan bukti menentukan kebutuhan diferensiasi pembelajaran pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek menentukan kebutuhan diferensiasi pembelajaran dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian menentukan kebutuhan diferensiasi pembelajaran, topik larutan penyangga dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa penyangga dianggap mempertahankan pH secara mutlak. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta analisis perubahan komposisi setelah penambahan asam atau basa. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa penyangga membatasi perubahan dalam kapasitas tertentu, bukan meniadakannya. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menempatkan diagnosis sebagai dasar keputusan pedagogis yang spesifik, proporsional, dan dapat ditindaklanjuti, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Kerangka operasional

Aspek	Bukti yang dicari	Risiko interpretasi	Tindak lanjut
Target konsep	Penjelasan tentang larutan penyangga	Target terlalu luas	Batasi subkonsep dan prasyarat
Respons	Kualitas model mental yang dieksternalisasi	Memperlakukan konsepsi sebagai sifat permanen	Menggunakan triangulasi bukti

Aspek	Bukti yang dicari	Risiko interpretasi	Tindak lanjut
Penalaran	Koherensi alasan	Alasan diperlakukan sekadar tambahan	Menyajikan tugas paralel
Konfirmasi	Wawancara semi-terstruktur atau tugas isomorfik	Kesimpulan dibuat dari satu konteks	Gunakan bukti kedua dan rediagnosis

Implikasinya, menentukan kebutuhan diferensiasi pembelajaran harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.2.9 Memberikan umpan balik yang spesifik dan dapat ditindaklanjuti

Aspek memberikan umpan balik yang spesifik dan dapat ditindaklanjuti menegaskan bahwa diagnosis adalah kegiatan ilmiah berskala kelas: guru merumuskan pertanyaan, mengumpulkan bukti, mempertimbangkan penjelasan alternatif, lalu menentukan kesimpulan yang paling masuk akal. Inferensi harus proporsional terhadap kualitas data. Semakin spesifik label yang diberikan, semakin kuat pula bukti yang diperlukan (Black & Wiliam, 2009; Taber, 2002a).

Secara operasional, guru dapat menggunakan gambar partikel untuk menelusuri respons terhadap probing. Data tersebut perlu dibaca bersama tingkat keyakinan, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mengabaikan ketidakpastian pengukuran. Untuk menguranginya, guru dapat menyusun tindak lanjut berbeda untuk kategori respons yang berbeda dan memeriksa konsep prasyarat. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai memberikan umpan balik yang spesifik dan dapat ditindaklanjuti, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan memberikan umpan balik yang spesifik dan dapat ditindaklanjuti melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan memberikan umpan balik yang spesifik dan dapat ditindaklanjuti;
- Bandingkan bukti memberikan umpan balik yang spesifik dan dapat ditindaklanjuti pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek memberikan umpan balik yang spesifik dan dapat ditindaklanjuti dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian memberikan umpan balik yang spesifik dan dapat ditindaklanjuti, topik katalis dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa katalis dianggap menambah jumlah produk kesetimbangan. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta analisis jalur energi dan laju maju-balik. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa katalis mempercepat tercapainya kesetimbangan tanpa mengubah K. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menempatkan diagnosis sebagai dasar keputusan pedagogis yang spesifik, proporsional, dan dapat ditindaklanjuti, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, memberikan umpan balik yang spesifik dan dapat ditindaklanjuti harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.2.10 Mendukung regulasi diri dan metakognisi peserta didik

Penerapan mendukung regulasi diri dan metakognisi peserta didik harus menghubungkan teori pengukuran dengan pengetahuan pedagogis kimia. Instrumen yang teknis tetapi tidak memahami struktur konsep akan menghasilkan kategori dangkal; sebaliknya, pengetahuan konten tanpa prosedur pengumpulan bukti dapat menghasilkan kesimpulan intuitif yang sulit diuji. Keduanya perlu dipadukan dalam satu alur yang transparan (Taber, 2002a; Shiddiqi et al., 2024).

Secara operasional, guru dapat menggunakan observasi praktikum untuk menelusuri ketepatan pilihan. Data tersebut perlu dibaca bersama konsistensi lintas butir, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menyimpulkan miskonsepsi hanya dari satu jawaban. Untuk mengurangnya, guru dapat membandingkan jawaban dengan alasan dan meminta translasi antarrepresentasi. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai mendukung regulasi diri dan metakognisi peserta didik, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan mendukung regulasi diri dan metakognisi peserta didik melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan mendukung regulasi diri dan metakognisi peserta didik;

- Bandingkan bukti mendukung regulasi diri dan metakognisi peserta didik pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek mendukung regulasi diri dan metakognisi peserta didik dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian mendukung regulasi diri dan metakognisi peserta didik, topik elektrolisis dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa anoda selalu dianggap bermuatan negatif. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta pembedaan definisi anoda-katoda berdasarkan proses oksidasi-reduksi. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa tanda elektroda bergantung jenis sel, tetapi anoda tetap lokasi oksidasi. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menempatkan diagnosis sebagai dasar keputusan pedagogis yang spesifik, proporsional, dan dapat ditindaklanjuti, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, mendukung regulasi diri dan metakognisi peserta didik harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.2.11 Mengevaluasi efektivitas penjelasan, analogi, dan model

Mengevaluasi efektivitas penjelasan, analogi, dan model tidak berdiri sebagai prosedur administratif. Ia menjadi bagian dari keputusan tentang apa yang perlu dijelaskan ulang, representasi apa yang perlu dibandingkan, dan siapa yang memerlukan dukungan berbeda. Oleh sebab itu, bukti diagnosis harus langsung terhubung dengan tindakan pembelajaran, bukan berhenti pada daftar kesalahan (Shiddiqi et al., 2024; Treagust, 1988).

Secara operasional, guru dapat menggunakan exit ticket untuk menelusuri koherensi alasan. Data tersebut perlu dibaca bersama konsistensi lintas representasi, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mencampur kesalahan matematika dengan kesalahan konsep. Untuk mengurangnya, guru dapat menggunakan pertanyaan tindak lanjut dan mengulang diagnosis setelah remediasi. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai mengevaluasi efektivitas penjelasan, analogi, dan model, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;

- Pilih tugas yang merepresentasikan mengevaluasi efektivitas penjelasan, analogi, dan model melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan mengevaluasi efektivitas penjelasan, analogi, dan model;
- Bandingkan bukti mengevaluasi efektivitas penjelasan, analogi, dan model pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek mengevaluasi efektivitas penjelasan, analogi, dan model dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian mengevaluasi efektivitas penjelasan, analogi, dan model, topik pengukuran laboratorium dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa angka lebih banyak selalu dianggap lebih akurat. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta perbedaan akurasi, presisi, resolusi, dan angka penting. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa kualitas data ditentukan oleh desain, kalibrasi, dan ketidakpastian. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menempatkan diagnosis sebagai dasar keputusan pedagogis yang spesifik, proporsional, dan dapat ditindaklanjuti, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, mengevaluasi efektivitas penjelasan, analogi, dan model harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.2.12 Mengevaluasi kualitas buku teks dan sumber belajar

Mengevaluasi kualitas buku teks dan sumber belajar perlu ditempatkan dalam kerangka diagnosis yang menempatkan diagnosis sebagai dasar keputusan pedagogis yang spesifik, proporsional, dan dapat ditindaklanjuti. Fokusnya bukan sekadar menghasilkan skor, melainkan memperoleh bukti yang cukup untuk menjelaskan bagaimana peserta didik membangun, menghubungkan, dan menggunakan konsep. Karena konsepsi tidak dapat diamati secara langsung, guru harus menarik inferensi dari respons verbal, tulisan, gambar, tindakan, dan perubahan strategi yang tampak selama penyelesaian tugas (Treagust, 1988; Black & Wiliam, 2009).

Secara operasional, guru dapat menggunakan tes pilihan ganda dua tingkat untuk menelusuri tingkat keyakinan. Data tersebut perlu dibaca bersama kemampuan transfer, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini

adalah menggunakan ambang kategori tanpa bukti. Untuk mengurangnya, guru dapat menyajikan tugas paralel dan mencatat derajat keyakinan. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai mengevaluasi kualitas buku teks dan sumber belajar, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan mengevaluasi kualitas buku teks dan sumber belajar melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan mengevaluasi kualitas buku teks dan sumber belajar;
- Bandingkan bukti mengevaluasi kualitas buku teks dan sumber belajar pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek mengevaluasi kualitas buku teks dan sumber belajar dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian mengevaluasi kualitas buku teks dan sumber belajar, topik ikatan ion dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa NaCl dipandang sebagai molekul diskrit. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta gambar partikel padatan NaCl dan penjelasan hubungan struktur-kisi dengan titik leleh. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa senyawa ionik membentuk kisi dan rumusnya menyatakan rasio ion. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menempatkan diagnosis sebagai dasar keputusan pedagogis yang spesifik, proporsional, dan dapat ditindaklanjuti, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Kerangka operasional

Aspek	Bukti yang dicari	Risiko interpretasi	Tindak lanjut
Target konsep	Penjelasan tentang ikatan ion	Target terlalu luas	Batasi subkonsep dan prasyarat
Respons	Tingkat keyakinan	Menggunakan ambang kategori tanpa bukti	Menyajikan tugas paralel
Penalaran	Kemampuan transfer	Alasan diperlakukan sekadar tambahan	Mencatat derajat keyakinan
Konfirmasi	Tes pilihan ganda dua tingkat atau tugas isomorfik	Kesimpulan dibuat dari satu konteks	Gunakan bukti kedua dan rediagnosis

Implikasinya, mengevaluasi kualitas buku teks dan sumber belajar harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.2.13 Mengidentifikasi miskonsepsi kolektif pada tingkat kelas

Pembahasan tentang mengidentifikasi miskonsepsi kolektif pada tingkat kelas berangkat dari prinsip bahwa data asesmen selalu merupakan jejak tidak langsung dari proses berpikir. Satu pilihan jawaban hanya menunjukkan produk akhir, sedangkan struktur pengetahuan yang melahirkannya dapat berbeda. Diagnosis yang kuat karena itu menggabungkan konsistensi lintas butir, stabilitas setelah jeda waktu, dan bukti kontekstual sehingga keputusan pedagogis tidak bergantung pada satu indikator saja (Black & William, 2009; Taber, 2002a).

Secara operasional, guru dapat menggunakan alasan terbuka untuk menelusuri konsistensi lintas butir. Data tersebut perlu dibaca bersama stabilitas setelah jeda waktu, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah memperlakukan konsepsi sebagai sifat permanen. Untuk mengurangnya, guru dapat memeriksa konsep prasyarat dan menelusuri langkah penyelesaian. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai mengidentifikasi miskonsepsi kolektif pada tingkat kelas, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan mengidentifikasi miskonsepsi kolektif pada tingkat kelas melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan mengidentifikasi miskonsepsi kolektif pada tingkat kelas;
- Bandingkan bukti mengidentifikasi miskonsepsi kolektif pada tingkat kelas pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek mengidentifikasi miskonsepsi kolektif pada tingkat kelas dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian mengidentifikasi miskonsepsi kolektif pada tingkat kelas, topik pereaksi pembatas dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa pereaksi dengan massa terkecil selalu dianggap pembatas. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta analisis rasio mol terhadap koefisien. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa status pembatas ditentukan oleh hubungan jumlah zat dan stoikiometri. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menempatkan diagnosis sebagai

dasar keputusan pedagogis yang spesifik, proporsional, dan dapat ditindaklanjuti, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, mengidentifikasi miskonsepsi kolektif pada tingkat kelas harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.2.14 Mendeteksi miskonsepsi yang khas secara individual

Dalam pembelajaran kimia, mendeteksi miskonsepsi yang khas secara individual memiliki arti strategis karena banyak jawaban dapat diperoleh melalui hafalan prosedur atau pengenalan pola. Peserta didik mungkin tampak berhasil tetapi tidak mampu menerjemahkan makna simbol, model, atau besaran yang digunakan. Prinsip diagnosis menuntut bukti tentang cara konsep dipakai, bukan hanya apakah hasil akhirnya sesuai kunci (Taber, 2002a; Shiddiqi et al., 2024).

Secara operasional, guru dapat menggunakan peta konsep untuk menelusuri konsistensi lintas representasi. Data tersebut perlu dibaca bersama urutan langkah penyelesaian, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mengabaikan ketidakpastian pengukuran. Untuk menguranginya, guru dapat meminta translasi antarrepresentasi dan menggunakan triangulasi bukti. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai mendeteksi miskonsepsi yang khas secara individual, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan mendeteksi miskonsepsi yang khas secara individual melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan mendeteksi miskonsepsi yang khas secara individual;
- Bandingkan bukti mendeteksi miskonsepsi yang khas secara individual pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek mendeteksi miskonsepsi yang khas secara individual dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian mendeteksi miskonsepsi yang khas secara individual, topik larutan penyangga dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa penyangga dianggap mempertahankan pH secara mutlak. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah

tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta analisis perubahan komposisi setelah penambahan asam atau basa. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa penyangga membatasi perubahan dalam kapasitas tertentu, bukan meniadakannya. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menempatkan diagnosis sebagai dasar keputusan pedagogis yang spesifik, proporsional, dan dapat ditindaklanjuti, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, mendeteksi miskonsepsi yang khas secara individual harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.2.15 Menilai perubahan konseptual dan kestabilannya

Aspek menilai perubahan konseptual dan kestabilannya menegaskan bahwa diagnosis adalah kegiatan ilmiah berskala kelas: guru merumuskan pertanyaan, mengumpulkan bukti, mempertimbangkan penjelasan alternatif, lalu menentukan kesimpulan yang paling masuk akal. Inferensi harus proporsional terhadap kualitas data. Semakin spesifik label yang diberikan, semakin kuat pula bukti yang diperlukan (Shiddiqi et al., 2024; Treagust, 1988).

Secara operasional, guru dapat menggunakan analisis respons tertulis untuk menelusuri kemampuan transfer. Data tersebut perlu dibaca bersama kualitas model mental yang dieksternalisasi, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menyimpulkan miskonsepsi hanya dari satu jawaban. Untuk mengurangnya, guru dapat mengulang diagnosis setelah remediasi dan menyusun tindak lanjut berbeda untuk kategori respons yang berbeda. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai perubahan konseptual dan kestabilannya, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan menilai perubahan konseptual dan kestabilannya melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan menilai perubahan konseptual dan kestabilannya;
- Bandingkan bukti menilai perubahan konseptual dan kestabilannya pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek menilai perubahan konseptual dan kestabilannya dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian menilai perubahan konseptual dan kestabilannya, topik katalis dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa katalis dianggap menambah jumlah produk kesetimbangan. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta analisis jalur energi dan laju maju-balik. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa katalis mempercepat tercapainya kesetimbangan tanpa mengubah K . Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menempatkan diagnosis sebagai dasar keputusan pedagogis yang spesifik, proporsional, dan dapat ditindaklanjuti, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, menilai perubahan konseptual dan kestabilannya harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.2.16 Menguji transfer pemahaman ke konteks baru

Penerapan menguji transfer pemahaman ke konteks baru harus menghubungkan teori pengukuran dengan pengetahuan pedagogis kimia. Instrumen yang teknis tetapi tidak memahami struktur konsep akan menghasilkan kategori dangkal; sebaliknya, pengetahuan konten tanpa prosedur pengumpulan bukti dapat menghasilkan kesimpulan intuitif yang sulit diuji. Keduanya perlu dipadukan dalam satu alur yang transparan (Treagust, 1988; Black & Wiliam, 2009).

Secara operasional, guru dapat menggunakan diskusi argumentatif untuk menelusuri stabilitas setelah jeda waktu. Data tersebut perlu dibaca bersama respons terhadap probing, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mencampur kesalahan matematika dengan kesalahan konsep. Untuk mengurangnya, guru dapat mencatat derajat keyakinan dan membandingkan jawaban dengan alasan. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai menguji transfer pemahaman ke konteks baru, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan menguji transfer pemahaman ke konteks baru melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan menguji transfer pemahaman ke konteks baru;

- Bandingkan bukti menguji transfer pemahaman ke konteks baru pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek menguji transfer pemahaman ke konteks baru dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian menguji transfer pemahaman ke konteks baru, topik elektrolisis dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa anoda selalu dianggap bermuatan negatif. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta pembedaan definisi anoda-katoda berdasarkan proses oksidasi-reduksi. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa tanda elektroda bergantung jenis sel, tetapi anoda tetap lokasi oksidasi. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menempatkan diagnosis sebagai dasar keputusan pedagogis yang spesifik, proporsional, dan dapat ditindaklanjuti, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Kerangka operasional

Aspek	Bukti yang dicari	Risiko interpretasi	Tindak lanjut
Target konsep	Penjelasan tentang elektrolisis	Target terlalu luas	Batasi subkonsep dan prasyarat
Respons	Stabilitas setelah jeda waktu	Mencampur kesalahan matematika dengan kesalahan konsep	Mencatat derajat keyakinan
Penalaran	Respons terhadap probing	Alasan diperlakukan sekadar tambahan	Membandingkan jawaban dengan alasan
Konfirmasi	Diskusi argumentatif atau tugas isomorfik	Kesimpulan dibuat dari satu konteks	Gunakan bukti kedua dan rediagnosis

Implikasinya, menguji transfer pemahaman ke konteks baru harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.2.17 Mengidentifikasi hubungan antara konsep dan prosedur

Mengidentifikasi hubungan antara konsep dan prosedur tidak berdiri sebagai prosedur administratif. Ia menjadi bagian dari keputusan tentang apa yang perlu dijelaskan ulang,

representasi apa yang perlu dibandingkan, dan siapa yang memerlukan dukungan berbeda. Oleh sebab itu, bukti diagnosis harus langsung terhubung dengan tindakan pembelajaran, bukan berhenti pada daftar kesalahan (Black & Wiliam, 2009; Taber, 2002a).

Secara operasional, guru dapat menggunakan analisis grafik untuk menelusuri urutan langkah penyelesaian. Data tersebut perlu dibaca bersama ketepatan pilihan, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menggunakan ambang kategori tanpa bukti. Untuk mengurangnya, guru dapat menelusuri langkah penyelesaian dan menggunakan pertanyaan tindak lanjut. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai mengidentifikasi hubungan antara konsep dan prosedur, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan mengidentifikasi hubungan antara konsep dan prosedur melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan mengidentifikasi hubungan antara konsep dan prosedur;
- Bandingkan bukti mengidentifikasi hubungan antara konsep dan prosedur pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek mengidentifikasi hubungan antara konsep dan prosedur dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian mengidentifikasi hubungan antara konsep dan prosedur, topik pengukuran laboratorium dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa angka lebih banyak selalu dianggap lebih akurat. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta perbedaan akurasi, presisi, resolusi, dan angka penting. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa kualitas data ditentukan oleh desain, kalibrasi, dan ketidakpastian. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menempatkan diagnosis sebagai dasar keputusan pedagogis yang spesifik, proporsional, dan dapat ditindaklanjuti, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, mengidentifikasi hubungan antara konsep dan prosedur harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.2.18 Mengurangi keputusan pembelajaran yang hanya berbasis skor

Mengurangi keputusan pembelajaran yang hanya berbasis skor perlu ditempatkan dalam kerangka diagnosis yang menempatkan diagnosis sebagai dasar keputusan pedagogis yang spesifik, proporsional, dan dapat ditindaklanjuti. Fokusnya bukan sekadar menghasilkan skor, melainkan memperoleh bukti yang cukup untuk menjelaskan bagaimana peserta didik membangun, menghubungkan, dan menggunakan konsep. Karena konsepsi tidak dapat diamati secara langsung, guru harus menarik inferensi dari respons verbal, tulisan, gambar, tindakan, dan perubahan strategi yang tampak selama penyelesaian tugas (Taber, 2002a; Shiddiqi et al., 2024).

Secara operasional, guru dapat menggunakan tes empat tingkat untuk menelusuri kualitas model mental yang dieksternalisasi. Data tersebut perlu dibaca bersama koherensi alasan, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah memperlakukan konsepsi sebagai sifat permanen. Untuk menguranginya, guru dapat menggunakan triangulasi bukti dan menyajikan tugas paralel. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai mengurangi keputusan pembelajaran yang hanya berbasis skor, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan mengurangi keputusan pembelajaran yang hanya berbasis skor melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan mengurangi keputusan pembelajaran yang hanya berbasis skor;
- Bandingkan bukti mengurangi keputusan pembelajaran yang hanya berbasis skor pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek mengurangi keputusan pembelajaran yang hanya berbasis skor dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian mengurangi keputusan pembelajaran yang hanya berbasis skor, topik ikatan ion dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa NaCl dipandang sebagai molekul diskrit. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta gambar partikel padatan NaCl dan penjelasan hubungan struktur-kisi dengan titik leleh. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa senyawa ionik membentuk kisi dan rumusnya menyatakan rasio ion. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menempatkan diagnosis sebagai dasar keputusan pedagogis yang spesifik, proporsional, dan dapat ditindaklanjuti, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, mengurangi keputusan pembelajaran yang hanya berbasis skor harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.2.19 Mendukung komunikasi guru, peserta didik, dan orang tua

Pembahasan tentang mendukung komunikasi guru, peserta didik, dan orang tua berangkat dari prinsip bahwa data asesmen selalu merupakan jejak tidak langsung dari proses berpikir. Satu pilihan jawaban hanya menunjukkan produk akhir, sedangkan struktur pengetahuan yang melahirkannya dapat berbeda. Diagnosis yang kuat karena itu menggabungkan respons terhadap probing, tingkat keyakinan, dan bukti kontekstual sehingga keputusan pedagogis tidak bergantung pada satu indikator saja (Shiddiqi et al., 2024; Treagust, 1988).

Secara operasional, guru dapat menggunakan think-aloud untuk menelusuri respons terhadap probing. Data tersebut perlu dibaca bersama tingkat keyakinan, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mengabaikan ketidakpastian pengukuran. Untuk menguranginya, guru dapat menyusun tindak lanjut berbeda untuk kategori respons yang berbeda dan memeriksa konsep prasyarat. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai mendukung komunikasi guru, peserta didik, dan orang tua, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan mendukung komunikasi guru, peserta didik, dan orang tua melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan mendukung komunikasi guru, peserta didik, dan orang tua;
- Bandingkan bukti mendukung komunikasi guru, peserta didik, dan orang tua pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek mendukung komunikasi guru, peserta didik, dan orang tua dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian mendukung komunikasi guru, peserta didik, dan orang tua, topik pereaksi pembatas dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa pereaksi dengan massa terkecil selalu dianggap pembatas. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta analisis rasio mol terhadap koefisien. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa status pembatas ditentukan oleh hubungan jumlah zat dan stoikiometri. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menempatkan diagnosis sebagai

dasar keputusan pedagogis yang spesifik, proporsional, dan dapat ditindaklanjuti, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, mendukung komunikasi guru, peserta didik, dan orang tua harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.2.20 Menyediakan data bagi pengembangan kurikulum

Dalam pembelajaran kimia, menyediakan data bagi pengembangan kurikulum memiliki arti strategis karena banyak jawaban dapat diperoleh melalui hafalan prosedur atau pengenalan pola. Peserta didik mungkin tampak berhasil tetapi tidak mampu menerjemahkan makna simbol, model, atau besaran yang digunakan. Prinsip diagnosis menuntut bukti tentang cara konsep dipakai, bukan hanya apakah hasil akhirnya sesuai kunci (Treagust, 1988; Black & Wiliam, 2009).

Secara operasional, guru dapat menggunakan predict-observe-explain untuk menelusuri ketepatan pilihan. Data tersebut perlu dibaca bersama konsistensi lintas butir, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menyimpulkan miskonsepsi hanya dari satu jawaban. Untuk mengurangnya, guru dapat membandingkan jawaban dengan alasan dan meminta translasi antarrepresentasi. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai menyediakan data bagi pengembangan kurikulum, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan menyediakan data bagi pengembangan kurikulum melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan menyediakan data bagi pengembangan kurikulum;
- Bandingkan bukti menyediakan data bagi pengembangan kurikulum pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek menyediakan data bagi pengembangan kurikulum dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian menyediakan data bagi pengembangan kurikulum, topik larutan penyangga dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa penyangga dianggap mempertahankan pH secara mutlak. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah

tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta analisis perubahan komposisi setelah penambahan asam atau basa. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa penyangga membatasi perubahan dalam kapasitas tertentu, bukan meniadakannya. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menempatkan diagnosis sebagai dasar keputusan pedagogis yang spesifik, proporsional, dan dapat ditindaklanjuti, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Kerangka operasional

Aspek	Bukti yang dicari	Risiko interpretasi	Tindak lanjut
Target konsep	Penjelasan tentang larutan penyangga	Target terlalu luas	Batasi subkonsep dan prasyarat
Respons	Ketepatan pilihan	Menyimpulkan miskonsepsi hanya dari satu jawaban	Membandingkan jawaban dengan alasan
Penalaran	Konsistensi lintas butir	Alasan diperlakukan sekadar tambahan	Meminta translasi antarrepresentasi
Konfirmasi	Predict-observe-explain atau tugas isomorfik	Kesimpulan dibuat dari satu konteks	Gunakan bukti kedua dan rediagnosis

Implikasinya, menyediakan data bagi pengembangan kurikulum harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.2.21 Mendukung penelitian tindakan kelas

Aspek mendukung penelitian tindakan kelas menegaskan bahwa diagnosis adalah kegiatan ilmiah berskala kelas: guru merumuskan pertanyaan, mengumpulkan bukti, mempertimbangkan penjelasan alternatif, lalu menentukan kesimpulan yang paling masuk akal. Inferensi harus proporsional terhadap kualitas data. Semakin spesifik label yang diberikan, semakin kuat pula bukti yang diperlukan (Black & Wiliam, 2009; Taber, 2002a).

Secara operasional, guru dapat menggunakan tugas isomorfik untuk menelusuri koherensi alasan. Data tersebut perlu dibaca bersama konsistensi lintas representasi, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mencampur kesalahan matematika dengan kesalahan konsep. Untuk mengurangnya,

guru dapat menggunakan pertanyaan tindak lanjut dan mengulang diagnosis setelah remediasi. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai mendukung penelitian tindakan kelas, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan mendukung penelitian tindakan kelas melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan mendukung penelitian tindakan kelas;
- Bandingkan bukti mendukung penelitian tindakan kelas pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek mendukung penelitian tindakan kelas dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian mendukung penelitian tindakan kelas, topik katalis dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa katalis dianggap menambah jumlah produk kesetimbangan. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta analisis jalur energi dan laju maju-balik. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa katalis mempercepat tercapainya kesetimbangan tanpa mengubah K. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menempatkan diagnosis sebagai dasar keputusan pedagogis yang spesifik, proporsional, dan dapat ditindaklanjuti, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, mendukung penelitian tindakan kelas harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.2.22 Menjaga fokus diagnosis pada perbaikan, bukan pelabelan

Penerapan menjaga fokus diagnosis pada perbaikan, bukan pelabelan harus menghubungkan teori pengukuran dengan pengetahuan pedagogis kimia. Instrumen yang teknis tetapi tidak memahami struktur konsep akan menghasilkan kategori dangkal; sebaliknya, pengetahuan konten tanpa prosedur pengumpulan bukti dapat menghasilkan kesimpulan intuitif yang sulit diuji. Keduanya perlu dipadukan dalam satu alur yang transparan (Taber, 2002a; Shiddiqi et al., 2024).

Secara operasional, guru dapat menggunakan kuis formatif digital untuk menelusuri tingkat keyakinan. Data tersebut perlu dibaca bersama kemampuan transfer, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menggunakan ambang kategori tanpa bukti. Untuk mengurangnya, guru dapat menyajikan tugas paralel dan mencatat derajat keyakinan. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai menjaga fokus diagnosis pada perbaikan, bukan pelabelan, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan menjaga fokus diagnosis pada perbaikan, bukan pelabelan melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan menjaga fokus diagnosis pada perbaikan, bukan pelabelan;
- Bandingkan bukti menjaga fokus diagnosis pada perbaikan, bukan pelabelan pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek menjaga fokus diagnosis pada perbaikan, bukan pelabelan dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian menjaga fokus diagnosis pada perbaikan, bukan pelabelan, topik elektrolisis dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa anoda selalu dianggap bermuatan negatif. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta pembedaan definisi anoda-katoda berdasarkan proses oksidasi-reduksi. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa tanda elektroda bergantung jenis sel, tetapi anoda tetap lokasi oksidasi. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menempatkan diagnosis sebagai dasar keputusan pedagogis yang spesifik, proporsional, dan dapat ditindaklanjuti, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, menjaga fokus diagnosis pada perbaikan, bukan pelabelan harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.2.23 Indikator ketercapaian tujuan diagnosis

Indikator ketercapaian tujuan diagnosis tidak berdiri sebagai prosedur administratif. Ia menjadi bagian dari keputusan tentang apa yang perlu dijelaskan ulang, representasi apa yang

perlu dibandingkan, dan siapa yang memerlukan dukungan berbeda. Oleh sebab itu, bukti diagnosis harus langsung terhubung dengan tindakan pembelajaran, bukan berhenti pada daftar kesalahan (Shiddiqi et al., 2024; Treagust, 1988).

Secara operasional, guru dapat menggunakan tes tiga tingkat untuk menelusuri konsistensi lintas butir. Data tersebut perlu dibaca bersama stabilitas setelah jeda waktu, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah memperlakukan konsepsi sebagai sifat permanen. Untuk mengurangnya, guru dapat memeriksa konsep prasyarat dan menelusuri langkah penyelesaian. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai indikator ketercapaian tujuan diagnosis, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan indikator ketercapaian tujuan diagnosis melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan indikator ketercapaian tujuan diagnosis;
- Bandingkan bukti indikator ketercapaian tujuan diagnosis pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek indikator ketercapaian tujuan diagnosis dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian indikator ketercapaian tujuan diagnosis, topik pengukuran laboratorium dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa angka lebih banyak selalu dianggap lebih akurat. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta pembedaan akurasi, presisi, resolusi, dan angka penting. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa kualitas data ditentukan oleh desain, kalibrasi, dan ketidakpastian. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menempatkan diagnosis sebagai dasar keputusan pedagogis yang spesifik, proporsional, dan dapat ditindaklanjuti, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, indikator ketercapaian tujuan diagnosis harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.2.24 Sintesis tujuan diagnosis miskonsepsi

Sintesis tujuan diagnosis miskonsepsi perlu ditempatkan dalam kerangka diagnosis yang menempatkan diagnosis sebagai dasar keputusan pedagogis yang spesifik, proporsional, dan dapat ditindaklanjuti. Fokusnya bukan sekadar menghasilkan skor, melainkan memperoleh bukti yang cukup untuk menjelaskan bagaimana peserta didik membangun, menghubungkan, dan menggunakan konsep. Karena konsepsi tidak dapat diamati secara langsung, guru harus menarik inferensi dari respons verbal, tulisan, gambar, tindakan, dan perubahan strategi yang tampak selama penyelesaian tugas (Treagust, 1988; Black & Wiliam, 2009).

Secara operasional, guru dapat menggunakan wawancara semi-terstruktur untuk menelusuri konsistensi lintas representasi. Data tersebut perlu dibaca bersama urutan langkah penyelesaian, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mengabaikan ketidakpastian pengukuran. Untuk menguranginya, guru dapat meminta translasi antarrepresentasi dan menggunakan triangulasi bukti. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai sintesis tujuan diagnosis miskonsepsi, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan sintesis tujuan diagnosis miskonsepsi melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan sintesis tujuan diagnosis miskonsepsi;
- Bandingkan bukti sintesis tujuan diagnosis miskonsepsi pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek sintesis tujuan diagnosis miskonsepsi dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian sintesis tujuan diagnosis miskonsepsi, topik ikatan ion dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa NaCl dipandang sebagai molekul diskrit. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta gambar partikel padatan NaCl dan penjelasan hubungan struktur-kisi dengan titik leleh. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa senyawa ionik membentuk kisi dan rumusnya menyatakan rasio ion. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menempatkan diagnosis sebagai dasar keputusan pedagogis yang spesifik, proporsional, dan dapat ditindaklanjuti, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, sintesis tujuan diagnosis miskonsepsi harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.2.25 Implikasi praktis bagi guru kimia

Implikasi praktis subbab 5.2 dapat dirangkum sebagai alur: tentukan konsep target, pilih bukti yang relevan, kumpulkan respons dalam lebih dari satu bentuk, analisis pola, nyatakan derajat kepastian, susun tindak lanjut, dan lakukan rediagnosis. Alur ini menjaga agar tujuan diagnosis miskonsepsi tidak berubah menjadi administrasi skor, tetapi berfungsi sebagai mekanisme perbaikan pembelajaran yang berulang.

8. Identifikasi konsep inti dan konsep prasyarat sebelum menyusun instrumen.
9. Gunakan pertanyaan yang meminta alasan, representasi, atau langkah kerja.
10. Bandingkan respons pada beberapa konteks dan waktu.
11. Pisahkan bukti faktual dari interpretasi guru.
12. Tentukan tindak lanjut yang berbeda untuk kategori kebutuhan yang berbeda.
13. Simpan data secukupnya dan lindungi kerahasiaan peserta didik.
14. Periksa kembali perubahan melalui tugas transfer dan retensi.

5.3 Diagnosis sebelum, selama, dan setelah Pembelajaran

Subbab ini mengintegrasikan diagnosis secara berkelanjutan pada fase awal, proses, dan akhir pembelajaran. Uraian disusun dari prinsip teoretis menuju prosedur kelas, contoh kimia, risiko interpretasi, dan tindak lanjut. Setiap bagian menekankan bahwa kesimpulan diagnostik harus didasarkan pada pola bukti, bukan pada satu respons yang terpisah.

5.3.1 Prinsip kontinuitas diagnosis

Pembahasan tentang prinsip kontinuitas diagnosis berangkat dari prinsip bahwa data asesmen selalu merupakan jejak tidak langsung dari proses berpikir. Satu pilihan jawaban hanya menunjukkan produk akhir, sedangkan struktur pengetahuan yang melahirkannya dapat berbeda. Diagnosis yang kuat karena itu menggabungkan koherensi alasan, konsistensi lintas representasi, dan bukti kontekstual sehingga keputusan pedagogis tidak bergantung pada satu indikator saja (Sadler, 1989; Anggraena et al., 2025).

Secara operasional, guru dapat menggunakan peta konsep untuk menelusuri koherensi alasan. Data tersebut perlu dibaca bersama konsistensi lintas representasi, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mencampur kesalahan matematika dengan kesalahan konsep. Untuk mengurangnya, guru dapat menggunakan pertanyaan tindak lanjut dan mengulang diagnosis setelah

remediasi. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai prinsip kontinuitas diagnosis, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan prinsip kontinuitas diagnosis melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan prinsip kontinuitas diagnosis;
- Bandingkan bukti prinsip kontinuitas diagnosis pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek prinsip kontinuitas diagnosis dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian prinsip kontinuitas diagnosis, topik gas dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa partikel gas dianggap mengembang ketika dipanaskan. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta gambar partikel sebelum dan sesudah pemanasan pada kondisi berbeda. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa ukuran partikel tidak berubah secara bermakna; distribusi kecepatan dan jarak berubah. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengintegrasikan diagnosis secara berkelanjutan pada fase awal, proses, dan akhir pembelajaran, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, prinsip kontinuitas diagnosis harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.3.2 Diagnosis sebelum pembelajaran sebagai pemetaan kesiapan

Dalam pembelajaran kimia, diagnosis sebelum pembelajaran sebagai pemetaan kesiapan memiliki arti strategis karena banyak jawaban dapat diperoleh melalui hafalan prosedur atau pengenalan pola. Peserta didik mungkin tampak berhasil tetapi tidak mampu menerjemahkan makna simbol, model, atau besaran yang digunakan. Prinsip diagnosis menuntut bukti tentang cara konsep dipakai, bukan hanya apakah hasil akhirnya sesuai kunci (Anggraena et al., 2025; Shute, 2008).

Secara operasional, guru dapat menggunakan analisis respons tertulis untuk menelusuri tingkat keyakinan. Data tersebut perlu dibaca bersama kemampuan transfer, karena pola yang

sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menggunakan ambang kategori tanpa bukti. Untuk mengurangnya, guru dapat menyajikan tugas paralel dan mencatat derajat keyakinan. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai diagnosis sebelum pembelajaran sebagai pemetaan kesiapan, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan diagnosis sebelum pembelajaran sebagai pemetaan kesiapan melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan diagnosis sebelum pembelajaran sebagai pemetaan kesiapan;
- Bandingkan bukti diagnosis sebelum pembelajaran sebagai pemetaan kesiapan pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek diagnosis sebelum pembelajaran sebagai pemetaan kesiapan dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian diagnosis sebelum pembelajaran sebagai pemetaan kesiapan, topik kesetimbangan dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa reaksi dianggap berhenti saat setimbang. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta hubungan grafik laju, konsentrasi, dan model partikel. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa kesetimbangan bersifat dinamis dengan laju maju dan balik yang sama. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengintegrasikan diagnosis secara berkelanjutan pada fase awal, proses, dan akhir pembelajaran, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, diagnosis sebelum pembelajaran sebagai pemetaan kesiapan harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.3.3 Menentukan konsep prasyarat yang harus diperiksa

Aspek menentukan konsep prasyarat yang harus diperiksa menegaskan bahwa diagnosis adalah kegiatan ilmiah berskala kelas: guru merumuskan pertanyaan, mengumpulkan bukti, mempertimbangkan penjelasan alternatif, lalu menentukan kesimpulan yang paling masuk

akal. Inferensi harus proporsional terhadap kualitas data. Semakin spesifik label yang diberikan, semakin kuat pula bukti yang diperlukan (Shute, 2008; Black & Wiliam, 2009).

Secara operasional, guru dapat menggunakan diskusi argumentatif untuk menelusuri konsistensi lintas butir. Data tersebut perlu dibaca bersama stabilitas setelah jeda waktu, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah memperlakukan konsepsi sebagai sifat permanen. Untuk menguranginya, guru dapat memeriksa konsep prasyarat dan menelusuri langkah penyelesaian. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai menentukan konsep prasyarat yang harus diperiksa, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan menentukan konsep prasyarat yang harus diperiksa melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan menentukan konsep prasyarat yang harus diperiksa;
- Bandingkan bukti menentukan konsep prasyarat yang harus diperiksa pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek menentukan konsep prasyarat yang harus diperiksa dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian menentukan konsep prasyarat yang harus diperiksa, topik termokimia dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa pemutusan ikatan dianggap melepaskan energi. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta analisis energi yang diserap saat memutuskan dan dilepas saat membentuk ikatan. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa perubahan entalpi bersih berasal dari selisih proses pemutusan dan pembentukan. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengintegrasikan diagnosis secara berkelanjutan pada fase awal, proses, dan akhir pembelajaran, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, menentukan konsep prasyarat yang harus diperiksa harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.3.4 Menyusun asesmen awal yang singkat tetapi bermakna

Penerapan menyusun asesmen awal yang singkat tetapi bermakna harus menghubungkan teori pengukuran dengan pengetahuan pedagogis kimia. Instrumen yang teknis tetapi tidak memahami struktur konsep akan menghasilkan kategori dangkal; sebaliknya, pengetahuan konten tanpa prosedur pengumpulan bukti dapat menghasilkan kesimpulan intuitif yang sulit diuji. Keduanya perlu dipadukan dalam satu alur yang transparan (Black & Wiliam, 2009; Sadler, 1989).

Secara operasional, guru dapat menggunakan analisis grafik untuk menelusuri konsistensi lintas representasi. Data tersebut perlu dibaca bersama urutan langkah penyelesaian, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mengabaikan ketidakpastian pengukuran. Untuk mengurangnya, guru dapat meminta translasi antarrepresentasi dan menggunakan triangulasi bukti. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai menyusun asesmen awal yang singkat tetapi bermakna, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan menyusun asesmen awal yang singkat tetapi bermakna melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan menyusun asesmen awal yang singkat tetapi bermakna;
- Bandingkan bukti menyusun asesmen awal yang singkat tetapi bermakna pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek menyusun asesmen awal yang singkat tetapi bermakna dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian menyusun asesmen awal yang singkat tetapi bermakna, topik hidrokarbon dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa isomer dipandang sebagai zat yang sama karena rumus molekulnya sama. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta hubungan struktur, konektivitas, dan sifat. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa rumus molekul tidak cukup menentukan struktur dan identitas senyawa. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengintegrasikan diagnosis secara berkelanjutan pada fase awal, proses, dan akhir pembelajaran, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Kerangka operasional

Aspek	Bukti yang dicari	Risiko interpretasi	Tindak lanjut
Target konsep	Penjelasan tentang hidrokarbon	Target terlalu luas	Batasi subkonsep dan prasyarat
Respons	Konsistensi lintas representasi	Mengabaikan ketidakpastian pengukuran	Meminta translasi antarrepresentasi
Penalaran	Urutan langkah penyelesaian	Alasan diperlakukan sekadar tambahan	Menggunakan triangulasi bukti
Konfirmasi	Analisis grafik atau tugas isomorfik	Kesimpulan dibuat dari satu konteks	Gunakan bukti kedua dan rediagnosis

Implikasinya, menyusun asesmen awal yang singkat tetapi bermakna harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.3.5 Menggunakan fenomena pemantik untuk elisitasi gagasan awal

Menggunakan fenomena pemantik untuk elisitasi gagasan awal tidak berdiri sebagai prosedur administratif. Ia menjadi bagian dari keputusan tentang apa yang perlu dijelaskan ulang, representasi apa yang perlu dibandingkan, dan siapa yang memerlukan dukungan berbeda. Oleh sebab itu, bukti diagnosis harus langsung terhubung dengan tindakan pembelajaran, bukan berhenti pada daftar kesalahan (Sadler, 1989; Anggraena et al., 2025).

Secara operasional, guru dapat menggunakan tes empat tingkat untuk menelusuri kemampuan transfer. Data tersebut perlu dibaca bersama kualitas model mental yang dieksternalisasi, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menyimpulkan miskonsepsi hanya dari satu jawaban. Untuk mengurangnya, guru dapat mengulang diagnosis setelah remediasi dan menyusun tindak lanjut berbeda untuk kategori respons yang berbeda. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai menggunakan fenomena pemantik untuk elisitasi gagasan awal, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan menggunakan fenomena pemantik untuk elisitasi gagasan awal melalui penggunaan konsep, bukan sekadar pengenalan istilah;

- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan menggunakan fenomena pemantik untuk elisitasi gagasan awal;
- Bandingkan bukti menggunakan fenomena pemantik untuk elisitasi gagasan awal pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek menggunakan fenomena pemantik untuk elisitasi gagasan awal dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian menggunakan fenomena pemantik untuk elisitasi gagasan awal, topik spektroskopi dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa spektrum dipandang sebagai gambar langsung molekul. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta hubungan sinyal, transisi energi, dan model struktur. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa spektrum adalah data yang ditafsirkan melalui model dan aturan seleksi. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengintegrasikan diagnosis secara berkelanjutan pada fase awal, proses, dan akhir pembelajaran, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, menggunakan fenomena pemantik untuk elisitasi gagasan awal harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.3.6 Diagnosis selama pembelajaran melalui pertanyaan berjenjang

Diagnosis selama pembelajaran melalui pertanyaan berjenjang perlu ditempatkan dalam kerangka diagnosis yang mengintegrasikan diagnosis secara berkelanjutan pada fase awal, proses, dan akhir pembelajaran. Fokusnya bukan sekadar menghasilkan skor, melainkan memperoleh bukti yang cukup untuk menjelaskan bagaimana peserta didik membangun, menghubungkan, dan menggunakan konsep. Karena konsepsi tidak dapat diamati secara langsung, guru harus menarik inferensi dari respons verbal, tulisan, gambar, tindakan, dan perubahan strategi yang tampak selama penyelesaian tugas (Anggraena et al., 2025; Shute, 2008).

Secara operasional, guru dapat menggunakan think-aloud untuk menelusuri stabilitas setelah jeda waktu. Data tersebut perlu dibaca bersama respons terhadap probing, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mencampur kesalahan matematika dengan kesalahan konsep. Untuk mengurangnya, guru dapat mencatat derajat keyakinan dan membandingkan jawaban dengan alasan.

Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai diagnosis selama pembelajaran melalui pertanyaan berjenjang, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan diagnosis selama pembelajaran melalui pertanyaan berjenjang melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan diagnosis selama pembelajaran melalui pertanyaan berjenjang;
- Bandingkan bukti diagnosis selama pembelajaran melalui pertanyaan berjenjang pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek diagnosis selama pembelajaran melalui pertanyaan berjenjang dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian diagnosis selama pembelajaran melalui pertanyaan berjenjang, topik ikatan kovalen dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa garis ikatan dipandang sebagai batang fisik. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta penjelasan energi potensial dan distribusi elektron pada jarak antarinti berbeda. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa representasi garis adalah konvensi yang merangkum pasangan elektron dan konektivitas. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengintegrasikan diagnosis secara berkelanjutan pada fase awal, proses, dan akhir pembelajaran, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, diagnosis selama pembelajaran melalui pertanyaan berjenjang harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.3.7 Diagnosis melalui diskusi kelas dan argumentasi

Pembahasan tentang diagnosis melalui diskusi kelas dan argumentasi berangkat dari prinsip bahwa data asesmen selalu merupakan jejak tidak langsung dari proses berpikir. Satu pilihan jawaban hanya menunjukkan produk akhir, sedangkan struktur pengetahuan yang melahirkannya dapat berbeda. Diagnosis yang kuat karena itu menggabungkan urutan langkah penyelesaian, ketepatan pilihan, dan bukti kontekstual sehingga keputusan pedagogis tidak bergantung pada satu indikator saja (Shute, 2008; Black & Wiliam, 2009).

Secara operasional, guru dapat menggunakan predict-observe-explain untuk menelusuri urutan langkah penyelesaian. Data tersebut perlu dibaca bersama ketepatan pilihan, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menggunakan ambang kategori tanpa bukti. Untuk menguranginya, guru dapat menelusuri langkah penyelesaian dan menggunakan pertanyaan tindak lanjut. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai diagnosis melalui diskusi kelas dan argumentasi, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan diagnosis melalui diskusi kelas dan argumentasi melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan diagnosis melalui diskusi kelas dan argumentasi;
- Bandingkan bukti diagnosis melalui diskusi kelas dan argumentasi pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek diagnosis melalui diskusi kelas dan argumentasi dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian diagnosis melalui diskusi kelas dan argumentasi, topik gas dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa partikel gas dianggap mengembang ketika dipanaskan. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta gambar partikel sebelum dan sesudah pemanasan pada kondisi berbeda. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa ukuran partikel tidak berubah secara bermakna; distribusi kecepatan dan jarak berubah. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengintegrasikan diagnosis secara berkelanjutan pada fase awal, proses, dan akhir pembelajaran, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, diagnosis melalui diskusi kelas dan argumentasi harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.3.8 Diagnosis melalui kerja kelompok dan interaksi sosial

Dalam pembelajaran kimia, diagnosis melalui kerja kelompok dan interaksi sosial memiliki arti strategis karena banyak jawaban dapat diperoleh melalui hafalan prosedur atau pengenalan

pola. Peserta didik mungkin tampak berhasil tetapi tidak mampu menerjemahkan makna simbol, model, atau besaran yang digunakan. Prinsip diagnosis menuntut bukti tentang cara konsep dipakai, bukan hanya apakah hasil akhirnya sesuai kunci (Black & Wiliam, 2009; Sadler, 1989).

Secara operasional, guru dapat menggunakan tugas isomorfik untuk menelusuri kualitas model mental yang dieksternalisasi. Data tersebut perlu dibaca bersama koherensi alasan, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah memperlakukan konsepsi sebagai sifat permanen. Untuk mengurangnya, guru dapat menggunakan triangulasi bukti dan menyajikan tugas paralel. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai diagnosis melalui kerja kelompok dan interaksi sosial, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan diagnosis melalui kerja kelompok dan interaksi sosial melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan diagnosis melalui kerja kelompok dan interaksi sosial;
- Bandingkan bukti diagnosis melalui kerja kelompok dan interaksi sosial pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek diagnosis melalui kerja kelompok dan interaksi sosial dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian diagnosis melalui kerja kelompok dan interaksi sosial, topik kesetimbangan dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa reaksi dianggap berhenti saat setimbang. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta hubungan grafik laju, konsentrasi, dan model partikel. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa kesetimbangan bersifat dinamis dengan laju maju dan balik yang sama. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengintegrasikan diagnosis secara berkelanjutan pada fase awal, proses, dan akhir pembelajaran, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Kerangka operasional

Aspek	Bukti yang dicari	Risiko interpretasi	Tindak lanjut
Target konsep	Penjelasan tentang kesetimbangan	Target terlalu luas	Batasi subkonsep dan prasyarat

Aspek	Bukti yang dicari	Risiko interpretasi	Tindak lanjut
Respons	Kualitas model mental yang dieksternalisasi	Memperlakukan konsepsi sebagai sifat permanen	Menggunakan triangulasi bukti
Penalaran	Koherensi alasan	Alasan diperlakukan sekadar tambahan	Menyajikan tugas paralel
Konfirmasi	Tugas isomorfik atau tugas isomorfik	Kesimpulan dibuat dari satu konteks	Gunakan bukti kedua dan rediagnosis

Implikasinya, diagnosis melalui kerja kelompok dan interaksi sosial harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.3.9 Diagnosis melalui praktikum dan demonstrasi

Aspek diagnosis melalui praktikum dan demonstrasi menegaskan bahwa diagnosis adalah kegiatan ilmiah berskala kelas: guru merumuskan pertanyaan, mengumpulkan bukti, mempertimbangkan penjelasan alternatif, lalu menentukan kesimpulan yang paling masuk akal. Inferensi harus proporsional terhadap kualitas data. Semakin spesifik label yang diberikan, semakin kuat pula bukti yang diperlukan (Sadler, 1989; Anggraena et al., 2025).

Secara operasional, guru dapat menggunakan kuis formatif digital untuk menelusuri respons terhadap probing. Data tersebut perlu dibaca bersama tingkat keyakinan, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mengabaikan ketidakpastian pengukuran. Untuk mengurangnya, guru dapat menyusun tindak lanjut berbeda untuk kategori respons yang berbeda dan memeriksa konsep prasyarat. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai diagnosis melalui praktikum dan demonstrasi, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan diagnosis melalui praktikum dan demonstrasi melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan diagnosis melalui praktikum dan demonstrasi;
- Bandingkan bukti diagnosis melalui praktikum dan demonstrasi pada lebih dari satu konteks, format tugas, atau bentuk representasi;

- Hubungkan kategori hasil pada aspek diagnosis melalui praktikum dan demonstrasi dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian diagnosis melalui praktikum dan demonstrasi, topik termokimia dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa pemutusan ikatan dianggap melepaskan energi. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta analisis energi yang diserap saat memutus dan dilepas saat membentuk ikatan. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa perubahan entalpi bersih berasal dari selisih proses pemutusan dan pembentukan. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengintegrasikan diagnosis secara berkelanjutan pada fase awal, proses, dan akhir pembelajaran, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, diagnosis melalui praktikum dan demonstrasi harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.3.10 Diagnosis melalui representasi yang berubah-ubah

Penerapan diagnosis melalui representasi yang berubah-ubah harus menghubungkan teori pengukuran dengan pengetahuan pedagogis kimia. Instrumen yang teknis tetapi tidak memahami struktur konsep akan menghasilkan kategori dangkal; sebaliknya, pengetahuan konten tanpa prosedur pengumpulan bukti dapat menghasilkan kesimpulan intuitif yang sulit diuji. Keduanya perlu dipadukan dalam satu alur yang transparan (Anggraena et al., 2025; Shute, 2008).

Secara operasional, guru dapat menggunakan tes tiga tingkat untuk menelusuri ketepatan pilihan. Data tersebut perlu dibaca bersama konsistensi lintas butir, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menyimpulkan miskonsepsi hanya dari satu jawaban. Untuk mengurangnya, guru dapat membandingkan jawaban dengan alasan dan meminta translasi antarrepresentasi. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai diagnosis melalui representasi yang berubah-ubah, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan diagnosis melalui representasi yang berubah-ubah melalui penggunaan konsep, bukan sekadar pengenalan istilah;

- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan diagnosis melalui representasi yang berubah-ubah;
- Bandingkan bukti diagnosis melalui representasi yang berubah-ubah pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek diagnosis melalui representasi yang berubah-ubah dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian diagnosis melalui representasi yang berubah-ubah, topik hidrokarbon dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa isomer dipandang sebagai zat yang sama karena rumus molekulnya sama. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta hubungan struktur, konektivitas, dan sifat. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa rumus molekul tidak cukup menentukan struktur dan identitas senyawa. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengintegrasikan diagnosis secara berkelanjutan pada fase awal, proses, dan akhir pembelajaran, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, diagnosis melalui representasi yang berubah-ubah harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.3.11 Exit ticket dan minute paper sebagai diagnosis cepat

Exit ticket dan minute paper sebagai diagnosis cepat tidak berdiri sebagai prosedur administratif. Ia menjadi bagian dari keputusan tentang apa yang perlu dijelaskan ulang, representasi apa yang perlu dibandingkan, dan siapa yang memerlukan dukungan berbeda. Oleh sebab itu, bukti diagnosis harus langsung terhubung dengan tindakan pembelajaran, bukan berhenti pada daftar kesalahan (Shute, 2008; Black & Wiliam, 2009).

Secara operasional, guru dapat menggunakan wawancara semi-terstruktur untuk menelusuri koherensi alasan. Data tersebut perlu dibaca bersama konsistensi lintas representasi, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mencampur kesalahan matematika dengan kesalahan konsep. Untuk mengurangnya, guru dapat menggunakan pertanyaan tindak lanjut dan mengulang diagnosis setelah remediasi. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai exit ticket dan minute paper sebagai diagnosis cepat, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan exit ticket dan minute paper sebagai diagnosis cepat melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan exit ticket dan minute paper sebagai diagnosis cepat;
- Bandingkan bukti exit ticket dan minute paper sebagai diagnosis cepat pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek exit ticket dan minute paper sebagai diagnosis cepat dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian exit ticket dan minute paper sebagai diagnosis cepat, topik spektroskopi dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa spektrum dipandang sebagai gambar langsung molekul. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta hubungan sinyal, transisi energi, dan model struktur. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa spektrum adalah data yang ditafsirkan melalui model dan aturan seleksi. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengintegrasikan diagnosis secara berkelanjutan pada fase awal, proses, dan akhir pembelajaran, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, exit ticket dan minute paper sebagai diagnosis cepat harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.3.12 Penggunaan kuis formatif berulang

Penggunaan kuis formatif berulang perlu ditempatkan dalam kerangka diagnosis yang mengintegrasikan diagnosis secara berkelanjutan pada fase awal, proses, dan akhir pembelajaran. Fokusnya bukan sekadar menghasilkan skor, melainkan memperoleh bukti yang cukup untuk menjelaskan bagaimana peserta didik membangun, menghubungkan, dan menggunakan konsep. Karena konsepsi tidak dapat diamati secara langsung, guru harus menarik inferensi dari respons verbal, tulisan, gambar, tindakan, dan perubahan strategi yang tampak selama penyelesaian tugas (Black & Wiliam, 2009; Sadler, 1989).

Secara operasional, guru dapat menggunakan gambar partikel untuk menelusuri tingkat keyakinan. Data tersebut perlu dibaca bersama kemampuan transfer, karena pola yang sama

pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menggunakan ambang kategori tanpa bukti. Untuk mengurangnya, guru dapat menyajikan tugas paralel dan mencatat derajat keyakinan. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai penggunaan kuis formatif berulang, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan penggunaan kuis formatif berulang melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan penggunaan kuis formatif berulang;
- Bandingkan bukti penggunaan kuis formatif berulang pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek penggunaan kuis formatif berulang dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian penggunaan kuis formatif berulang, topik ikatan kovalen dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa garis ikatan dipandang sebagai batang fisik. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta penjelasan energi potensial dan distribusi elektron pada jarak antarinti berbeda. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa representasi garis adalah konvensi yang merangkum pasangan elektron dan konektivitas. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengintegrasikan diagnosis secara berkelanjutan pada fase awal, proses, dan akhir pembelajaran, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Kerangka operasional

Aspek	Bukti yang dicari	Risiko interpretasi	Tindak lanjut
Target konsep	Penjelasan tentang ikatan kovalen	Target terlalu luas	Batasi subkonsep dan prasyarat
Respons	Tingkat keyakinan	Menggunakan ambang kategori tanpa bukti	Menyajikan tugas paralel
Penalaran	Kemampuan transfer	Alasan diperlakukan sekadar tambahan	Mencatat derajat keyakinan
Konfirmasi	Gambar partikel atau tugas isomorfik	Kesimpulan dibuat dari satu konteks	Gunakan bukti kedua dan rediagnosis

Implikasinya, penggunaan kuis formatif berulang harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.3.13 Diagnosis setelah pembelajaran untuk memeriksa perubahan konsep

Pembahasan tentang diagnosis setelah pembelajaran untuk memeriksa perubahan konsep berangkat dari prinsip bahwa data asesmen selalu merupakan jejak tidak langsung dari proses berpikir. Satu pilihan jawaban hanya menunjukkan produk akhir, sedangkan struktur pengetahuan yang melahirkannya dapat berbeda. Diagnosis yang kuat karena itu menggabungkan konsistensi lintas butir, stabilitas setelah jeda waktu, dan bukti kontekstual sehingga keputusan pedagogis tidak bergantung pada satu indikator saja (Sadler, 1989; Anggraena et al., 2025).

Secara operasional, guru dapat menggunakan observasi praktikum untuk menelusuri konsistensi lintas butir. Data tersebut perlu dibaca bersama stabilitas setelah jeda waktu, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah memperlakukan konsepsi sebagai sifat permanen. Untuk mengurangnya, guru dapat memeriksa konsep prasyarat dan menelusuri langkah penyelesaian. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai diagnosis setelah pembelajaran untuk memeriksa perubahan konsep, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan diagnosis setelah pembelajaran untuk memeriksa perubahan konsep melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan diagnosis setelah pembelajaran untuk memeriksa perubahan konsep;
- Bandingkan bukti diagnosis setelah pembelajaran untuk memeriksa perubahan konsep pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek diagnosis setelah pembelajaran untuk memeriksa perubahan konsep dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian diagnosis setelah pembelajaran untuk memeriksa perubahan konsep, topik gas dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa partikel gas dianggap mengembang ketika dipanaskan. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta gambar partikel sebelum dan sesudah pemanasan pada kondisi berbeda. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa ukuran partikel tidak berubah secara

bermakna; distribusi kecepatan dan jarak berubah. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengintegrasikan diagnosis secara berkelanjutan pada fase awal, proses, dan akhir pembelajaran, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, diagnosis setelah pembelajaran untuk memeriksa perubahan konsep harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.3.14 Membedakan peningkatan skor dan perubahan konseptual

Dalam pembelajaran kimia, membedakan peningkatan skor dan perubahan konseptual memiliki arti strategis karena banyak jawaban dapat diperoleh melalui hafalan prosedur atau pengenalan pola. Peserta didik mungkin tampak berhasil tetapi tidak mampu menerjemahkan makna simbol, model, atau besaran yang digunakan. Prinsip diagnosis menuntut bukti tentang cara konsep dipakai, bukan hanya apakah hasil akhirnya sesuai kunci (Anggraena et al., 2025; Shute, 2008).

Secara operasional, guru dapat menggunakan exit ticket untuk menelusuri konsistensi lintas representasi. Data tersebut perlu dibaca bersama urutan langkah penyelesaian, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mengabaikan ketidakpastian pengukuran. Untuk mengurangnya, guru dapat meminta translasi antarrepresentasi dan menggunakan triangulasi bukti. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai membedakan peningkatan skor dan perubahan konseptual, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan membedakan peningkatan skor dan perubahan konseptual melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan membedakan peningkatan skor dan perubahan konseptual;
- Bandingkan bukti membedakan peningkatan skor dan perubahan konseptual pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek membedakan peningkatan skor dan perubahan konseptual dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian membedakan peningkatan skor dan perubahan konseptual, topik kesetimbangan dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis.

Konsepsi yang sering muncul ialah bahwa reaksi dianggap berhenti saat setimbang. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta hubungan grafik laju, konsentrasi, dan model partikel. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa kesetimbangan bersifat dinamis dengan laju maju dan balik yang sama. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengintegrasikan diagnosis secara berkelanjutan pada fase awal, proses, dan akhir pembelajaran, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, membedakan peningkatan skor dan perubahan konseptual harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.3.15 Diagnosis tertunda untuk menguji retensi

Aspek diagnosis tertunda untuk menguji retensi menegaskan bahwa diagnosis adalah kegiatan ilmiah berskala kelas: guru merumuskan pertanyaan, mengumpulkan bukti, mempertimbangkan penjelasan alternatif, lalu menentukan kesimpulan yang paling masuk akal. Inferensi harus proporsional terhadap kualitas data. Semakin spesifik label yang diberikan, semakin kuat pula bukti yang diperlukan (Shute, 2008; Black & Wiliam, 2009).

Secara operasional, guru dapat menggunakan tes pilihan ganda dua tingkat untuk menelusuri kemampuan transfer. Data tersebut perlu dibaca bersama kualitas model mental yang dieksternalisasi, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menyimpulkan miskonsepsi hanya dari satu jawaban. Untuk mengurangnya, guru dapat mengulang diagnosis setelah remediasi dan menyusun tindak lanjut berbeda untuk kategori respons yang berbeda. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai diagnosis tertunda untuk menguji retensi, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan diagnosis tertunda untuk menguji retensi melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan diagnosis tertunda untuk menguji retensi;
- Bandingkan bukti diagnosis tertunda untuk menguji retensi pada lebih dari satu konteks, format tugas, atau bentuk representasi;

- Hubungkan kategori hasil pada aspek diagnosis tertunda untuk menguji retensi dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian diagnosis tertunda untuk menguji retensi, topik termokimia dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa pemutusan ikatan dianggap melepaskan energi. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta analisis energi yang diserap saat memutuskan dan dilepas saat membentuk ikatan. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa perubahan entalpi bersih berasal dari selisih proses pemutusan dan pembentukan. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengintegrasikan diagnosis secara berkelanjutan pada fase awal, proses, dan akhir pembelajaran, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, diagnosis tertunda untuk menguji retensi harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.3.16 Diagnosis transfer pada konteks baru

Penerapan diagnosis transfer pada konteks baru harus menghubungkan teori pengukuran dengan pengetahuan pedagogis kimia. Instrumen yang teknis tetapi tidak memahami struktur konsep akan menghasilkan kategori dangkal; sebaliknya, pengetahuan konten tanpa prosedur pengumpulan bukti dapat menghasilkan kesimpulan intuitif yang sulit diuji. Keduanya perlu dipadukan dalam satu alur yang transparan (Black & Wiliam, 2009; Sadler, 1989).

Secara operasional, guru dapat menggunakan alasan terbuka untuk menelusuri stabilitas setelah jeda waktu. Data tersebut perlu dibaca bersama respons terhadap probing, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mencampur kesalahan matematika dengan kesalahan konsep. Untuk mengurangnya, guru dapat mencatat derajat keyakinan dan membandingkan jawaban dengan alasan. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai diagnosis transfer pada konteks baru, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan diagnosis transfer pada konteks baru melalui penggunaan konsep, bukan sekadar pengenalan istilah;

- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan diagnosis transfer pada konteks baru;
- Bandingkan bukti diagnosis transfer pada konteks baru pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek diagnosis transfer pada konteks baru dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian diagnosis transfer pada konteks baru, topik hidrokarbon dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa isomer dipandang sebagai zat yang sama karena rumus molekulnya sama. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta hubungan struktur, konektivitas, dan sifat. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa rumus molekul tidak cukup menentukan struktur dan identitas senyawa. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengintegrasikan diagnosis secara berkelanjutan pada fase awal, proses, dan akhir pembelajaran, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Kerangka operasional

Aspek	Bukti yang dicari	Risiko interpretasi	Tindak lanjut
Target konsep	Penjelasan tentang hidrokarbon	Target terlalu luas	Batasi subkonsep dan prasyarat
Respons	Stabilitas setelah jeda waktu	Mencampur kesalahan matematika dengan kesalahan konsep	Mencatat derajat keyakinan
Penalaran	Respons terhadap probing	Alasan diperlakukan sekadar tambahan	Membandingkan jawaban dengan alasan
Konfirmasi	Alasan terbuka atau tugas isomorfik	Kesimpulan dibuat dari satu konteks	Gunakan bukti kedua dan rediagnosis

Implikasinya, diagnosis transfer pada konteks baru harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.3.17 Membandingkan bukti sebelum dan sesudah pembelajaran

Membandingkan bukti sebelum dan sesudah pembelajaran tidak berdiri sebagai prosedur administratif. Ia menjadi bagian dari keputusan tentang apa yang perlu dijelaskan ulang, representasi apa yang perlu dibandingkan, dan siapa yang memerlukan dukungan berbeda. Oleh sebab itu, bukti diagnosis harus langsung terhubung dengan tindakan pembelajaran, bukan berhenti pada daftar kesalahan (Sadler, 1989; Anggraena et al., 2025).

Secara operasional, guru dapat menggunakan peta konsep untuk menelusuri urutan langkah penyelesaian. Data tersebut perlu dibaca bersama ketepatan pilihan, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menggunakan ambang kategori tanpa bukti. Untuk mengurangnya, guru dapat menelusuri langkah penyelesaian dan menggunakan pertanyaan tindak lanjut. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai membandingkan bukti sebelum dan sesudah pembelajaran, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan membandingkan bukti sebelum dan sesudah pembelajaran melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan membandingkan bukti sebelum dan sesudah pembelajaran;
- Bandingkan bukti membandingkan bukti sebelum dan sesudah pembelajaran pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek membandingkan bukti sebelum dan sesudah pembelajaran dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian membandingkan bukti sebelum dan sesudah pembelajaran, topik spektroskopi dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa spektrum dipandang sebagai gambar langsung molekul. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta hubungan sinyal, transisi energi, dan model struktur. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa spektrum adalah data yang ditafsirkan melalui model dan aturan seleksi. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengintegrasikan diagnosis secara berkelanjutan pada fase awal, proses, dan akhir pembelajaran, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, membandingkan bukti sebelum dan sesudah pembelajaran harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta

didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.3.18 Menentukan tindak lanjut pascadiagnosis

Menentukan tindak lanjut pascadiagnosis perlu ditempatkan dalam kerangka diagnosis yang mengintegrasikan diagnosis secara berkelanjutan pada fase awal, proses, dan akhir pembelajaran. Fokusnya bukan sekadar menghasilkan skor, melainkan memperoleh bukti yang cukup untuk menjelaskan bagaimana peserta didik membangun, menghubungkan, dan menggunakan konsep. Karena konsepsi tidak dapat diamati secara langsung, guru harus menarik inferensi dari respons verbal, tulisan, gambar, tindakan, dan perubahan strategi yang tampak selama penyelesaian tugas (Anggraena et al., 2025; Shute, 2008).

Secara operasional, guru dapat menggunakan analisis respons tertulis untuk menelusuri kualitas model mental yang dieksternalisasi. Data tersebut perlu dibaca bersama koherensi alasan, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah memperlakukan konsepsi sebagai sifat permanen. Untuk mengurangnya, guru dapat menggunakan triangulasi bukti dan menyajikan tugas paralel. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai menentukan tindak lanjut pascadiagnosis, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan menentukan tindak lanjut pascadiagnosis melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan menentukan tindak lanjut pascadiagnosis;
- Bandingkan bukti menentukan tindak lanjut pascadiagnosis pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek menentukan tindak lanjut pascadiagnosis dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian menentukan tindak lanjut pascadiagnosis, topik ikatan kovalen dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa garis ikatan dipandang sebagai batang fisik. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta penjelasan energi potensial dan distribusi elektron pada jarak antarinti berbeda. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa representasi garis adalah konvensi yang merangkum pasangan elektron dan konektivitas. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengintegrasikan diagnosis secara berkelanjutan pada fase awal, proses, dan akhir pembelajaran, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik

mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, menentukan tindak lanjut pascadiagnosis harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.3.19 Siklus diagnosis-remediasi-rediagnosis

Pembahasan tentang siklus diagnosis-remediasi-rediagnosis berangkat dari prinsip bahwa data asesmen selalu merupakan jejak tidak langsung dari proses berpikir. Satu pilihan jawaban hanya menunjukkan produk akhir, sedangkan struktur pengetahuan yang melahirkannya dapat berbeda. Diagnosis yang kuat karena itu menggabungkan respons terhadap probing, tingkat keyakinan, dan bukti kontekstual sehingga keputusan pedagogis tidak bergantung pada satu indikator saja (Shute, 2008; Black & Wiliam, 2009).

Secara operasional, guru dapat menggunakan diskusi argumentatif untuk menelusuri respons terhadap probing. Data tersebut perlu dibaca bersama tingkat keyakinan, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mengabaikan ketidakpastian pengukuran. Untuk menguranginya, guru dapat menyusun tindak lanjut berbeda untuk kategori respons yang berbeda dan memeriksa konsep prasyarat. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai siklus diagnosis-remediasi-rediagnosis, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan siklus diagnosis-remediasi-rediagnosis melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan siklus diagnosis-remediasi-rediagnosis;
- Bandingkan bukti siklus diagnosis-remediasi-rediagnosis pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek siklus diagnosis-remediasi-rediagnosis dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian siklus diagnosis-remediasi-rediagnosis, topik gas dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa partikel gas dianggap mengembang ketika dipanaskan. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta gambar partikel sebelum dan sesudah pemanasan pada kondisi berbeda. Bukti yang diharapkan adalah kemampuan peserta didik

menjelaskan bahwa ukuran partikel tidak berubah secara bermakna; distribusi kecepatan dan jarak berubah. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengintegrasikan diagnosis secara berkelanjutan pada fase awal, proses, dan akhir pembelajaran, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, siklus diagnosis-remediasi-rediagnosis harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.3.20 Pengelolaan waktu dalam diagnosis berkelanjutan

Dalam pembelajaran kimia, pengelolaan waktu dalam diagnosis berkelanjutan memiliki arti strategis karena banyak jawaban dapat diperoleh melalui hafalan prosedur atau pengenalan pola. Peserta didik mungkin tampak berhasil tetapi tidak mampu menerjemahkan makna simbol, model, atau besaran yang digunakan. Prinsip diagnosis menuntut bukti tentang cara konsep dipakai, bukan hanya apakah hasil akhirnya sesuai kunci (Black & Wiliam, 2009; Sadler, 1989).

Secara operasional, guru dapat menggunakan analisis grafik untuk menelusuri ketepatan pilihan. Data tersebut perlu dibaca bersama konsistensi lintas butir, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menyimpulkan miskonsepsi hanya dari satu jawaban. Untuk mengurangnya, guru dapat membandingkan jawaban dengan alasan dan meminta translasi antarrepresentasi. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai pengelolaan waktu dalam diagnosis berkelanjutan, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan pengelolaan waktu dalam diagnosis berkelanjutan melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan pengelolaan waktu dalam diagnosis berkelanjutan;
- Bandingkan bukti pengelolaan waktu dalam diagnosis berkelanjutan pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek pengelolaan waktu dalam diagnosis berkelanjutan dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian pengelolaan waktu dalam diagnosis berkelanjutan, topik kesetimbangan dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering

muncul ialah bahwa reaksi dianggap berhenti saat setimbang. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta hubungan grafik laju, konsentrasi, dan model partikel. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa kesetimbangan bersifat dinamis dengan laju maju dan balik yang sama. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengintegrasikan diagnosis secara berkelanjutan pada fase awal, proses, dan akhir pembelajaran, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Kerangka operasional

Aspek	Bukti yang dicari	Risiko interpretasi	Tindak lanjut
Target konsep	Penjelasan tentang kesetimbangan	Target terlalu luas	Batasi subkonsep dan prasyarat
Respons	Ketepatan pilihan	Menyimpulkan miskonsepsi hanya dari satu jawaban	Membandingkan jawaban dengan alasan
Penalaran	Konsistensi lintas butir	Alasan diperlakukan sekadar tambahan	Meminta translasi antarrepresentasi
Konfirmasi	Analisis grafik atau tugas isomorfik	Kesimpulan dibuat dari satu konteks	Gunakan bukti kedua dan rediagnosis

Implikasinya, pengelolaan waktu dalam diagnosis berkelanjutan harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.3.21 Dokumentasi longitudinal perkembangan konsep

Aspek dokumentasi longitudinal perkembangan konsep menegaskan bahwa diagnosis adalah kegiatan ilmiah berskala kelas: guru merumuskan pertanyaan, mengumpulkan bukti, mempertimbangkan penjelasan alternatif, lalu menentukan kesimpulan yang paling masuk akal. Inferensi harus proporsional terhadap kualitas data. Semakin spesifik label yang diberikan, semakin kuat pula bukti yang diperlukan (Sadler, 1989; Anggraena et al., 2025).

Secara operasional, guru dapat menggunakan tes empat tingkat untuk menelusuri koherensi alasan. Data tersebut perlu dibaca bersama konsistensi lintas representasi, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini

adalah mencampur kesalahan matematika dengan kesalahan konsep. Untuk mengurangnya, guru dapat menggunakan pertanyaan tindak lanjut dan mengulang diagnosis setelah remediasi. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai dokumentasi longitudinal perkembangan konsep, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan dokumentasi longitudinal perkembangan konsep melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan dokumentasi longitudinal perkembangan konsep;
- Bandingkan bukti dokumentasi longitudinal perkembangan konsep pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek dokumentasi longitudinal perkembangan konsep dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian dokumentasi longitudinal perkembangan konsep, topik termokimia dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa pemutusan ikatan dianggap melepaskan energi. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta analisis energi yang diserap saat memutus dan dilepas saat membentuk ikatan. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa perubahan entalpi bersih berasal dari selisih proses pemutusan dan pembentukan. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengintegrasikan diagnosis secara berkelanjutan pada fase awal, proses, dan akhir pembelajaran, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, dokumentasi longitudinal perkembangan konsep harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.3.22 Contoh alur diagnosis pada satu unit kimia

Penerapan contoh alur diagnosis pada satu unit kimia harus menghubungkan teori pengukuran dengan pengetahuan pedagogis kimia. Instrumen yang teknis tetapi tidak memahami struktur konsep akan menghasilkan kategori dangkal; sebaliknya, pengetahuan konten tanpa prosedur pengumpulan bukti dapat menghasilkan kesimpulan intuitif yang sulit

diuji. Keduanya perlu dipadukan dalam satu alur yang transparan (Anggraena et al., 2025; Shute, 2008).

Secara operasional, guru dapat menggunakan think-aloud untuk menelusuri tingkat keyakinan. Data tersebut perlu dibaca bersama kemampuan transfer, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menggunakan ambang kategori tanpa bukti. Untuk mengurangnya, guru dapat menyajikan tugas paralel dan mencatat derajat keyakinan. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai contoh alur diagnosis pada satu unit kimia, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan contoh alur diagnosis pada satu unit kimia melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan contoh alur diagnosis pada satu unit kimia;
- Bandingkan bukti contoh alur diagnosis pada satu unit kimia pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek contoh alur diagnosis pada satu unit kimia dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian contoh alur diagnosis pada satu unit kimia, topik hidrokarbon dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa isomer dipandang sebagai zat yang sama karena rumus molekulnya sama. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta hubungan struktur, konektivitas, dan sifat. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa rumus molekul tidak cukup menentukan struktur dan identitas senyawa. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengintegrasikan diagnosis secara berkelanjutan pada fase awal, proses, dan akhir pembelajaran, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, contoh alur diagnosis pada satu unit kimia harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.3.23 Analisis kritis terhadap model pretest-posttest sederhana

Analisis kritis terhadap model pretest-posttest sederhana tidak berdiri sebagai prosedur administratif. Ia menjadi bagian dari keputusan tentang apa yang perlu dijelaskan ulang, representasi apa yang perlu dibandingkan, dan siapa yang memerlukan dukungan berbeda. Oleh sebab itu, bukti diagnosis harus langsung terhubung dengan tindakan pembelajaran, bukan berhenti pada daftar kesalahan (Shute, 2008; Black & Wiliam, 2009).

Secara operasional, guru dapat menggunakan predict-observe-explain untuk menelusuri konsistensi lintas butir. Data tersebut perlu dibaca bersama stabilitas setelah jeda waktu, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah memperlakukan konsepsi sebagai sifat permanen. Untuk mengurangnya, guru dapat memeriksa konsep prasyarat dan menelusuri langkah penyelesaian. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai analisis kritis terhadap model pretest-posttest sederhana, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan analisis kritis terhadap model pretest-posttest sederhana melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan analisis kritis terhadap model pretest-posttest sederhana;
- Bandingkan bukti analisis kritis terhadap model pretest-posttest sederhana pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek analisis kritis terhadap model pretest-posttest sederhana dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian analisis kritis terhadap model pretest-posttest sederhana, topik spektroskopi dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa spektrum dipandang sebagai gambar langsung molekul. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta hubungan sinyal, transisi energi, dan model struktur. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa spektrum adalah data yang ditafsirkan melalui model dan aturan seleksi. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengintegrasikan diagnosis secara berkelanjutan pada fase awal, proses, dan akhir pembelajaran, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, analisis kritis terhadap model pretest-posttest sederhana harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta

didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.3.24 Sintesis diagnosis sepanjang pembelajaran

Sintesis diagnosis sepanjang pembelajaran perlu ditempatkan dalam kerangka diagnosis yang mengintegrasikan diagnosis secara berkelanjutan pada fase awal, proses, dan akhir pembelajaran. Fokusnya bukan sekadar menghasilkan skor, melainkan memperoleh bukti yang cukup untuk menjelaskan bagaimana peserta didik membangun, menghubungkan, dan menggunakan konsep. Karena konsepsi tidak dapat diamati secara langsung, guru harus menarik inferensi dari respons verbal, tulisan, gambar, tindakan, dan perubahan strategi yang tampak selama penyelesaian tugas (Black & Wiliam, 2009; Sadler, 1989).

Secara operasional, guru dapat menggunakan tugas isomorfik untuk menelusuri konsistensi lintas representasi. Data tersebut perlu dibaca bersama urutan langkah penyelesaian, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mengabaikan ketidakpastian pengukuran. Untuk menguranginya, guru dapat meminta translasi antarrepresentasi dan menggunakan triangulasi bukti. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai sintesis diagnosis sepanjang pembelajaran, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan sintesis diagnosis sepanjang pembelajaran melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan sintesis diagnosis sepanjang pembelajaran;
- Bandingkan bukti sintesis diagnosis sepanjang pembelajaran pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek sintesis diagnosis sepanjang pembelajaran dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian sintesis diagnosis sepanjang pembelajaran, topik ikatan kovalen dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa garis ikatan dipandang sebagai batang fisik. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta penjelasan energi potensial dan distribusi elektron pada jarak antarinti berbeda. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa representasi garis adalah konvensi yang merangkum pasangan elektron dan konektivitas. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengintegrasikan diagnosis secara berkelanjutan pada fase awal, proses, dan akhir pembelajaran, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik

mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, sintesis diagnosis sepanjang pembelajaran harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.3.25 Implikasi praktis bagi guru kimia

Implikasi praktis subbab 5.3 dapat dirangkum sebagai alur: tentukan konsep target, pilih bukti yang relevan, kumpulkan respons dalam lebih dari satu bentuk, analisis pola, nyatakan derajat kepastian, susun tindak lanjut, dan lakukan rediagnosis. Alur ini menjaga agar diagnosis sebelum, selama, dan setelah pembelajaran tidak berubah menjadi administrasi skor, tetapi berfungsi sebagai mekanisme perbaikan pembelajaran yang berulang.

15. Identifikasi konsep inti dan konsep prasyarat sebelum menyusun instrumen.
16. Gunakan pertanyaan yang meminta alasan, representasi, atau langkah kerja.
17. Bandingkan respons pada beberapa konteks dan waktu.
18. Pisahkan bukti faktual dari interpretasi guru.
19. Tentukan tindak lanjut yang berbeda untuk kategori kebutuhan yang berbeda.
20. Simpan data secukupnya dan lindungi kerahasiaan peserta didik.
21. Periksa kembali perubahan melalui tugas transfer dan retensi.

5.4 Membedakan Kesalahan dan Miskonsepsi

Subbab ini membangun dasar inferensi yang membedakan kekeliruan sesaat dari pola konseptual yang relatif sistematis. Uraian disusun dari prinsip teoretis menuju prosedur kelas, contoh kimia, risiko interpretasi, dan tindak lanjut. Setiap bagian menekankan bahwa kesimpulan diagnostik harus didasarkan pada pola bukti, bukan pada satu respons yang terpisah.

5.4.1 Mengapa setiap jawaban salah bukan miskonsepsi

Pembahasan tentang mengapa setiap jawaban salah bukan miskonsepsi berangkat dari prinsip bahwa data asesmen selalu merupakan jejak tidak langsung dari proses berpikir. Satu pilihan jawaban hanya menunjukkan produk akhir, sedangkan struktur pengetahuan yang melahirkannya dapat berbeda. Diagnosis yang kuat karena itu menggabungkan koherensi alasan, konsistensi lintas representasi, dan bukti kontekstual sehingga keputusan pedagogis tidak bergantung pada satu indikator saja (Smith et al., 1994; Yan & Subramaniam, 2018).

Secara operasional, guru dapat menggunakan gambar partikel untuk menelusuri koherensi alasan. Data tersebut perlu dibaca bersama konsistensi lintas representasi, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mencampur kesalahan matematika dengan kesalahan konsep. Untuk mengurangnya, guru dapat menggunakan pertanyaan tindak lanjut dan mengulang diagnosis setelah remediasi. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai mengapa setiap jawaban salah bukan miskonsepsi, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan mengapa setiap jawaban salah bukan miskonsepsi melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan mengapa setiap jawaban salah bukan miskonsepsi;
- Bandingkan bukti mengapa setiap jawaban salah bukan miskonsepsi pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek mengapa setiap jawaban salah bukan miskonsepsi dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian mengapa setiap jawaban salah bukan miskonsepsi, topik larutan dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa zat terlarut dianggap menghilang ketika larut. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta representasi makro, partikel, dan simbolik proses pelarutan. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa pelarutan mempertahankan materi sambil mengubah interaksi dan penyebaran partikel. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus membangun dasar inferensi yang membedakan kekeliruan sesaat dari pola konseptual yang relatif sistematis, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, mengapa setiap jawaban salah bukan miskonsepsi harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.4.2 Kesalahan karena tidak tahu

Dalam pembelajaran kimia, kesalahan karena tidak tahu memiliki arti strategis karena banyak jawaban dapat diperoleh melalui hafalan prosedur atau pengenalan pola. Peserta didik

mungkin tampak berhasil tetapi tidak mampu menerjemahkan makna simbol, model, atau besaran yang digunakan. Prinsip diagnosis menuntut bukti tentang cara konsep dipakai, bukan hanya apakah hasil akhirnya sesuai kunci (Yan & Subramaniam, 2018; Brandriet & Bretz, 2014).

Secara operasional, guru dapat menggunakan observasi praktikum untuk menelusuri tingkat keyakinan. Data tersebut perlu dibaca bersama kemampuan transfer, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menggunakan ambang kategori tanpa bukti. Untuk mengurangnya, guru dapat menyajikan tugas paralel dan mencatat derajat keyakinan. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai kesalahan karena tidak tahu, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan kesalahan karena tidak tahu melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan kesalahan karena tidak tahu;
- Bandingkan bukti kesalahan karena tidak tahu pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek kesalahan karena tidak tahu dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian kesalahan karena tidak tahu, topik asas Le Chatelier dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa setiap gangguan dianggap otomatis menggeser sistem ke sisi yang mengurangi gangguan tanpa analisis. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta penggunaan kuosien reaksi dan tetapan kesetimbangan. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa prediksi pergeseran harus terkait perubahan Q terhadap K dan kondisi sistem. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus membangun dasar inferensi yang membedakan kekeliruan sesaat dari pola konseptual yang relatif sistematis, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, kesalahan karena tidak tahu harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.4.3 Kesalahan karena lupa atau gagal memanggil pengetahuan

Aspek kesalahan karena lupa atau gagal memanggil pengetahuan menegaskan bahwa diagnosis adalah kegiatan ilmiah berskala kelas: guru merumuskan pertanyaan, mengumpulkan bukti, mempertimbangkan penjelasan alternatif, lalu menentukan kesimpulan yang paling masuk akal. Inferensi harus proporsional terhadap kualitas data. Semakin spesifik label yang diberikan, semakin kuat pula bukti yang diperlukan (Brandriet & Bretz, 2014; Taber, 2002a).

Secara operasional, guru dapat menggunakan exit ticket untuk menelusuri konsistensi lintas butir. Data tersebut perlu dibaca bersama stabilitas setelah jeda waktu, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah memperlakukan konsepsi sebagai sifat permanen. Untuk mengurangnya, guru dapat memeriksa konsep prasyarat dan menelusuri langkah penyelesaian. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai kesalahan karena lupa atau gagal memanggil pengetahuan, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan kesalahan karena lupa atau gagal memanggil pengetahuan melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan kesalahan karena lupa atau gagal memanggil pengetahuan;
- Bandingkan bukti kesalahan karena lupa atau gagal memanggil pengetahuan pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek kesalahan karena lupa atau gagal memanggil pengetahuan dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian kesalahan karena lupa atau gagal memanggil pengetahuan, topik kalor dan suhu dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa kalor dipandang sebagai zat yang tersimpan. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta penjelasan perpindahan energi dan perubahan suhu. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa kalor adalah energi dalam transit akibat perbedaan suhu. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus membangun dasar inferensi yang membedakan kekeliruan sesaat dari pola konseptual yang relatif sistematis, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, kesalahan karena lupa atau gagal memanggil pengetahuan harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian

klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.4.4 Kesalahan membaca dan memahami bahasa soal

Penerapan kesalahan membaca dan memahami bahasa soal harus menghubungkan teori pengukuran dengan pengetahuan pedagogis kimia. Instrumen yang teknis tetapi tidak memahami struktur konsep akan menghasilkan kategori dangkal; sebaliknya, pengetahuan konten tanpa prosedur pengumpulan bukti dapat menghasilkan kesimpulan intuitif yang sulit diuji. Keduanya perlu dipadukan dalam satu alur yang transparan (Taber, 2002a; Smith et al., 1994).

Secara operasional, guru dapat menggunakan tes pilihan ganda dua tingkat untuk menelusuri konsistensi lintas representasi. Data tersebut perlu dibaca bersama urutan langkah penyelesaian, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mengabaikan ketidakpastian pengukuran. Untuk menguranginya, guru dapat meminta translasi antarrepresentasi dan menggunakan triangulasi bukti. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai kesalahan membaca dan memahami bahasa soal, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan kesalahan membaca dan memahami bahasa soal melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan kesalahan membaca dan memahami bahasa soal;
- Bandingkan bukti kesalahan membaca dan memahami bahasa soal pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek kesalahan membaca dan memahami bahasa soal dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian kesalahan membaca dan memahami bahasa soal, topik polimer dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa polimer dipandang sekadar molekul besar tanpa kaitan struktur-sifat. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta analisis monomer, ikatan silang, kristalinitas, dan sifat mekanik. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa sifat polimer muncul dari struktur pada beberapa skala. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus membangun dasar inferensi yang membedakan kekeliruan sesaat dari pola konseptual yang relatif sistematis, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban

setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Kerangka operasional

Aspek	Bukti yang dicari	Risiko interpretasi	Tindak lanjut
Target konsep	Penjelasan tentang polimer	Target terlalu luas	Batasi subkonsep dan prasyarat
Respons	Konsistensi lintas representasi	Mengabaikan ketidakpastian pengukuran	Meminta translasi antarrepresentasi
Penalaran	Urutan langkah penyelesaian	Alasan diperlakukan sekadar tambahan	Menggunakan triangulasi bukti
Konfirmasi	Tes pilihan ganda dua tingkat atau tugas isomorfik	Kesimpulan dibuat dari satu konteks	Gunakan bukti kedua dan rediagnosis

Implikasinya, kesalahan membaca dan memahami bahasa soal harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.4.5 Kesalahan aritmetika dan manipulasi aljabar

Kesalahan aritmetika dan manipulasi aljabar tidak berdiri sebagai prosedur administratif. Ia menjadi bagian dari keputusan tentang apa yang perlu dijelaskan ulang, representasi apa yang perlu dibandingkan, dan siapa yang memerlukan dukungan berbeda. Oleh sebab itu, bukti diagnosis harus langsung terhubung dengan tindakan pembelajaran, bukan berhenti pada daftar kesalahan (Smith et al., 1994; Yan & Subramaniam, 2018).

Secara operasional, guru dapat menggunakan alasan terbuka untuk menelusuri kemampuan transfer. Data tersebut perlu dibaca bersama kualitas model mental yang dieksternalisasi, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menyimpulkan miskonsepsi hanya dari satu jawaban. Untuk mengurangnya, guru dapat mengulang diagnosis setelah remediasi dan menyusun tindak lanjut berbeda untuk kategori respons yang berbeda. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai kesalahan aritmetika dan manipulasi aljabar, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan kesalahan aritmetika dan manipulasi aljabar melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan kesalahan aritmetika dan manipulasi aljabar;
- Bandingkan bukti kesalahan aritmetika dan manipulasi aljabar pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek kesalahan aritmetika dan manipulasi aljabar dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian kesalahan aritmetika dan manipulasi aljabar, topik kimia lingkungan dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa zat alami dianggap selalu aman dan zat sintetis selalu berbahaya. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta penilaian berdasarkan dosis, paparan, sifat, dan konteks. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa risiko kimia bergantung pada bahaya, dosis, jalur, dan durasi paparan. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus membangun dasar inferensi yang membedakan kekeliruan sesaat dari pola konseptual yang relatif sistematis, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, kesalahan aritmetika dan manipulasi aljabar harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.4.6 Kesalahan simbolik dan notasi kimia

Kesalahan simbolik dan notasi kimia perlu ditempatkan dalam kerangka diagnosis yang membangun dasar inferensi yang membedakan kekeliruan sesaat dari pola konseptual yang relatif sistematis. Fokusnya bukan sekadar menghasilkan skor, melainkan memperoleh bukti yang cukup untuk menjelaskan bagaimana peserta didik membangun, menghubungkan, dan menggunakan konsep. Karena konsepsi tidak dapat diamati secara langsung, guru harus menarik inferensi dari respons verbal, tulisan, gambar, tindakan, dan perubahan strategi yang tampak selama penyelesaian tugas (Yan & Subramaniam, 2018; Brandriet & Bretz, 2014).

Secara operasional, guru dapat menggunakan peta konsep untuk menelusuri stabilitas setelah jeda waktu. Data tersebut perlu dibaca bersama respons terhadap probing, karena pola yang

sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mencampur kesalahan matematika dengan kesalahan konsep. Untuk menguranginya, guru dapat mencatat derajat keyakinan dan membandingkan jawaban dengan alasan. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai kesalahan simbolik dan notasi kimia, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan kesalahan simbolik dan notasi kimia melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan kesalahan simbolik dan notasi kimia;
- Bandingkan bukti kesalahan simbolik dan notasi kimia pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek kesalahan simbolik dan notasi kimia dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian kesalahan simbolik dan notasi kimia, topik gaya antarmolekul dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa pendidihan dianggap memutus ikatan kovalen dalam molekul. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta peserta didik membedakan perubahan intramolekul dan antarmolekul. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa perubahan fase terutama mengubah organisasi dan interaksi antarmolekul. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus membangun dasar inferensi yang membedakan kekeliruan sesaat dari pola konseptual yang relatif sistematis, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, kesalahan simbolik dan notasi kimia harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.4.7 Kesalahan prosedural dalam algoritma penyelesaian

Pembahasan tentang kesalahan prosedural dalam algoritma penyelesaian berangkat dari prinsip bahwa data asesmen selalu merupakan jejak tidak langsung dari proses berpikir. Satu pilihan jawaban hanya menunjukkan produk akhir, sedangkan struktur pengetahuan yang melahirkannya dapat berbeda. Diagnosis yang kuat karena itu menggabungkan urutan

langkah penyelesaian, ketepatan pilihan, dan bukti kontekstual sehingga keputusan pedagogis tidak bergantung pada satu indikator saja (Brandriet & Bretz, 2014; Taber, 2002a).

Secara operasional, guru dapat menggunakan analisis respons tertulis untuk menelusuri urutan langkah penyelesaian. Data tersebut perlu dibaca bersama ketepatan pilihan, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menggunakan ambang kategori tanpa bukti. Untuk mengurangnya, guru dapat menelusuri langkah penyelesaian dan menggunakan pertanyaan tindak lanjut. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai kesalahan prosedural dalam algoritma penyelesaian, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan kesalahan prosedural dalam algoritma penyelesaian melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan kesalahan prosedural dalam algoritma penyelesaian;
- Bandingkan bukti kesalahan prosedural dalam algoritma penyelesaian pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek kesalahan prosedural dalam algoritma penyelesaian dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian kesalahan prosedural dalam algoritma penyelesaian, topik larutan dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa zat terlarut dianggap menghilang ketika larut. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta representasi makro, partikel, dan simbolik proses pelarutan. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa pelarutan mempertahankan materi sambil mengubah interaksi dan penyebaran partikel. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus membangun dasar inferensi yang membedakan kekeliruan sesaat dari pola konseptual yang relatif sistematis, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, kesalahan prosedural dalam algoritma penyelesaian harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.4.8 Kesalahan akibat beban kognitif

Dalam pembelajaran kimia, kesalahan akibat beban kognitif memiliki arti strategis karena banyak jawaban dapat diperoleh melalui hafalan prosedur atau pengenalan pola. Peserta didik mungkin tampak berhasil tetapi tidak mampu menerjemahkan makna simbol, model, atau besaran yang digunakan. Prinsip diagnosis menuntut bukti tentang cara konsep dipakai, bukan hanya apakah hasil akhirnya sesuai kunci (Taber, 2002a; Smith et al., 1994).

Secara operasional, guru dapat menggunakan diskusi argumentatif untuk menelusuri kualitas model mental yang dieksternalisasi. Data tersebut perlu dibaca bersama koherensi alasan, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah memperlakukan konsepsi sebagai sifat permanen. Untuk mengurangnya, guru dapat menggunakan triangulasi bukti dan menyajikan tugas paralel. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai kesalahan akibat beban kognitif, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan kesalahan akibat beban kognitif melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan kesalahan akibat beban kognitif;
- Bandingkan bukti kesalahan akibat beban kognitif pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek kesalahan akibat beban kognitif dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian kesalahan akibat beban kognitif, topik asas Le Chatelier dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa setiap gangguan dianggap otomatis menggeser sistem ke sisi yang mengurangi gangguan tanpa analisis. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta penggunaan kuosien reaksi dan tetapan kesetimbangan. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa prediksi pergeseran harus terkait perubahan Q terhadap K dan kondisi sistem. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus membangun dasar inferensi yang membedakan kekeliruan sesaat dari pola konseptual yang relatif sistematis, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Kerangka operasional

Aspek	Bukti yang dicari	Risiko interpretasi	Tindak lanjut
Target konsep	Penjelasan tentang asas Le Chatelier	Target terlalu luas	Batasi subkonsep dan prasyarat
Respons	Kualitas model mental yang dieksternalisasi	Memperlakukan konsepsi sebagai sifat permanen	Menggunakan triangulasi bukti
Penalaran	Koherensi alasan	Alasan diperlakukan sekadar tambahan	Menyajikan tugas paralel
Konfirmasi	Diskusi argumentatif atau tugas isomorfik	Kesimpulan dibuat dari satu konteks	Gunakan bukti kedua dan rediagnosis

Implikasinya, kesalahan akibat beban kognitif harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.4.9 Kesalahan akibat ketidaktelitian

Aspek kesalahan akibat ketidaktelitian menegaskan bahwa diagnosis adalah kegiatan ilmiah berskala kelas: guru merumuskan pertanyaan, mengumpulkan bukti, mempertimbangkan penjelasan alternatif, lalu menentukan kesimpulan yang paling masuk akal. Inferensi harus proporsional terhadap kualitas data. Semakin spesifik label yang diberikan, semakin kuat pula bukti yang diperlukan (Smith et al., 1994; Yan & Subramaniam, 2018).

Secara operasional, guru dapat menggunakan analisis grafik untuk menelusuri respons terhadap probing. Data tersebut perlu dibaca bersama tingkat keyakinan, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mengabaikan ketidakpastian pengukuran. Untuk mengurangnya, guru dapat menyusun tindak lanjut berbeda untuk kategori respons yang berbeda dan memeriksa konsep prasyarat. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai kesalahan akibat ketidaktelitian, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan kesalahan akibat ketidaktelitian melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan kesalahan akibat ketidaktelitian;

- Bandingkan bukti kesalahan akibat ketidaktelitian pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek kesalahan akibat ketidaktelitian dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian kesalahan akibat ketidaktelitian, topik kalor dan suhu dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa kalor dipandang sebagai zat yang tersimpan. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta penjelasan perpindahan energi dan perubahan suhu. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa kalor adalah energi dalam transit akibat perbedaan suhu. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus membangun dasar inferensi yang membedakan kekeliruan sesaat dari pola konseptual yang relatif sistematis, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, kesalahan akibat ketidaktelitian harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.4.10 Kesalahan karena strategi tebakan

Penerapan kesalahan karena strategi tebakan harus menghubungkan teori pengukuran dengan pengetahuan pedagogis kimia. Instrumen yang teknis tetapi tidak memahami struktur konsep akan menghasilkan kategori dangkal; sebaliknya, pengetahuan konten tanpa prosedur pengumpulan bukti dapat menghasilkan kesimpulan intuitif yang sulit diuji. Keduanya perlu dipadukan dalam satu alur yang transparan (Yan & Subramaniam, 2018; Brandriet & Bretz, 2014).

Secara operasional, guru dapat menggunakan tes empat tingkat untuk menelusuri ketepatan pilihan. Data tersebut perlu dibaca bersama konsistensi lintas butir, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menyimpulkan miskonsepsi hanya dari satu jawaban. Untuk mengurangnya, guru dapat membandingkan jawaban dengan alasan dan meminta translasi antarrepresentasi. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai kesalahan karena strategi tebakan, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;

- Pilih tugas yang merepresentasikan kesalahan karena strategi tebakan melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan kesalahan karena strategi tebakan;
- Bandingkan bukti kesalahan karena strategi tebakan pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek kesalahan karena strategi tebakan dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian kesalahan karena strategi tebakan, topik polimer dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa polimer dipandang sekadar molekul besar tanpa kaitan struktur-sifat. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta analisis monomer, ikatan silang, kristalinitas, dan sifat mekanik. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa sifat polimer muncul dari struktur pada beberapa skala. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus membangun dasar inferensi yang membedakan kekeliruan sesaat dari pola konseptual yang relatif sistematis, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, kesalahan karena strategi tebakan harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.4.11 Kesalahan kontekstual dan ketidakstabilan respons

Kesalahan kontekstual dan ketidakstabilan respons tidak berdiri sebagai prosedur administratif. Ia menjadi bagian dari keputusan tentang apa yang perlu dijelaskan ulang, representasi apa yang perlu dibandingkan, dan siapa yang memerlukan dukungan berbeda. Oleh sebab itu, bukti diagnosis harus langsung terhubung dengan tindakan pembelajaran, bukan berhenti pada daftar kesalahan (Brandriet & Bretz, 2014; Taber, 2002a).

Secara operasional, guru dapat menggunakan think-aloud untuk menelusuri koherensi alasan. Data tersebut perlu dibaca bersama konsistensi lintas representasi, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mencampur kesalahan matematika dengan kesalahan konsep. Untuk mengurangnya, guru dapat menggunakan pertanyaan tindak lanjut dan mengulang diagnosis setelah remediasi.

Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai kesalahan kontekstual dan ketidakstabilan respons, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan kesalahan kontekstual dan ketidakstabilan respons melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan kesalahan kontekstual dan ketidakstabilan respons;
- Bandingkan bukti kesalahan kontekstual dan ketidakstabilan respons pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek kesalahan kontekstual dan ketidakstabilan respons dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian kesalahan kontekstual dan ketidakstabilan respons, topik kimia lingkungan dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa zat alami dianggap selalu aman dan zat sintetis selalu berbahaya. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta penilaian berdasarkan dosis, paparan, sifat, dan konteks. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa risiko kimia bergantung pada bahaya, dosis, jalur, dan durasi paparan. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus membangun dasar inferensi yang membedakan kekeliruan sesaat dari pola konseptual yang relatif sistematis, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, kesalahan kontekstual dan ketidakstabilan respons harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.4.12 Miskonsepsi sebagai pola penalaran yang sistematis

Miskonsepsi sebagai pola penalaran yang sistematis perlu ditempatkan dalam kerangka diagnosis yang membangun dasar inferensi yang membedakan kekeliruan sesaat dari pola konseptual yang relatif sistematis. Fokusnya bukan sekadar menghasilkan skor, melainkan memperoleh bukti yang cukup untuk menjelaskan bagaimana peserta didik membangun, menghubungkan, dan menggunakan konsep. Karena konsepsi tidak dapat diamati secara langsung, guru harus menarik inferensi dari respons verbal, tulisan, gambar, tindakan, dan perubahan strategi yang tampak selama penyelesaian tugas (Taber, 2002a; Smith et al., 1994).

Secara operasional, guru dapat menggunakan predict-observe-explain untuk menelusuri tingkat keyakinan. Data tersebut perlu dibaca bersama kemampuan transfer, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menggunakan ambang kategori tanpa bukti. Untuk mengurangnya, guru dapat menyajikan tugas paralel dan mencatat derajat keyakinan. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai miskonsepsi sebagai pola penalaran yang sistematis, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan miskonsepsi sebagai pola penalaran yang sistematis melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan miskonsepsi sebagai pola penalaran yang sistematis;
- Bandingkan bukti miskonsepsi sebagai pola penalaran yang sistematis pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek miskonsepsi sebagai pola penalaran yang sistematis dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian miskonsepsi sebagai pola penalaran yang sistematis, topik gaya antarmolekul dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa pendidihan dianggap memutus ikatan kovalen dalam molekul. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta peserta didik membedakan perubahan intramolekul dan antarmolekul. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa perubahan fase terutama mengubah organisasi dan interaksi antarmolekul. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus membangun dasar inferensi yang membedakan kekeliruan sesaat dari pola konseptual yang relatif sistematis, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Kerangka operasional

Aspek	Bukti yang dicari	Risiko interpretasi	Tindak lanjut
Target konsep	Penjelasan tentang gaya antarmolekul	Target terlalu luas	Batasi subkonsep dan prasyarat
Respons	Tingkat keyakinan	Menggunakan ambang kategori tanpa bukti	Menyajikan tugas paralel
Penalaran	Kemampuan transfer	Alasan diperlakukan sekadar tambahan	Mencatat derajat keyakinan

Aspek	Bukti yang dicari	Risiko interpretasi	Tindak lanjut
Konfirmasi	Predict-observe-explain atau tugas isomorfik	Kesimpulan dibuat dari satu konteks	Gunakan bukti kedua dan rediagnosis

Implikasinya, miskonsepsi sebagai pola penalaran yang sistematis harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.4.13 Konsistensi lintas butir sebagai indikator

Pembahasan tentang konsistensi lintas butir sebagai indikator berangkat dari prinsip bahwa data asesmen selalu merupakan jejak tidak langsung dari proses berpikir. Satu pilihan jawaban hanya menunjukkan produk akhir, sedangkan struktur pengetahuan yang melahirkannya dapat berbeda. Diagnosis yang kuat karena itu menggabungkan konsistensi lintas butir, stabilitas setelah jeda waktu, dan bukti kontekstual sehingga keputusan pedagogis tidak bergantung pada satu indikator saja (Smith et al., 1994; Yan & Subramaniam, 2018).

Secara operasional, guru dapat menggunakan tugas isomorfik untuk menelusuri konsistensi lintas butir. Data tersebut perlu dibaca bersama stabilitas setelah jeda waktu, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah memperlakukan konsepsi sebagai sifat permanen. Untuk mengurangnya, guru dapat memeriksa konsep prasyarat dan menelusuri langkah penyelesaian. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai konsistensi lintas butir sebagai indikator, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan konsistensi lintas butir sebagai indikator melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan konsistensi lintas butir sebagai indikator;
- Bandingkan bukti konsistensi lintas butir sebagai indikator pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek konsistensi lintas butir sebagai indikator dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian konsistensi lintas butir sebagai indikator, topik larutan dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa zat

terlarut dianggap menghilang ketika larut. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta representasi makro, partikel, dan simbolik proses pelarutan. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa pelarutan mempertahankan materi sambil mengubah interaksi dan penyebaran partikel. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus membangun dasar inferensi yang membedakan kekeliruan sesaat dari pola konseptual yang relatif sistematis, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, konsistensi lintas butir sebagai indikator harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.4.14 Keyakinan tinggi sebagai bukti tambahan

Dalam pembelajaran kimia, keyakinan tinggi sebagai bukti tambahan memiliki arti strategis karena banyak jawaban dapat diperoleh melalui hafalan prosedur atau pengenalan pola. Peserta didik mungkin tampak berhasil tetapi tidak mampu menerjemahkan makna simbol, model, atau besaran yang digunakan. Prinsip diagnosis menuntut bukti tentang cara konsep dipakai, bukan hanya apakah hasil akhirnya sesuai kunci (Yan & Subramaniam, 2018; Brandriet & Bretz, 2014).

Secara operasional, guru dapat menggunakan kuis formatif digital untuk menelusuri konsistensi lintas representasi. Data tersebut perlu dibaca bersama urutan langkah penyelesaian, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mengabaikan ketidakpastian pengukuran. Untuk mengurangnya, guru dapat meminta translasi antarrepresentasi dan menggunakan triangulasi bukti. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai keyakinan tinggi sebagai bukti tambahan, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan keyakinan tinggi sebagai bukti tambahan melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan keyakinan tinggi sebagai bukti tambahan;
- Bandingkan bukti keyakinan tinggi sebagai bukti tambahan pada lebih dari satu konteks, format tugas, atau bentuk representasi;

- Hubungkan kategori hasil pada aspek keyakinan tinggi sebagai bukti tambahan dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian keyakinan tinggi sebagai bukti tambahan, topik asas Le Chatelier dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa setiap gangguan dianggap otomatis menggeser sistem ke sisi yang mengurangi gangguan tanpa analisis. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta penggunaan kuosien reaksi dan tetapan kesetimbangan. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa prediksi pergeseran harus terkait perubahan Q terhadap K dan kondisi sistem. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus membangun dasar inferensi yang membedakan kekeliruan sesaat dari pola konseptual yang relatif sistematis, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, keyakinan tinggi sebagai bukti tambahan harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.4.15 Resistensi terhadap koreksi sebagai indikator

Aspek resistensi terhadap koreksi sebagai indikator menegaskan bahwa diagnosis adalah kegiatan ilmiah berskala kelas: guru merumuskan pertanyaan, mengumpulkan bukti, mempertimbangkan penjelasan alternatif, lalu menentukan kesimpulan yang paling masuk akal. Inferensi harus proporsional terhadap kualitas data. Semakin spesifik label yang diberikan, semakin kuat pula bukti yang diperlukan (Brandriet & Bretz, 2014; Taber, 2002a).

Secara operasional, guru dapat menggunakan tes tiga tingkat untuk menelusuri kemampuan transfer. Data tersebut perlu dibaca bersama kualitas model mental yang dieksternalisasi, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menyimpulkan miskonsepsi hanya dari satu jawaban. Untuk menguranginya, guru dapat mengulang diagnosis setelah remediasi dan menyusun tindak lanjut berbeda untuk kategori respons yang berbeda. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai resistensi terhadap koreksi sebagai indikator, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;

- Pilih tugas yang merepresentasikan resistensi terhadap koreksi sebagai indikator melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan resistensi terhadap koreksi sebagai indikator;
- Bandingkan bukti resistensi terhadap koreksi sebagai indikator pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek resistensi terhadap koreksi sebagai indikator dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian resistensi terhadap koreksi sebagai indikator, topik kalor dan suhu dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa kalor dipandang sebagai zat yang tersimpan. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta penjelasan perpindahan energi dan perubahan suhu. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa kalor adalah energi dalam transit akibat perbedaan suhu. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus membangun dasar inferensi yang membedakan kekeliruan sesaat dari pola konseptual yang relatif sistematis, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, resistensi terhadap koreksi sebagai indikator harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.4.16 Peran wawancara dalam diferensiasi kesalahan

Penerapan peran wawancara dalam diferensiasi kesalahan harus menghubungkan teori pengukuran dengan pengetahuan pedagogis kimia. Instrumen yang teknis tetapi tidak memahami struktur konsep akan menghasilkan kategori dangkal; sebaliknya, pengetahuan konten tanpa prosedur pengumpulan bukti dapat menghasilkan kesimpulan intuitif yang sulit diuji. Keduanya perlu dipadukan dalam satu alur yang transparan (Taber, 2002a; Smith et al., 1994).

Secara operasional, guru dapat menggunakan wawancara semi-terstruktur untuk menelusuri stabilitas setelah jeda waktu. Data tersebut perlu dibaca bersama respons terhadap probing, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mencampur kesalahan matematika dengan kesalahan konsep. Untuk mengurangnya, guru dapat mencatat derajat keyakinan dan membandingkan jawaban

dengan alasan. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai peran wawancara dalam diferensiasi kesalahan, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan peran wawancara dalam diferensiasi kesalahan melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan peran wawancara dalam diferensiasi kesalahan;
- Bandingkan bukti peran wawancara dalam diferensiasi kesalahan pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek peran wawancara dalam diferensiasi kesalahan dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian peran wawancara dalam diferensiasi kesalahan, topik polimer dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa polimer dipandang sekadar molekul besar tanpa kaitan struktur-sifat. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta analisis monomer, ikatan silang, kristalinitas, dan sifat mekanik. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa sifat polimer muncul dari struktur pada beberapa skala. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus membangun dasar inferensi yang membedakan kekeliruan sesaat dari pola konseptual yang relatif sistematis, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Kerangka operasional

Aspek	Bukti yang dicari	Risiko interpretasi	Tindak lanjut
Target konsep	Penjelasan tentang polimer	Target terlalu luas	Batasi subkonsep dan prasyarat
Respons	Stabilitas setelah jeda waktu	Mencampur kesalahan matematika dengan kesalahan konsep	Mencatat derajat keyakinan
Penalaran	Respons terhadap probing	Alasan diperlakukan sekadar tambahan	Membandingkan jawaban dengan alasan
Konfirmasi	Wawancara semi-terstruktur atau tugas isomorfik	Kesimpulan dibuat dari satu konteks	Gunakan bukti kedua dan rediagnosis

Implikasinya, peran wawancara dalam diferensiasi kesalahan harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.4.17 Peran tugas paralel dan isomorfik

Peran tugas paralel dan isomorfik tidak berdiri sebagai prosedur administratif. Ia menjadi bagian dari keputusan tentang apa yang perlu dijelaskan ulang, representasi apa yang perlu dibandingkan, dan siapa yang memerlukan dukungan berbeda. Oleh sebab itu, bukti diagnosis harus langsung terhubung dengan tindakan pembelajaran, bukan berhenti pada daftar kesalahan (Smith et al., 1994; Yan & Subramaniam, 2018).

Secara operasional, guru dapat menggunakan gambar partikel untuk menelusuri urutan langkah penyelesaian. Data tersebut perlu dibaca bersama ketepatan pilihan, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menggunakan ambang kategori tanpa bukti. Untuk mengurangnya, guru dapat menelusuri langkah penyelesaian dan menggunakan pertanyaan tindak lanjut. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai peran tugas paralel dan isomorfik, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan peran tugas paralel dan isomorfik melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan peran tugas paralel dan isomorfik;
- Bandingkan bukti peran tugas paralel dan isomorfik pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek peran tugas paralel dan isomorfik dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian peran tugas paralel dan isomorfik, topik kimia lingkungan dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa zat alami dianggap selalu aman dan zat sintesis selalu berbahaya. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta penilaian berdasarkan dosis, paparan, sifat, dan konteks. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa risiko kimia bergantung pada bahaya, dosis, jalur, dan durasi paparan. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus membangun dasar inferensi yang membedakan kekeliruan sesaat dari pola konseptual yang relatif sistematis, bukan diperlakukan sebagai skor yang

berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, peran tugas paralel dan isomorfik harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.4.18 Analisis jejak kerja dan langkah penyelesaian

Analisis jejak kerja dan langkah penyelesaian perlu ditempatkan dalam kerangka diagnosis yang membangun dasar inferensi yang membedakan kekeliruan sesaat dari pola konseptual yang relatif sistematis. Fokusnya bukan sekadar menghasilkan skor, melainkan memperoleh bukti yang cukup untuk menjelaskan bagaimana peserta didik membangun, menghubungkan, dan menggunakan konsep. Karena konsepsi tidak dapat diamati secara langsung, guru harus menarik inferensi dari respons verbal, tulisan, gambar, tindakan, dan perubahan strategi yang tampak selama penyelesaian tugas (Yan & Subramaniam, 2018; Brandriet & Bretz, 2014).

Secara operasional, guru dapat menggunakan observasi praktikum untuk menelusuri kualitas model mental yang dieksternalisasi. Data tersebut perlu dibaca bersama koherensi alasan, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah memperlakukan konsepsi sebagai sifat permanen. Untuk mengurangnya, guru dapat menggunakan triangulasi bukti dan menyajikan tugas paralel. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai analisis jejak kerja dan langkah penyelesaian, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan analisis jejak kerja dan langkah penyelesaian melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan analisis jejak kerja dan langkah penyelesaian;
- Bandingkan bukti analisis jejak kerja dan langkah penyelesaian pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek analisis jejak kerja dan langkah penyelesaian dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian analisis jejak kerja dan langkah penyelesaian, topik gaya antarmolekul dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa pendidihan dianggap memutus ikatan kovalen dalam molekul. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat

istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta peserta didik membedakan perubahan intramolekul dan antarmolekul. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa perubahan fase terutama mengubah organisasi dan interaksi antarmolekul. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus membangun dasar inferensi yang membedakan kekeliruan sesaat dari pola konseptual yang relatif sistematis, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, analisis jejak kerja dan langkah penyelesaian harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.4.19 Matriks klasifikasi kesalahan dan miskonsepsi

Pembahasan tentang matriks klasifikasi kesalahan dan miskonsepsi berangkat dari prinsip bahwa data asesmen selalu merupakan jejak tidak langsung dari proses berpikir. Satu pilihan jawaban hanya menunjukkan produk akhir, sedangkan struktur pengetahuan yang melahirkannya dapat berbeda. Diagnosis yang kuat karena itu menggabungkan respons terhadap probing, tingkat keyakinan, dan bukti kontekstual sehingga keputusan pedagogis tidak bergantung pada satu indikator saja (Brandriet & Bretz, 2014; Taber, 2002a).

Secara operasional, guru dapat menggunakan exit ticket untuk menelusuri respons terhadap probing. Data tersebut perlu dibaca bersama tingkat keyakinan, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mengabaikan ketidakpastian pengukuran. Untuk menguranginya, guru dapat menyusun tindak lanjut berbeda untuk kategori respons yang berbeda dan memeriksa konsep prasyarat. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai matriks klasifikasi kesalahan dan miskonsepsi, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan matriks klasifikasi kesalahan dan miskonsepsi melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan matriks klasifikasi kesalahan dan miskonsepsi;
- Bandingkan bukti matriks klasifikasi kesalahan dan miskonsepsi pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek matriks klasifikasi kesalahan dan miskonsepsi dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian matriks klasifikasi kesalahan dan miskonsepsi, topik larutan dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa zat terlarut dianggap menghilang ketika larut. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta representasi makro, partikel, dan simbolik proses pelarutan. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa pelarutan mempertahankan materi sambil mengubah interaksi dan penyebaran partikel. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus membangun dasar inferensi yang membedakan kekeliruan sesaat dari pola konseptual yang relatif sistematis, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, matriks klasifikasi kesalahan dan miskonsepsi harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.4.20 Risiko false positive dalam pelabelan miskonsepsi

Dalam pembelajaran kimia, risiko false positive dalam pelabelan miskonsepsi memiliki arti strategis karena banyak jawaban dapat diperoleh melalui hafalan prosedur atau pengenalan pola. Peserta didik mungkin tampak berhasil tetapi tidak mampu menerjemahkan makna simbol, model, atau besaran yang digunakan. Prinsip diagnosis menuntut bukti tentang cara konsep dipakai, bukan hanya apakah hasil akhirnya sesuai kunci (Taber, 2002a; Smith et al., 1994).

Secara operasional, guru dapat menggunakan tes pilihan ganda dua tingkat untuk menelusuri ketepatan pilihan. Data tersebut perlu dibaca bersama konsistensi lintas butir, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menyimpulkan miskonsepsi hanya dari satu jawaban. Untuk mengurangnya, guru dapat membandingkan jawaban dengan alasan dan meminta translasi antarrepresentasi. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai risiko false positive dalam pelabelan miskonsepsi, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan risiko false positive dalam pelabelan miskonsepsi melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan risiko false positive dalam pelabelan miskonsepsi;

- Bandingkan bukti risiko false positive dalam pelabelan miskonsepsi pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek risiko false positive dalam pelabelan miskonsepsi dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian risiko false positive dalam pelabelan miskonsepsi, topik asas Le Chatelier dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa setiap gangguan dianggap otomatis menggeser sistem ke sisi yang mengurangi gangguan tanpa analisis. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta penggunaan kuosien reaksi dan tetapan kesetimbangan. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa prediksi pergeseran harus terkait perubahan Q terhadap K dan kondisi sistem. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus membangun dasar inferensi yang membedakan kekeliruan sesaat dari pola konseptual yang relatif sistematis, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Kerangka operasional

Aspek	Bukti yang dicari	Risiko interpretasi	Tindak lanjut
Target konsep	Penjelasan tentang asas Le Chatelier	Target terlalu luas	Batasi subkonsep dan prasyarat
Respons	Ketepatan pilihan	Menyimpulkan miskonsepsi hanya dari satu jawaban	Membandingkan jawaban dengan alasan
Penalaran	Konsistensi lintas butir	Alasan diperlakukan sekadar tambahan	Meminta translasi antarrepresentasi
Konfirmasi	Tes pilihan ganda dua tingkat atau tugas isomorfik	Kesimpulan dibuat dari satu konteks	Gunakan bukti kedua dan rediagnosis

Implikasinya, risiko false positive dalam pelabelan miskonsepsi harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.4.21 Risiko false negative dalam diagnosis

Aspek risiko false negative dalam diagnosis menegaskan bahwa diagnosis adalah kegiatan ilmiah berskala kelas: guru merumuskan pertanyaan, mengumpulkan bukti, mempertimbangkan penjelasan alternatif, lalu menentukan kesimpulan yang paling masuk akal. Inferensi harus proporsional terhadap kualitas data. Semakin spesifik label yang diberikan, semakin kuat pula bukti yang diperlukan (Smith et al., 1994; Yan & Subramaniam, 2018).

Secara operasional, guru dapat menggunakan alasan terbuka untuk menelusuri koherensi alasan. Data tersebut perlu dibaca bersama konsistensi lintas representasi, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mencampur kesalahan matematika dengan kesalahan konsep. Untuk mengurangnya, guru dapat menggunakan pertanyaan tindak lanjut dan mengulang diagnosis setelah remediasi. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai risiko false negative dalam diagnosis, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan risiko false negative dalam diagnosis melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan risiko false negative dalam diagnosis;
- Bandingkan bukti risiko false negative dalam diagnosis pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek risiko false negative dalam diagnosis dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian risiko false negative dalam diagnosis, topik kalor dan suhu dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa kalor dipandang sebagai zat yang tersimpan. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta penjelasan perpindahan energi dan perubahan suhu. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa kalor adalah energi dalam transit akibat perbedaan suhu. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus membangun dasar inferensi yang membedakan kekeliruan sesaat dari pola konseptual yang relatif sistematis, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, risiko false negative dalam diagnosis harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta

rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.4.22 Contoh diferensiasi pada stoikiometri

Penerapan contoh diferensiasi pada stoikiometri harus menghubungkan teori pengukuran dengan pengetahuan pedagogis kimia. Instrumen yang teknis tetapi tidak memahami struktur konsep akan menghasilkan kategori dangkal; sebaliknya, pengetahuan konten tanpa prosedur pengumpulan bukti dapat menghasilkan kesimpulan intuitif yang sulit diuji. Keduanya perlu dipadukan dalam satu alur yang transparan (Yan & Subramaniam, 2018; Brandriet & Bretz, 2014).

Secara operasional, guru dapat menggunakan peta konsep untuk menelusuri tingkat keyakinan. Data tersebut perlu dibaca bersama kemampuan transfer, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menggunakan ambang kategori tanpa bukti. Untuk mengurangnya, guru dapat menyajikan tugas paralel dan mencatat derajat keyakinan. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai contoh diferensiasi pada stoikiometri, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan contoh diferensiasi pada stoikiometri melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan contoh diferensiasi pada stoikiometri;
- Bandingkan bukti contoh diferensiasi pada stoikiometri pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek contoh diferensiasi pada stoikiometri dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian contoh diferensiasi pada stoikiometri, topik polimer dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa polimer dipandang sekadar molekul besar tanpa kaitan struktur-sifat. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta analisis monomer, ikatan silang, kristalinitas, dan sifat mekanik. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa sifat polimer muncul dari struktur pada beberapa skala. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus membangun dasar inferensi yang membedakan kekeliruan sesaat dari pola konseptual yang relatif sistematis, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan

klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, contoh diferensiasi pada stoikiometri harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.4.23 Contoh diferensiasi pada kesetimbangan kimia

Contoh diferensiasi pada kesetimbangan kimia tidak berdiri sebagai prosedur administratif. Ia menjadi bagian dari keputusan tentang apa yang perlu dijelaskan ulang, representasi apa yang perlu dibandingkan, dan siapa yang memerlukan dukungan berbeda. Oleh sebab itu, bukti diagnosis harus langsung terhubung dengan tindakan pembelajaran, bukan berhenti pada daftar kesalahan (Brandriet & Bretz, 2014; Taber, 2002a).

Secara operasional, guru dapat menggunakan analisis respons tertulis untuk menelusuri konsistensi lintas butir. Data tersebut perlu dibaca bersama stabilitas setelah jeda waktu, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah memperlakukan konsepsi sebagai sifat permanen. Untuk mengurangnya, guru dapat memeriksa konsep prasyarat dan menelusuri langkah penyelesaian. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai contoh diferensiasi pada kesetimbangan kimia, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan contoh diferensiasi pada kesetimbangan kimia melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan contoh diferensiasi pada kesetimbangan kimia;
- Bandingkan bukti contoh diferensiasi pada kesetimbangan kimia pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek contoh diferensiasi pada kesetimbangan kimia dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian contoh diferensiasi pada kesetimbangan kimia, topik kimia lingkungan dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa zat alami dianggap selalu aman dan zat sintetis selalu berbahaya. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta penilaian berdasarkan dosis, paparan, sifat, dan konteks. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa risiko kimia bergantung pada bahaya, dosis,

jalur, dan durasi paparan. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus membangun dasar inferensi yang membedakan kekeliruan sesaat dari pola konseptual yang relatif sistematis, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, contoh diferensiasi pada kesetimbangan kimia harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.4.24 Sintesis pembedaan kesalahan dan miskonsepsi

Sintesis pembedaan kesalahan dan miskonsepsi perlu ditempatkan dalam kerangka diagnosis yang membangun dasar inferensi yang membedakan kekeliruan sesaat dari pola konseptual yang relatif sistematis. Fokusnya bukan sekadar menghasilkan skor, melainkan memperoleh bukti yang cukup untuk menjelaskan bagaimana peserta didik membangun, menghubungkan, dan menggunakan konsep. Karena konsepsi tidak dapat diamati secara langsung, guru harus menarik inferensi dari respons verbal, tulisan, gambar, tindakan, dan perubahan strategi yang tampak selama penyelesaian tugas (Taber, 2002a; Smith et al., 1994).

Secara operasional, guru dapat menggunakan diskusi argumentatif untuk menelusuri konsistensi lintas representasi. Data tersebut perlu dibaca bersama urutan langkah penyelesaian, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mengabaikan ketidakpastian pengukuran. Untuk menguranginya, guru dapat meminta translasi antarrepresentasi dan menggunakan triangulasi bukti. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai sintesis pembedaan kesalahan dan miskonsepsi, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan sintesis pembedaan kesalahan dan miskonsepsi melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan sintesis pembedaan kesalahan dan miskonsepsi;
- Bandingkan bukti sintesis pembedaan kesalahan dan miskonsepsi pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek sintesis pembedaan kesalahan dan miskonsepsi dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian sintesis pembedaan kesalahan dan miskonsepsi, topik gaya antarmolekul dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering

muncul ialah bahwa pendidihan dianggap memutus ikatan kovalen dalam molekul. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta peserta didik membedakan perubahan intramolekul dan antarmolekul. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa perubahan fase terutama mengubah organisasi dan interaksi antarmolekul. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus membangun dasar inferensi yang membedakan kekeliruan sesaat dari pola konseptual yang relatif sistematis, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, sintesis pembedaan kesalahan dan miskonsepsi harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.4.25 Implikasi praktis bagi guru kimia

Implikasi praktis subbab 5.4 dapat dirangkum sebagai alur: tentukan konsep target, pilih bukti yang relevan, kumpulkan respons dalam lebih dari satu bentuk, analisis pola, nyatakan derajat kepastian, susun tindak lanjut, dan lakukan rediagnosis. Alur ini menjaga agar membedakan kesalahan dan miskonsepsi tidak berubah menjadi administrasi skor, tetapi berfungsi sebagai mekanisme perbaikan pembelajaran yang berulang.

22. Identifikasi konsep inti dan konsep prasyarat sebelum menyusun instrumen.
23. Gunakan pertanyaan yang meminta alasan, representasi, atau langkah kerja.
24. Bandingkan respons pada beberapa konteks dan waktu.
25. Pisahkan bukti faktual dari interpretasi guru.
26. Tentukan tindak lanjut yang berbeda untuk kategori kebutuhan yang berbeda.
27. Simpan data secukupnya dan lindungi kerahasiaan peserta didik.
28. Periksa kembali perubahan melalui tugas transfer dan retensi.

5.5 Hubungan Jawaban, Alasan, dan Keyakinan

Subbab ini menggabungkan ketepatan jawaban, kualitas alasan, dan derajat keyakinan untuk memperkaya klasifikasi. Uraian disusun dari prinsip teoretis menuju prosedur kelas, contoh kimia, risiko interpretasi, dan tindak lanjut. Setiap bagian menekankan bahwa kesimpulan diagnostik harus didasarkan pada pola bukti, bukan pada satu respons yang terpisah.

5.5.1 Jawaban sebagai bukti yang belum lengkap

Pembahasan tentang jawaban sebagai bukti yang belum lengkap berangkat dari prinsip bahwa data asesmen selalu merupakan jejak tidak langsung dari proses berpikir. Satu pilihan jawaban hanya menunjukkan produk akhir, sedangkan struktur pengetahuan yang melahirkannya dapat berbeda. Diagnosis yang kuat karena itu menggabungkan koherensi alasan, konsistensi lintas representasi, dan bukti kontekstual sehingga keputusan pedagogis tidak bergantung pada satu indikator saja (Koriat, 1997; Brandriet & Bretz, 2014).

Secara operasional, guru dapat menggunakan observasi praktikum untuk menelusuri koherensi alasan. Data tersebut perlu dibaca bersama konsistensi lintas representasi, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mencampur kesalahan matematika dengan kesalahan konsep. Untuk mengurangnya, guru dapat menggunakan pertanyaan tindak lanjut dan mengulang diagnosis setelah remediasi. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai jawaban sebagai bukti yang belum lengkap, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan jawaban sebagai bukti yang belum lengkap melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan jawaban sebagai bukti yang belum lengkap;
- Bandingkan bukti jawaban sebagai bukti yang belum lengkap pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek jawaban sebagai bukti yang belum lengkap dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian jawaban sebagai bukti yang belum lengkap, topik konsentrasi dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa konsentrasi disamakan dengan jumlah zat terlarut tanpa memperhatikan volume. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta perbandingan beberapa larutan dengan jumlah mol dan volume berbeda. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa konsentrasi adalah besaran rasio yang bergantung pada pembilang dan penyebut. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menggabungkan ketepatan jawaban, kualitas alasan, dan derajat keyakinan untuk memperkaya klasifikasi, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, jawaban sebagai bukti yang belum lengkap harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.5.2 Alasan sebagai jendela struktur pengetahuan

Dalam pembelajaran kimia, alasan sebagai jendela struktur pengetahuan memiliki arti strategis karena banyak jawaban dapat diperoleh melalui hafalan prosedur atau pengenalan pola. Peserta didik mungkin tampak berhasil tetapi tidak mampu menerjemahkan makna simbol, model, atau besaran yang digunakan. Prinsip diagnosis menuntut bukti tentang cara konsep dipakai, bukan hanya apakah hasil akhirnya sesuai kunci (Brandriet & Bretz, 2014; Yan & Subramaniam, 2018).

Secara operasional, guru dapat menggunakan exit ticket untuk menelusuri tingkat keyakinan. Data tersebut perlu dibaca bersama kemampuan transfer, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menggunakan ambang kategori tanpa bukti. Untuk mengurangnya, guru dapat menyajikan tugas paralel dan mencatat derajat keyakinan. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai alasan sebagai jendela struktur pengetahuan, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan alasan sebagai jendela struktur pengetahuan melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan alasan sebagai jendela struktur pengetahuan;
- Bandingkan bukti alasan sebagai jendela struktur pengetahuan pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek alasan sebagai jendela struktur pengetahuan dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian alasan sebagai jendela struktur pengetahuan, topik kelarutan dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa larutan jenuh dianggap tidak mengandung proses pelarutan lagi. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta model partikel pelarutan dan pengendapan pada keadaan jenuh. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa kejenuhan adalah keadaan kesetimbangan dinamis. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menggabungkan ketepatan jawaban, kualitas

alasan, dan derajat keyakinan untuk memperkaya klasifikasi, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, alasan sebagai jendela struktur pengetahuan harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.5.3 Keyakinan sebagai dimensi metakognitif

Aspek keyakinan sebagai dimensi metakognitif menegaskan bahwa diagnosis adalah kegiatan ilmiah berskala kelas: guru merumuskan pertanyaan, mengumpulkan bukti, mempertimbangkan penjelasan alternatif, lalu menentukan kesimpulan yang paling masuk akal. Inferensi harus proporsional terhadap kualitas data. Semakin spesifik label yang diberikan, semakin kuat pula bukti yang diperlukan (Yan & Subramaniam, 2018; Hasan et al., 1999).

Secara operasional, guru dapat menggunakan tes pilihan ganda dua tingkat untuk menelusuri konsistensi lintas butir. Data tersebut perlu dibaca bersama stabilitas setelah jeda waktu, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah memperlakukan konsepsi sebagai sifat permanen. Untuk mengurangnya, guru dapat memeriksa konsep prasyarat dan menelusuri langkah penyelesaian. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai keyakinan sebagai dimensi metakognitif, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan keyakinan sebagai dimensi metakognitif melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan keyakinan sebagai dimensi metakognitif;
- Bandingkan bukti keyakinan sebagai dimensi metakognitif pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek keyakinan sebagai dimensi metakognitif dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian keyakinan sebagai dimensi metakognitif, topik redoks dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa oksidasi selalu dianggap melibatkan oksigen. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi

konseptual yang tepat. Guru dapat meminta analisis transfer elektron dan perubahan bilangan oksidasi. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa pengikatan oksigen merupakan kasus khusus dari kerangka redoks yang lebih luas. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menggabungkan ketepatan jawaban, kualitas alasan, dan derajat keyakinan untuk memperkaya klasifikasi, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, keyakinan sebagai dimensi metakognitif harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.5.4 Empat pola dasar jawaban dan alasan

Penerapan empat pola dasar jawaban dan alasan harus menghubungkan teori pengukuran dengan pengetahuan pedagogis kimia. Instrumen yang teknis tetapi tidak memahami struktur konsep akan menghasilkan kategori dangkal; sebaliknya, pengetahuan konten tanpa prosedur pengumpulan bukti dapat menghasilkan kesimpulan intuitif yang sulit diuji. Keduanya perlu dipadukan dalam satu alur yang transparan (Hasan et al., 1999; Koriati, 1997).

Secara operasional, guru dapat menggunakan alasan terbuka untuk menelusuri konsistensi lintas representasi. Data tersebut perlu dibaca bersama urutan langkah penyelesaian, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mengabaikan ketidakpastian pengukuran. Untuk menguranginya, guru dapat meminta translasi antarrepresentasi dan menggunakan triangulasi bukti. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai empat pola dasar jawaban dan alasan, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan empat pola dasar jawaban dan alasan melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan empat pola dasar jawaban dan alasan;
- Bandingkan bukti empat pola dasar jawaban dan alasan pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek empat pola dasar jawaban dan alasan dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian empat pola dasar jawaban dan alasan, topik kimia organik dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa reaksi dipahami hanya sebagai penggantian simbol tanpa mekanisme. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta penelusuran pusat elektron dan perubahan ikatan. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa mekanisme menjelaskan urutan perpindahan elektron dan perubahan struktur. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menggabungkan ketepatan jawaban, kualitas alasan, dan derajat keyakinan untuk memperkaya klasifikasi, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Kerangka operasional

Aspek	Bukti yang dicari	Risiko interpretasi	Tindak lanjut
Target konsep	Penjelasan tentang kimia organik	Target terlalu luas	Batasi subkonsep dan prasyarat
Respons	Konsistensi lintas representasi	Mengabaikan ketidakpastian pengukuran	Meminta translasi antarrepresentasi
Penalaran	Urutan langkah penyelesaian	Alasan diperlakukan sekadar tambahan	Menggunakan triangulasi bukti
Konfirmasi	Alasan terbuka atau tugas isomorfik	Kesimpulan dibuat dari satu konteks	Gunakan bukti kedua dan rediagnosis

Implikasinya, empat pola dasar jawaban dan alasan harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.5.5 Jawaban benar dengan alasan benar

Jawaban benar dengan alasan benar tidak berdiri sebagai prosedur administratif. Ia menjadi bagian dari keputusan tentang apa yang perlu dijelaskan ulang, representasi apa yang perlu dibandingkan, dan siapa yang memerlukan dukungan berbeda. Oleh sebab itu, bukti diagnosis harus langsung terhubung dengan tindakan pembelajaran, bukan berhenti pada daftar kesalahan (Koriat, 1997; Brandriet & Bretz, 2014).

Secara operasional, guru dapat menggunakan peta konsep untuk menelusuri kemampuan transfer. Data tersebut perlu dibaca bersama kualitas model mental yang dieksternalisasi, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menyimpulkan miskonsepsi hanya dari satu jawaban. Untuk mengurangnya, guru dapat mengulang diagnosis setelah remediasi dan menyusun tindak lanjut berbeda untuk kategori respons yang berbeda. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai jawaban benar dengan alasan benar, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan jawaban benar dengan alasan benar melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan jawaban benar dengan alasan benar;
- Bandingkan bukti jawaban benar dengan alasan benar pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek jawaban benar dengan alasan benar dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian jawaban benar dengan alasan benar, topik struktur atom dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa elektron dipandang bergerak pada lintasan planet yang tetap. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta peserta didik membandingkan model Bohr dan model orbital serta menjelaskan batas setiap model. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa model atom adalah representasi teoretis dengan domain penggunaan tertentu. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menggabungkan ketepatan jawaban, kualitas alasan, dan derajat keyakinan untuk memperkaya klasifikasi, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, jawaban benar dengan alasan benar harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.5.6 Jawaban benar dengan alasan salah

Jawaban benar dengan alasan salah perlu ditempatkan dalam kerangka diagnosis yang menggabungkan ketepatan jawaban, kualitas alasan, dan derajat keyakinan untuk memperkaya klasifikasi. Fokusnya bukan sekadar menghasilkan skor, melainkan memperoleh bukti yang cukup untuk menjelaskan bagaimana peserta didik membangun, menghubungkan, dan menggunakan konsep. Karena konsepsi tidak dapat diamati secara langsung, guru harus menarik inferensi dari respons verbal, tulisan, gambar, tindakan, dan perubahan strategi yang tampak selama penyelesaian tugas (Brandriet & Bretz, 2014; Yan & Subramaniam, 2018).

Secara operasional, guru dapat menggunakan analisis respons tertulis untuk menelusuri stabilitas setelah jeda waktu. Data tersebut perlu dibaca bersama respons terhadap probing, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mencampur kesalahan matematika dengan kesalahan konsep. Untuk mengurangnya, guru dapat mencatat derajat keyakinan dan membandingkan jawaban dengan alasan. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai jawaban benar dengan alasan salah, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan jawaban benar dengan alasan salah melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan jawaban benar dengan alasan salah;
- Bandingkan bukti jawaban benar dengan alasan salah pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek jawaban benar dengan alasan salah dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian jawaban benar dengan alasan salah, topik mol dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa satu mol semua zat dianggap memiliki massa yang sama. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta perbandingan jumlah partikel dan massa satu mol beberapa zat. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa mol menyatakan jumlah entitas, sedangkan massa bergantung pada massa molar. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menggabungkan ketepatan jawaban, kualitas alasan, dan derajat keyakinan untuk memperkaya klasifikasi, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, jawaban benar dengan alasan salah harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.5.7 Jawaban salah dengan alasan benar sebagian

Pembahasan tentang jawaban salah dengan alasan benar sebagian berangkat dari prinsip bahwa data asesmen selalu merupakan jejak tidak langsung dari proses berpikir. Satu pilihan jawaban hanya menunjukkan produk akhir, sedangkan struktur pengetahuan yang melahirkannya dapat berbeda. Diagnosis yang kuat karena itu menggabungkan urutan langkah penyelesaian, ketepatan pilihan, dan bukti kontekstual sehingga keputusan pedagogis tidak bergantung pada satu indikator saja (Yan & Subramaniam, 2018; Hasan et al., 1999).

Secara operasional, guru dapat menggunakan diskusi argumentatif untuk menelusuri urutan langkah penyelesaian. Data tersebut perlu dibaca bersama ketepatan pilihan, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menggunakan ambang kategori tanpa bukti. Untuk mengurangnya, guru dapat menelusuri langkah penyelesaian dan menggunakan pertanyaan tindak lanjut. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai jawaban salah dengan alasan benar sebagian, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan jawaban salah dengan alasan benar sebagian melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan jawaban salah dengan alasan benar sebagian;
- Bandingkan bukti jawaban salah dengan alasan benar sebagian pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek jawaban salah dengan alasan benar sebagian dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian jawaban salah dengan alasan benar sebagian, topik konsentrasi dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa konsentrasi disamakan dengan jumlah zat terlarut tanpa memperhatikan volume. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta perbandingan beberapa larutan dengan jumlah mol dan volume berbeda. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa konsentrasi adalah besaran rasio yang bergantung pada pembilang dan penyebut. Kualitas bukti tersebut kemudian

ditafsirkan sesuai fokus menggabungkan ketepatan jawaban, kualitas alasan, dan derajat keyakinan untuk memperkaya klasifikasi, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, jawaban salah dengan alasan benar sebagian harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.5.8 Jawaban salah dengan alasan salah

Dalam pembelajaran kimia, jawaban salah dengan alasan salah memiliki arti strategis karena banyak jawaban dapat diperoleh melalui hafalan prosedur atau pengenalan pola. Peserta didik mungkin tampak berhasil tetapi tidak mampu menerjemahkan makna simbol, model, atau besaran yang digunakan. Prinsip diagnosis menuntut bukti tentang cara konsep dipakai, bukan hanya apakah hasil akhirnya sesuai kunci (Hasan et al., 1999; Koriati, 1997).

Secara operasional, guru dapat menggunakan analisis grafik untuk menelusuri kualitas model mental yang dieksternalisasi. Data tersebut perlu dibaca bersama koherensi alasan, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah memperlakukan konsepsi sebagai sifat permanen. Untuk mengurangnya, guru dapat menggunakan triangulasi bukti dan menyajikan tugas paralel. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai jawaban salah dengan alasan salah, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan jawaban salah dengan alasan salah melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan jawaban salah dengan alasan salah;
- Bandingkan bukti jawaban salah dengan alasan salah pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek jawaban salah dengan alasan salah dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian jawaban salah dengan alasan salah, topik kelarutan dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa larutan jenuh dianggap tidak mengandung proses pelarutan lagi. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa

menggunakan relasi konseptual yang tepat. Guru dapat meminta model partikel pelarutan dan pengendapan pada keadaan jenuh. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa kejenuhan adalah keadaan kesetimbangan dinamis. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menggabungkan ketepatan jawaban, kualitas alasan, dan derajat keyakinan untuk memperkaya klasifikasi, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Kerangka operasional

Aspek	Bukti yang dicari	Risiko interpretasi	Tindak lanjut
Target konsep	Penjelasan tentang kelarutan	Target terlalu luas	Batasi subkonsep dan prasyarat
Respons	Kualitas model mental yang dieksternalisasi	Memperlakukan konsepsi sebagai sifat permanen	Menggunakan triangulasi bukti
Penalaran	Koherensi alasan	Alasan diperlakukan sekadar tambahan	Menyajikan tugas paralel
Konfirmasi	Analisis grafik atau tugas isomorfik	Kesimpulan dibuat dari satu konteks	Gunakan bukti kedua dan rediagnosis

Implikasinya, jawaban salah dengan alasan salah harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.5.9 Makna keyakinan tinggi dan rendah

Aspek makna keyakinan tinggi dan rendah menegaskan bahwa diagnosis adalah kegiatan ilmiah berskala kelas: guru merumuskan pertanyaan, mengumpulkan bukti, mempertimbangkan penjelasan alternatif, lalu menentukan kesimpulan yang paling masuk akal. Inferensi harus proporsional terhadap kualitas data. Semakin spesifik label yang diberikan, semakin kuat pula bukti yang diperlukan (Koriat, 1997; Brandriet & Bretz, 2014).

Secara operasional, guru dapat menggunakan tes empat tingkat untuk menelusuri respons terhadap probing. Data tersebut perlu dibaca bersama tingkat keyakinan, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mengabaikan ketidakpastian pengukuran. Untuk mengurangnya, guru dapat menyusun tindak lanjut berbeda untuk kategori respons yang berbeda dan memeriksa konsep

prasyarat. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai makna keyakinan tinggi dan rendah, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan makna keyakinan tinggi dan rendah melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan makna keyakinan tinggi dan rendah;
- Bandingkan bukti makna keyakinan tinggi dan rendah pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek makna keyakinan tinggi dan rendah dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian makna keyakinan tinggi dan rendah, topik redoks dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa oksidasi selalu dianggap melibatkan oksigen. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta analisis transfer elektron dan perubahan bilangan oksidasi. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa pengikatan oksigen merupakan kasus khusus dari kerangka redoks yang lebih luas. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menggabungkan ketepatan jawaban, kualitas alasan, dan derajat keyakinan untuk memperkaya klasifikasi, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, makna keyakinan tinggi dan rendah harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.5.10 Certainty of Response Index dan penggunaannya

Penerapan certainty of response index dan penggunaannya harus menghubungkan teori pengukuran dengan pengetahuan pedagogis kimia. Instrumen yang teknis tetapi tidak memahami struktur konsep akan menghasilkan kategori dangkal; sebaliknya, pengetahuan konten tanpa prosedur pengumpulan bukti dapat menghasilkan kesimpulan intuitif yang sulit diuji. Keduanya perlu dipadukan dalam satu alur yang transparan (Brandriet & Bretz, 2014; Yan & Subramaniam, 2018).

Secara operasional, guru dapat menggunakan think-aloud untuk menelusuri ketepatan pilihan. Data tersebut perlu dibaca bersama konsistensi lintas butir, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menyimpulkan miskonsepsi hanya dari satu jawaban. Untuk mengurangnya, guru dapat membandingkan jawaban dengan alasan dan meminta translasi antarrepresentasi. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai certainty of response index dan penggunaannya, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan certainty of response index dan penggunaannya melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan certainty of response index dan penggunaannya;
- Bandingkan bukti certainty of response index dan penggunaannya pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek certainty of response index dan penggunaannya dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian certainty of response index dan penggunaannya, topik kimia organik dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa reaksi dipahami hanya sebagai penggantian simbol tanpa mekanisme. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta penelusuran pusat elektron dan perubahan ikatan. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa mekanisme menjelaskan urutan perpindahan elektron dan perubahan struktur. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menggabungkan ketepatan jawaban, kualitas alasan, dan derajat keyakinan untuk memperkaya klasifikasi, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, certainty of response index dan penggunaannya harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.5.11 Tes dua tingkat untuk diagnosis konseptual

Tes dua tingkat untuk diagnosis konseptual tidak berdiri sebagai prosedur administratif. Ia menjadi bagian dari keputusan tentang apa yang perlu dijelaskan ulang, representasi apa yang

perlu dibandingkan, dan siapa yang memerlukan dukungan berbeda. Oleh sebab itu, bukti diagnosis harus langsung terhubung dengan tindakan pembelajaran, bukan berhenti pada daftar kesalahan (Yan & Subramaniam, 2018; Hasan et al., 1999).

Secara operasional, guru dapat menggunakan predict-observe-explain untuk menelusuri koherensi alasan. Data tersebut perlu dibaca bersama konsistensi lintas representasi, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mencampur kesalahan matematika dengan kesalahan konsep. Untuk mengurangnya, guru dapat menggunakan pertanyaan tindak lanjut dan mengulang diagnosis setelah remediasi. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai tes dua tingkat untuk diagnosis konseptual, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan tes dua tingkat untuk diagnosis konseptual melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan tes dua tingkat untuk diagnosis konseptual;
- Bandingkan bukti tes dua tingkat untuk diagnosis konseptual pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek tes dua tingkat untuk diagnosis konseptual dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian tes dua tingkat untuk diagnosis konseptual, topik struktur atom dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa elektron dipandang bergerak pada lintasan planet yang tetap. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta peserta didik membandingkan model Bohr dan model orbital serta menjelaskan batas setiap model. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa model atom adalah representasi teoretis dengan domain penggunaan tertentu. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menggabungkan ketepatan jawaban, kualitas alasan, dan derajat keyakinan untuk memperkaya klasifikasi, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, tes dua tingkat untuk diagnosis konseptual harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.5.12 Tes tiga tingkat dan penambahan keyakinan

Tes tiga tingkat dan penambahan keyakinan perlu ditempatkan dalam kerangka diagnosis yang menggabungkan ketepatan jawaban, kualitas alasan, dan derajat keyakinan untuk memperkaya klasifikasi. Fokusnya bukan sekadar menghasilkan skor, melainkan memperoleh bukti yang cukup untuk menjelaskan bagaimana peserta didik membangun, menghubungkan, dan menggunakan konsep. Karena konsepsi tidak dapat diamati secara langsung, guru harus menarik inferensi dari respons verbal, tulisan, gambar, tindakan, dan perubahan strategi yang tampak selama penyelesaian tugas (Hasan et al., 1999; Koriati, 1997).

Secara operasional, guru dapat menggunakan tugas isomorfik untuk menelusuri tingkat keyakinan. Data tersebut perlu dibaca bersama kemampuan transfer, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menggunakan ambang kategori tanpa bukti. Untuk menguranginya, guru dapat menyajikan tugas paralel dan mencatat derajat keyakinan. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai tes tiga tingkat dan penambahan keyakinan, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan tes tiga tingkat dan penambahan keyakinan melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan tes tiga tingkat dan penambahan keyakinan;
- Bandingkan bukti tes tiga tingkat dan penambahan keyakinan pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek tes tiga tingkat dan penambahan keyakinan dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian tes tiga tingkat dan penambahan keyakinan, topik mol dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa satu mol semua zat dianggap memiliki massa yang sama. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta perbandingan jumlah partikel dan massa satu mol beberapa zat. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa mol menyatakan jumlah entitas, sedangkan massa bergantung pada massa molar. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menggabungkan ketepatan jawaban, kualitas alasan, dan derajat keyakinan untuk memperkaya klasifikasi, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Kerangka operasional

Aspek	Bukti yang dicari	Risiko interpretasi	Tindak lanjut
Target konsep	Penjelasan tentang mol	Target terlalu luas	Batasi subkonsep dan prasyarat
Respons	Tingkat keyakinan	Menggunakan ambang kategori tanpa bukti	Menyajikan tugas paralel
Penalaran	Kemampuan transfer	Alasan diperlakukan sekadar tambahan	Mencatat derajat keyakinan
Konfirmasi	Tugas isomorfik atau tugas isomorfik	Kesimpulan dibuat dari satu konteks	Gunakan bukti kedua dan rediagnosis

Implikasinya, tes tiga tingkat dan penambahan keyakinan harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.5.13 Tes empat tingkat dan pemisahan keyakinan jawaban-alasan

Pembahasan tentang tes empat tingkat dan pemisahan keyakinan jawaban-alasan berangkat dari prinsip bahwa data asesmen selalu merupakan jejak tidak langsung dari proses berpikir. Satu pilihan jawaban hanya menunjukkan produk akhir, sedangkan struktur pengetahuan yang melahirkannya dapat berbeda. Diagnosis yang kuat karena itu menggabungkan konsistensi lintas butir, stabilitas setelah jeda waktu, dan bukti kontekstual sehingga keputusan pedagogis tidak bergantung pada satu indikator saja (Koriat, 1997; Brandriet & Bretz, 2014).

Secara operasional, guru dapat menggunakan kuis formatif digital untuk menelusuri konsistensi lintas butir. Data tersebut perlu dibaca bersama stabilitas setelah jeda waktu, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah memperlakukan konsepsi sebagai sifat permanen. Untuk mengurangnya, guru dapat memeriksa konsep prasyarat dan menelusuri langkah penyelesaian. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai tes empat tingkat dan pemisahan keyakinan jawaban-alasan, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan tes empat tingkat dan pemisahan keyakinan jawaban-alasan melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan tes empat tingkat dan pemisahan keyakinan jawaban-alasan;

- Bandingkan bukti tes empat tingkat dan pemisahan keyakinan jawaban-alasan pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek tes empat tingkat dan pemisahan keyakinan jawaban-alasan dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian tes empat tingkat dan pemisahan keyakinan jawaban-alasan, topik konsentrasi dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa konsentrasi disamakan dengan jumlah zat terlarut tanpa memperhatikan volume. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta perbandingan beberapa larutan dengan jumlah mol dan volume berbeda. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa konsentrasi adalah besaran rasio yang bergantung pada pembilang dan penyebut. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menggabungkan ketepatan jawaban, kualitas alasan, dan derajat keyakinan untuk memperkaya klasifikasi, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, tes empat tingkat dan pemisahan keyakinan jawaban-alasan harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.5.14 Alasan terbuka dan alasan pilihan ganda

Dalam pembelajaran kimia, alasan terbuka dan alasan pilihan ganda memiliki arti strategis karena banyak jawaban dapat diperoleh melalui hafalan prosedur atau pengenalan pola. Peserta didik mungkin tampak berhasil tetapi tidak mampu menerjemahkan makna simbol, model, atau besaran yang digunakan. Prinsip diagnosis menuntut bukti tentang cara konsep dipakai, bukan hanya apakah hasil akhirnya sesuai kunci (Brandriet & Bretz, 2014; Yan & Subramaniam, 2018).

Secara operasional, guru dapat menggunakan tes tiga tingkat untuk menelusuri konsistensi lintas representasi. Data tersebut perlu dibaca bersama urutan langkah penyelesaian, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mengabaikan ketidakpastian pengukuran. Untuk menguranginya, guru dapat meminta translasi antarrepresentasi dan menggunakan triangulasi bukti. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai alasan terbuka dan alasan pilihan ganda, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan alasan terbuka dan alasan pilihan ganda melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan alasan terbuka dan alasan pilihan ganda;
- Bandingkan bukti alasan terbuka dan alasan pilihan ganda pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek alasan terbuka dan alasan pilihan ganda dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian alasan terbuka dan alasan pilihan ganda, topik kelarutan dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa larutan jenuh dianggap tidak mengandung proses pelarutan lagi. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta model partikel pelarutan dan pengendapan pada keadaan jenuh. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa kejenuhan adalah keadaan kesetimbangan dinamis. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menggabungkan ketepatan jawaban, kualitas alasan, dan derajat keyakinan untuk memperkaya klasifikasi, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, alasan terbuka dan alasan pilihan ganda harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.5.15 Kualitas distraktor sebagai representasi konsepsi alternatif

Aspek kualitas distraktor sebagai representasi konsepsi alternatif menegaskan bahwa diagnosis adalah kegiatan ilmiah berskala kelas: guru merumuskan pertanyaan, mengumpulkan bukti, mempertimbangkan penjelasan alternatif, lalu menentukan kesimpulan yang paling masuk akal. Inferensi harus proporsional terhadap kualitas data. Semakin spesifik label yang diberikan, semakin kuat pula bukti yang diperlukan (Yan & Subramaniam, 2018; Hasan et al., 1999).

Secara operasional, guru dapat menggunakan wawancara semi-terstruktur untuk menelusuri kemampuan transfer. Data tersebut perlu dibaca bersama kualitas model mental yang dieksternalisasi, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda.

Bahaya utama pada tahap ini adalah menyimpulkan miskonsepsi hanya dari satu jawaban. Untuk mengurangnya, guru dapat mengulang diagnosis setelah remediasi dan menyusun tindak lanjut berbeda untuk kategori respons yang berbeda. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai kualitas distraktor sebagai representasi konsepsi alternatif, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan kualitas distraktor sebagai representasi konsepsi alternatif melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan kualitas distraktor sebagai representasi konsepsi alternatif;
- Bandingkan bukti kualitas distraktor sebagai representasi konsepsi alternatif pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek kualitas distraktor sebagai representasi konsepsi alternatif dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian kualitas distraktor sebagai representasi konsepsi alternatif, topik redoks dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa oksidasi selalu dianggap melibatkan oksigen. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta analisis transfer elektron dan perubahan bilangan oksidasi. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa pengikatan oksigen merupakan kasus khusus dari kerangka redoks yang lebih luas. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menggabungkan ketepatan jawaban, kualitas alasan, dan derajat keyakinan untuk memperkaya klasifikasi, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, kualitas distraktor sebagai representasi konsepsi alternatif harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.5.16 Kalibrasi keyakinan dan meta-ignorance

Penerapan kalibrasi keyakinan dan meta-ignorance harus menghubungkan teori pengukuran dengan pengetahuan pedagogis kimia. Instrumen yang teknis tetapi tidak memahami struktur konsep akan menghasilkan kategori dangkal; sebaliknya, pengetahuan konten tanpa prosedur

pengumpulan bukti dapat menghasilkan kesimpulan intuitif yang sulit diuji. Keduanya perlu dipadukan dalam satu alur yang transparan (Hasan et al., 1999; Koriati, 1997).

Secara operasional, guru dapat menggunakan gambar partikel untuk menelusuri stabilitas setelah jeda waktu. Data tersebut perlu dibaca bersama respons terhadap probing, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mencampur kesalahan matematika dengan kesalahan konsep. Untuk mengurangnya, guru dapat mencatat derajat keyakinan dan membandingkan jawaban dengan alasan. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai kalibrasi keyakinan dan meta-ignorance, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan kalibrasi keyakinan dan meta-ignorance melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan kalibrasi keyakinan dan meta-ignorance;
- Bandingkan bukti kalibrasi keyakinan dan meta-ignorance pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek kalibrasi keyakinan dan meta-ignorance dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian kalibrasi keyakinan dan meta-ignorance, topik kimia organik dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa reaksi dipahami hanya sebagai penggantian simbol tanpa mekanisme. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta penelusuran pusat elektron dan perubahan ikatan. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa mekanisme menjelaskan urutan perpindahan elektron dan perubahan struktur. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menggabungkan ketepatan jawaban, kualitas alasan, dan derajat keyakinan untuk memperkaya klasifikasi, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Kerangka operasional

Aspek	Bukti yang dicari	Risiko interpretasi	Tindak lanjut
Target konsep	Penjelasan tentang kimia organik	Target terlalu luas	Batasi subkonsep dan prasyarat
Respons	Stabilitas setelah jeda waktu	Mencampur kesalahan matematika dengan kesalahan konsep	Mencatat derajat keyakinan

Aspek	Bukti yang dicari	Risiko interpretasi	Tindak lanjut
Penalaran	Respons terhadap probing	Alasan diperlakukan sekadar tambahan	Membandingkan jawaban dengan alasan
Konfirmasi	Gambar partikel atau tugas isomorfik	Kesimpulan dibuat dari satu konteks	Gunakan bukti kedua dan rediagnosis

Implikasinya, kalibrasi keyakinan dan meta-ignorance harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.5.17 Ketidaksesuaian antara performa dan keyakinan

Ketidaksesuaian antara performa dan keyakinan tidak berdiri sebagai prosedur administratif. Ia menjadi bagian dari keputusan tentang apa yang perlu dijelaskan ulang, representasi apa yang perlu dibandingkan, dan siapa yang memerlukan dukungan berbeda. Oleh sebab itu, bukti diagnosis harus langsung terhubung dengan tindakan pembelajaran, bukan berhenti pada daftar kesalahan (Koriat, 1997; Brandriet & Bretz, 2014).

Secara operasional, guru dapat menggunakan observasi praktikum untuk menelusuri urutan langkah penyelesaian. Data tersebut perlu dibaca bersama ketepatan pilihan, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menggunakan ambang kategori tanpa bukti. Untuk mengurangnya, guru dapat menelusuri langkah penyelesaian dan menggunakan pertanyaan tindak lanjut. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai ketidaksesuaian antara performa dan keyakinan, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan ketidaksesuaian antara performa dan keyakinan melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan ketidaksesuaian antara performa dan keyakinan;
- Bandingkan bukti ketidaksesuaian antara performa dan keyakinan pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek ketidaksesuaian antara performa dan keyakinan dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian ketidaksesuaian antara performa dan keyakinan, topik struktur atom dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering

muncul ialah bahwa elektron dipandang bergerak pada lintasan planet yang tetap. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta peserta didik membandingkan model Bohr dan model orbital serta menjelaskan batas setiap model. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa model atom adalah representasi teoretis dengan domain penggunaan tertentu. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menggabungkan ketepatan jawaban, kualitas alasan, dan derajat keyakinan untuk memperkaya klasifikasi, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, ketidaksesuaian antara performa dan keyakinan harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.5.18 Analisis pola respons per butir

Analisis pola respons per butir perlu ditempatkan dalam kerangka diagnosis yang menggabungkan ketepatan jawaban, kualitas alasan, dan derajat keyakinan untuk memperkaya klasifikasi. Fokusnya bukan sekadar menghasilkan skor, melainkan memperoleh bukti yang cukup untuk menjelaskan bagaimana peserta didik membangun, menghubungkan, dan menggunakan konsep. Karena konsepsi tidak dapat diamati secara langsung, guru harus menarik inferensi dari respons verbal, tulisan, gambar, tindakan, dan perubahan strategi yang tampak selama penyelesaian tugas (Brandriet & Bretz, 2014; Yan & Subramaniam, 2018).

Secara operasional, guru dapat menggunakan exit ticket untuk menelusuri kualitas model mental yang dieksternalisasi. Data tersebut perlu dibaca bersama koherensi alasan, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah memperlakukan konsepsi sebagai sifat permanen. Untuk mengurangnya, guru dapat menggunakan triangulasi bukti dan menyajikan tugas paralel. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai analisis pola respons per butir, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan analisis pola respons per butir melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan analisis pola respons per butir;

- Bandingkan bukti analisis pola respons per butir pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek analisis pola respons per butir dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian analisis pola respons per butir, topik mol dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa satu mol semua zat dianggap memiliki massa yang sama. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta perbandingan jumlah partikel dan massa satu mol beberapa zat. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa mol menyatakan jumlah entitas, sedangkan massa bergantung pada massa molar. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menggabungkan ketepatan jawaban, kualitas alasan, dan derajat keyakinan untuk memperkaya klasifikasi, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, analisis pola respons per butir harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.5.19 Analisis pola respons per peserta didik

Pembahasan tentang analisis pola respons per peserta didik berangkat dari prinsip bahwa data asesmen selalu merupakan jejak tidak langsung dari proses berpikir. Satu pilihan jawaban hanya menunjukkan produk akhir, sedangkan struktur pengetahuan yang melahirkannya dapat berbeda. Diagnosis yang kuat karena itu menggabungkan respons terhadap probing, tingkat keyakinan, dan bukti kontekstual sehingga keputusan pedagogis tidak bergantung pada satu indikator saja (Yan & Subramaniam, 2018; Hasan et al., 1999).

Secara operasional, guru dapat menggunakan tes pilihan ganda dua tingkat untuk menelusuri respons terhadap probing. Data tersebut perlu dibaca bersama tingkat keyakinan, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mengabaikan ketidakpastian pengukuran. Untuk menguranginya, guru dapat menyusun tindak lanjut berbeda untuk kategori respons yang berbeda dan memeriksa konsep prasyarat. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai analisis pola respons per peserta didik, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;

- Pilih tugas yang merepresentasikan analisis pola respons per peserta didik melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan analisis pola respons per peserta didik;
- Bandingkan bukti analisis pola respons per peserta didik pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek analisis pola respons per peserta didik dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian analisis pola respons per peserta didik, topik konsentrasi dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa konsentrasi disamakan dengan jumlah zat terlarut tanpa memperhatikan volume. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta perbandingan beberapa larutan dengan jumlah mol dan volume berbeda. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa konsentrasi adalah besaran rasio yang bergantung pada pembilang dan penyebut. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menggabungkan ketepatan jawaban, kualitas alasan, dan derajat keyakinan untuk memperkaya klasifikasi, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, analisis pola respons per peserta didik harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.5.20 Batas penggunaan skor keyakinan

Dalam pembelajaran kimia, batas penggunaan skor keyakinan memiliki arti strategis karena banyak jawaban dapat diperoleh melalui hafalan prosedur atau pengenalan pola. Peserta didik mungkin tampak berhasil tetapi tidak mampu menerjemahkan makna simbol, model, atau besaran yang digunakan. Prinsip diagnosis menuntut bukti tentang cara konsep dipakai, bukan hanya apakah hasil akhirnya sesuai kunci (Hasan et al., 1999; Koriati, 1997).

Secara operasional, guru dapat menggunakan alasan terbuka untuk menelusuri ketepatan pilihan. Data tersebut perlu dibaca bersama konsistensi lintas butir, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menyimpulkan miskonsepsi hanya dari satu jawaban. Untuk menguranginya, guru dapat membandingkan jawaban dengan alasan dan meminta translasi antarrepresentasi. Prosedur

tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai batas penggunaan skor keyakinan, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan batas penggunaan skor keyakinan melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan batas penggunaan skor keyakinan;
- Bandingkan bukti batas penggunaan skor keyakinan pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek batas penggunaan skor keyakinan dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian batas penggunaan skor keyakinan, topik kelarutan dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa larutan jenuh dianggap tidak mengandung proses pelarutan lagi. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta model partikel pelarutan dan pengendapan pada keadaan jenuh. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa kejenuhan adalah keadaan kesetimbangan dinamis. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menggabungkan ketepatan jawaban, kualitas alasan, dan derajat keyakinan untuk memperkaya klasifikasi, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Kerangka operasional

Aspek	Bukti yang dicari	Risiko interpretasi	Tindak lanjut
Target konsep	Penjelasan tentang kelarutan	Target terlalu luas	Batasi subkonsep dan prasyarat
Respons	Ketepatan pilihan	Menyimpulkan miskonsepsi hanya dari satu jawaban	Membandingkan jawaban dengan alasan
Penalaran	Konsistensi lintas butir	Alasan diperlakukan sekadar tambahan	Meminta translasi antarrepresentasi
Konfirmasi	Alasan terbuka atau tugas isomorfik	Kesimpulan dibuat dari satu konteks	Gunakan bukti kedua dan rediagnosis

Implikasinya, batas penggunaan skor keyakinan harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.5.21 Wawancara konfirmasi terhadap klasifikasi tier

Aspek wawancara konfirmasi terhadap klasifikasi tier menegaskan bahwa diagnosis adalah kegiatan ilmiah berskala kelas: guru merumuskan pertanyaan, mengumpulkan bukti, mempertimbangkan penjelasan alternatif, lalu menentukan kesimpulan yang paling masuk akal. Inferensi harus proporsional terhadap kualitas data. Semakin spesifik label yang diberikan, semakin kuat pula bukti yang diperlukan (Koriat, 1997; Brandriet & Bretz, 2014).

Secara operasional, guru dapat menggunakan peta konsep untuk menelusuri koherensi alasan. Data tersebut perlu dibaca bersama konsistensi lintas representasi, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mencampur kesalahan matematika dengan kesalahan konsep. Untuk mengurangnya, guru dapat menggunakan pertanyaan tindak lanjut dan mengulang diagnosis setelah remediasi. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai wawancara konfirmasi terhadap klasifikasi tier, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan wawancara konfirmasi terhadap klasifikasi tier melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan wawancara konfirmasi terhadap klasifikasi tier;
- Bandingkan bukti wawancara konfirmasi terhadap klasifikasi tier pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek wawancara konfirmasi terhadap klasifikasi tier dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian wawancara konfirmasi terhadap klasifikasi tier, topik redoks dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa oksidasi selalu dianggap melibatkan oksigen. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta analisis transfer elektron dan perubahan bilangan oksidasi. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa pengikatan oksigen merupakan kasus khusus dari kerangka redoks yang lebih luas. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menggabungkan ketepatan jawaban, kualitas alasan, dan derajat keyakinan untuk memperkaya klasifikasi, bukan diperlakukan sebagai skor

yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, wawancara konfirmasi terhadap klasifikasi tier harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.5.22 Contoh matriks jawaban-alasan-keyakinan

Penerapan contoh matriks jawaban-alasan-keyakinan harus menghubungkan teori pengukuran dengan pengetahuan pedagogis kimia. Instrumen yang teknis tetapi tidak memahami struktur konsep akan menghasilkan kategori dangkal; sebaliknya, pengetahuan konten tanpa prosedur pengumpulan bukti dapat menghasilkan kesimpulan intuitif yang sulit diuji. Keduanya perlu dipadukan dalam satu alur yang transparan (Brandriet & Bretz, 2014; Yan & Subramaniam, 2018).

Secara operasional, guru dapat menggunakan analisis respons tertulis untuk menelusuri tingkat keyakinan. Data tersebut perlu dibaca bersama kemampuan transfer, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menggunakan ambang kategori tanpa bukti. Untuk mengurangnya, guru dapat menyajikan tugas paralel dan mencatat derajat keyakinan. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai contoh matriks jawaban-alasan-keyakinan, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan contoh matriks jawaban-alasan-keyakinan melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan contoh matriks jawaban-alasan-keyakinan;
- Bandingkan bukti contoh matriks jawaban-alasan-keyakinan pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek contoh matriks jawaban-alasan-keyakinan dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian contoh matriks jawaban-alasan-keyakinan, topik kimia organik dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa reaksi dipahami hanya sebagai penggantian simbol tanpa mekanisme. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta penelusuran pusat

elektron dan perubahan ikatan. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa mekanisme menjelaskan urutan perpindahan elektron dan perubahan struktur. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menggabungkan ketepatan jawaban, kualitas alasan, dan derajat keyakinan untuk memperkaya klasifikasi, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, contoh matriks jawaban-alasan-keyakinan harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.5.23 Implikasi bagi umpan balik dan remediasi

Implikasi bagi umpan balik dan remediasi tidak berdiri sebagai prosedur administratif. Ia menjadi bagian dari keputusan tentang apa yang perlu dijelaskan ulang, representasi apa yang perlu dibandingkan, dan siapa yang memerlukan dukungan berbeda. Oleh sebab itu, bukti diagnosis harus langsung terhubung dengan tindakan pembelajaran, bukan berhenti pada daftar kesalahan (Yan & Subramaniam, 2018; Hasan et al., 1999).

Secara operasional, guru dapat menggunakan diskusi argumentatif untuk menelusuri konsistensi lintas butir. Data tersebut perlu dibaca bersama stabilitas setelah jeda waktu, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah memperlakukan konsepsi sebagai sifat permanen. Untuk mengurangnya, guru dapat memeriksa konsep prasyarat dan menelusuri langkah penyelesaian. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai implikasi bagi umpan balik dan remediasi, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan implikasi bagi umpan balik dan remediasi melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan implikasi bagi umpan balik dan remediasi;
- Bandingkan bukti implikasi bagi umpan balik dan remediasi pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek implikasi bagi umpan balik dan remediasi dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian implikasi bagi umpan balik dan remediasi, topik struktur atom dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah

bahwa elektron dipandang bergerak pada lintasan planet yang tetap. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta peserta didik membandingkan model Bohr dan model orbital serta menjelaskan batas setiap model. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa model atom adalah representasi teoretis dengan domain penggunaan tertentu. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menggabungkan ketepatan jawaban, kualitas alasan, dan derajat keyakinan untuk memperkaya klasifikasi, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, implikasi bagi umpan balik dan remediasi harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.5.24 Sintesis hubungan jawaban, alasan, dan keyakinan

Sintesis hubungan jawaban, alasan, dan keyakinan perlu ditempatkan dalam kerangka diagnosis yang menggabungkan ketepatan jawaban, kualitas alasan, dan derajat keyakinan untuk memperkaya klasifikasi. Fokusnya bukan sekadar menghasilkan skor, melainkan memperoleh bukti yang cukup untuk menjelaskan bagaimana peserta didik membangun, menghubungkan, dan menggunakan konsep. Karena konsepsi tidak dapat diamati secara langsung, guru harus menarik inferensi dari respons verbal, tulisan, gambar, tindakan, dan perubahan strategi yang tampak selama penyelesaian tugas (Hasan et al., 1999; Koriati, 1997).

Secara operasional, guru dapat menggunakan analisis grafik untuk menelusuri konsistensi lintas representasi. Data tersebut perlu dibaca bersama urutan langkah penyelesaian, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mengabaikan ketidakpastian pengukuran. Untuk mengurangnya, guru dapat meminta translasi antarrepresentasi dan menggunakan triangulasi bukti. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai sintesis hubungan jawaban, alasan, dan keyakinan, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan sintesis hubungan jawaban, alasan, dan keyakinan melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan sintesis hubungan jawaban, alasan, dan keyakinan;

- Bandingkan bukti sintesis hubungan jawaban, alasan, dan keyakinan pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek sintesis hubungan jawaban, alasan, dan keyakinan dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian sintesis hubungan jawaban, alasan, dan keyakinan, topik mol dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa satu mol semua zat dianggap memiliki massa yang sama. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta perbandingan jumlah partikel dan massa satu mol beberapa zat. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa mol menyatakan jumlah entitas, sedangkan massa bergantung pada massa molar. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menggabungkan ketepatan jawaban, kualitas alasan, dan derajat keyakinan untuk memperkaya klasifikasi, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, sintesis hubungan jawaban, alasan, dan keyakinan harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.5.25 Implikasi praktis bagi guru kimia

Implikasi praktis subbab 5.5 dapat dirangkum sebagai alur: tentukan konsep target, pilih bukti yang relevan, kumpulkan respons dalam lebih dari satu bentuk, analisis pola, nyatakan derajat kepastian, susun tindak lanjut, dan lakukan rediagnosis. Alur ini menjaga agar hubungan jawaban, alasan, dan keyakinan tidak berubah menjadi administrasi skor, tetapi berfungsi sebagai mekanisme perbaikan pembelajaran yang berulang.

29. Identifikasi konsep inti dan konsep prasyarat sebelum menyusun instrumen.
30. Gunakan pertanyaan yang meminta alasan, representasi, atau langkah kerja.
31. Bandingkan respons pada beberapa konteks dan waktu.
32. Pisahkan bukti faktual dari interpretasi guru.
33. Tentukan tindak lanjut yang berbeda untuk kategori kebutuhan yang berbeda.
34. Simpan data secukupnya dan lindungi kerahasiaan peserta didik.
35. Periksa kembali perubahan melalui tugas transfer dan retensi.

5.6 Prinsip Validitas Diagnosis

Subbab ini menjamin bahwa interpretasi hasil diagnosis didukung bukti yang memadai dan sesuai tujuan penggunaannya. Uraian disusun dari prinsip teoretis menuju prosedur kelas, contoh kimia, risiko interpretasi, dan tindak lanjut. Setiap bagian menekankan bahwa kesimpulan diagnostik harus didasarkan pada pola bukti, bukan pada satu respons yang terpisah.

5.6.1 Makna validitas dalam diagnosis miskonsepsi

Pembahasan tentang makna validitas dalam diagnosis miskonsepsi berangkat dari prinsip bahwa data asesmen selalu merupakan jejak tidak langsung dari proses berpikir. Satu pilihan jawaban hanya menunjukkan produk akhir, sedangkan struktur pengetahuan yang melahirkannya dapat berbeda. Diagnosis yang kuat karena itu menggabungkan koherensi alasan, konsistensi lintas representasi, dan bukti kontekstual sehingga keputusan pedagogis tidak bergantung pada satu indikator saja (Messick, 1995; Kane, 2013).

Secara operasional, guru dapat menggunakan kuis formatif digital untuk menelusuri koherensi alasan. Data tersebut perlu dibaca bersama konsistensi lintas representasi, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mencampur kesalahan matematika dengan kesalahan konsep. Untuk mengurangnya, guru dapat menggunakan pertanyaan tindak lanjut dan mengulang diagnosis setelah remediasi. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai makna validitas dalam diagnosis miskonsepsi, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan makna validitas dalam diagnosis miskonsepsi melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan makna validitas dalam diagnosis miskonsepsi;
- Bandingkan bukti makna validitas dalam diagnosis miskonsepsi pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek makna validitas dalam diagnosis miskonsepsi dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian makna validitas dalam diagnosis miskonsepsi, topik asam-basa dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa asam kuat dianggap selalu berkonsentrasi tinggi. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta pemisahan konsep derajat ionisasi, K_a , konsentrasi, dan pH. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa kekuatan asam dan konsentrasi adalah dimensi berbeda. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menjamin bahwa interpretasi hasil diagnosis didukung bukti

yang memadai dan sesuai tujuan penggunaannya, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, makna validitas dalam diagnosis miskonsepsi harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.6.2 Validitas melekat pada interpretasi dan penggunaan hasil

Dalam pembelajaran kimia, validitas melekat pada interpretasi dan penggunaan hasil memiliki arti strategis karena banyak jawaban dapat diperoleh melalui hafalan prosedur atau pengenalan pola. Peserta didik mungkin tampak berhasil tetapi tidak mampu menerjemahkan makna simbol, model, atau besaran yang digunakan. Prinsip diagnosis menuntut bukti tentang cara konsep dipakai, bukan hanya apakah hasil akhirnya sesuai kunci (Kane, 2013; Downing, 2003).

Secara operasional, guru dapat menggunakan tes tiga tingkat untuk menelusuri tingkat keyakinan. Data tersebut perlu dibaca bersama kemampuan transfer, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menggunakan ambang kategori tanpa bukti. Untuk menguranginya, guru dapat menyajikan tugas paralel dan mencatat derajat keyakinan. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai validitas melekat pada interpretasi dan penggunaan hasil, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan validitas melekat pada interpretasi dan penggunaan hasil melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan validitas melekat pada interpretasi dan penggunaan hasil;
- Bandingkan bukti validitas melekat pada interpretasi dan penggunaan hasil pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek validitas melekat pada interpretasi dan penggunaan hasil dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian validitas melekat pada interpretasi dan penggunaan hasil, topik kinetika dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa reaksi eksoterm dianggap selalu cepat. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan

relasi konseptual yang tepat. Guru dapat meminta perbandingan profil energi dan faktor kinetik. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa termodinamika dan kinetika menjawab pertanyaan berbeda. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menjamin bahwa interpretasi hasil diagnosis didukung bukti yang memadai dan sesuai tujuan penggunaannya, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, validitas melekat pada interpretasi dan penggunaan hasil harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.6.3 Argumentasi interpretasi dan penggunaan skor

Aspek argumentasi interpretasi dan penggunaan skor menegaskan bahwa diagnosis adalah kegiatan ilmiah berskala kelas: guru merumuskan pertanyaan, mengumpulkan bukti, mempertimbangkan penjelasan alternatif, lalu menentukan kesimpulan yang paling masuk akal. Inferensi harus proporsional terhadap kualitas data. Semakin spesifik label yang diberikan, semakin kuat pula bukti yang diperlukan (Downing, 2003; AERA et al., 2014).

Secara operasional, guru dapat menggunakan wawancara semi-terstruktur untuk menelusuri konsistensi lintas butir. Data tersebut perlu dibaca bersama stabilitas setelah jeda waktu, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah memperlakukan konsepsi sebagai sifat permanen. Untuk mengurangnya, guru dapat memeriksa konsep prasyarat dan menelusuri langkah penyelesaian. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai argumentasi interpretasi dan penggunaan skor, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan argumentasi interpretasi dan penggunaan skor melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan argumentasi interpretasi dan penggunaan skor;
- Bandingkan bukti argumentasi interpretasi dan penggunaan skor pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek argumentasi interpretasi dan penggunaan skor dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian argumentasi interpretasi dan penggunaan skor, topik elektrokimia dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa elektron dianggap bergerak melalui jembatan garam. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta penelusuran jalur elektron dan ion pada sel galvanik. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa elektron mengalir pada rangkaian luar, sedangkan ion menjaga kenetralan muatan. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menjamin bahwa interpretasi hasil diagnosis didukung bukti yang memadai dan sesuai tujuan penggunaannya, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, argumentasi interpretasi dan penggunaan skor harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.6.4 Representasi konstruk miskonsepsi

Penerapan representasi konstruk miskonsepsi harus menghubungkan teori pengukuran dengan pengetahuan pedagogis kimia. Instrumen yang teknis tetapi tidak memahami struktur konsep akan menghasilkan kategori dangkal; sebaliknya, pengetahuan konten tanpa prosedur pengumpulan bukti dapat menghasilkan kesimpulan intuitif yang sulit diuji. Keduanya perlu dipadukan dalam satu alur yang transparan (AERA et al., 2014; Messick, 1995).

Secara operasional, guru dapat menggunakan gambar partikel untuk menelusuri konsistensi lintas representasi. Data tersebut perlu dibaca bersama urutan langkah penyelesaian, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mengabaikan ketidakpastian pengukuran. Untuk menguranginya, guru dapat meminta translasi antarrepresentasi dan menggunakan triangulasi bukti. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai representasi konstruk miskonsepsi, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan representasi konstruk miskonsepsi melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan representasi konstruk miskonsepsi;

- Bandingkan bukti representasi konstruk miskonsepsi pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek representasi konstruk miskonsepsi dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian representasi konstruk miskonsepsi, topik praktikum titrasi dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa titik akhir dianggap selalu identik dengan titik ekuivalen. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta analisis indikator, kurva titrasi, dan ketidakpastian. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa titik akhir adalah observasi operasional yang mendekati titik ekuivalen. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menjamin bahwa interpretasi hasil diagnosis didukung bukti yang memadai dan sesuai tujuan penggunaannya, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Kerangka operasional

Aspek	Bukti yang dicari	Risiko interpretasi	Tindak lanjut
Target konsep	Penjelasan tentang praktikum titrasi	Target terlalu luas	Batasi subkonsep dan prasyarat
Respons	Konsistensi lintas representasi	Mengabaikan ketidakpastian pengukuran	Meminta translasi antarrepresentasi
Penalaran	Urutan langkah penyelesaian	Alasan diperlakukan sekadar tambahan	Menggunakan triangulasi bukti
Konfirmasi	Gambar partikel atau tugas isomorfik	Kesimpulan dibuat dari satu konteks	Gunakan bukti kedua dan rediagnosis

Implikasinya, representasi konstruk miskonsepsi harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.6.5 Batas domain konsep yang didiagnosis

Batas domain konsep yang didiagnosis tidak berdiri sebagai prosedur administratif. Ia menjadi bagian dari keputusan tentang apa yang perlu dijelaskan ulang, representasi apa yang perlu

dibandingkan, dan siapa yang memerlukan dukungan berbeda. Oleh sebab itu, bukti diagnosis harus langsung terhubung dengan tindakan pembelajaran, bukan berhenti pada daftar kesalahan (Messick, 1995; Kane, 2013).

Secara operasional, guru dapat menggunakan observasi praktikum untuk menelusuri kemampuan transfer. Data tersebut perlu dibaca bersama kualitas model mental yang dieksternalisasi, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menyimpulkan miskonsepsi hanya dari satu jawaban. Untuk mengurangnya, guru dapat mengulang diagnosis setelah remediasi dan menyusun tindak lanjut berbeda untuk kategori respons yang berbeda. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai batas domain konsep yang didiagnosis, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan batas domain konsep yang didiagnosis melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan batas domain konsep yang didiagnosis;
- Bandingkan bukti batas domain konsep yang didiagnosis pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek batas domain konsep yang didiagnosis dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian batas domain konsep yang didiagnosis, topik sistem periodik dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa jari-jari atom dianggap selalu bertambah bersama nomor atom. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta prediksi tren disertai penjelasan muatan inti efektif dan kulit elektron. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa tren periodik perlu dijelaskan melalui struktur dan interaksi, bukan hafalan arah panah. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menjamin bahwa interpretasi hasil diagnosis didukung bukti yang memadai dan sesuai tujuan penggunaannya, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, batas domain konsep yang didiagnosis harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.6.6 Analisis konsep sebagai dasar kisi-kisi

Analisis konsep sebagai dasar kisi-kisi perlu ditempatkan dalam kerangka diagnosis yang menjamin bahwa interpretasi hasil diagnosis didukung bukti yang memadai dan sesuai tujuan penggunaannya. Fokusnya bukan sekadar menghasilkan skor, melainkan memperoleh bukti yang cukup untuk menjelaskan bagaimana peserta didik membangun, menghubungkan, dan menggunakan konsep. Karena konsepsi tidak dapat diamati secara langsung, guru harus menarik inferensi dari respons verbal, tulisan, gambar, tindakan, dan perubahan strategi yang tampak selama penyelesaian tugas (Kane, 2013; Downing, 2003).

Secara operasional, guru dapat menggunakan exit ticket untuk menelusuri stabilitas setelah jeda waktu. Data tersebut perlu dibaca bersama respons terhadap probing, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mencampur kesalahan matematika dengan kesalahan konsep. Untuk mengurangnya, guru dapat mencatat derajat keyakinan dan membandingkan jawaban dengan alasan. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai analisis konsep sebagai dasar kisi-kisi, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan analisis konsep sebagai dasar kisi-kisi melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan analisis konsep sebagai dasar kisi-kisi;
- Bandingkan bukti analisis konsep sebagai dasar kisi-kisi pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek analisis konsep sebagai dasar kisi-kisi dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian analisis konsep sebagai dasar kisi-kisi, topik stoikiometri dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa koefisien reaksi diperlakukan sebagai perbandingan massa langsung. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta translasi persamaan dari partikel ke mol dan massa. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa koefisien terutama menyatakan rasio jumlah entitas atau mol. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menjamin bahwa interpretasi hasil diagnosis didukung bukti yang memadai dan sesuai tujuan penggunaannya, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, analisis konsep sebagai dasar kisi-kisi harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.6.7 Bukti validitas berbasis isi

Pembahasan tentang bukti validitas berbasis isi berangkat dari prinsip bahwa data asesmen selalu merupakan jejak tidak langsung dari proses berpikir. Satu pilihan jawaban hanya menunjukkan produk akhir, sedangkan struktur pengetahuan yang melahirkannya dapat berbeda. Diagnosis yang kuat karena itu menggabungkan urutan langkah penyelesaian, ketepatan pilihan, dan bukti kontekstual sehingga keputusan pedagogis tidak bergantung pada satu indikator saja (Downing, 2003; AERA et al., 2014).

Secara operasional, guru dapat menggunakan tes pilihan ganda dua tingkat untuk menelusuri urutan langkah penyelesaian. Data tersebut perlu dibaca bersama ketepatan pilihan, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menggunakan ambang kategori tanpa bukti. Untuk mengurangnya, guru dapat menelusuri langkah penyelesaian dan menggunakan pertanyaan tindak lanjut. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai bukti validitas berbasis isi, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan bukti validitas berbasis isi melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan bukti validitas berbasis isi;
- Bandingkan bukti bukti validitas berbasis isi pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek bukti validitas berbasis isi dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian bukti validitas berbasis isi, topik asam-basa dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa asam kuat dianggap selalu berkonsentrasi tinggi. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta pemisahan konsep derajat ionisasi, K_a , konsentrasi, dan pH. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa kekuatan asam dan konsentrasi adalah dimensi berbeda. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menjamin bahwa interpretasi hasil diagnosis didukung bukti

yang memadai dan sesuai tujuan penggunaannya, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, bukti validitas berbasis isi harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.6.8 Bukti validitas berbasis proses respons

Dalam pembelajaran kimia, bukti validitas berbasis proses respons memiliki arti strategis karena banyak jawaban dapat diperoleh melalui hafalan prosedur atau pengenalan pola. Peserta didik mungkin tampak berhasil tetapi tidak mampu menerjemahkan makna simbol, model, atau besaran yang digunakan. Prinsip diagnosis menuntut bukti tentang cara konsep dipakai, bukan hanya apakah hasil akhirnya sesuai kunci (AERA et al., 2014; Messick, 1995).

Secara operasional, guru dapat menggunakan alasan terbuka untuk menelusuri kualitas model mental yang dieksternalisasi. Data tersebut perlu dibaca bersama koherensi alasan, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah memperlakukan konsepsi sebagai sifat permanen. Untuk mengurangnya, guru dapat menggunakan triangulasi bukti dan menyajikan tugas paralel. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai bukti validitas berbasis proses respons, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan bukti validitas berbasis proses respons melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan bukti validitas berbasis proses respons;
- Bandingkan bukti bukti validitas berbasis proses respons pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek bukti validitas berbasis proses respons dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian bukti validitas berbasis proses respons, topik kinetika dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa reaksi eksoterm dianggap selalu cepat. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta perbandingan profil energi dan faktor kinetik. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa termodinamika dan kinetika

menjawab pertanyaan berbeda. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menjamin bahwa interpretasi hasil diagnosis didukung bukti yang memadai dan sesuai tujuan penggunaannya, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Kerangka operasional

Aspek	Bukti yang dicari	Risiko interpretasi	Tindak lanjut
Target konsep	Penjelasan tentang kinetika	Target terlalu luas	Batasi subkonsep dan prasyarat
Respons	Kualitas model mental yang dieksternalisasi	Memperlakukan konsepsi sebagai sifat permanen	Menggunakan triangulasi bukti
Penalaran	Koherensi alasan	Alasan diperlakukan sekadar tambahan	Menyajikan tugas paralel
Konfirmasi	Alasan terbuka atau tugas isomorfik	Kesimpulan dibuat dari satu konteks	Gunakan bukti kedua dan rediagnosis

Implikasinya, bukti validitas berbasis proses respons harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.6.9 Bukti validitas berbasis struktur internal

Aspek bukti validitas berbasis struktur internal menegaskan bahwa diagnosis adalah kegiatan ilmiah berskala kelas: guru merumuskan pertanyaan, mengumpulkan bukti, mempertimbangkan penjelasan alternatif, lalu menentukan kesimpulan yang paling masuk akal. Inferensi harus proporsional terhadap kualitas data. Semakin spesifik label yang diberikan, semakin kuat pula bukti yang diperlukan (Messick, 1995; Kane, 2013).

Secara operasional, guru dapat menggunakan peta konsep untuk menelusuri respons terhadap probing. Data tersebut perlu dibaca bersama tingkat keyakinan, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mengabaikan ketidakpastian pengukuran. Untuk menguranginya, guru dapat menyusun tindak lanjut berbeda untuk kategori respons yang berbeda dan memeriksa konsep prasyarat. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai bukti validitas berbasis struktur internal, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan bukti validitas berbasis struktur internal melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan bukti validitas berbasis struktur internal;
- Bandingkan bukti bukti validitas berbasis struktur internal pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek bukti validitas berbasis struktur internal dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian bukti validitas berbasis struktur internal, topik elektrokimia dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa elektron dianggap bergerak melalui jembatan garam. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta penelusuran jalur elektron dan ion pada sel galvanik. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa elektron mengalir pada rangkaian luar, sedangkan ion menjaga kenetralan muatan. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menjamin bahwa interpretasi hasil diagnosis didukung bukti yang memadai dan sesuai tujuan penggunaannya, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, bukti validitas berbasis struktur internal harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.6.10 Bukti validitas berbasis hubungan dengan variabel lain

Penerapan bukti validitas berbasis hubungan dengan variabel lain harus menghubungkan teori pengukuran dengan pengetahuan pedagogis kimia. Instrumen yang teknis tetapi tidak memahami struktur konsep akan menghasilkan kategori dangkal; sebaliknya, pengetahuan konten tanpa prosedur pengumpulan bukti dapat menghasilkan kesimpulan intuitif yang sulit diuji. Keduanya perlu dipadukan dalam satu alur yang transparan (Kane, 2013; Downing, 2003).

Secara operasional, guru dapat menggunakan analisis respons tertulis untuk menelusuri ketepatan pilihan. Data tersebut perlu dibaca bersama konsistensi lintas butir, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap

ini adalah menyimpulkan miskonsepsi hanya dari satu jawaban. Untuk mengurangnya, guru dapat membandingkan jawaban dengan alasan dan meminta translasi antarrepresentasi. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai bukti validitas berbasis hubungan dengan variabel lain, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan bukti validitas berbasis hubungan dengan variabel lain melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan bukti validitas berbasis hubungan dengan variabel lain;
- Bandingkan bukti validitas berbasis hubungan dengan variabel lain pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek bukti validitas berbasis hubungan dengan variabel lain dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian bukti validitas berbasis hubungan dengan variabel lain, topik praktikum titrasi dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa titik akhir dianggap selalu identik dengan titik ekuivalen. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta analisis indikator, kurva titrasi, dan ketidakpastian. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa titik akhir adalah observasi operasional yang mendekati titik ekuivalen. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menjamin bahwa interpretasi hasil diagnosis didukung bukti yang memadai dan sesuai tujuan penggunaannya, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, bukti validitas berbasis hubungan dengan variabel lain harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.6.11 Bukti validitas berbasis konsekuensi penggunaan

Bukti validitas berbasis konsekuensi penggunaan tidak berdiri sebagai prosedur administratif. Ia menjadi bagian dari keputusan tentang apa yang perlu dijelaskan ulang, representasi apa yang perlu dibandingkan, dan siapa yang memerlukan dukungan berbeda. Oleh sebab itu, bukti diagnosis harus langsung terhubung dengan tindakan pembelajaran, bukan berhenti pada daftar kesalahan (Downing, 2003; AERA et al., 2014).

Secara operasional, guru dapat menggunakan diskusi argumentatif untuk menelusuri koherensi alasan. Data tersebut perlu dibaca bersama konsistensi lintas representasi, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mencampur kesalahan matematika dengan kesalahan konsep. Untuk mengurangnya, guru dapat menggunakan pertanyaan tindak lanjut dan mengulang diagnosis setelah remediasi. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai bukti validitas berbasis konsekuensi penggunaan, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan bukti validitas berbasis konsekuensi penggunaan melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan bukti validitas berbasis konsekuensi penggunaan;
- Bandingkan bukti validitas berbasis konsekuensi penggunaan pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek bukti validitas berbasis konsekuensi penggunaan dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian bukti validitas berbasis konsekuensi penggunaan, topik sistem periodik dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa jari-jari atom dianggap selalu bertambah bersama nomor atom. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta prediksi tren disertai penjelasan muatan inti efektif dan kulit elektron. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa tren periodik perlu dijelaskan melalui struktur dan interaksi, bukan hafalan arah panah. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menjamin bahwa interpretasi hasil diagnosis didukung bukti yang memadai dan sesuai tujuan penggunaannya, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, bukti validitas berbasis konsekuensi penggunaan harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.6.12 Validitas distraktor konseptual

Validitas distraktor konseptual perlu ditempatkan dalam kerangka diagnosis yang menjamin bahwa interpretasi hasil diagnosis didukung bukti yang memadai dan sesuai tujuan

penggunaannya. Fokusnya bukan sekadar menghasilkan skor, melainkan memperoleh bukti yang cukup untuk menjelaskan bagaimana peserta didik membangun, menghubungkan, dan menggunakan konsep. Karena konsepsi tidak dapat diamati secara langsung, guru harus menarik inferensi dari respons verbal, tulisan, gambar, tindakan, dan perubahan strategi yang tampak selama penyelesaian tugas (AERA et al., 2014; Messick, 1995).

Secara operasional, guru dapat menggunakan analisis grafik untuk menelusuri tingkat keyakinan. Data tersebut perlu dibaca bersama kemampuan transfer, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menggunakan ambang kategori tanpa bukti. Untuk mengurangnya, guru dapat menyajikan tugas paralel dan mencatat derajat keyakinan. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai validitas distraktor konseptual, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan validitas distraktor konseptual melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan validitas distraktor konseptual;
- Bandingkan bukti validitas distraktor konseptual pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek validitas distraktor konseptual dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian validitas distraktor konseptual, topik stoikiometri dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa koefisien reaksi diperlakukan sebagai perbandingan massa langsung. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta translasi persamaan dari partikel ke mol dan massa. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa koefisien terutama menyatakan rasio jumlah entitas atau mol. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menjamin bahwa interpretasi hasil diagnosis didukung bukti yang memadai dan sesuai tujuan penggunaannya, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Kerangka operasional

Aspek	Bukti yang dicari	Risiko interpretasi	Tindak lanjut
Target konsep	Penjelasan tentang stoikiometri	Target terlalu luas	Batasi subkonsep dan prasyarat

Aspek	Bukti yang dicari	Risiko interpretasi	Tindak lanjut
Respons	Tingkat keyakinan	Menggunakan ambang kategori tanpa bukti	Menyajikan tugas paralel
Penalaran	Kemampuan transfer	Alasan diperlakukan sekadar tambahan	Mencatat derajat keyakinan
Konfirmasi	Analisis grafik atau tugas isomorfik	Kesimpulan dibuat dari satu konteks	Gunakan bukti kedua dan rediagnosis

Implikasinya, validitas distraktor konseptual harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.6.13 Validitas alasan dan tingkat keyakinan

Pembahasan tentang validitas alasan dan tingkat keyakinan berangkat dari prinsip bahwa data asesmen selalu merupakan jejak tidak langsung dari proses berpikir. Satu pilihan jawaban hanya menunjukkan produk akhir, sedangkan struktur pengetahuan yang melahirkannya dapat berbeda. Diagnosis yang kuat karena itu menggabungkan konsistensi lintas butir, stabilitas setelah jeda waktu, dan bukti kontekstual sehingga keputusan pedagogis tidak bergantung pada satu indikator saja (Messick, 1995; Kane, 2013).

Secara operasional, guru dapat menggunakan tes empat tingkat untuk menelusuri konsistensi lintas butir. Data tersebut perlu dibaca bersama stabilitas setelah jeda waktu, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah memperlakukan konsepsi sebagai sifat permanen. Untuk mengurangnya, guru dapat memeriksa konsep prasyarat dan menelusuri langkah penyelesaian. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai validitas alasan dan tingkat keyakinan, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan validitas alasan dan tingkat keyakinan melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan validitas alasan dan tingkat keyakinan;
- Bandingkan bukti validitas alasan dan tingkat keyakinan pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek validitas alasan dan tingkat keyakinan dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian validitas alasan dan tingkat keyakinan, topik asam-basa dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa asam kuat dianggap selalu berkonsentrasi tinggi. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta pemisahan konsep derajat ionisasi, K_a , konsentrasi, dan pH. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa kekuatan asam dan konsentrasi adalah dimensi berbeda. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menjamin bahwa interpretasi hasil diagnosis didukung bukti yang memadai dan sesuai tujuan penggunaannya, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, validitas alasan dan tingkat keyakinan harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.6.14 Validitas wawancara diagnostik

Dalam pembelajaran kimia, validitas wawancara diagnostik memiliki arti strategis karena banyak jawaban dapat diperoleh melalui hafalan prosedur atau pengenalan pola. Peserta didik mungkin tampak berhasil tetapi tidak mampu menerjemahkan makna simbol, model, atau besaran yang digunakan. Prinsip diagnosis menuntut bukti tentang cara konsep dipakai, bukan hanya apakah hasil akhirnya sesuai kunci (Kane, 2013; Downing, 2003).

Secara operasional, guru dapat menggunakan think-aloud untuk menelusuri konsistensi lintas representasi. Data tersebut perlu dibaca bersama urutan langkah penyelesaian, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mengabaikan ketidakpastian pengukuran. Untuk menguranginya, guru dapat meminta translasi antarrepresentasi dan menggunakan triangulasi bukti. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai validitas wawancara diagnostik, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan validitas wawancara diagnostik melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan validitas wawancara diagnostik;

- Bandingkan bukti validitas wawancara diagnostik pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek validitas wawancara diagnostik dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian validitas wawancara diagnostik, topik kinetika dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa reaksi eksoterm dianggap selalu cepat. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta perbandingan profil energi dan faktor kinetik. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa termodinamika dan kinetika menjawab pertanyaan berbeda. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menjamin bahwa interpretasi hasil diagnosis didukung bukti yang memadai dan sesuai tujuan penggunaannya, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, validitas wawancara diagnostik harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.6.15 Validitas tugas representasional

Aspek validitas tugas representasional menegaskan bahwa diagnosis adalah kegiatan ilmiah berskala kelas: guru merumuskan pertanyaan, mengumpulkan bukti, mempertimbangkan penjelasan alternatif, lalu menentukan kesimpulan yang paling masuk akal. Inferensi harus proporsional terhadap kualitas data. Semakin spesifik label yang diberikan, semakin kuat pula bukti yang diperlukan (Downing, 2003; AERA et al., 2014).

Secara operasional, guru dapat menggunakan predict-observe-explain untuk menelusuri kemampuan transfer. Data tersebut perlu dibaca bersama kualitas model mental yang dieksternalisasi, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menyimpulkan miskonsepsi hanya dari satu jawaban. Untuk mengurangnya, guru dapat mengulang diagnosis setelah remediasi dan menyusun tindak lanjut berbeda untuk kategori respons yang berbeda. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai validitas tugas representasional, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan validitas tugas representasional melalui penggunaan konsep, bukan sekadar pengenalan istilah;

- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan validitas tugas representasional;
- Bandingkan bukti validitas tugas representasional pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek validitas tugas representasional dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian validitas tugas representasional, topik elektrokimia dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa elektron dianggap bergerak melalui jembatan garam. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta penelusuran jalur elektron dan ion pada sel galvanik. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa elektron mengalir pada rangkaian luar, sedangkan ion menjaga kenetralan muatan. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menjamin bahwa interpretasi hasil diagnosis didukung bukti yang memadai dan sesuai tujuan penggunaannya, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, validitas tugas representasional harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.6.16 Validitas lintas konteks dan transfer

Penerapan validitas lintas konteks dan transfer harus menghubungkan teori pengukuran dengan pengetahuan pedagogis kimia. Instrumen yang teknis tetapi tidak memahami struktur konsep akan menghasilkan kategori dangkal; sebaliknya, pengetahuan konten tanpa prosedur pengumpulan bukti dapat menghasilkan kesimpulan intuitif yang sulit diuji. Keduanya perlu dipadukan dalam satu alur yang transparan (AERA et al., 2014; Messick, 1995).

Secara operasional, guru dapat menggunakan tugas isomorfik untuk menelusuri stabilitas setelah jeda waktu. Data tersebut perlu dibaca bersama respons terhadap probing, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mencampur kesalahan matematika dengan kesalahan konsep. Untuk menguranginya, guru dapat mencatat derajat keyakinan dan membandingkan jawaban dengan alasan. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai validitas lintas konteks dan transfer, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;

- Pilih tugas yang merepresentasikan validitas lintas konteks dan transfer melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan validitas lintas konteks dan transfer;
- Bandingkan bukti validitas lintas konteks dan transfer pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek validitas lintas konteks dan transfer dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian validitas lintas konteks dan transfer, topik praktikum titrasi dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa titik akhir dianggap selalu identik dengan titik ekuivalen. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta analisis indikator, kurva titrasi, dan ketidakpastian. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa titik akhir adalah observasi operasional yang mendekati titik ekuivalen. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menjamin bahwa interpretasi hasil diagnosis didukung bukti yang memadai dan sesuai tujuan penggunaannya, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Kerangka operasional

Aspek	Bukti yang dicari	Risiko interpretasi	Tindak lanjut
Target konsep	Penjelasan tentang praktikum titrasi	Target terlalu luas	Batasi subkonsep dan prasyarat
Respons	Stabilitas setelah jeda waktu	Mencampur kesalahan matematika dengan kesalahan konsep	Mencatat derajat keyakinan
Penalaran	Respons terhadap probing	Alasan diperlakukan sekadar tambahan	Membandingkan jawaban dengan alasan
Konfirmasi	Tugas isomorfik atau tugas isomorfik	Kesimpulan dibuat dari satu konteks	Gunakan bukti kedua dan rediagnosis

Implikasinya, validitas lintas konteks dan transfer harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.6.17 Validitas lintas kelompok peserta didik

Validitas lintas kelompok peserta didik tidak berdiri sebagai prosedur administratif. Ia menjadi bagian dari keputusan tentang apa yang perlu dijelaskan ulang, representasi apa yang perlu dibandingkan, dan siapa yang memerlukan dukungan berbeda. Oleh sebab itu, bukti diagnosis harus langsung terhubung dengan tindakan pembelajaran, bukan berhenti pada daftar kesalahan (Messick, 1995; Kane, 2013).

Secara operasional, guru dapat menggunakan kuis formatif digital untuk menelusuri urutan langkah penyelesaian. Data tersebut perlu dibaca bersama ketepatan pilihan, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menggunakan ambang kategori tanpa bukti. Untuk mengurangnya, guru dapat menelusuri langkah penyelesaian dan menggunakan pertanyaan tindak lanjut. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai validitas lintas kelompok peserta didik, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan validitas lintas kelompok peserta didik melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan validitas lintas kelompok peserta didik;
- Bandingkan bukti validitas lintas kelompok peserta didik pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek validitas lintas kelompok peserta didik dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian validitas lintas kelompok peserta didik, topik sistem periodik dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa jari-jari atom dianggap selalu bertambah bersama nomor atom. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta prediksi tren disertai penjelasan muatan inti efektif dan kulit elektron. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa tren periodik perlu dijelaskan melalui struktur dan interaksi, bukan hafalan arah panah. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menjamin bahwa interpretasi hasil diagnosis didukung bukti yang memadai dan sesuai tujuan penggunaannya, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, validitas lintas kelompok peserta didik harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik

'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.6.18 Ancaman underrepresentation konstruk

Ancaman underrepresentation konstruk perlu ditempatkan dalam kerangka diagnosis yang menjamin bahwa interpretasi hasil diagnosis didukung bukti yang memadai dan sesuai tujuan penggunaannya. Fokusnya bukan sekadar menghasilkan skor, melainkan memperoleh bukti yang cukup untuk menjelaskan bagaimana peserta didik membangun, menghubungkan, dan menggunakan konsep. Karena konsepsi tidak dapat diamati secara langsung, guru harus menarik inferensi dari respons verbal, tulisan, gambar, tindakan, dan perubahan strategi yang tampak selama penyelesaian tugas (Kane, 2013; Downing, 2003).

Secara operasional, guru dapat menggunakan tes tiga tingkat untuk menelusuri kualitas model mental yang dieksternalisasi. Data tersebut perlu dibaca bersama koherensi alasan, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah memperlakukan konsepsi sebagai sifat permanen. Untuk mengurangnya, guru dapat menggunakan triangulasi bukti dan menyajikan tugas paralel. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai ancaman underrepresentation konstruk, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan ancaman underrepresentation konstruk melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan ancaman underrepresentation konstruk;
- Bandingkan bukti ancaman underrepresentation konstruk pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek ancaman underrepresentation konstruk dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian ancaman underrepresentation konstruk, topik stoikiometri dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa koefisien reaksi diperlakukan sebagai perbandingan massa langsung. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta translasi persamaan dari partikel ke mol dan massa. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa koefisien terutama menyatakan rasio jumlah entitas atau mol. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menjamin bahwa interpretasi hasil diagnosis didukung bukti yang memadai dan sesuai tujuan penggunaannya, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan

klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, ancaman underrepresentation konstruk harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.6.19 Ancaman construct-irrelevant variance

Pembahasan tentang ancaman construct-irrelevant variance berangkat dari prinsip bahwa data asesmen selalu merupakan jejak tidak langsung dari proses berpikir. Satu pilihan jawaban hanya menunjukkan produk akhir, sedangkan struktur pengetahuan yang melahirkannya dapat berbeda. Diagnosis yang kuat karena itu menggabungkan respons terhadap probing, tingkat keyakinan, dan bukti kontekstual sehingga keputusan pedagogis tidak bergantung pada satu indikator saja (Downing, 2003; AERA et al., 2014).

Secara operasional, guru dapat menggunakan wawancara semi-terstruktur untuk menelusuri respons terhadap probing. Data tersebut perlu dibaca bersama tingkat keyakinan, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mengabaikan ketidakpastian pengukuran. Untuk menguranginya, guru dapat menyusun tindak lanjut berbeda untuk kategori respons yang berbeda dan memeriksa konsep prasyarat. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai ancaman construct-irrelevant variance, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan ancaman construct-irrelevant variance melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan ancaman construct-irrelevant variance;
- Bandingkan bukti ancaman construct-irrelevant variance pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek ancaman construct-irrelevant variance dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian ancaman construct-irrelevant variance, topik asam-basa dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa asam kuat dianggap selalu berkonsentrasi tinggi. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta pemisahan konsep derajat ionisasi, K_a , konsentrasi, dan pH. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan

bahwa kekuatan asam dan konsentrasi adalah dimensi berbeda. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menjamin bahwa interpretasi hasil diagnosis didukung bukti yang memadai dan sesuai tujuan penggunaannya, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, ancaman construct-irrelevant variance harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.6.20 Bahasa, keterbacaan, dan validitas

Dalam pembelajaran kimia, bahasa, keterbacaan, dan validitas memiliki arti strategis karena banyak jawaban dapat diperoleh melalui hafalan prosedur atau pengenalan pola. Peserta didik mungkin tampak berhasil tetapi tidak mampu menerjemahkan makna simbol, model, atau besaran yang digunakan. Prinsip diagnosis menuntut bukti tentang cara konsep dipakai, bukan hanya apakah hasil akhirnya sesuai kunci (AERA et al., 2014; Messick, 1995).

Secara operasional, guru dapat menggunakan gambar partikel untuk menelusuri ketepatan pilihan. Data tersebut perlu dibaca bersama konsistensi lintas butir, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menyimpulkan miskonsepsi hanya dari satu jawaban. Untuk mengurangnya, guru dapat membandingkan jawaban dengan alasan dan meminta translasi antarrepresentasi. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai bahasa, keterbacaan, dan validitas, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan bahasa, keterbacaan, dan validitas melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan bahasa, keterbacaan, dan validitas;
- Bandingkan bukti bahasa, keterbacaan, dan validitas pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek bahasa, keterbacaan, dan validitas dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian bahasa, keterbacaan, dan validitas, topik kinetika dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa reaksi eksoterm dianggap selalu cepat. Diagnosis tidak cukup dilakukan dengan menanyakan

definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta perbandingan profil energi dan faktor kinetik. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa termodinamika dan kinetika menjawab pertanyaan berbeda. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menjamin bahwa interpretasi hasil diagnosis didukung bukti yang memadai dan sesuai tujuan penggunaannya, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Kerangka operasional

Aspek	Bukti yang dicari	Risiko interpretasi	Tindak lanjut
Target konsep	Penjelasan tentang kinetika	Target terlalu luas	Batasi subkonsep dan prasyarat
Respons	Ketepatan pilihan	Menyimpulkan miskonsepsi hanya dari satu jawaban	Membandingkan jawaban dengan alasan
Penalaran	Konsistensi lintas butir	Alasan diperlakukan sekadar tambahan	Meminta translasi antarrepresentasi
Konfirmasi	Gambar partikel atau tugas isomorfik	Kesimpulan dibuat dari satu konteks	Gunakan bukti kedua dan rediagnosis

Implikasinya, bahasa, keterbacaan, dan validitas harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.6.21 Bias budaya dan pengalaman kontekstual

Aspek bias budaya dan pengalaman kontekstual menegaskan bahwa diagnosis adalah kegiatan ilmiah berskala kelas: guru merumuskan pertanyaan, mengumpulkan bukti, mempertimbangkan penjelasan alternatif, lalu menentukan kesimpulan yang paling masuk akal. Inferensi harus proporsional terhadap kualitas data. Semakin spesifik label yang diberikan, semakin kuat pula bukti yang diperlukan (Messick, 1995; Kane, 2013).

Secara operasional, guru dapat menggunakan observasi praktikum untuk menelusuri koherensi alasan. Data tersebut perlu dibaca bersama konsistensi lintas representasi, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mencampur kesalahan matematika dengan kesalahan konsep. Untuk mengurangnya, guru dapat menggunakan pertanyaan tindak lanjut dan mengulang diagnosis

setelah remediasi. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai bias budaya dan pengalaman kontekstual, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan bias budaya dan pengalaman kontekstual melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan bias budaya dan pengalaman kontekstual;
- Bandingkan bukti bias budaya dan pengalaman kontekstual pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek bias budaya dan pengalaman kontekstual dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian bias budaya dan pengalaman kontekstual, topik elektrokimia dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa elektron dianggap bergerak melalui jembatan garam. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta penelusuran jalur elektron dan ion pada sel galvanik. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa elektron mengalir pada rangkaian luar, sedangkan ion menjaga kenetralan muatan. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menjamin bahwa interpretasi hasil diagnosis didukung bukti yang memadai dan sesuai tujuan penggunaannya, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, bias budaya dan pengalaman kontekstual harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.6.22 Expert judgment dan indeks validitas isi

Penerapan expert judgment dan indeks validitas isi harus menghubungkan teori pengukuran dengan pengetahuan pedagogis kimia. Instrumen yang teknis tetapi tidak memahami struktur konsep akan menghasilkan kategori dangkal; sebaliknya, pengetahuan konten tanpa prosedur pengumpulan bukti dapat menghasilkan kesimpulan intuitif yang sulit diuji. Keduanya perlu dipadukan dalam satu alur yang transparan (Kane, 2013; Downing, 2003).

Secara operasional, guru dapat menggunakan exit ticket untuk menelusuri tingkat keyakinan. Data tersebut perlu dibaca bersama kemampuan transfer, karena pola yang sama pada

permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menggunakan ambang kategori tanpa bukti. Untuk menguranginya, guru dapat menyajikan tugas paralel dan mencatat derajat keyakinan. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai expert judgment dan indeks validitas isi, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan expert judgment dan indeks validitas isi melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan expert judgment dan indeks validitas isi;
- Bandingkan bukti expert judgment dan indeks validitas isi pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek expert judgment dan indeks validitas isi dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian expert judgment dan indeks validitas isi, topik praktikum titrasi dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa titik akhir dianggap selalu identik dengan titik ekuivalen. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta analisis indikator, kurva titrasi, dan ketidakpastian. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa titik akhir adalah observasi operasional yang mendekati titik ekuivalen. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menjamin bahwa interpretasi hasil diagnosis didukung bukti yang memadai dan sesuai tujuan penggunaannya, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, expert judgment dan indeks validitas isi harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.6.23 Uji coba kognitif dan revisi butir

Uji coba kognitif dan revisi butir tidak berdiri sebagai prosedur administratif. Ia menjadi bagian dari keputusan tentang apa yang perlu dijelaskan ulang, representasi apa yang perlu dibandingkan, dan siapa yang memerlukan dukungan berbeda. Oleh sebab itu, bukti diagnosis

harus langsung terhubung dengan tindakan pembelajaran, bukan berhenti pada daftar kesalahan (Downing, 2003; AERA et al., 2014).

Secara operasional, guru dapat menggunakan tes pilihan ganda dua tingkat untuk menelusuri konsistensi lintas butir. Data tersebut perlu dibaca bersama stabilitas setelah jeda waktu, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah memperlakukan konsepsi sebagai sifat permanen. Untuk menguranginya, guru dapat memeriksa konsep prasyarat dan menelusuri langkah penyelesaian. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai uji coba kognitif dan revisi butir, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan uji coba kognitif dan revisi butir melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan uji coba kognitif dan revisi butir;
- Bandingkan bukti uji coba kognitif dan revisi butir pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek uji coba kognitif dan revisi butir dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian uji coba kognitif dan revisi butir, topik sistem periodik dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa jari-jari atom dianggap selalu bertambah bersama nomor atom. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta prediksi tren disertai penjelasan muatan inti efektif dan kulit elektron. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa tren periodik perlu dijelaskan melalui struktur dan interaksi, bukan hafalan arah panah. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menjamin bahwa interpretasi hasil diagnosis didukung bukti yang memadai dan sesuai tujuan penggunaannya, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, uji coba kognitif dan revisi butir harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.6.24 Triangulasi sebagai penguat argumentasi validitas

Triangulasi sebagai penguat argumentasi validitas perlu ditempatkan dalam kerangka diagnosis yang menjamin bahwa interpretasi hasil diagnosis didukung bukti yang memadai dan sesuai tujuan penggunaannya. Fokusnya bukan sekadar menghasilkan skor, melainkan memperoleh bukti yang cukup untuk menjelaskan bagaimana peserta didik membangun, menghubungkan, dan menggunakan konsep. Karena konsepsi tidak dapat diamati secara langsung, guru harus menarik inferensi dari respons verbal, tulisan, gambar, tindakan, dan perubahan strategi yang tampak selama penyelesaian tugas (AERA et al., 2014; Messick, 1995).

Secara operasional, guru dapat menggunakan alasan terbuka untuk menelusuri konsistensi lintas representasi. Data tersebut perlu dibaca bersama urutan langkah penyelesaian, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mengabaikan ketidakpastian pengukuran. Untuk menguranginya, guru dapat meminta translasi antarrepresentasi dan menggunakan triangulasi bukti. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai triangulasi sebagai penguat argumentasi validitas, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan triangulasi sebagai penguat argumentasi validitas melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan triangulasi sebagai penguat argumentasi validitas;
- Bandingkan bukti triangulasi sebagai penguat argumentasi validitas pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek triangulasi sebagai penguat argumentasi validitas dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian triangulasi sebagai penguat argumentasi validitas, topik stoikiometri dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa koefisien reaksi diperlakukan sebagai perbandingan massa langsung. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta translasi persamaan dari partikel ke mol dan massa. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa koefisien terutama menyatakan rasio jumlah entitas atau mol. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menjamin bahwa interpretasi hasil diagnosis didukung bukti yang memadai dan sesuai tujuan penggunaannya, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Kerangka operasional

Aspek	Bukti yang dicari	Risiko interpretasi	Tindak lanjut
Target konsep	Penjelasan tentang stoikiometri	Target terlalu luas	Batasi subkonsep dan prasyarat
Respons	Konsistensi lintas representasi	Mengabaikan ketidakpastian pengukuran	Meminta translasi antarrepresentasi
Penalaran	Urutan langkah penyelesaian	Alasan diperlakukan sekadar tambahan	Menggunakan triangulasi bukti
Konfirmasi	Alasan terbuka atau tugas isomorfik	Kesimpulan dibuat dari satu konteks	Gunakan bukti kedua dan rediagnosis

Implikasinya, triangulasi sebagai penguat argumentasi validitas harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.6.25 Dokumentasi bukti validitas

Pembahasan tentang dokumentasi bukti validitas berangkat dari prinsip bahwa data asesmen selalu merupakan jejak tidak langsung dari proses berpikir. Satu pilihan jawaban hanya menunjukkan produk akhir, sedangkan struktur pengetahuan yang melahirkannya dapat berbeda. Diagnosis yang kuat karena itu menggabungkan kemampuan transfer, kualitas model mental yang dieksternalisasi, dan bukti kontekstual sehingga keputusan pedagogis tidak bergantung pada satu indikator saja (Messick, 1995; Kane, 2013).

Secara operasional, guru dapat menggunakan peta konsep untuk menelusuri kemampuan transfer. Data tersebut perlu dibaca bersama kualitas model mental yang dieksternalisasi, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menyimpulkan miskonsepsi hanya dari satu jawaban. Untuk mengurangnya, guru dapat mengulang diagnosis setelah remediasi dan menyusun tindak lanjut berbeda untuk kategori respons yang berbeda. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai dokumentasi bukti validitas, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan dokumentasi bukti validitas melalui penggunaan konsep, bukan sekadar pengenalan istilah;

- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan dokumentasi bukti validitas;
- Bandingkan bukti dokumentasi bukti validitas pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek dokumentasi bukti validitas dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian dokumentasi bukti validitas, topik asam-basa dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa asam kuat dianggap selalu berkonsentrasi tinggi. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta pemisahan konsep derajat ionisasi, K_a , konsentrasi, dan pH. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa kekuatan asam dan konsentrasi adalah dimensi berbeda. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menjamin bahwa interpretasi hasil diagnosis didukung bukti yang memadai dan sesuai tujuan penggunaannya, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, dokumentasi bukti validitas harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.6.26 Sintesis prinsip validitas diagnosis

Dalam pembelajaran kimia, sintesis prinsip validitas diagnosis memiliki arti strategis karena banyak jawaban dapat diperoleh melalui hafalan prosedur atau pengenalan pola. Peserta didik mungkin tampak berhasil tetapi tidak mampu menerjemahkan makna simbol, model, atau besaran yang digunakan. Prinsip diagnosis menuntut bukti tentang cara konsep dipakai, bukan hanya apakah hasil akhirnya sesuai kunci (Kane, 2013; Downing, 2003).

Secara operasional, guru dapat menggunakan analisis respons tertulis untuk menelusuri stabilitas setelah jeda waktu. Data tersebut perlu dibaca bersama respons terhadap probing, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mencampur kesalahan matematika dengan kesalahan konsep. Untuk mengurangnya, guru dapat mencatat derajat keyakinan dan membandingkan jawaban dengan alasan. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai sintesis prinsip validitas diagnosis, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;

- Pilih tugas yang merepresentasikan sintesis prinsip validitas diagnosis melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan sintesis prinsip validitas diagnosis;
- Bandingkan bukti sintesis prinsip validitas diagnosis pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek sintesis prinsip validitas diagnosis dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian sintesis prinsip validitas diagnosis, topik kinetika dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa reaksi eksoterm dianggap selalu cepat. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta perbandingan profil energi dan faktor kinetik. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa termodinamika dan kinetika menjawab pertanyaan berbeda. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menjamin bahwa interpretasi hasil diagnosis didukung bukti yang memadai dan sesuai tujuan penggunaannya, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, sintesis prinsip validitas diagnosis harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.6.27 Implikasi praktis bagi guru kimia

Implikasi praktis subbab 5.6 dapat dirangkum sebagai alur: tentukan konsep target, pilih bukti yang relevan, kumpulkan respons dalam lebih dari satu bentuk, analisis pola, nyatakan derajat kepastian, susun tindak lanjut, dan lakukan rediagnosis. Alur ini menjaga agar prinsip validitas diagnosis tidak berubah menjadi administrasi skor, tetapi berfungsi sebagai mekanisme perbaikan pembelajaran yang berulang.

36. Identifikasi konsep inti dan konsep prasyarat sebelum menyusun instrumen.
37. Gunakan pertanyaan yang meminta alasan, representasi, atau langkah kerja.
38. Bandingkan respons pada beberapa konteks dan waktu.
39. Pisahkan bukti faktual dari interpretasi guru.
40. Tentukan tindak lanjut yang berbeda untuk kategori kebutuhan yang berbeda.
41. Simpan data secukupnya dan lindungi kerahasiaan peserta didik.
42. Periksa kembali perubahan melalui tugas transfer dan retensi.

5.7 Prinsip Reliabilitas Instrumen

Subbab ini mengendalikan ketidakpastian pengukuran dan konsistensi klasifikasi tanpa mereduksi kompleksitas konsepsi. Uraian disusun dari prinsip teoretis menuju prosedur kelas, contoh kimia, risiko interpretasi, dan tindak lanjut. Setiap bagian menekankan bahwa kesimpulan diagnostik harus didasarkan pada pola bukti, bukan pada satu respons yang terpisah.

5.7.1 Makna reliabilitas dalam diagnosis konseptual

Pembahasan tentang makna reliabilitas dalam diagnosis konseptual berangkat dari prinsip bahwa data asesmen selalu merupakan jejak tidak langsung dari proses berpikir. Satu pilihan jawaban hanya menunjukkan produk akhir, sedangkan struktur pengetahuan yang melahirkannya dapat berbeda. Diagnosis yang kuat karena itu menggabungkan koherensi alasan, konsistensi lintas representasi, dan bukti kontekstual sehingga keputusan pedagogis tidak bergantung pada satu indikator saja (AERA et al., 2014; DeVellis & Thorpe, 2021).

Secara operasional, guru dapat menggunakan tes tiga tingkat untuk menelusuri koherensi alasan. Data tersebut perlu dibaca bersama konsistensi lintas representasi, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mencampur kesalahan matematika dengan kesalahan konsep. Untuk mengurangnya, guru dapat menggunakan pertanyaan tindak lanjut dan mengulang diagnosis setelah remediasi. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai makna reliabilitas dalam diagnosis konseptual, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan makna reliabilitas dalam diagnosis konseptual melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan makna reliabilitas dalam diagnosis konseptual;
- Bandingkan bukti makna reliabilitas dalam diagnosis konseptual pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek makna reliabilitas dalam diagnosis konseptual dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian makna reliabilitas dalam diagnosis konseptual, topik larutan penyangga dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa penyangga dianggap mempertahankan pH secara mutlak. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta analisis perubahan komposisi setelah penambahan asam atau basa. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa penyangga membatasi perubahan dalam kapasitas tertentu, bukan meniadakannya. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus

mengendalikan ketidakpastian pengukuran dan konsistensi klasifikasi tanpa mereduksi kompleksitas konsepsi, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, makna reliabilitas dalam diagnosis konseptual harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.7.2 Reliabilitas sebagai konsistensi bukti

Dalam pembelajaran kimia, reliabilitas sebagai konsistensi bukti memiliki arti strategis karena banyak jawaban dapat diperoleh melalui hafalan prosedur atau pengenalan pola. Peserta didik mungkin tampak berhasil tetapi tidak mampu menerjemahkan makna simbol, model, atau besaran yang digunakan. Prinsip diagnosis menuntut bukti tentang cara konsep dipakai, bukan hanya apakah hasil akhirnya sesuai kunci (DeVellis & Thorpe, 2021; Wren & Barbera, 2014).

Secara operasional, guru dapat menggunakan wawancara semi-terstruktur untuk menelusuri tingkat keyakinan. Data tersebut perlu dibaca bersama kemampuan transfer, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menggunakan ambang kategori tanpa bukti. Untuk menguranginya, guru dapat menyajikan tugas paralel dan mencatat derajat keyakinan. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai reliabilitas sebagai konsistensi bukti, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan reliabilitas sebagai konsistensi bukti melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan reliabilitas sebagai konsistensi bukti;
- Bandingkan bukti reliabilitas sebagai konsistensi bukti pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek reliabilitas sebagai konsistensi bukti dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian reliabilitas sebagai konsistensi bukti, topik katalis dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa katalis dianggap menambah jumlah produk kesetimbangan. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan

relasi konseptual yang tepat. Guru dapat meminta analisis jalur energi dan laju maju-balik. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa katalis mempercepat tercapainya kesetimbangan tanpa mengubah K . Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengendalikan ketidakpastian pengukuran dan konsistensi klasifikasi tanpa mereduksi kompleksitas konsepsi, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, reliabilitas sebagai konsistensi bukti harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.7.3 Sumber kesalahan pengukuran

Aspek sumber kesalahan pengukuran menegaskan bahwa diagnosis adalah kegiatan ilmiah berskala kelas: guru merumuskan pertanyaan, mengumpulkan bukti, mempertimbangkan penjelasan alternatif, lalu menentukan kesimpulan yang paling masuk akal. Inferensi harus proporsional terhadap kualitas data. Semakin spesifik label yang diberikan, semakin kuat pula bukti yang diperlukan (Wren & Barbera, 2014; Cronbach, 1951).

Secara operasional, guru dapat menggunakan gambar partikel untuk menelusuri konsistensi lintas butir. Data tersebut perlu dibaca bersama stabilitas setelah jeda waktu, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah memperlakukan konsepsi sebagai sifat permanen. Untuk mengurangnya, guru dapat memeriksa konsep prasyarat dan menelusuri langkah penyelesaian. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai sumber kesalahan pengukuran, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan sumber kesalahan pengukuran melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan sumber kesalahan pengukuran;
- Bandingkan bukti sumber kesalahan pengukuran pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek sumber kesalahan pengukuran dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian sumber kesalahan pengukuran, topik elektrolisis dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa anoda selalu dianggap bermuatan negatif. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta pembedaan definisi anoda-katoda berdasarkan proses oksidasi-reduksi. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa tanda elektroda bergantung jenis sel, tetapi anoda tetap lokasi oksidasi. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengendalikan ketidakpastian pengukuran dan konsistensi klasifikasi tanpa mereduksi kompleksitas konsepsi, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, sumber kesalahan pengukuran harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.7.4 Konsistensi internal dan keterbatasannya

Penerapan konsistensi internal dan keterbatasannya harus menghubungkan teori pengukuran dengan pengetahuan pedagogis kimia. Instrumen yang teknis tetapi tidak memahami struktur konsep akan menghasilkan kategori dangkal; sebaliknya, pengetahuan konten tanpa prosedur pengumpulan bukti dapat menghasilkan kesimpulan intuitif yang sulit diuji. Keduanya perlu dipadukan dalam satu alur yang transparan (Cronbach, 1951; AERA et al., 2014).

Secara operasional, guru dapat menggunakan observasi praktikum untuk menelusuri konsistensi lintas representasi. Data tersebut perlu dibaca bersama urutan langkah penyelesaian, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mengabaikan ketidakpastian pengukuran. Untuk menguranginya, guru dapat meminta translasi antarrepresentasi dan menggunakan triangulasi bukti. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai konsistensi internal dan keterbatasannya, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan konsistensi internal dan keterbatasannya melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan konsistensi internal dan keterbatasannya;
- Bandingkan bukti konsistensi internal dan keterbatasannya pada lebih dari satu konteks, format tugas, atau bentuk representasi;

- Hubungkan kategori hasil pada aspek konsistensi internal dan keterbatasannya dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian konsistensi internal dan keterbatasannya, topik pengukuran laboratorium dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa angka lebih banyak selalu dianggap lebih akurat. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta pembedaan akurasi, presisi, resolusi, dan angka penting. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa kualitas data ditentukan oleh desain, kalibrasi, dan ketidakpastian. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengendalikan ketidakpastian pengukuran dan konsistensi klasifikasi tanpa mereduksi kompleksitas konsepsi, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Kerangka operasional

Aspek	Bukti yang dicari	Risiko interpretasi	Tindak lanjut
Target konsep	Penjelasan tentang pengukuran laboratorium	Target terlalu luas	Batasi subkonsep dan prasyarat
Respons	Konsistensi lintas representasi	Mengabaikan ketidakpastian pengukuran	Meminta translasi antarrepresentasi
Penalaran	Urutan langkah penyelesaian	Alasan diperlakukan sekadar tambahan	Menggunakan triangulasi bukti
Konfirmasi	Observasi praktikum atau tugas isomorfik	Kesimpulan dibuat dari satu konteks	Gunakan bukti kedua dan rediagnosis

Implikasinya, konsistensi internal dan keterbatasannya harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.7.5 Koefisien alpha dan asumsi dasarnya

Koefisien alpha dan asumsi dasarnya tidak berdiri sebagai prosedur administratif. Ia menjadi bagian dari keputusan tentang apa yang perlu dijelaskan ulang, representasi apa yang perlu dibandingkan, dan siapa yang memerlukan dukungan berbeda. Oleh sebab itu, bukti diagnosis

harus langsung terhubung dengan tindakan pembelajaran, bukan berhenti pada daftar kesalahan (AERA et al., 2014; DeVellis & Thorpe, 2021).

Secara operasional, guru dapat menggunakan exit ticket untuk menelusuri kemampuan transfer. Data tersebut perlu dibaca bersama kualitas model mental yang dieksternalisasi, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menyimpulkan miskonsepsi hanya dari satu jawaban. Untuk mengurangnya, guru dapat mengulang diagnosis setelah remediasi dan menyusun tindak lanjut berbeda untuk kategori respons yang berbeda. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai koefisien alpha dan asumsi dasarnya, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan koefisien alpha dan asumsi dasarnya melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan koefisien alpha dan asumsi dasarnya;
- Bandingkan bukti koefisien alpha dan asumsi dasarnya pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek koefisien alpha dan asumsi dasarnya dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian koefisien alpha dan asumsi dasarnya, topik ikatan ion dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa NaCl dipandang sebagai molekul diskrit. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta gambar partikel padatan NaCl dan penjelasan hubungan struktur-kisi dengan titik leleh. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa senyawa ionik membentuk kisi dan rumusnya menyatakan rasio ion. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengendalikan ketidakpastian pengukuran dan konsistensi klasifikasi tanpa mereduksi kompleksitas konsepsi, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, koefisien alpha dan asumsi dasarnya harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.7.6 Alternatif koefisien omega dan reliabilitas komposit

Alternatif koefisien omega dan reliabilitas komposit perlu ditempatkan dalam kerangka diagnosis yang mengendalikan ketidakpastian pengukuran dan konsistensi klasifikasi tanpa mereduksi kompleksitas konsepsi. Fokusnya bukan sekadar menghasilkan skor, melainkan memperoleh bukti yang cukup untuk menjelaskan bagaimana peserta didik membangun, menghubungkan, dan menggunakan konsep. Karena konsepsi tidak dapat diamati secara langsung, guru harus menarik inferensi dari respons verbal, tulisan, gambar, tindakan, dan perubahan strategi yang tampak selama penyelesaian tugas (DeVellis & Thorpe, 2021; Wren & Barbera, 2014).

Secara operasional, guru dapat menggunakan tes pilihan ganda dua tingkat untuk menelusuri stabilitas setelah jeda waktu. Data tersebut perlu dibaca bersama respons terhadap probing, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mencampur kesalahan matematika dengan kesalahan konsep. Untuk mengurangnya, guru dapat mencatat derajat keyakinan dan membandingkan jawaban dengan alasan. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai alternatif koefisien omega dan reliabilitas komposit, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan alternatif koefisien omega dan reliabilitas komposit melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan alternatif koefisien omega dan reliabilitas komposit;
- Bandingkan bukti alternatif koefisien omega dan reliabilitas komposit pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek alternatif koefisien omega dan reliabilitas komposit dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian alternatif koefisien omega dan reliabilitas komposit, topik pereaksi pembatas dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa pereaksi dengan massa terkecil selalu dianggap pembatas. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta analisis rasio mol terhadap koefisien. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa status pembatas ditentukan oleh hubungan jumlah zat dan stoikiometri. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengendalikan ketidakpastian pengukuran dan konsistensi klasifikasi tanpa mereduksi kompleksitas konsepsi, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, alternatif koefisien omega dan reliabilitas komposit harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.7.7 Reliabilitas skor dikotomis

Pembahasan tentang reliabilitas skor dikotomis berangkat dari prinsip bahwa data asesmen selalu merupakan jejak tidak langsung dari proses berpikir. Satu pilihan jawaban hanya menunjukkan produk akhir, sedangkan struktur pengetahuan yang melahirkannya dapat berbeda. Diagnosis yang kuat karena itu menggabungkan urutan langkah penyelesaian, ketepatan pilihan, dan bukti kontekstual sehingga keputusan pedagogis tidak bergantung pada satu indikator saja (Wren & Barbera, 2014; Cronbach, 1951).

Secara operasional, guru dapat menggunakan alasan terbuka untuk menelusuri urutan langkah penyelesaian. Data tersebut perlu dibaca bersama ketepatan pilihan, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menggunakan ambang kategori tanpa bukti. Untuk mengurangnya, guru dapat menelusuri langkah penyelesaian dan menggunakan pertanyaan tindak lanjut. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai reliabilitas skor dikotomis, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan reliabilitas skor dikotomis melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan reliabilitas skor dikotomis;
- Bandingkan bukti reliabilitas skor dikotomis pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek reliabilitas skor dikotomis dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian reliabilitas skor dikotomis, topik larutan penyangga dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa penyangga dianggap mempertahankan pH secara mutlak. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta analisis perubahan komposisi setelah penambahan asam atau basa. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa penyangga membatasi perubahan dalam kapasitas tertentu, bukan meniadakannya. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengendalikan

ketidakpastian pengukuran dan konsistensi klasifikasi tanpa mereduksi kompleksitas konsepsi, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, reliabilitas skor dikotomis harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.7.8 Reliabilitas skor politomus dan tier

Dalam pembelajaran kimia, reliabilitas skor politomus dan tier memiliki arti strategis karena banyak jawaban dapat diperoleh melalui hafalan prosedur atau pengenalan pola. Peserta didik mungkin tampak berhasil tetapi tidak mampu menerjemahkan makna simbol, model, atau besaran yang digunakan. Prinsip diagnosis menuntut bukti tentang cara konsep dipakai, bukan hanya apakah hasil akhirnya sesuai kunci (Cronbach, 1951; AERA et al., 2014).

Secara operasional, guru dapat menggunakan peta konsep untuk menelusuri kualitas model mental yang dieksternalisasi. Data tersebut perlu dibaca bersama koherensi alasan, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah memperlakukan konsepsi sebagai sifat permanen. Untuk mengurangnya, guru dapat menggunakan triangulasi bukti dan menyajikan tugas paralel. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai reliabilitas skor politomus dan tier, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan reliabilitas skor politomus dan tier melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan reliabilitas skor politomus dan tier;
- Bandingkan bukti reliabilitas skor politomus dan tier pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek reliabilitas skor politomus dan tier dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian reliabilitas skor politomus dan tier, topik katalis dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa katalis dianggap menambah jumlah produk kesetimbangan. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta analisis jalur energi dan laju maju-balik. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa katalis

mempercepat tercapainya kesetimbangan tanpa mengubah K. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengendalikan ketidakpastian pengukuran dan konsistensi klasifikasi tanpa mereduksi kompleksitas konsepsi, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Kerangka operasional

Aspek	Bukti yang dicari	Risiko interpretasi	Tindak lanjut
Target konsep	Penjelasan tentang katalis	Target terlalu luas	Batasi subkonsep dan prasyarat
Respons	Kualitas model mental yang dieksternalisasi	Memperlakukan konsepsi sebagai sifat permanen	Menggunakan triangulasi bukti
Penalaran	Koherensi alasan	Alasan diperlakukan sekadar tambahan	Menyajikan tugas paralel
Konfirmasi	Peta konsep atau tugas isomorfik	Kesimpulan dibuat dari satu konteks	Gunakan bukti kedua dan rediagnosis

Implikasinya, reliabilitas skor politomus dan tier harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.7.9 Stabilitas waktu melalui test-retest

Aspek stabilitas waktu melalui test-retest menegaskan bahwa diagnosis adalah kegiatan ilmiah berskala kelas: guru merumuskan pertanyaan, mengumpulkan bukti, mempertimbangkan penjelasan alternatif, lalu menentukan kesimpulan yang paling masuk akal. Inferensi harus proporsional terhadap kualitas data. Semakin spesifik label yang diberikan, semakin kuat pula bukti yang diperlukan (AERA et al., 2014; DeVellis & Thorpe, 2021).

Secara operasional, guru dapat menggunakan analisis respons tertulis untuk menelusuri respons terhadap probing. Data tersebut perlu dibaca bersama tingkat keyakinan, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mengabaikan ketidakpastian pengukuran. Untuk menguranginya, guru dapat menyusun tindak lanjut berbeda untuk kategori respons yang berbeda dan memeriksa konsep prasyarat. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai stabilitas waktu melalui test-retest, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan stabilitas waktu melalui test-retest melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan stabilitas waktu melalui test-retest;
- Bandingkan bukti stabilitas waktu melalui test-retest pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek stabilitas waktu melalui test-retest dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian stabilitas waktu melalui test-retest, topik elektrolisis dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa anoda selalu dianggap bermuatan negatif. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta pembedaan definisi anoda-katoda berdasarkan proses oksidasi-reduksi. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa tanda elektroda bergantung jenis sel, tetapi anoda tetap lokasi oksidasi. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengendalikan ketidakpastian pengukuran dan konsistensi klasifikasi tanpa mereduksi kompleksitas konsepsi, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, stabilitas waktu melalui test-retest harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.7.10 Kesetaraan bentuk paralel

Penerapan kesetaraan bentuk paralel harus menghubungkan teori pengukuran dengan pengetahuan pedagogis kimia. Instrumen yang teknis tetapi tidak memahami struktur konsep akan menghasilkan kategori dangkal; sebaliknya, pengetahuan konten tanpa prosedur pengumpulan bukti dapat menghasilkan kesimpulan intuitif yang sulit diuji. Keduanya perlu dipadukan dalam satu alur yang transparan (DeVellis & Thorpe, 2021; Wren & Barbera, 2014).

Secara operasional, guru dapat menggunakan diskusi argumentatif untuk menelusuri ketepatan pilihan. Data tersebut perlu dibaca bersama konsistensi lintas butir, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menyimpulkan miskonsepsi hanya dari satu jawaban. Untuk menguranginya, guru

dapat membandingkan jawaban dengan alasan dan meminta translasi antarrepresentasi. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai kesetaraan bentuk paralel, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan kesetaraan bentuk paralel melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan kesetaraan bentuk paralel;
- Bandingkan bukti kesetaraan bentuk paralel pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek kesetaraan bentuk paralel dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian kesetaraan bentuk paralel, topik pengukuran laboratorium dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa angka lebih banyak selalu dianggap lebih akurat. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta pembedaan akurasi, presisi, resolusi, dan angka penting. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa kualitas data ditentukan oleh desain, kalibrasi, dan ketidakpastian. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengendalikan ketidakpastian pengukuran dan konsistensi klasifikasi tanpa mereduksi kompleksitas konsepsi, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, kesetaraan bentuk paralel harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.7.11 Reliabilitas antarpemilai

Reliabilitas antarpemilai tidak berdiri sebagai prosedur administratif. Ia menjadi bagian dari keputusan tentang apa yang perlu dijelaskan ulang, representasi apa yang perlu dibandingkan, dan siapa yang memerlukan dukungan berbeda. Oleh sebab itu, bukti diagnosis harus langsung terhubung dengan tindakan pembelajaran, bukan berhenti pada daftar kesalahan (Wren & Barbera, 2014; Cronbach, 1951).

Secara operasional, guru dapat menggunakan analisis grafik untuk menelusuri koherensi alasan. Data tersebut perlu dibaca bersama konsistensi lintas representasi, karena pola yang

sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mencampur kesalahan matematika dengan kesalahan konsep. Untuk menguranginya, guru dapat menggunakan pertanyaan tindak lanjut dan mengulang diagnosis setelah remediasi. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai reliabilitas antarpemilai, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan reliabilitas antarpemilai melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan reliabilitas antarpemilai;
- Bandingkan bukti reliabilitas antarpemilai pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek reliabilitas antarpemilai dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian reliabilitas antarpemilai, topik ikatan ion dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa NaCl dipandang sebagai molekul diskrit. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta gambar partikel padatan NaCl dan penjelasan hubungan struktur-kisi dengan titik leleh. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa senyawa ionik membentuk kisi dan rumusnya menyatakan rasio ion. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengendalikan ketidakpastian pengukuran dan konsistensi klasifikasi tanpa mereduksi kompleksitas konsepsi, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, reliabilitas antarpemilai harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.7.12 Reliabilitas pengodean alasan terbuka

Reliabilitas pengodean alasan terbuka perlu ditempatkan dalam kerangka diagnosis yang mengendalikan ketidakpastian pengukuran dan konsistensi klasifikasi tanpa mereduksi kompleksitas konsepsi. Fokusnya bukan sekadar menghasilkan skor, melainkan memperoleh bukti yang cukup untuk menjelaskan bagaimana peserta didik membangun, menghubungkan, dan menggunakan konsep. Karena konsepsi tidak dapat diamati secara langsung, guru harus

menarik inferensi dari respons verbal, tulisan, gambar, tindakan, dan perubahan strategi yang tampak selama penyelesaian tugas (Cronbach, 1951; AERA et al., 2014).

Secara operasional, guru dapat menggunakan tes empat tingkat untuk menelusuri tingkat keyakinan. Data tersebut perlu dibaca bersama kemampuan transfer, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menggunakan ambang kategori tanpa bukti. Untuk mengurangnya, guru dapat menyajikan tugas paralel dan mencatat derajat keyakinan. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai reliabilitas pengodean alasan terbuka, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan reliabilitas pengodean alasan terbuka melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan reliabilitas pengodean alasan terbuka;
- Bandingkan bukti reliabilitas pengodean alasan terbuka pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek reliabilitas pengodean alasan terbuka dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian reliabilitas pengodean alasan terbuka, topik pereaksi pembatas dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa pereaksi dengan massa terkecil selalu dianggap pembatas. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta analisis rasio mol terhadap koefisien. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa status pembatas ditentukan oleh hubungan jumlah zat dan stoikiometri. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengendalikan ketidakpastian pengukuran dan konsistensi klasifikasi tanpa mereduksi kompleksitas konsepsi, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Kerangka operasional

Aspek	Bukti yang dicari	Risiko interpretasi	Tindak lanjut
Target konsep	Penjelasan tentang pereaksi pembatas	Target terlalu luas	Batasi subkonsep dan prasyarat
Respons	Tingkat keyakinan	Menggunakan ambang kategori tanpa bukti	Menyajikan tugas paralel

Aspek	Bukti yang dicari	Risiko interpretasi	Tindak lanjut
Penalaran	Kemampuan transfer	Alasan diperlakukan sekadar tambahan	Mencatat derajat keyakinan
Konfirmasi	Tes empat tingkat atau tugas isomorfik	Kesimpulan dibuat dari satu konteks	Gunakan bukti kedua dan rediagnosis

Implikasinya, reliabilitas pengodean alasan terbuka harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.7.13 Reliabilitas wawancara diagnostik

Pembahasan tentang reliabilitas wawancara diagnostik berangkat dari prinsip bahwa data asesmen selalu merupakan jejak tidak langsung dari proses berpikir. Satu pilihan jawaban hanya menunjukkan produk akhir, sedangkan struktur pengetahuan yang melahirkannya dapat berbeda. Diagnosis yang kuat karena itu menggabungkan konsistensi lintas butir, stabilitas setelah jeda waktu, dan bukti kontekstual sehingga keputusan pedagogis tidak bergantung pada satu indikator saja (AERA et al., 2014; DeVellis & Thorpe, 2021).

Secara operasional, guru dapat menggunakan think-aloud untuk menelusuri konsistensi lintas butir. Data tersebut perlu dibaca bersama stabilitas setelah jeda waktu, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah memperlakukan konsepsi sebagai sifat permanen. Untuk mengurangnya, guru dapat memeriksa konsep prasyarat dan menelusuri langkah penyelesaian. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai reliabilitas wawancara diagnostik, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan reliabilitas wawancara diagnostik melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan reliabilitas wawancara diagnostik;
- Bandingkan bukti reliabilitas wawancara diagnostik pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek reliabilitas wawancara diagnostik dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian reliabilitas wawancara diagnostik, topik larutan penyangga dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa penyangga dianggap mempertahankan pH secara mutlak. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta analisis perubahan komposisi setelah penambahan asam atau basa. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa penyangga membatasi perubahan dalam kapasitas tertentu, bukan meniadakannya. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengendalikan ketidakpastian pengukuran dan konsistensi klasifikasi tanpa mereduksi kompleksitas konsepsi, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, reliabilitas wawancara diagnostik harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.7.14 Reliabilitas observasi dan rubrik

Dalam pembelajaran kimia, reliabilitas observasi dan rubrik memiliki arti strategis karena banyak jawaban dapat diperoleh melalui hafalan prosedur atau pengenalan pola. Peserta didik mungkin tampak berhasil tetapi tidak mampu menerjemahkan makna simbol, model, atau besaran yang digunakan. Prinsip diagnosis menuntut bukti tentang cara konsep dipakai, bukan hanya apakah hasil akhirnya sesuai kunci (DeVellis & Thorpe, 2021; Wren & Barbera, 2014).

Secara operasional, guru dapat menggunakan predict-observe-explain untuk menelusuri konsistensi lintas representasi. Data tersebut perlu dibaca bersama urutan langkah penyelesaian, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mengabaikan ketidakpastian pengukuran. Untuk mengurangnya, guru dapat meminta translasi antarrepresentasi dan menggunakan triangulasi bukti. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai reliabilitas observasi dan rubrik, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan reliabilitas observasi dan rubrik melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan reliabilitas observasi dan rubrik;

- Bandingkan bukti reliabilitas observasi dan rubrik pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek reliabilitas observasi dan rubrik dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian reliabilitas observasi dan rubrik, topik katalis dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa katalis dianggap menambah jumlah produk kesetimbangan. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta analisis jalur energi dan laju maju-balik. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa katalis mempercepat tercapainya kesetimbangan tanpa mengubah K. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengendalikan ketidakpastian pengukuran dan konsistensi klasifikasi tanpa mereduksi kompleksitas konsepsi, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, reliabilitas observasi dan rubrik harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.7.15 Reliabilitas pada sampel kecil

Aspek reliabilitas pada sampel kecil menegaskan bahwa diagnosis adalah kegiatan ilmiah berskala kelas: guru merumuskan pertanyaan, mengumpulkan bukti, mempertimbangkan penjelasan alternatif, lalu menentukan kesimpulan yang paling masuk akal. Inferensi harus proporsional terhadap kualitas data. Semakin spesifik label yang diberikan, semakin kuat pula bukti yang diperlukan (Wren & Barbera, 2014; Cronbach, 1951).

Secara operasional, guru dapat menggunakan tugas isomorfik untuk menelusuri kemampuan transfer. Data tersebut perlu dibaca bersama kualitas model mental yang dieksternalisasi, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menyimpulkan miskonsepsi hanya dari satu jawaban. Untuk mengurangnya, guru dapat mengulang diagnosis setelah remediasi dan menyusun tindak lanjut berbeda untuk kategori respons yang berbeda. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai reliabilitas pada sampel kecil, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;

- Pilih tugas yang merepresentasikan reliabilitas pada sampel kecil melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan reliabilitas pada sampel kecil;
- Bandingkan bukti reliabilitas pada sampel kecil pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek reliabilitas pada sampel kecil dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian reliabilitas pada sampel kecil, topik elektrolisis dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa anoda selalu dianggap bermuatan negatif. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta pembedaan definisi anoda-katoda berdasarkan proses oksidasi-reduksi. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa tanda elektroda bergantung jenis sel, tetapi anoda tetap lokasi oksidasi. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengendalikan ketidakpastian pengukuran dan konsistensi klasifikasi tanpa mereduksi kompleksitas konsepsi, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, reliabilitas pada sampel kecil harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.7.16 Pengaruh panjang tes terhadap reliabilitas

Penerapan pengaruh panjang tes terhadap reliabilitas harus menghubungkan teori pengukuran dengan pengetahuan pedagogis kimia. Instrumen yang teknis tetapi tidak memahami struktur konsep akan menghasilkan kategori dangkal; sebaliknya, pengetahuan konten tanpa prosedur pengumpulan bukti dapat menghasilkan kesimpulan intuitif yang sulit diuji. Keduanya perlu dipadukan dalam satu alur yang transparan (Cronbach, 1951; AERA et al., 2014).

Secara operasional, guru dapat menggunakan kuis formatif digital untuk menelusuri stabilitas setelah jeda waktu. Data tersebut perlu dibaca bersama respons terhadap probing, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mencampur kesalahan matematika dengan kesalahan konsep. Untuk menguranginya, guru dapat mencatat derajat keyakinan dan membandingkan jawaban

dengan alasan. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai pengaruh panjang tes terhadap reliabilitas, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan pengaruh panjang tes terhadap reliabilitas melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan pengaruh panjang tes terhadap reliabilitas;
- Bandingkan bukti pengaruh panjang tes terhadap reliabilitas pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek pengaruh panjang tes terhadap reliabilitas dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian pengaruh panjang tes terhadap reliabilitas, topik pengukuran laboratorium dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa angka lebih banyak selalu dianggap lebih akurat. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta pembedaan akurasi, presisi, resolusi, dan angka penting. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa kualitas data ditentukan oleh desain, kalibrasi, dan ketidakpastian. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengendalikan ketidakpastian pengukuran dan konsistensi klasifikasi tanpa mereduksi kompleksitas konsepsi, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Kerangka operasional

Aspek	Bukti yang dicari	Risiko interpretasi	Tindak lanjut
Target konsep	Penjelasan tentang pengukuran laboratorium	Target terlalu luas	Batasi subkonsep dan prasyarat
Respons	Stabilitas setelah jeda waktu	Mencampur kesalahan matematika dengan kesalahan konsep	Mencatat derajat keyakinan
Penalaran	Respons terhadap probing	Alasan diperlakukan sekadar tambahan	Membandingkan jawaban dengan alasan
Konfirmasi	Kuis formatif digital atau tugas isomorfik	Kesimpulan dibuat dari satu konteks	Gunakan bukti kedua dan rediagnosis

Implikasinya, pengaruh panjang tes terhadap reliabilitas harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.7.17 Pengaruh heterogenitas peserta didik

Pengaruh heterogenitas peserta didik tidak berdiri sebagai prosedur administratif. Ia menjadi bagian dari keputusan tentang apa yang perlu dijelaskan ulang, representasi apa yang perlu dibandingkan, dan siapa yang memerlukan dukungan berbeda. Oleh sebab itu, bukti diagnosis harus langsung terhubung dengan tindakan pembelajaran, bukan berhenti pada daftar kesalahan (AERA et al., 2014; DeVellis & Thorpe, 2021).

Secara operasional, guru dapat menggunakan tes tiga tingkat untuk menelusuri urutan langkah penyelesaian. Data tersebut perlu dibaca bersama ketepatan pilihan, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menggunakan ambang kategori tanpa bukti. Untuk mengurangnya, guru dapat menelusuri langkah penyelesaian dan menggunakan pertanyaan tindak lanjut. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai pengaruh heterogenitas peserta didik, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan pengaruh heterogenitas peserta didik melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan pengaruh heterogenitas peserta didik;
- Bandingkan bukti pengaruh heterogenitas peserta didik pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek pengaruh heterogenitas peserta didik dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian pengaruh heterogenitas peserta didik, topik ikatan ion dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa NaCl dipandang sebagai molekul diskrit. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta gambar partikel padatan NaCl dan penjelasan hubungan struktur-kisi dengan titik leleh. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa senyawa ionik membentuk kisi dan rumusnya menyatakan rasio ion. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengendalikan ketidakpastian pengukuran dan konsistensi klasifikasi tanpa mereduksi kompleksitas konsepsi, bukan

diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, pengaruh heterogenitas peserta didik harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.7.18 Analisis butir klasik

Analisis butir klasik perlu ditempatkan dalam kerangka diagnosis yang mengendalikan ketidakpastian pengukuran dan konsistensi klasifikasi tanpa mereduksi kompleksitas konsepsi. Fokusnya bukan sekadar menghasilkan skor, melainkan memperoleh bukti yang cukup untuk menjelaskan bagaimana peserta didik membangun, menghubungkan, dan menggunakan konsep. Karena konsepsi tidak dapat diamati secara langsung, guru harus menarik inferensi dari respons verbal, tulisan, gambar, tindakan, dan perubahan strategi yang tampak selama penyelesaian tugas (DeVellis & Thorpe, 2021; Wren & Barbera, 2014).

Secara operasional, guru dapat menggunakan wawancara semi-terstruktur untuk menelusuri kualitas model mental yang dieksternalisasi. Data tersebut perlu dibaca bersama koherensi alasan, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah memperlakukan konsepsi sebagai sifat permanen. Untuk mengurangnya, guru dapat menggunakan triangulasi bukti dan menyajikan tugas paralel. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai analisis butir klasik, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan analisis butir klasik melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan analisis butir klasik;
- Bandingkan bukti analisis butir klasik pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek analisis butir klasik dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian analisis butir klasik, topik pereaksi pembatas dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa pereaksi dengan massa terkecil selalu dianggap pembatas. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan

relasi konseptual yang tepat. Guru dapat meminta analisis rasio mol terhadap koefisien. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa status pembatas ditentukan oleh hubungan jumlah zat dan stoikiometri. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengendalikan ketidakpastian pengukuran dan konsistensi klasifikasi tanpa mereduksi kompleksitas konsepsi, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, analisis butir klasik harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.7.19 Model Rasch dan fungsi informasi

Pembahasan tentang model rasch dan fungsi informasi berangkat dari prinsip bahwa data asesmen selalu merupakan jejak tidak langsung dari proses berpikir. Satu pilihan jawaban hanya menunjukkan produk akhir, sedangkan struktur pengetahuan yang melahirkannya dapat berbeda. Diagnosis yang kuat karena itu menggabungkan respons terhadap probing, tingkat keyakinan, dan bukti kontekstual sehingga keputusan pedagogis tidak bergantung pada satu indikator saja (Wren & Barbera, 2014; Cronbach, 1951).

Secara operasional, guru dapat menggunakan gambar partikel untuk menelusuri respons terhadap probing. Data tersebut perlu dibaca bersama tingkat keyakinan, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mengabaikan ketidakpastian pengukuran. Untuk mengurangnya, guru dapat menyusun tindak lanjut berbeda untuk kategori respons yang berbeda dan memeriksa konsep prasyarat. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai model rasch dan fungsi informasi, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan model rasch dan fungsi informasi melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan model rasch dan fungsi informasi;
- Bandingkan bukti model rasch dan fungsi informasi pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek model rasch dan fungsi informasi dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian model rasch dan fungsi informasi, topik larutan penyangga dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa penyangga dianggap mempertahankan pH secara mutlak. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta analisis perubahan komposisi setelah penambahan asam atau basa. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa penyangga membatasi perubahan dalam kapasitas tertentu, bukan meniadakannya. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengendalikan ketidakpastian pengukuran dan konsistensi klasifikasi tanpa mereduksi kompleksitas konsepsi, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, model rasch dan fungsi informasi harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.7.20 Person reliability dan item reliability

Dalam pembelajaran kimia, person reliability dan item reliability memiliki arti strategis karena banyak jawaban dapat diperoleh melalui hafalan prosedur atau pengenalan pola. Peserta didik mungkin tampak berhasil tetapi tidak mampu menerjemahkan makna simbol, model, atau besaran yang digunakan. Prinsip diagnosis menuntut bukti tentang cara konsep dipakai, bukan hanya apakah hasil akhirnya sesuai kunci (Cronbach, 1951; AERA et al., 2014).

Secara operasional, guru dapat menggunakan observasi praktikum untuk menelusuri ketepatan pilihan. Data tersebut perlu dibaca bersama konsistensi lintas butir, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menyimpulkan miskonsepsi hanya dari satu jawaban. Untuk mengurangnya, guru dapat membandingkan jawaban dengan alasan dan meminta translasi antarrepresentasi. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai person reliability dan item reliability, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan person reliability dan item reliability melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan person reliability dan item reliability;

- Bandingkan bukti person reliability dan item reliability pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek person reliability dan item reliability dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian person reliability dan item reliability, topik katalis dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa katalis dianggap menambah jumlah produk kesetimbangan. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta analisis jalur energi dan laju maju-balik. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa katalis mempercepat tercapainya kesetimbangan tanpa mengubah K. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengendalikan ketidakpastian pengukuran dan konsistensi klasifikasi tanpa mereduksi kompleksitas konsepsi, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Kerangka operasional

Aspek	Bukti yang dicari	Risiko interpretasi	Tindak lanjut
Target konsep	Penjelasan tentang katalis	Target terlalu luas	Batasi subkonsep dan prasyarat
Respons	Ketepatan pilihan	Menyimpulkan miskonsepsi hanya dari satu jawaban	Membandingkan jawaban dengan alasan
Penalaran	Konsistensi lintas butir	Alasan diperlakukan sekadar tambahan	Meminta translasi antarrepresentasi
Konfirmasi	Observasi praktikum atau tugas isomorfik	Kesimpulan dibuat dari satu konteks	Gunakan bukti kedua dan rediagnosis

Implikasinya, person reliability dan item reliability harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.7.21 Differential item functioning

Aspek differential item functioning menegaskan bahwa diagnosis adalah kegiatan ilmiah berskala kelas: guru merumuskan pertanyaan, mengumpulkan bukti, mempertimbangkan

penjelasan alternatif, lalu menentukan kesimpulan yang paling masuk akal. Inferensi harus proporsional terhadap kualitas data. Semakin spesifik label yang diberikan, semakin kuat pula bukti yang diperlukan (AERA et al., 2014; DeVellis & Thorpe, 2021).

Secara operasional, guru dapat menggunakan exit ticket untuk menelusuri koherensi alasan. Data tersebut perlu dibaca bersama konsistensi lintas representasi, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mencampur kesalahan matematika dengan kesalahan konsep. Untuk mengurangnya, guru dapat menggunakan pertanyaan tindak lanjut dan mengulang diagnosis setelah remediasi. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai differential item functioning, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan differential item functioning melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan differential item functioning;
- Bandingkan bukti differential item functioning pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek differential item functioning dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian differential item functioning, topik elektrolisis dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa anoda selalu dianggap bermuatan negatif. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta pembedaan definisi anoda-katoda berdasarkan proses oksidasi-reduksi. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa tanda elektroda bergantung jenis sel, tetapi anoda tetap lokasi oksidasi. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengendalikan ketidakpastian pengukuran dan konsistensi klasifikasi tanpa mereduksi kompleksitas konsepsi, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, differential item functioning harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.7.22 Standard error pada skor individual

Penerapan standard error pada skor individual harus menghubungkan teori pengukuran dengan pengetahuan pedagogis kimia. Instrumen yang teknis tetapi tidak memahami struktur konsep akan menghasilkan kategori dangkal; sebaliknya, pengetahuan konten tanpa prosedur pengumpulan bukti dapat menghasilkan kesimpulan intuitif yang sulit diuji. Keduanya perlu dipadukan dalam satu alur yang transparan (DeVellis & Thorpe, 2021; Wren & Barbera, 2014).

Secara operasional, guru dapat menggunakan tes pilihan ganda dua tingkat untuk menelusuri tingkat keyakinan. Data tersebut perlu dibaca bersama kemampuan transfer, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menggunakan ambang kategori tanpa bukti. Untuk mengurangnya, guru dapat menyajikan tugas paralel dan mencatat derajat keyakinan. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai standard error pada skor individual, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan standard error pada skor individual melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan standard error pada skor individual;
- Bandingkan bukti standard error pada skor individual pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek standard error pada skor individual dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian standard error pada skor individual, topik pengukuran laboratorium dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa angka lebih banyak selalu dianggap lebih akurat. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta pembedaan akurasi, presisi, resolusi, dan angka penting. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa kualitas data ditentukan oleh desain, kalibrasi, dan ketidakpastian. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengendalikan ketidakpastian pengukuran dan konsistensi klasifikasi tanpa mereduksi kompleksitas konsepsi, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, standard error pada skor individual harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik

'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.7.23 Keputusan kategori dan konsistensi klasifikasi

Keputusan kategori dan konsistensi klasifikasi tidak berdiri sebagai prosedur administratif. Ia menjadi bagian dari keputusan tentang apa yang perlu dijelaskan ulang, representasi apa yang perlu dibandingkan, dan siapa yang memerlukan dukungan berbeda. Oleh sebab itu, bukti diagnosis harus langsung terhubung dengan tindakan pembelajaran, bukan berhenti pada daftar kesalahan (Wren & Barbera, 2014; Cronbach, 1951).

Secara operasional, guru dapat menggunakan alasan terbuka untuk menelusuri konsistensi lintas butir. Data tersebut perlu dibaca bersama stabilitas setelah jeda waktu, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah memperlakukan konsepsi sebagai sifat permanen. Untuk menguranginya, guru dapat memeriksa konsep prasyarat dan menelusuri langkah penyelesaian. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai keputusan kategori dan konsistensi klasifikasi, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan keputusan kategori dan konsistensi klasifikasi melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan keputusan kategori dan konsistensi klasifikasi;
- Bandingkan bukti keputusan kategori dan konsistensi klasifikasi pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek keputusan kategori dan konsistensi klasifikasi dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian keputusan kategori dan konsistensi klasifikasi, topik ikatan ion dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa NaCl dipandang sebagai molekul diskrit. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta gambar partikel padatan NaCl dan penjelasan hubungan struktur-kisi dengan titik leleh. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa senyawa ionik membentuk kisi dan rumusnya menyatakan rasio ion. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengendalikan ketidakpastian pengukuran dan konsistensi klasifikasi tanpa mereduksi kompleksitas konsepsi, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, keputusan kategori dan konsistensi klasifikasi harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.7.24 Menyeimbangkan reliabilitas dan kedalaman diagnosis

Menyeimbangkan reliabilitas dan kedalaman diagnosis perlu ditempatkan dalam kerangka diagnosis yang mengendalikan ketidakpastian pengukuran dan konsistensi klasifikasi tanpa mereduksi kompleksitas konsepsi. Fokusnya bukan sekadar menghasilkan skor, melainkan memperoleh bukti yang cukup untuk menjelaskan bagaimana peserta didik membangun, menghubungkan, dan menggunakan konsep. Karena konsepsi tidak dapat diamati secara langsung, guru harus menarik inferensi dari respons verbal, tulisan, gambar, tindakan, dan perubahan strategi yang tampak selama penyelesaian tugas (Cronbach, 1951; AERA et al., 2014).

Secara operasional, guru dapat menggunakan peta konsep untuk menelusuri konsistensi lintas representasi. Data tersebut perlu dibaca bersama urutan langkah penyelesaian, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mengabaikan ketidakpastian pengukuran. Untuk mengurangnya, guru dapat meminta translasi antarrepresentasi dan menggunakan triangulasi bukti. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai menyeimbangkan reliabilitas dan kedalaman diagnosis, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan menyeimbangkan reliabilitas dan kedalaman diagnosis melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan menyeimbangkan reliabilitas dan kedalaman diagnosis;
- Bandingkan bukti menyeimbangkan reliabilitas dan kedalaman diagnosis pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek menyeimbangkan reliabilitas dan kedalaman diagnosis dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian menyeimbangkan reliabilitas dan kedalaman diagnosis, topik pereaksi pembatas dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa pereaksi dengan massa terkecil selalu dianggap pembatas. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta analisis rasio mol terhadap koefisien. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan

bahwa status pembatas ditentukan oleh hubungan jumlah zat dan stoikiometri. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengendalikan ketidakpastian pengukuran dan konsistensi klasifikasi tanpa mereduksi kompleksitas konsepsi, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Kerangka operasional

Aspek	Bukti yang dicari	Risiko interpretasi	Tindak lanjut
Target konsep	Penjelasan tentang pereaksi pembatas	Target terlalu luas	Batasi subkonsep dan prasyarat
Respons	Konsistensi lintas representasi	Mengabaikan ketidakpastian pengukuran	Meminta translasi antarrepresentasi
Penalaran	Urutan langkah penyelesaian	Alasan diperlakukan sekadar tambahan	Menggunakan triangulasi bukti
Konfirmasi	Peta konsep atau tugas isomorfik	Kesimpulan dibuat dari satu konteks	Gunakan bukti kedua dan rediagnosis

Implikasinya, menyeimbangkan reliabilitas dan kedalaman diagnosis harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.7.25 Pelaporan reliabilitas secara bertanggung jawab

Pembahasan tentang pelaporan reliabilitas secara bertanggung jawab berangkat dari prinsip bahwa data asesmen selalu merupakan jejak tidak langsung dari proses berpikir. Satu pilihan jawaban hanya menunjukkan produk akhir, sedangkan struktur pengetahuan yang melahirkannya dapat berbeda. Diagnosis yang kuat karena itu menggabungkan kemampuan transfer, kualitas model mental yang dieksternalisasi, dan bukti kontekstual sehingga keputusan pedagogis tidak bergantung pada satu indikator saja (AERA et al., 2014; DeVellis & Thorpe, 2021).

Secara operasional, guru dapat menggunakan analisis respons tertulis untuk menelusuri kemampuan transfer. Data tersebut perlu dibaca bersama kualitas model mental yang dieksternalisasi, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menyimpulkan miskonsepsi hanya dari satu jawaban. Untuk mengurangnya, guru dapat mengulang diagnosis setelah remediasi dan menyusun

tindak lanjut berbeda untuk kategori respons yang berbeda. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai pelaporan reliabilitas secara bertanggung jawab, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan pelaporan reliabilitas secara bertanggung jawab melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan pelaporan reliabilitas secara bertanggung jawab;
- Bandingkan bukti pelaporan reliabilitas secara bertanggung jawab pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek pelaporan reliabilitas secara bertanggung jawab dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian pelaporan reliabilitas secara bertanggung jawab, topik larutan penyangga dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa penyangga dianggap mempertahankan pH secara mutlak. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta analisis perubahan komposisi setelah penambahan asam atau basa. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa penyangga membatasi perubahan dalam kapasitas tertentu, bukan meniadakannya. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengendalikan ketidakpastian pengukuran dan konsistensi klasifikasi tanpa mereduksi kompleksitas konsepsi, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, pelaporan reliabilitas secara bertanggung jawab harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.7.26 Sintesis prinsip reliabilitas instrumen

Dalam pembelajaran kimia, sintesis prinsip reliabilitas instrumen memiliki arti strategis karena banyak jawaban dapat diperoleh melalui hafalan prosedur atau pengenalan pola. Peserta didik mungkin tampak berhasil tetapi tidak mampu menerjemahkan makna simbol, model, atau besaran yang digunakan. Prinsip diagnosis menuntut bukti tentang cara konsep dipakai, bukan hanya apakah hasil akhirnya sesuai kunci (DeVellis & Thorpe, 2021; Wren & Barbera, 2014).

Secara operasional, guru dapat menggunakan diskusi argumentatif untuk menelusuri stabilitas setelah jeda waktu. Data tersebut perlu dibaca bersama respons terhadap probing, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mencampur kesalahan matematika dengan kesalahan konsep. Untuk mengurangnya, guru dapat mencatat derajat keyakinan dan membandingkan jawaban dengan alasan. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai sintesis prinsip reliabilitas instrumen, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan sintesis prinsip reliabilitas instrumen melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan sintesis prinsip reliabilitas instrumen;
- Bandingkan bukti sintesis prinsip reliabilitas instrumen pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek sintesis prinsip reliabilitas instrumen dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian sintesis prinsip reliabilitas instrumen, topik katalis dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa katalis dianggap menambah jumlah produk kesetimbangan. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta analisis jalur energi dan laju maju-balik. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa katalis mempercepat tercapainya kesetimbangan tanpa mengubah K . Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengendalikan ketidakpastian pengukuran dan konsistensi klasifikasi tanpa mereduksi kompleksitas konsepsi, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, sintesis prinsip reliabilitas instrumen harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.7.27 Implikasi praktis bagi guru kimia

Implikasi praktis subbab 5.7 dapat dirangkum sebagai alur: tentukan konsep target, pilih bukti yang relevan, kumpulkan respons dalam lebih dari satu bentuk, analisis pola, nyatakan derajat

kepastian, susun tindak lanjut, dan lakukan rediagnosis. Alur ini menjaga agar prinsip reliabilitas instrumen tidak berubah menjadi administrasi skor, tetapi berfungsi sebagai mekanisme perbaikan pembelajaran yang berulang.

43. Identifikasi konsep inti dan konsep prasyarat sebelum menyusun instrumen.
44. Gunakan pertanyaan yang meminta alasan, representasi, atau langkah kerja.
45. Bandingkan respons pada beberapa konteks dan waktu.
46. Pisahkan bukti faktual dari interpretasi guru.
47. Tentukan tindak lanjut yang berbeda untuk kategori kebutuhan yang berbeda.
48. Simpan data secukupnya dan lindungi kerahasiaan peserta didik.
49. Periksa kembali perubahan melalui tugas transfer dan retensi.

5.8 Diagnosis Individual dan Diagnosis Kelas

Subbab ini mengubah data respons menjadi gambaran kebutuhan individual sekaligus pola pembelajaran tingkat kelas. Uraian disusun dari prinsip teoretis menuju prosedur kelas, contoh kimia, risiko interpretasi, dan tindak lanjut. Setiap bagian menekankan bahwa kesimpulan diagnostik harus didasarkan pada pola bukti, bukan pada satu respons yang terpisah.

5.8.1 Perbedaan unit analisis individual dan kelas

Pembahasan tentang perbedaan unit analisis individual dan kelas berangkat dari prinsip bahwa data asesmen selalu merupakan jejak tidak langsung dari proses berpikir. Satu pilihan jawaban hanya menunjukkan produk akhir, sedangkan struktur pengetahuan yang melahirkannya dapat berbeda. Diagnosis yang kuat karena itu menggabungkan koherensi alasan, konsistensi lintas representasi, dan bukti kontekstual sehingga keputusan pedagogis tidak bergantung pada satu indikator saja (Taber, 2002a; Anggraena et al., 2025).

Secara operasional, guru dapat menggunakan tugas isomorfik untuk menelusuri koherensi alasan. Data tersebut perlu dibaca bersama konsistensi lintas representasi, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mencampur kesalahan matematika dengan kesalahan konsep. Untuk mengurangnya, guru dapat menggunakan pertanyaan tindak lanjut dan mengulang diagnosis setelah remediasi. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai perbedaan unit analisis individual dan kelas, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan perbedaan unit analisis individual dan kelas melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan perbedaan unit analisis individual dan kelas;

- Bandingkan bukti perbedaan unit analisis individual dan kelas pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek perbedaan unit analisis individual dan kelas dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian perbedaan unit analisis individual dan kelas, topik kesetimbangan dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa reaksi dianggap berhenti saat setimbang. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta hubungan grafik laju, konsentrasi, dan model partikel. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa kesetimbangan bersifat dinamis dengan laju maju dan balik yang sama. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengubah data respons menjadi gambaran kebutuhan individual sekaligus pola pembelajaran tingkat kelas, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, perbedaan unit analisis individual dan kelas harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.8.2 Tujuan diagnosis individual

Dalam pembelajaran kimia, tujuan diagnosis individual memiliki arti strategis karena banyak jawaban dapat diperoleh melalui hafalan prosedur atau pengenalan pola. Peserta didik mungkin tampak berhasil tetapi tidak mampu menerjemahkan makna simbol, model, atau besaran yang digunakan. Prinsip diagnosis menuntut bukti tentang cara konsep dipakai, bukan hanya apakah hasil akhirnya sesuai kunci (Anggraena et al., 2025; Hoe & Subramaniam, 2016).

Secara operasional, guru dapat menggunakan kuis formatif digital untuk menelusuri tingkat keyakinan. Data tersebut perlu dibaca bersama kemampuan transfer, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menggunakan ambang kategori tanpa bukti. Untuk mengurangnya, guru dapat menyajikan tugas paralel dan mencatat derajat keyakinan. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai tujuan diagnosis individual, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;

- Pilih tugas yang merepresentasikan tujuan diagnosis individual melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan tujuan diagnosis individual;
- Bandingkan bukti tujuan diagnosis individual pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek tujuan diagnosis individual dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian tujuan diagnosis individual, topik termokimia dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa pemutusan ikatan dianggap melepaskan energi. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta analisis energi yang diserap saat memutus dan dilepas saat membentuk ikatan. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa perubahan entalpi bersih berasal dari selisih proses pemutusan dan pembentukan. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengubah data respons menjadi gambaran kebutuhan individual sekaligus pola pembelajaran tingkat kelas, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, tujuan diagnosis individual harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.8.3 Tujuan diagnosis kelas

Aspek tujuan diagnosis kelas menegaskan bahwa diagnosis adalah kegiatan ilmiah berskala kelas: guru merumuskan pertanyaan, mengumpulkan bukti, mempertimbangkan penjelasan alternatif, lalu menentukan kesimpulan yang paling masuk akal. Inferensi harus proporsional terhadap kualitas data. Semakin spesifik label yang diberikan, semakin kuat pula bukti yang diperlukan (Hoe & Subramaniam, 2016; Treagust, 1988).

Secara operasional, guru dapat menggunakan tes tiga tingkat untuk menelusuri konsistensi lintas butir. Data tersebut perlu dibaca bersama stabilitas setelah jeda waktu, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah memperlakukan konsepsi sebagai sifat permanen. Untuk menguranginya, guru dapat memeriksa konsep prasyarat dan menelusuri langkah penyelesaian. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai tujuan diagnosis kelas, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan tujuan diagnosis kelas melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan tujuan diagnosis kelas;
- Bandingkan bukti tujuan diagnosis kelas pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek tujuan diagnosis kelas dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian tujuan diagnosis kelas, topik hidrokarbon dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa isomer dipandang sebagai zat yang sama karena rumus molekulnya sama. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta hubungan struktur, konektivitas, dan sifat. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa rumus molekul tidak cukup menentukan struktur dan identitas senyawa. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengubah data respons menjadi gambaran kebutuhan individual sekaligus pola pembelajaran tingkat kelas, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, tujuan diagnosis kelas harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.8.4 Profil respons individual

Penerapan profil respons individual harus menghubungkan teori pengukuran dengan pengetahuan pedagogis kimia. Instrumen yang teknis tetapi tidak memahami struktur konsep akan menghasilkan kategori dangkal; sebaliknya, pengetahuan konten tanpa prosedur pengumpulan bukti dapat menghasilkan kesimpulan intuitif yang sulit diuji. Keduanya perlu dipadukan dalam satu alur yang transparan (Treagust, 1988; Taber, 2002a).

Secara operasional, guru dapat menggunakan wawancara semi-terstruktur untuk menelusuri konsistensi lintas representasi. Data tersebut perlu dibaca bersama urutan langkah penyelesaian, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mengabaikan ketidakpastian pengukuran. Untuk menguranginya, guru dapat meminta translasi antarrepresentasi dan menggunakan

triangulasi bukti. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai profil respons individual, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan profil respons individual melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan profil respons individual;
- Bandingkan bukti profil respons individual pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek profil respons individual dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian profil respons individual, topik spektroskopi dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa spektrum dipandang sebagai gambar langsung molekul. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta hubungan sinyal, transisi energi, dan model struktur. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa spektrum adalah data yang ditafsirkan melalui model dan aturan seleksi. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengubah data respons menjadi gambaran kebutuhan individual sekaligus pola pembelajaran tingkat kelas, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Kerangka operasional

Aspek	Bukti yang dicari	Risiko interpretasi	Tindak lanjut
Target konsep	Penjelasan tentang spektroskopi	Target terlalu luas	Batasi subkonsep dan prasyarat
Respons	Konsistensi lintas representasi	Mengabaikan ketidakpastian pengukuran	Meminta translasi antarrepresentasi
Penalaran	Urutan langkah penyelesaian	Alasan diperlakukan sekadar tambahan	Menggunakan triangulasi bukti
Konfirmasi	Wawancara semi-terstruktur atau tugas isomorfik	Kesimpulan dibuat dari satu konteks	Gunakan bukti kedua dan rediagnosis

Implikasinya, profil respons individual harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.8.5 Pola dominan pada tingkat kelas

Pola dominan pada tingkat kelas tidak berdiri sebagai prosedur administratif. Ia menjadi bagian dari keputusan tentang apa yang perlu dijelaskan ulang, representasi apa yang perlu dibandingkan, dan siapa yang memerlukan dukungan berbeda. Oleh sebab itu, bukti diagnosis harus langsung terhubung dengan tindakan pembelajaran, bukan berhenti pada daftar kesalahan (Taber, 2002a; Anggraena et al., 2025).

Secara operasional, guru dapat menggunakan gambar partikel untuk menelusuri kemampuan transfer. Data tersebut perlu dibaca bersama kualitas model mental yang dieksternalisasi, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menyimpulkan miskonsepsi hanya dari satu jawaban. Untuk mengurangnya, guru dapat mengulang diagnosis setelah remediasi dan menyusun tindak lanjut berbeda untuk kategori respons yang berbeda. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai pola dominan pada tingkat kelas, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan pola dominan pada tingkat kelas melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan pola dominan pada tingkat kelas;
- Bandingkan bukti pola dominan pada tingkat kelas pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek pola dominan pada tingkat kelas dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian pola dominan pada tingkat kelas, topik ikatan kovalen dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa garis ikatan dipandang sebagai batang fisik. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta penjelasan energi potensial dan distribusi elektron pada jarak antarinti berbeda. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa representasi garis adalah konvensi yang merangkum pasangan elektron dan konektivitas. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengubah data respons menjadi gambaran kebutuhan individual sekaligus pola pembelajaran

tingkat kelas, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, pola dominan pada tingkat kelas harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.8.6 Prevalensi miskonsepsi dan maknanya

Prevalensi miskonsepsi dan maknanya perlu ditempatkan dalam kerangka diagnosis yang mengubah data respons menjadi gambaran kebutuhan individual sekaligus pola pembelajaran tingkat kelas. Fokusnya bukan sekadar menghasilkan skor, melainkan memperoleh bukti yang cukup untuk menjelaskan bagaimana peserta didik membangun, menghubungkan, dan menggunakan konsep. Karena konsepsi tidak dapat diamati secara langsung, guru harus menarik inferensi dari respons verbal, tulisan, gambar, tindakan, dan perubahan strategi yang tampak selama penyelesaian tugas (Anggraena et al., 2025; Hoe & Subramaniam, 2016).

Secara operasional, guru dapat menggunakan observasi praktikum untuk menelusuri stabilitas setelah jeda waktu. Data tersebut perlu dibaca bersama respons terhadap probing, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mencampur kesalahan matematika dengan kesalahan konsep. Untuk mengurangnya, guru dapat mencatat derajat keyakinan dan membandingkan jawaban dengan alasan. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai prevalensi miskonsepsi dan maknanya, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan prevalensi miskonsepsi dan maknanya melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan prevalensi miskonsepsi dan maknanya;
- Bandingkan bukti prevalensi miskonsepsi dan maknanya pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek prevalensi miskonsepsi dan maknanya dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian prevalensi miskonsepsi dan maknanya, topik gas dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa partikel gas dianggap mengembang ketika dipanaskan. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta gambar partikel sebelum dan sesudah

pemanasan pada kondisi berbeda. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa ukuran partikel tidak berubah secara bermakna; distribusi kecepatan dan jarak berubah. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengubah data respons menjadi gambaran kebutuhan individual sekaligus pola pembelajaran tingkat kelas, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, prevalensi miskonsepsi dan maknanya harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.8.7 Distribusi keyakinan dalam kelas

Pembahasan tentang distribusi keyakinan dalam kelas berangkat dari prinsip bahwa data asesmen selalu merupakan jejak tidak langsung dari proses berpikir. Satu pilihan jawaban hanya menunjukkan produk akhir, sedangkan struktur pengetahuan yang melahirkannya dapat berbeda. Diagnosis yang kuat karena itu menggabungkan urutan langkah penyelesaian, ketepatan pilihan, dan bukti kontekstual sehingga keputusan pedagogis tidak bergantung pada satu indikator saja (Hoe & Subramaniam, 2016; Treagust, 1988).

Secara operasional, guru dapat menggunakan exit ticket untuk menelusuri urutan langkah penyelesaian. Data tersebut perlu dibaca bersama ketepatan pilihan, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menggunakan ambang kategori tanpa bukti. Untuk mengurangnya, guru dapat menelusuri langkah penyelesaian dan menggunakan pertanyaan tindak lanjut. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai distribusi keyakinan dalam kelas, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan distribusi keyakinan dalam kelas melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan distribusi keyakinan dalam kelas;
- Bandingkan bukti distribusi keyakinan dalam kelas pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek distribusi keyakinan dalam kelas dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian distribusi keyakinan dalam kelas, topik kesetimbangan dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa reaksi dianggap berhenti saat setimbang. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta hubungan grafik laju, konsentrasi, dan model partikel. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa kesetimbangan bersifat dinamis dengan laju maju dan balik yang sama. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengubah data respons menjadi gambaran kebutuhan individual sekaligus pola pembelajaran tingkat kelas, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, distribusi keyakinan dalam kelas harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.8.8 Kelompok konseptual dan diferensiasi

Dalam pembelajaran kimia, kelompok konseptual dan diferensiasi memiliki arti strategis karena banyak jawaban dapat diperoleh melalui hafalan prosedur atau pengenalan pola. Peserta didik mungkin tampak berhasil tetapi tidak mampu menerjemahkan makna simbol, model, atau besaran yang digunakan. Prinsip diagnosis menuntut bukti tentang cara konsep dipakai, bukan hanya apakah hasil akhirnya sesuai kunci (Treagust, 1988; Taber, 2002a).

Secara operasional, guru dapat menggunakan tes pilihan ganda dua tingkat untuk menelusuri kualitas model mental yang dieksternalisasi. Data tersebut perlu dibaca bersama koherensi alasan, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah memperlakukan konsepsi sebagai sifat permanen. Untuk mengurangnya, guru dapat menggunakan triangulasi bukti dan menyajikan tugas paralel. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai kelompok konseptual dan diferensiasi, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan kelompok konseptual dan diferensiasi melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan kelompok konseptual dan diferensiasi;
- Bandingkan bukti kelompok konseptual dan diferensiasi pada lebih dari satu konteks, format tugas, atau bentuk representasi;

- Hubungkan kategori hasil pada aspek kelompok konseptual dan diferensiasi dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian kelompok konseptual dan diferensiasi, topik termokimia dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa pemutusan ikatan dianggap melepaskan energi. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta analisis energi yang diserap saat memutus dan dilepas saat membentuk ikatan. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa perubahan entalpi bersih berasal dari selisih proses pemutusan dan pembentukan. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengubah data respons menjadi gambaran kebutuhan individual sekaligus pola pembelajaran tingkat kelas, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Kerangka operasional

Aspek	Bukti yang dicari	Risiko interpretasi	Tindak lanjut
Target konsep	Penjelasan tentang termokimia	Target terlalu luas	Batasi subkonsep dan prasyarat
Respons	Kualitas model mental yang dieksternalisasi	Memperlakukan konsepsi sebagai sifat permanen	Menggunakan triangulasi bukti
Penalaran	Koherensi alasan	Alasan diperlakukan sekadar tambahan	Menyajikan tugas paralel
Konfirmasi	Tes pilihan ganda dua tingkat atau tugas isomorfik	Kesimpulan dibuat dari satu konteks	Gunakan bukti kedua dan rediagnosis

Implikasinya, kelompok konseptual dan diferensiasi harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.8.9 Diagnosis peserta didik dengan respons tidak konsisten

Aspek diagnosis peserta didik dengan respons tidak konsisten menegaskan bahwa diagnosis adalah kegiatan ilmiah berskala kelas: guru merumuskan pertanyaan, mengumpulkan bukti, mempertimbangkan penjelasan alternatif, lalu menentukan kesimpulan yang paling masuk

akal. Inferensi harus proporsional terhadap kualitas data. Semakin spesifik label yang diberikan, semakin kuat pula bukti yang diperlukan (Taber, 2002a; Anggraena et al., 2025).

Secara operasional, guru dapat menggunakan alasan terbuka untuk menelusuri respons terhadap probing. Data tersebut perlu dibaca bersama tingkat keyakinan, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mengabaikan ketidakpastian pengukuran. Untuk mengurangnya, guru dapat menyusun tindak lanjut berbeda untuk kategori respons yang berbeda dan memeriksa konsep prasyarat. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai diagnosis peserta didik dengan respons tidak konsisten, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan diagnosis peserta didik dengan respons tidak konsisten melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan diagnosis peserta didik dengan respons tidak konsisten;
- Bandingkan bukti diagnosis peserta didik dengan respons tidak konsisten pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek diagnosis peserta didik dengan respons tidak konsisten dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian diagnosis peserta didik dengan respons tidak konsisten, topik hidrokarbon dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa isomer dipandang sebagai zat yang sama karena rumus molekulnya sama. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta hubungan struktur, konektivitas, dan sifat. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa rumus molekul tidak cukup menentukan struktur dan identitas senyawa. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengubah data respons menjadi gambaran kebutuhan individual sekaligus pola pembelajaran tingkat kelas, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, diagnosis peserta didik dengan respons tidak konsisten harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.8.10 Diagnosis peserta didik berkemampuan tinggi

Penerapan diagnosis peserta didik berkemampuan tinggi harus menghubungkan teori pengukuran dengan pengetahuan pedagogis kimia. Instrumen yang teknis tetapi tidak memahami struktur konsep akan menghasilkan kategori dangkal; sebaliknya, pengetahuan konten tanpa prosedur pengumpulan bukti dapat menghasilkan kesimpulan intuitif yang sulit diuji. Keduanya perlu dipadukan dalam satu alur yang transparan (Anggraena et al., 2025; Hoe & Subramaniam, 2016).

Secara operasional, guru dapat menggunakan peta konsep untuk menelusuri ketepatan pilihan. Data tersebut perlu dibaca bersama konsistensi lintas butir, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menyimpulkan miskonsepsi hanya dari satu jawaban. Untuk mengurangnya, guru dapat membandingkan jawaban dengan alasan dan meminta translasi antarrepresentasi. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai diagnosis peserta didik berkemampuan tinggi, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan diagnosis peserta didik berkemampuan tinggi melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan diagnosis peserta didik berkemampuan tinggi;
- Bandingkan bukti diagnosis peserta didik berkemampuan tinggi pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek diagnosis peserta didik berkemampuan tinggi dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian diagnosis peserta didik berkemampuan tinggi, topik spektroskopi dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa spektrum dipandang sebagai gambar langsung molekul. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta hubungan sinyal, transisi energi, dan model struktur. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa spektrum adalah data yang ditafsirkan melalui model dan aturan seleksi. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengubah data respons menjadi gambaran kebutuhan individual sekaligus pola pembelajaran tingkat kelas, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, diagnosis peserta didik berkemampuan tinggi harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi,

serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.8.11 Diagnosis peserta didik yang mengalami hambatan belajar

Diagnosis peserta didik yang mengalami hambatan belajar tidak berdiri sebagai prosedur administratif. Ia menjadi bagian dari keputusan tentang apa yang perlu dijelaskan ulang, representasi apa yang perlu dibandingkan, dan siapa yang memerlukan dukungan berbeda. Oleh sebab itu, bukti diagnosis harus langsung terhubung dengan tindakan pembelajaran, bukan berhenti pada daftar kesalahan (Hoe & Subramaniam, 2016; Treagust, 1988).

Secara operasional, guru dapat menggunakan analisis respons tertulis untuk menelusuri koherensi alasan. Data tersebut perlu dibaca bersama konsistensi lintas representasi, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mencampur kesalahan matematika dengan kesalahan konsep. Untuk mengurangnya, guru dapat menggunakan pertanyaan tindak lanjut dan mengulang diagnosis setelah remediasi. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai diagnosis peserta didik yang mengalami hambatan belajar, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan diagnosis peserta didik yang mengalami hambatan belajar melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan diagnosis peserta didik yang mengalami hambatan belajar;
- Bandingkan bukti diagnosis peserta didik yang mengalami hambatan belajar pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek diagnosis peserta didik yang mengalami hambatan belajar dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian diagnosis peserta didik yang mengalami hambatan belajar, topik ikatan kovalen dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa garis ikatan dipandang sebagai batang fisik. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta penjelasan energi potensial dan distribusi elektron pada jarak antarinti berbeda. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa representasi garis adalah konvensi yang merangkum pasangan elektron dan konektivitas. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengubah data respons menjadi gambaran kebutuhan individual sekaligus pola pembelajaran tingkat kelas, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin

berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, diagnosis peserta didik yang mengalami hambatan belajar harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.8.12 Membaca variasi di balik rerata kelas

Membaca variasi di balik rerata kelas perlu ditempatkan dalam kerangka diagnosis yang mengubah data respons menjadi gambaran kebutuhan individual sekaligus pola pembelajaran tingkat kelas. Fokusnya bukan sekadar menghasilkan skor, melainkan memperoleh bukti yang cukup untuk menjelaskan bagaimana peserta didik membangun, menghubungkan, dan menggunakan konsep. Karena konsepsi tidak dapat diamati secara langsung, guru harus menarik inferensi dari respons verbal, tulisan, gambar, tindakan, dan perubahan strategi yang tampak selama penyelesaian tugas (Treagust, 1988; Taber, 2002a).

Secara operasional, guru dapat menggunakan diskusi argumentatif untuk menelusuri tingkat keyakinan. Data tersebut perlu dibaca bersama kemampuan transfer, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menggunakan ambang kategori tanpa bukti. Untuk menguranginya, guru dapat menyajikan tugas paralel dan mencatat derajat keyakinan. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai membaca variasi di balik rerata kelas, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan membaca variasi di balik rerata kelas melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan membaca variasi di balik rerata kelas;
- Bandingkan bukti membaca variasi di balik rerata kelas pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek membaca variasi di balik rerata kelas dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian membaca variasi di balik rerata kelas, topik gas dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa partikel gas dianggap mengembang ketika dipanaskan. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta gambar partikel sebelum dan sesudah

pemanasan pada kondisi berbeda. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa ukuran partikel tidak berubah secara bermakna; distribusi kecepatan dan jarak berubah. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengubah data respons menjadi gambaran kebutuhan individual sekaligus pola pembelajaran tingkat kelas, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Kerangka operasional

Aspek	Bukti yang dicari	Risiko interpretasi	Tindak lanjut
Target konsep	Penjelasan tentang gas	Target terlalu luas	Batasi subkonsep dan prasyarat
Respons	Tingkat keyakinan	Menggunakan ambang kategori tanpa bukti	Menyajikan tugas paralel
Penalaran	Kemampuan transfer	Alasan diperlakukan sekadar tambahan	Mencatat derajat keyakinan
Konfirmasi	Diskusi argumentatif atau tugas isomorfik	Kesimpulan dibuat dari satu konteks	Gunakan bukti kedua dan rediagnosis

Implikasinya, membaca variasi di balik rerata kelas harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.8.13 Analisis butir untuk pemetaan kelas

Pembahasan tentang analisis butir untuk pemetaan kelas berangkat dari prinsip bahwa data asesmen selalu merupakan jejak tidak langsung dari proses berpikir. Satu pilihan jawaban hanya menunjukkan produk akhir, sedangkan struktur pengetahuan yang melahirkannya dapat berbeda. Diagnosis yang kuat karena itu menggabungkan konsistensi lintas butir, stabilitas setelah jeda waktu, dan bukti kontekstual sehingga keputusan pedagogis tidak bergantung pada satu indikator saja (Taber, 2002a; Anggraena et al., 2025).

Secara operasional, guru dapat menggunakan analisis grafik untuk menelusuri konsistensi lintas butir. Data tersebut perlu dibaca bersama stabilitas setelah jeda waktu, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah memperlakukan konsepsi sebagai sifat permanen. Untuk menguranginya, guru dapat memeriksa konsep prasyarat dan menelusuri langkah penyelesaian. Prosedur tersebut

membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai analisis butir untuk pemetaan kelas, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan analisis butir untuk pemetaan kelas melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan analisis butir untuk pemetaan kelas;
- Bandingkan bukti analisis butir untuk pemetaan kelas pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek analisis butir untuk pemetaan kelas dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian analisis butir untuk pemetaan kelas, topik kesetimbangan dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa reaksi dianggap berhenti saat setimbang. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta hubungan grafik laju, konsentrasi, dan model partikel. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa kesetimbangan bersifat dinamis dengan laju maju dan balik yang sama. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengubah data respons menjadi gambaran kebutuhan individual sekaligus pola pembelajaran tingkat kelas, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, analisis butir untuk pemetaan kelas harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.8.14 Heatmap konsepsi dan matriks peserta didik-konsep

Dalam pembelajaran kimia, heatmap konsepsi dan matriks peserta didik-konsep memiliki arti strategis karena banyak jawaban dapat diperoleh melalui hafalan prosedur atau pengenalan pola. Peserta didik mungkin tampak berhasil tetapi tidak mampu menerjemahkan makna simbol, model, atau besaran yang digunakan. Prinsip diagnosis menuntut bukti tentang cara konsep dipakai, bukan hanya apakah hasil akhirnya sesuai kunci (Anggraena et al., 2025; Hoe & Subramaniam, 2016).

Secara operasional, guru dapat menggunakan tes empat tingkat untuk menelusuri konsistensi lintas representasi. Data tersebut perlu dibaca bersama urutan langkah penyelesaian, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mengabaikan ketidakpastian pengukuran. Untuk menguranginya, guru dapat meminta translasi antarrepresentasi dan menggunakan triangulasi bukti. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai heatmap konsepsi dan matriks peserta didik-konsep, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan heatmap konsepsi dan matriks peserta didik-konsep melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan heatmap konsepsi dan matriks peserta didik-konsep;
- Bandingkan bukti heatmap konsepsi dan matriks peserta didik-konsep pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek heatmap konsepsi dan matriks peserta didik-konsep dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian heatmap konsepsi dan matriks peserta didik-konsep, topik termokimia dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa pemutusan ikatan dianggap melepaskan energi. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta analisis energi yang diserap saat memutus dan dilepas saat membentuk ikatan. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa perubahan entalpi bersih berasal dari selisih proses pemutusan dan pembentukan. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengubah data respons menjadi gambaran kebutuhan individual sekaligus pola pembelajaran tingkat kelas, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, heatmap konsepsi dan matriks peserta didik-konsep harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.8.15 Penggunaan peta konsep kelas

Aspek penggunaan peta konsep kelas menegaskan bahwa diagnosis adalah kegiatan ilmiah berskala kelas: guru merumuskan pertanyaan, mengumpulkan bukti, mempertimbangkan

penjelasan alternatif, lalu menentukan kesimpulan yang paling masuk akal. Inferensi harus proporsional terhadap kualitas data. Semakin spesifik label yang diberikan, semakin kuat pula bukti yang diperlukan (Hoe & Subramaniam, 2016; Treagust, 1988).

Secara operasional, guru dapat menggunakan think-aloud untuk menelusuri kemampuan transfer. Data tersebut perlu dibaca bersama kualitas model mental yang dieksternalisasi, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menyimpulkan miskonsepsi hanya dari satu jawaban. Untuk mengurangnya, guru dapat mengulang diagnosis setelah remediasi dan menyusun tindak lanjut berbeda untuk kategori respons yang berbeda. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai penggunaan peta konsep kelas, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan penggunaan peta konsep kelas melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan penggunaan peta konsep kelas;
- Bandingkan bukti penggunaan peta konsep kelas pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek penggunaan peta konsep kelas dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian penggunaan peta konsep kelas, topik hidrokarbon dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa isomer dipandang sebagai zat yang sama karena rumus molekulnya sama. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta hubungan struktur, konektivitas, dan sifat. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa rumus molekul tidak cukup menentukan struktur dan identitas senyawa. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengubah data respons menjadi gambaran kebutuhan individual sekaligus pola pembelajaran tingkat kelas, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, penggunaan peta konsep kelas harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.8.16 Wawancara sampel untuk konfirmasi kelas

Penerapan wawancara sampel untuk konfirmasi kelas harus menghubungkan teori pengukuran dengan pengetahuan pedagogis kimia. Instrumen yang teknis tetapi tidak memahami struktur konsep akan menghasilkan kategori dangkal; sebaliknya, pengetahuan konten tanpa prosedur pengumpulan bukti dapat menghasilkan kesimpulan intuitif yang sulit diuji. Keduanya perlu dipadukan dalam satu alur yang transparan (Treagust, 1988; Taber, 2002a).

Secara operasional, guru dapat menggunakan predict-observe-explain untuk menelusuri stabilitas setelah jeda waktu. Data tersebut perlu dibaca bersama respons terhadap probing, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mencampur kesalahan matematika dengan kesalahan konsep. Untuk mengurangnya, guru dapat mencatat derajat keyakinan dan membandingkan jawaban dengan alasan. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai wawancara sampel untuk konfirmasi kelas, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan wawancara sampel untuk konfirmasi kelas melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan wawancara sampel untuk konfirmasi kelas;
- Bandingkan bukti wawancara sampel untuk konfirmasi kelas pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek wawancara sampel untuk konfirmasi kelas dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian wawancara sampel untuk konfirmasi kelas, topik spektroskopi dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa spektrum dipandang sebagai gambar langsung molekul. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta hubungan sinyal, transisi energi, dan model struktur. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa spektrum adalah data yang ditafsirkan melalui model dan aturan seleksi. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengubah data respons menjadi gambaran kebutuhan individual sekaligus pola pembelajaran tingkat kelas, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Kerangka operasional

Aspek	Bukti yang dicari	Risiko interpretasi	Tindak lanjut
Target konsep	Penjelasan tentang spektroskopi	Target terlalu luas	Batasi subkonsep dan prasyarat
Respons	Stabilitas setelah jeda waktu	Mencampur kesalahan matematika dengan kesalahan konsep	Mencatat derajat keyakinan
Penalaran	Respons terhadap probing	Alasan diperlakukan sekadar tambahan	Membandingkan jawaban dengan alasan
Konfirmasi	Predict-observe-explain atau tugas isomorfik	Kesimpulan dibuat dari satu konteks	Gunakan bukti kedua dan rediagnosis

Implikasinya, wawancara sampel untuk konfirmasi kelas harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.8.17 Pemilihan kasus tipikal dan kasus menyimpang

Pemilihan kasus tipikal dan kasus menyimpang tidak berdiri sebagai prosedur administratif. Ia menjadi bagian dari keputusan tentang apa yang perlu dijelaskan ulang, representasi apa yang perlu dibandingkan, dan siapa yang memerlukan dukungan berbeda. Oleh sebab itu, bukti diagnosis harus langsung terhubung dengan tindakan pembelajaran, bukan berhenti pada daftar kesalahan (Taber, 2002a; Anggraena et al., 2025).

Secara operasional, guru dapat menggunakan tugas isomorfik untuk menelusuri urutan langkah penyelesaian. Data tersebut perlu dibaca bersama ketepatan pilihan, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menggunakan ambang kategori tanpa bukti. Untuk mengurangnya, guru dapat menelusuri langkah penyelesaian dan menggunakan pertanyaan tindak lanjut. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai pemilihan kasus tipikal dan kasus menyimpang, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan pemilihan kasus tipikal dan kasus menyimpang melalui penggunaan konsep, bukan sekadar pengenalan istilah;

- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan pemilihan kasus tipikal dan kasus menyimpang;
- Bandingkan bukti pemilihan kasus tipikal dan kasus menyimpang pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek pemilihan kasus tipikal dan kasus menyimpang dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian pemilihan kasus tipikal dan kasus menyimpang, topik ikatan kovalen dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa garis ikatan dipandang sebagai batang fisik. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta penjelasan energi potensial dan distribusi elektron pada jarak antarinti berbeda. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa representasi garis adalah konvensi yang merangkum pasangan elektron dan konektivitas. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengubah data respons menjadi gambaran kebutuhan individual sekaligus pola pembelajaran tingkat kelas, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, pemilihan kasus tipikal dan kasus menyimpang harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.8.18 Tindak lanjut individual

Tindak lanjut individual perlu ditempatkan dalam kerangka diagnosis yang mengubah data respons menjadi gambaran kebutuhan individual sekaligus pola pembelajaran tingkat kelas. Fokusnya bukan sekadar menghasilkan skor, melainkan memperoleh bukti yang cukup untuk menjelaskan bagaimana peserta didik membangun, menghubungkan, dan menggunakan konsep. Karena konsepsi tidak dapat diamati secara langsung, guru harus menarik inferensi dari respons verbal, tulisan, gambar, tindakan, dan perubahan strategi yang tampak selama penyelesaian tugas (Anggraena et al., 2025; Hoe & Subramaniam, 2016).

Secara operasional, guru dapat menggunakan kuis formatif digital untuk menelusuri kualitas model mental yang dieksternalisasi. Data tersebut perlu dibaca bersama koherensi alasan, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah memperlakukan konsepsi sebagai sifat permanen. Untuk mengurangnya, guru dapat menggunakan triangulasi bukti dan menyajikan tugas paralel.

Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai tindak lanjut individual, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan tindak lanjut individual melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan tindak lanjut individual;
- Bandingkan bukti tindak lanjut individual pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek tindak lanjut individual dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian tindak lanjut individual, topik gas dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa partikel gas dianggap mengembang ketika dipanaskan. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta gambar partikel sebelum dan sesudah pemanasan pada kondisi berbeda. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa ukuran partikel tidak berubah secara bermakna; distribusi kecepatan dan jarak berubah. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengubah data respons menjadi gambaran kebutuhan individual sekaligus pola pembelajaran tingkat kelas, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, tindak lanjut individual harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.8.19 Tindak lanjut kelompok kecil

Pembahasan tentang tindak lanjut kelompok kecil berangkat dari prinsip bahwa data asesmen selalu merupakan jejak tidak langsung dari proses berpikir. Satu pilihan jawaban hanya menunjukkan produk akhir, sedangkan struktur pengetahuan yang melahirkannya dapat berbeda. Diagnosis yang kuat karena itu menggabungkan respons terhadap probing, tingkat keyakinan, dan bukti kontekstual sehingga keputusan pedagogis tidak bergantung pada satu indikator saja (Hoe & Subramaniam, 2016; Treagust, 1988).

Secara operasional, guru dapat menggunakan tes tiga tingkat untuk menelusuri respons terhadap probing. Data tersebut perlu dibaca bersama tingkat keyakinan, karena pola yang

sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mengabaikan ketidakpastian pengukuran. Untuk menguranginya, guru dapat menyusun tindak lanjut berbeda untuk kategori respons yang berbeda dan memeriksa konsep prasyarat. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai tindak lanjut kelompok kecil, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan tindak lanjut kelompok kecil melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan tindak lanjut kelompok kecil;
- Bandingkan bukti tindak lanjut kelompok kecil pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek tindak lanjut kelompok kecil dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian tindak lanjut kelompok kecil, topik kesetimbangan dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa reaksi dianggap berhenti saat setimbang. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta hubungan grafik laju, konsentrasi, dan model partikel. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa kesetimbangan bersifat dinamis dengan laju maju dan balik yang sama. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengubah data respons menjadi gambaran kebutuhan individual sekaligus pola pembelajaran tingkat kelas, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, tindak lanjut kelompok kecil harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.8.20 Tindak lanjut pembelajaran klasikal

Dalam pembelajaran kimia, tindak lanjut pembelajaran klasikal memiliki arti strategis karena banyak jawaban dapat diperoleh melalui hafalan prosedur atau pengenalan pola. Peserta didik mungkin tampak berhasil tetapi tidak mampu menerjemahkan makna simbol, model, atau besaran yang digunakan. Prinsip diagnosis menuntut bukti tentang cara konsep dipakai, bukan hanya apakah hasil akhirnya sesuai kunci (Treagust, 1988; Taber, 2002a).

Secara operasional, guru dapat menggunakan wawancara semi-terstruktur untuk menelusuri ketepatan pilihan. Data tersebut perlu dibaca bersama konsistensi lintas butir, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menyimpulkan miskonsepsi hanya dari satu jawaban. Untuk mengurangnya, guru dapat membandingkan jawaban dengan alasan dan meminta translasi antarrepresentasi. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai tindak lanjut pembelajaran klasikal, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan tindak lanjut pembelajaran klasikal melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan tindak lanjut pembelajaran klasikal;
- Bandingkan bukti tindak lanjut pembelajaran klasikal pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek tindak lanjut pembelajaran klasikal dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian tindak lanjut pembelajaran klasikal, topik termokimia dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa pemutusan ikatan dianggap melepaskan energi. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta analisis energi yang diserap saat memutus dan dilepas saat membentuk ikatan. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa perubahan entalpi bersih berasal dari selisih proses pemutusan dan pembentukan. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengubah data respons menjadi gambaran kebutuhan individual sekaligus pola pembelajaran tingkat kelas, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Kerangka operasional

Aspek	Bukti yang dicari	Risiko interpretasi	Tindak lanjut
Target konsep	Penjelasan tentang termokimia	Target terlalu luas	Batasi subkonsep dan prasyarat
Respons	Ketepatan pilihan	Menyimpulkan miskonsepsi hanya dari satu jawaban	Membandingkan jawaban dengan alasan
Penalaran	Konsistensi lintas butir	Alasan diperlakukan sekadar tambahan	Meminta translasi antarrepresentasi

Aspek	Bukti yang dicari	Risiko interpretasi	Tindak lanjut
Konfirmasi	Wawancara semi-terstruktur atau tugas isomorfik	Kesimpulan dibuat dari satu konteks	Gunakan bukti kedua dan rediagnosis

Implikasinya, tindak lanjut pembelajaran klasikal harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.8.21 Pengelolaan data diagnosis pada kelas besar

Aspek pengelolaan data diagnosis pada kelas besar menegaskan bahwa diagnosis adalah kegiatan ilmiah berskala kelas: guru merumuskan pertanyaan, mengumpulkan bukti, mempertimbangkan penjelasan alternatif, lalu menentukan kesimpulan yang paling masuk akal. Inferensi harus proporsional terhadap kualitas data. Semakin spesifik label yang diberikan, semakin kuat pula bukti yang diperlukan (Taber, 2002a; Anggraena et al., 2025).

Secara operasional, guru dapat menggunakan gambar partikel untuk menelusuri koherensi alasan. Data tersebut perlu dibaca bersama konsistensi lintas representasi, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mencampur kesalahan matematika dengan kesalahan konsep. Untuk mengurangnya, guru dapat menggunakan pertanyaan tindak lanjut dan mengulang diagnosis setelah remediasi. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai pengelolaan data diagnosis pada kelas besar, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan pengelolaan data diagnosis pada kelas besar melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan pengelolaan data diagnosis pada kelas besar;
- Bandingkan bukti pengelolaan data diagnosis pada kelas besar pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek pengelolaan data diagnosis pada kelas besar dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian pengelolaan data diagnosis pada kelas besar, topik hidrokarbon dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa isomer dipandang sebagai zat yang sama karena rumus molekulnya sama. Diagnosis

tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta hubungan struktur, konektivitas, dan sifat. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa rumus molekul tidak cukup menentukan struktur dan identitas senyawa. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengubah data respons menjadi gambaran kebutuhan individual sekaligus pola pembelajaran tingkat kelas, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, pengelolaan data diagnosis pada kelas besar harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.8.22 Komunikasi hasil tanpa mempermalukan peserta didik

Penerapan komunikasi hasil tanpa mempermalukan peserta didik harus menghubungkan teori pengukuran dengan pengetahuan pedagogis kimia. Instrumen yang teknis tetapi tidak memahami struktur konsep akan menghasilkan kategori dangkal; sebaliknya, pengetahuan konten tanpa prosedur pengumpulan bukti dapat menghasilkan kesimpulan intuitif yang sulit diuji. Keduanya perlu dipadukan dalam satu alur yang transparan (Anggraena et al., 2025; Hoe & Subramaniam, 2016).

Secara operasional, guru dapat menggunakan observasi praktikum untuk menelusuri tingkat keyakinan. Data tersebut perlu dibaca bersama kemampuan transfer, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menggunakan ambang kategori tanpa bukti. Untuk mengurangnya, guru dapat menyajikan tugas paralel dan mencatat derajat keyakinan. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai komunikasi hasil tanpa mempermalukan peserta didik, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan komunikasi hasil tanpa mempermalukan peserta didik melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan komunikasi hasil tanpa mempermalukan peserta didik;
- Bandingkan bukti komunikasi hasil tanpa mempermalukan peserta didik pada lebih dari satu konteks, format tugas, atau bentuk representasi;

- Hubungkan kategori hasil pada aspek komunikasi hasil tanpa mempermalukan peserta didik dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian komunikasi hasil tanpa mempermalukan peserta didik, topik spektroskopi dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa spektrum dipandang sebagai gambar langsung molekul. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta hubungan sinyal, transisi energi, dan model struktur. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa spektrum adalah data yang ditafsirkan melalui model dan aturan seleksi. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengubah data respons menjadi gambaran kebutuhan individual sekaligus pola pembelajaran tingkat kelas, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, komunikasi hasil tanpa mempermalukan peserta didik harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.8.23 Monitoring perkembangan individual dan kelas

Monitoring perkembangan individual dan kelas tidak berdiri sebagai prosedur administratif. Ia menjadi bagian dari keputusan tentang apa yang perlu dijelaskan ulang, representasi apa yang perlu dibandingkan, dan siapa yang memerlukan dukungan berbeda. Oleh sebab itu, bukti diagnosis harus langsung terhubung dengan tindakan pembelajaran, bukan berhenti pada daftar kesalahan (Hoe & Subramaniam, 2016; Treagust, 1988).

Secara operasional, guru dapat menggunakan exit ticket untuk menelusuri konsistensi lintas butir. Data tersebut perlu dibaca bersama stabilitas setelah jeda waktu, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah memperlakukan konsepsi sebagai sifat permanen. Untuk mengurangnya, guru dapat memeriksa konsep prasyarat dan menelusuri langkah penyelesaian. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai monitoring perkembangan individual dan kelas, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan monitoring perkembangan individual dan kelas melalui penggunaan konsep, bukan sekadar pengenalan istilah;

- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan monitoring perkembangan individual dan kelas;
- Bandingkan bukti monitoring perkembangan individual dan kelas pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek monitoring perkembangan individual dan kelas dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian monitoring perkembangan individual dan kelas, topik ikatan kovalen dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa garis ikatan dipandang sebagai batang fisik. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta penjelasan energi potensial dan distribusi elektron pada jarak antarinti berbeda. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa representasi garis adalah konvensi yang merangkum pasangan elektron dan konektivitas. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengubah data respons menjadi gambaran kebutuhan individual sekaligus pola pembelajaran tingkat kelas, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, monitoring perkembangan individual dan kelas harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.8.24 Contoh dashboard diagnosis sederhana

Contoh dashboard diagnosis sederhana perlu ditempatkan dalam kerangka diagnosis yang mengubah data respons menjadi gambaran kebutuhan individual sekaligus pola pembelajaran tingkat kelas. Fokusnya bukan sekadar menghasilkan skor, melainkan memperoleh bukti yang cukup untuk menjelaskan bagaimana peserta didik membangun, menghubungkan, dan menggunakan konsep. Karena konsepsi tidak dapat diamati secara langsung, guru harus menarik inferensi dari respons verbal, tulisan, gambar, tindakan, dan perubahan strategi yang tampak selama penyelesaian tugas (Treagust, 1988; Taber, 2002a).

Secara operasional, guru dapat menggunakan tes pilihan ganda dua tingkat untuk menelusuri konsistensi lintas representasi. Data tersebut perlu dibaca bersama urutan langkah penyelesaian, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mengabaikan ketidakpastian pengukuran. Untuk menguranginya, guru dapat meminta translasi antarrepresentasi dan menggunakan

triangulasi bukti. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai contoh dashboard diagnosis sederhana, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan contoh dashboard diagnosis sederhana melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan contoh dashboard diagnosis sederhana;
- Bandingkan bukti contoh dashboard diagnosis sederhana pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek contoh dashboard diagnosis sederhana dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian contoh dashboard diagnosis sederhana, topik gas dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa partikel gas dianggap mengembang ketika dipanaskan. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta gambar partikel sebelum dan sesudah pemanasan pada kondisi berbeda. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa ukuran partikel tidak berubah secara bermakna; distribusi kecepatan dan jarak berubah. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengubah data respons menjadi gambaran kebutuhan individual sekaligus pola pembelajaran tingkat kelas, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Kerangka operasional

Aspek	Bukti yang dicari	Risiko interpretasi	Tindak lanjut
Target konsep	Penjelasan tentang gas	Target terlalu luas	Batasi subkonsep dan prasyarat
Respons	Konsistensi lintas representasi	Mengabaikan ketidakpastian pengukuran	Meminta translasi antarrepresentasi
Penalaran	Urutan langkah penyelesaian	Alasan diperlakukan sekadar tambahan	Menggunakan triangulasi bukti
Konfirmasi	Tes pilihan ganda dua tingkat atau tugas isomorfik	Kesimpulan dibuat dari satu konteks	Gunakan bukti kedua dan rediagnosis

Implikasinya, contoh dashboard diagnosis sederhana harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.8.25 Sintesis diagnosis individual dan kelas

Pembahasan tentang sintesis diagnosis individual dan kelas berangkat dari prinsip bahwa data asesmen selalu merupakan jejak tidak langsung dari proses berpikir. Satu pilihan jawaban hanya menunjukkan produk akhir, sedangkan struktur pengetahuan yang melahirkannya dapat berbeda. Diagnosis yang kuat karena itu menggabungkan kemampuan transfer, kualitas model mental yang dieksternalisasi, dan bukti kontekstual sehingga keputusan pedagogis tidak bergantung pada satu indikator saja (Taber, 2002a; Anggraena et al., 2025).

Secara operasional, guru dapat menggunakan alasan terbuka untuk menelusuri kemampuan transfer. Data tersebut perlu dibaca bersama kualitas model mental yang dieksternalisasi, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menyimpulkan miskonsepsi hanya dari satu jawaban. Untuk mengurangnya, guru dapat mengulang diagnosis setelah remediasi dan menyusun tindak lanjut berbeda untuk kategori respons yang berbeda. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai sintesis diagnosis individual dan kelas, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan sintesis diagnosis individual dan kelas melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan sintesis diagnosis individual dan kelas;
- Bandingkan bukti sintesis diagnosis individual dan kelas pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek sintesis diagnosis individual dan kelas dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian sintesis diagnosis individual dan kelas, topik kesetimbangan dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa reaksi dianggap berhenti saat setimbang. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta hubungan grafik laju, konsentrasi, dan model partikel. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa kesetimbangan bersifat dinamis dengan laju maju dan balik yang sama. Kualitas bukti tersebut

kemudian ditafsirkan sesuai fokus mengubah data respons menjadi gambaran kebutuhan individual sekaligus pola pembelajaran tingkat kelas, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, sintesis diagnosis individual dan kelas harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.8.26 Implikasi praktis bagi guru kimia

Implikasi praktis subbab 5.8 dapat dirangkum sebagai alur: tentukan konsep target, pilih bukti yang relevan, kumpulkan respons dalam lebih dari satu bentuk, analisis pola, nyatakan derajat kepastian, susun tindak lanjut, dan lakukan rediagnosis. Alur ini menjaga agar diagnosis individual dan diagnosis kelas tidak berubah menjadi administrasi skor, tetapi berfungsi sebagai mekanisme perbaikan pembelajaran yang berulang.

50. Identifikasi konsep inti dan konsep prasyarat sebelum menyusun instrumen.
51. Gunakan pertanyaan yang meminta alasan, representasi, atau langkah kerja.
52. Bandingkan respons pada beberapa konteks dan waktu.
53. Pisahkan bukti faktual dari interpretasi guru.
54. Tentukan tindak lanjut yang berbeda untuk kategori kebutuhan yang berbeda.
55. Simpan data secukupnya dan lindungi kerahasiaan peserta didik.
56. Periksa kembali perubahan melalui tugas transfer dan retensi.

5.9 Diagnosis Berbasis Representasi

Subbab ini menilai kemampuan peserta didik menghubungkan fenomena, model partikel, simbol, grafik, dan matematika. Uraian disusun dari prinsip teoretis menuju prosedur kelas, contoh kimia, risiko interpretasi, dan tindak lanjut. Setiap bagian menekankan bahwa kesimpulan diagnostik harus didasarkan pada pola bukti, bukan pada satu respons yang terpisah.

5.9.1 Representasi sebagai bagian dari pengetahuan kimia

Pembahasan tentang representasi sebagai bagian dari pengetahuan kimia berangkat dari prinsip bahwa data asesmen selalu merupakan jejak tidak langsung dari proses berpikir. Satu pilihan jawaban hanya menunjukkan produk akhir, sedangkan struktur pengetahuan yang melahirkannya dapat berbeda. Diagnosis yang kuat karena itu menggabungkan koherensi

alasan, konsistensi lintas representasi, dan bukti kontekstual sehingga keputusan pedagogis tidak bergantung pada satu indikator saja (Gilbert & Treagust, 2009; Ainsworth, 2006).

Secara operasional, guru dapat menggunakan tes empat tingkat untuk menelusuri koherensi alasan. Data tersebut perlu dibaca bersama konsistensi lintas representasi, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mencampur kesalahan matematika dengan kesalahan konsep. Untuk mengurangnya, guru dapat menggunakan pertanyaan tindak lanjut dan mengulang diagnosis setelah remediasi. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai representasi sebagai bagian dari pengetahuan kimia, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan representasi sebagai bagian dari pengetahuan kimia melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan representasi sebagai bagian dari pengetahuan kimia;
- Bandingkan bukti representasi sebagai bagian dari pengetahuan kimia pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek representasi sebagai bagian dari pengetahuan kimia dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian representasi sebagai bagian dari pengetahuan kimia, topik asas Le Chatelier dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa setiap gangguan dianggap otomatis menggeser sistem ke sisi yang mengurangi gangguan tanpa analisis. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta penggunaan kuosien reaksi dan tetapan kesetimbangan. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa prediksi pergeseran harus terkait perubahan Q terhadap K dan kondisi sistem. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menilai kemampuan peserta didik menghubungkan fenomena, model partikel, simbol, grafik, dan matematika, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, representasi sebagai bagian dari pengetahuan kimia harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.9.2 Tingkat makroskopik dalam diagnosis

Dalam pembelajaran kimia, tingkat makroskopik dalam diagnosis memiliki arti strategis karena banyak jawaban dapat diperoleh melalui hafalan prosedur atau pengenalan pola. Peserta didik mungkin tampak berhasil tetapi tidak mampu menerjemahkan makna simbol, model, atau besaran yang digunakan. Prinsip diagnosis menuntut bukti tentang cara konsep dipakai, bukan hanya apakah hasil akhirnya sesuai kunci (Ainsworth, 2006; Chandrasegaran et al., 2007).

Secara operasional, guru dapat menggunakan think-aloud untuk menelusuri tingkat keyakinan. Data tersebut perlu dibaca bersama kemampuan transfer, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menggunakan ambang kategori tanpa bukti. Untuk mengurangnya, guru dapat menyajikan tugas paralel dan mencatat derajat keyakinan. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai tingkat makroskopik dalam diagnosis, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan tingkat makroskopik dalam diagnosis melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan tingkat makroskopik dalam diagnosis;
- Bandingkan bukti tingkat makroskopik dalam diagnosis pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek tingkat makroskopik dalam diagnosis dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian tingkat makroskopik dalam diagnosis, topik kalor dan suhu dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa kalor dipandang sebagai zat yang tersimpan. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta penjelasan perpindahan energi dan perubahan suhu. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa kalor adalah energi dalam transit akibat perbedaan suhu. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menilai kemampuan peserta didik menghubungkan fenomena, model partikel, simbol, grafik, dan matematika, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, tingkat makroskopik dalam diagnosis harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta

rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.9.3 Tingkat submikroskopik dalam diagnosis

Aspek tingkat submikroskopik dalam diagnosis menegaskan bahwa diagnosis adalah kegiatan ilmiah berskala kelas: guru merumuskan pertanyaan, mengumpulkan bukti, mempertimbangkan penjelasan alternatif, lalu menentukan kesimpulan yang paling masuk akal. Inferensi harus proporsional terhadap kualitas data. Semakin spesifik label yang diberikan, semakin kuat pula bukti yang diperlukan (Chandrasegaran et al., 2007; Johnstone, 2000).

Secara operasional, guru dapat menggunakan predict-observe-explain untuk menelusuri konsistensi lintas butir. Data tersebut perlu dibaca bersama stabilitas setelah jeda waktu, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah memperlakukan konsepsi sebagai sifat permanen. Untuk mengurangnya, guru dapat memeriksa konsep prasyarat dan menelusuri langkah penyelesaian. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai tingkat submikroskopik dalam diagnosis, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan tingkat submikroskopik dalam diagnosis melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan tingkat submikroskopik dalam diagnosis;
- Bandingkan bukti tingkat submikroskopik dalam diagnosis pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek tingkat submikroskopik dalam diagnosis dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian tingkat submikroskopik dalam diagnosis, topik polimer dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa polimer dipandang sekadar molekul besar tanpa kaitan struktur-sifat. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta analisis monomer, ikatan silang, kristalinitas, dan sifat mekanik. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa sifat polimer muncul dari struktur pada beberapa skala. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menilai kemampuan peserta didik menghubungkan fenomena, model partikel, simbol, grafik, dan matematika, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban

setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, tingkat submikroskopik dalam diagnosis harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.9.4 Tingkat simbolik dalam diagnosis

Penerapan tingkat simbolik dalam diagnosis harus menghubungkan teori pengukuran dengan pengetahuan pedagogis kimia. Instrumen yang teknis tetapi tidak memahami struktur konsep akan menghasilkan kategori dangkal; sebaliknya, pengetahuan konten tanpa prosedur pengumpulan bukti dapat menghasilkan kesimpulan intuitif yang sulit diuji. Keduanya perlu dipadukan dalam satu alur yang transparan (Johnstone, 2000; Gilbert & Treagust, 2009).

Secara operasional, guru dapat menggunakan tugas isomorfik untuk menelusuri konsistensi lintas representasi. Data tersebut perlu dibaca bersama urutan langkah penyelesaian, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mengabaikan ketidakpastian pengukuran. Untuk menguranginya, guru dapat meminta translasi antarrepresentasi dan menggunakan triangulasi bukti. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai tingkat simbolik dalam diagnosis, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan tingkat simbolik dalam diagnosis melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan tingkat simbolik dalam diagnosis;
- Bandingkan bukti tingkat simbolik dalam diagnosis pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek tingkat simbolik dalam diagnosis dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian tingkat simbolik dalam diagnosis, topik kimia lingkungan dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa zat alami dianggap selalu aman dan zat sintesis selalu berbahaya. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta penilaian berdasarkan dosis, paparan, sifat, dan konteks. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa risiko kimia bergantung pada bahaya, dosis, jalur, dan durasi paparan. Kualitas bukti tersebut

kemudian ditafsirkan sesuai fokus menilai kemampuan peserta didik menghubungkan fenomena, model partikel, simbol, grafik, dan matematika, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Kerangka operasional

Aspek	Bukti yang dicari	Risiko interpretasi	Tindak lanjut
Target konsep	Penjelasan tentang kimia lingkungan	Target terlalu luas	Batasi subkonsep dan prasyarat
Respons	Konsistensi lintas representasi	Mengabaikan ketidakpastian pengukuran	Meminta translasi antarrepresentasi
Penalaran	Urutan langkah penyelesaian	Alasan diperlakukan sekadar tambahan	Menggunakan triangulasi bukti
Konfirmasi	Tugas isomorfik atau tugas isomorfik	Kesimpulan dibuat dari satu konteks	Gunakan bukti kedua dan rediagnosis

Implikasinya, tingkat simbolik dalam diagnosis harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.9.5 Representasi matematis dan kuantitatif

Representasi matematis dan kuantitatif tidak berdiri sebagai prosedur administratif. Ia menjadi bagian dari keputusan tentang apa yang perlu dijelaskan ulang, representasi apa yang perlu dibandingkan, dan siapa yang memerlukan dukungan berbeda. Oleh sebab itu, bukti diagnosis harus langsung terhubung dengan tindakan pembelajaran, bukan berhenti pada daftar kesalahan (Gilbert & Treagust, 2009; Ainsworth, 2006).

Secara operasional, guru dapat menggunakan kuis formatif digital untuk menelusuri kemampuan transfer. Data tersebut perlu dibaca bersama kualitas model mental yang dieksternalisasi, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menyimpulkan miskonsepsi hanya dari satu jawaban. Untuk mengurangnya, guru dapat mengulang diagnosis setelah remediasi dan menyusun tindak lanjut berbeda untuk kategori respons yang berbeda. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai representasi matematis dan kuantitatif, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan representasi matematis dan kuantitatif melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan representasi matematis dan kuantitatif;
- Bandingkan bukti representasi matematis dan kuantitatif pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek representasi matematis dan kuantitatif dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian representasi matematis dan kuantitatif, topik gaya antarmolekul dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa pendidihan dianggap memutus ikatan kovalen dalam molekul. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta peserta didik membedakan perubahan intramolekul dan antarmolekul. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa perubahan fase terutama mengubah organisasi dan interaksi antarmolekul. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menilai kemampuan peserta didik menghubungkan fenomena, model partikel, simbol, grafik, dan matematika, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, representasi matematis dan kuantitatif harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.9.6 Representasi grafis dan diagram

Representasi grafis dan diagram perlu ditempatkan dalam kerangka diagnosis yang menilai kemampuan peserta didik menghubungkan fenomena, model partikel, simbol, grafik, dan matematika. Fokusnya bukan sekadar menghasilkan skor, melainkan memperoleh bukti yang cukup untuk menjelaskan bagaimana peserta didik membangun, menghubungkan, dan menggunakan konsep. Karena konsepsi tidak dapat diamati secara langsung, guru harus menarik inferensi dari respons verbal, tulisan, gambar, tindakan, dan perubahan strategi yang tampak selama penyelesaian tugas (Ainsworth, 2006; Chandrasegaran et al., 2007).

Secara operasional, guru dapat menggunakan tes tiga tingkat untuk menelusuri stabilitas setelah jeda waktu. Data tersebut perlu dibaca bersama respons terhadap probing, karena

pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mencampur kesalahan matematika dengan kesalahan konsep. Untuk menguranginya, guru dapat mencatat derajat keyakinan dan membandingkan jawaban dengan alasan. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai representasi grafis dan diagram, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan representasi grafis dan diagram melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan representasi grafis dan diagram;
- Bandingkan bukti representasi grafis dan diagram pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek representasi grafis dan diagram dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian representasi grafis dan diagram, topik larutan dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa zat terlarut dianggap menghilang ketika larut. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta representasi makro, partikel, dan simbolik proses pelarutan. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa pelarutan mempertahankan materi sambil mengubah interaksi dan penyebaran partikel. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menilai kemampuan peserta didik menghubungkan fenomena, model partikel, simbol, grafik, dan matematika, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, representasi grafis dan diagram harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.9.7 Representasi verbal dan bahasa ilmiah

Pembahasan tentang representasi verbal dan bahasa ilmiah berangkat dari prinsip bahwa data asesmen selalu merupakan jejak tidak langsung dari proses berpikir. Satu pilihan jawaban hanya menunjukkan produk akhir, sedangkan struktur pengetahuan yang melahirkannya dapat berbeda. Diagnosis yang kuat karena itu menggabungkan urutan langkah penyelesaian,

ketepatan pilihan, dan bukti kontekstual sehingga keputusan pedagogis tidak bergantung pada satu indikator saja (Chandrasegaran et al., 2007; Johnstone, 2000).

Secara operasional, guru dapat menggunakan wawancara semi-terstruktur untuk menelusuri urutan langkah penyelesaian. Data tersebut perlu dibaca bersama ketepatan pilihan, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menggunakan ambang kategori tanpa bukti. Untuk mengurangnya, guru dapat menelusuri langkah penyelesaian dan menggunakan pertanyaan tindak lanjut. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai representasi verbal dan bahasa ilmiah, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan representasi verbal dan bahasa ilmiah melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan representasi verbal dan bahasa ilmiah;
- Bandingkan bukti representasi verbal dan bahasa ilmiah pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek representasi verbal dan bahasa ilmiah dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian representasi verbal dan bahasa ilmiah, topik asas Le Chatelier dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa setiap gangguan dianggap otomatis menggeser sistem ke sisi yang mengurangi gangguan tanpa analisis. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta penggunaan kuosien reaksi dan tetapan kesetimbangan. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa prediksi pergeseran harus terkait perubahan Q terhadap K dan kondisi sistem. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menilai kemampuan peserta didik menghubungkan fenomena, model partikel, simbol, grafik, dan matematika, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, representasi verbal dan bahasa ilmiah harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.9.8 Koordinasi antarrepresentasi

Dalam pembelajaran kimia, koordinasi antarrepresentasi memiliki arti strategis karena banyak jawaban dapat diperoleh melalui hafalan prosedur atau pengenalan pola. Peserta didik mungkin tampak berhasil tetapi tidak mampu menerjemahkan makna simbol, model, atau besaran yang digunakan. Prinsip diagnosis menuntut bukti tentang cara konsep dipakai, bukan hanya apakah hasil akhirnya sesuai kunci (Johnstone, 2000; Gilbert & Treagust, 2009).

Secara operasional, guru dapat menggunakan gambar partikel untuk menelusuri kualitas model mental yang dieksternalisasi. Data tersebut perlu dibaca bersama koherensi alasan, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah memperlakukan konsepsi sebagai sifat permanen. Untuk mengurangnya, guru dapat menggunakan triangulasi bukti dan menyajikan tugas paralel. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai koordinasi antarrepresentasi, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan koordinasi antarrepresentasi melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan koordinasi antarrepresentasi;
- Bandingkan bukti koordinasi antarrepresentasi pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek koordinasi antarrepresentasi dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian koordinasi antarrepresentasi, topik kalor dan suhu dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa kalor dipandang sebagai zat yang tersimpan. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta penjelasan perpindahan energi dan perubahan suhu. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa kalor adalah energi dalam transit akibat perbedaan suhu. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menilai kemampuan peserta didik menghubungkan fenomena, model partikel, simbol, grafik, dan matematika, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Kerangka operasional

Aspek	Bukti yang dicari	Risiko interpretasi	Tindak lanjut
Target konsep	Penjelasan tentang kalor dan suhu	Target terlalu luas	Batasi subkonsep dan prasyarat
Respons	Kualitas model mental yang dieksternalisasi	Memperlakukan konsepsi sebagai sifat permanen	Menggunakan triangulasi bukti
Penalaran	Koherensi alasan	Alasan diperlakukan sekadar tambahan	Menyajikan tugas paralel
Konfirmasi	Gambar partikel atau tugas isomorfik	Kesimpulan dibuat dari satu konteks	Gunakan bukti kedua dan rediagnosis

Implikasinya, koordinasi antarrepresentasi harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.9.9 Translasi dari fenomena ke model partikel

Aspek translasi dari fenomena ke model partikel menegaskan bahwa diagnosis adalah kegiatan ilmiah berskala kelas: guru merumuskan pertanyaan, mengumpulkan bukti, mempertimbangkan penjelasan alternatif, lalu menentukan kesimpulan yang paling masuk akal. Inferensi harus proporsional terhadap kualitas data. Semakin spesifik label yang diberikan, semakin kuat pula bukti yang diperlukan (Gilbert & Treagust, 2009; Ainsworth, 2006).

Secara operasional, guru dapat menggunakan observasi praktikum untuk menelusuri respons terhadap probing. Data tersebut perlu dibaca bersama tingkat keyakinan, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mengabaikan ketidakpastian pengukuran. Untuk mengurangnya, guru dapat menyusun tindak lanjut berbeda untuk kategori respons yang berbeda dan memeriksa konsep prasyarat. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai translasi dari fenomena ke model partikel, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan translasi dari fenomena ke model partikel melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan translasi dari fenomena ke model partikel;

- Bandingkan bukti translasi dari fenomena ke model partikel pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek translasi dari fenomena ke model partikel dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian translasi dari fenomena ke model partikel, topik polimer dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa polimer dipandang sekadar molekul besar tanpa kaitan struktur-sifat. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta analisis monomer, ikatan silang, kristalinitas, dan sifat mekanik. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa sifat polimer muncul dari struktur pada beberapa skala. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menilai kemampuan peserta didik menghubungkan fenomena, model partikel, simbol, grafik, dan matematika, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, translasi dari fenomena ke model partikel harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.9.10 Translasi dari model partikel ke persamaan

Penerapan translasi dari model partikel ke persamaan harus menghubungkan teori pengukuran dengan pengetahuan pedagogis kimia. Instrumen yang teknis tetapi tidak memahami struktur konsep akan menghasilkan kategori dangkal; sebaliknya, pengetahuan konten tanpa prosedur pengumpulan bukti dapat menghasilkan kesimpulan intuitif yang sulit diuji. Keduanya perlu dipadukan dalam satu alur yang transparan (Ainsworth, 2006; Chandrasegaran et al., 2007).

Secara operasional, guru dapat menggunakan exit ticket untuk menelusuri ketepatan pilihan. Data tersebut perlu dibaca bersama konsistensi lintas butir, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menyimpulkan miskonsepsi hanya dari satu jawaban. Untuk menguranginya, guru dapat membandingkan jawaban dengan alasan dan meminta translasi antarrepresentasi. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai translasi dari model partikel ke persamaan, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;

- Pilih tugas yang merepresentasikan translasi dari model partikel ke persamaan melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan translasi dari model partikel ke persamaan;
- Bandingkan bukti translasi dari model partikel ke persamaan pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek translasi dari model partikel ke persamaan dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian translasi dari model partikel ke persamaan, topik kimia lingkungan dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa zat alami dianggap selalu aman dan zat sintetis selalu berbahaya. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta penilaian berdasarkan dosis, paparan, sifat, dan konteks. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa risiko kimia bergantung pada bahaya, dosis, jalur, dan durasi paparan. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menilai kemampuan peserta didik menghubungkan fenomena, model partikel, simbol, grafik, dan matematika, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, translasi dari model partikel ke persamaan harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.9.11 Translasi dari persamaan ke makna kuantitatif

Translasi dari persamaan ke makna kuantitatif tidak berdiri sebagai prosedur administratif. Ia menjadi bagian dari keputusan tentang apa yang perlu dijelaskan ulang, representasi apa yang perlu dibandingkan, dan siapa yang memerlukan dukungan berbeda. Oleh sebab itu, bukti diagnosis harus langsung terhubung dengan tindakan pembelajaran, bukan berhenti pada daftar kesalahan (Chandrasegaran et al., 2007; Johnstone, 2000).

Secara operasional, guru dapat menggunakan tes pilihan ganda dua tingkat untuk menelusuri koherensi alasan. Data tersebut perlu dibaca bersama konsistensi lintas representasi, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mencampur kesalahan matematika dengan kesalahan konsep. Untuk menguranginya, guru dapat menggunakan pertanyaan tindak lanjut dan mengulang diagnosis

setelah remediasi. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai translasi dari persamaan ke makna kuantitatif, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan translasi dari persamaan ke makna kuantitatif melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan translasi dari persamaan ke makna kuantitatif;
- Bandingkan bukti translasi dari persamaan ke makna kuantitatif pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek translasi dari persamaan ke makna kuantitatif dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian translasi dari persamaan ke makna kuantitatif, topik gaya antarmolekul dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa pendidihan dianggap memutus ikatan kovalen dalam molekul. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta peserta didik membedakan perubahan intramolekul dan antarmolekul. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa perubahan fase terutama mengubah organisasi dan interaksi antarmolekul. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menilai kemampuan peserta didik menghubungkan fenomena, model partikel, simbol, grafik, dan matematika, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, translasi dari persamaan ke makna kuantitatif harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.9.12 Konsistensi representasional sebagai bukti pemahaman

Konsistensi representasional sebagai bukti pemahaman perlu ditempatkan dalam kerangka diagnosis yang menilai kemampuan peserta didik menghubungkan fenomena, model partikel, simbol, grafik, dan matematika. Fokusnya bukan sekadar menghasilkan skor, melainkan memperoleh bukti yang cukup untuk menjelaskan bagaimana peserta didik membangun, menghubungkan, dan menggunakan konsep. Karena konsepsi tidak dapat diamati secara langsung, guru harus menarik inferensi dari respons verbal, tulisan, gambar, tindakan, dan

perubahan strategi yang tampak selama penyelesaian tugas (Johnstone, 2000; Gilbert & Treagust, 2009).

Secara operasional, guru dapat menggunakan alasan terbuka untuk menelusuri tingkat keyakinan. Data tersebut perlu dibaca bersama kemampuan transfer, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menggunakan ambang kategori tanpa bukti. Untuk menguranginya, guru dapat menyajikan tugas paralel dan mencatat derajat keyakinan. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai konsistensi representasional sebagai bukti pemahaman, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan konsistensi representasional sebagai bukti pemahaman melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan konsistensi representasional sebagai bukti pemahaman;
- Bandingkan bukti konsistensi representasional sebagai bukti pemahaman pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek konsistensi representasional sebagai bukti pemahaman dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian konsistensi representasional sebagai bukti pemahaman, topik larutan dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa zat terlarut dianggap menghilang ketika larut. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta representasi makro, partikel, dan simbolik proses pelarutan. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa pelarutan mempertahankan materi sambil mengubah interaksi dan penyebaran partikel. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menilai kemampuan peserta didik menghubungkan fenomena, model partikel, simbol, grafik, dan matematika, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Kerangka operasional

Aspek	Bukti yang dicari	Risiko interpretasi	Tindak lanjut
Target konsep	Penjelasan tentang larutan	Target terlalu luas	Batasi subkonsep dan prasyarat
Respons	Tingkat keyakinan	Menggunakan ambang kategori tanpa bukti	Menyajikan tugas paralel

Aspek	Bukti yang dicari	Risiko interpretasi	Tindak lanjut
Penalaran	Kemampuan transfer	Alasan diperlakukan sekadar tambahan	Mencatat derajat keyakinan
Konfirmasi	Alasan terbuka atau tugas isomorfik	Kesimpulan dibuat dari satu konteks	Gunakan bukti kedua dan rediagnosis

Implikasinya, konsistensi representasional sebagai bukti pemahaman harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.9.13 Miskonsepsi akibat pencampuran tingkat representasi

Pembahasan tentang miskonsepsi akibat pencampuran tingkat representasi berangkat dari prinsip bahwa data asesmen selalu merupakan jejak tidak langsung dari proses berpikir. Satu pilihan jawaban hanya menunjukkan produk akhir, sedangkan struktur pengetahuan yang melahirkannya dapat berbeda. Diagnosis yang kuat karena itu menggabungkan konsistensi lintas butir, stabilitas setelah jeda waktu, dan bukti kontekstual sehingga keputusan pedagogis tidak bergantung pada satu indikator saja (Gilbert & Treagust, 2009; Ainsworth, 2006).

Secara operasional, guru dapat menggunakan peta konsep untuk menelusuri konsistensi lintas butir. Data tersebut perlu dibaca bersama stabilitas setelah jeda waktu, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah memperlakukan konsepsi sebagai sifat permanen. Untuk mengurangnya, guru dapat memeriksa konsep prasyarat dan menelusuri langkah penyelesaian. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai miskonsepsi akibat pencampuran tingkat representasi, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan miskonsepsi akibat pencampuran tingkat representasi melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan miskonsepsi akibat pencampuran tingkat representasi;
- Bandingkan bukti miskonsepsi akibat pencampuran tingkat representasi pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek miskonsepsi akibat pencampuran tingkat representasi dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian miskonsepsi akibat pencampuran tingkat representasi, topik asas Le Chatelier dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa setiap gangguan dianggap otomatis menggeser sistem ke sisi yang mengurangi gangguan tanpa analisis. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta penggunaan kuosien reaksi dan tetapan kesetimbangan. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa prediksi pergeseran harus terkait perubahan Q terhadap K dan kondisi sistem. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menilai kemampuan peserta didik menghubungkan fenomena, model partikel, simbol, grafik, dan matematika, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, miskonsepsi akibat pencampuran tingkat representasi harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.9.14 Gambar partikel sebagai tugas diagnostik

Dalam pembelajaran kimia, gambar partikel sebagai tugas diagnostik memiliki arti strategis karena banyak jawaban dapat diperoleh melalui hafalan prosedur atau pengenalan pola. Peserta didik mungkin tampak berhasil tetapi tidak mampu menerjemahkan makna simbol, model, atau besaran yang digunakan. Prinsip diagnosis menuntut bukti tentang cara konsep dipakai, bukan hanya apakah hasil akhirnya sesuai kunci (Ainsworth, 2006; Chandrasegaran et al., 2007).

Secara operasional, guru dapat menggunakan analisis respons tertulis untuk menelusuri konsistensi lintas representasi. Data tersebut perlu dibaca bersama urutan langkah penyelesaian, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mengabaikan ketidakpastian pengukuran. Untuk menguranginya, guru dapat meminta translasi antarrepresentasi dan menggunakan triangulasi bukti. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai gambar partikel sebagai tugas diagnostik, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan gambar partikel sebagai tugas diagnostik melalui penggunaan konsep, bukan sekadar pengenalan istilah;

- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan gambar partikel sebagai tugas diagnostik;
- Bandingkan bukti gambar partikel sebagai tugas diagnostik pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek gambar partikel sebagai tugas diagnostik dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian gambar partikel sebagai tugas diagnostik, topik kalor dan suhu dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa kalor dipandang sebagai zat yang tersimpan. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta penjelasan perpindahan energi dan perubahan suhu. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa kalor adalah energi dalam transit akibat perbedaan suhu. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menilai kemampuan peserta didik menghubungkan fenomena, model partikel, simbol, grafik, dan matematika, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, gambar partikel sebagai tugas diagnostik harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.9.15 Grafik sebagai tugas diagnostik

Aspek grafik sebagai tugas diagnostik menegaskan bahwa diagnosis adalah kegiatan ilmiah berskala kelas: guru merumuskan pertanyaan, mengumpulkan bukti, mempertimbangkan penjelasan alternatif, lalu menentukan kesimpulan yang paling masuk akal. Inferensi harus proporsional terhadap kualitas data. Semakin spesifik label yang diberikan, semakin kuat pula bukti yang diperlukan (Chandrasegaran et al., 2007; Johnstone, 2000).

Secara operasional, guru dapat menggunakan diskusi argumentatif untuk menelusuri kemampuan transfer. Data tersebut perlu dibaca bersama kualitas model mental yang dieksternalisasi, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menyimpulkan miskonsepsi hanya dari satu jawaban. Untuk mengurangnya, guru dapat mengulang diagnosis setelah remediasi dan menyusun tindak lanjut berbeda untuk kategori respons yang berbeda. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai grafik sebagai tugas diagnostik, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan grafik sebagai tugas diagnostik melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan grafik sebagai tugas diagnostik;
- Bandingkan bukti grafik sebagai tugas diagnostik pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek grafik sebagai tugas diagnostik dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian grafik sebagai tugas diagnostik, topik polimer dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa polimer dipandang sekadar molekul besar tanpa kaitan struktur-sifat. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta analisis monomer, ikatan silang, kristalinitas, dan sifat mekanik. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa sifat polimer muncul dari struktur pada beberapa skala. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menilai kemampuan peserta didik menghubungkan fenomena, model partikel, simbol, grafik, dan matematika, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, grafik sebagai tugas diagnostik harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.9.16 Model molekul dan keterbatasannya

Penerapan model molekul dan keterbatasannya harus menghubungkan teori pengukuran dengan pengetahuan pedagogis kimia. Instrumen yang teknis tetapi tidak memahami struktur konsep akan menghasilkan kategori dangkal; sebaliknya, pengetahuan konten tanpa prosedur pengumpulan bukti dapat menghasilkan kesimpulan intuitif yang sulit diuji. Keduanya perlu dipadukan dalam satu alur yang transparan (Johnstone, 2000; Gilbert & Treagust, 2009).

Secara operasional, guru dapat menggunakan analisis grafik untuk menelusuri stabilitas setelah jeda waktu. Data tersebut perlu dibaca bersama respons terhadap probing, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mencampur kesalahan matematika dengan kesalahan konsep. Untuk menguranginya, guru dapat mencatat derajat keyakinan dan membandingkan jawaban

dengan alasan. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai model molekul dan keterbatasannya, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan model molekul dan keterbatasannya melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan model molekul dan keterbatasannya;
- Bandingkan bukti model molekul dan keterbatasannya pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek model molekul dan keterbatasannya dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian model molekul dan keterbatasannya, topik kimia lingkungan dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa zat alami dianggap selalu aman dan zat sintetis selalu berbahaya. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta penilaian berdasarkan dosis, paparan, sifat, dan konteks. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa risiko kimia bergantung pada bahaya, dosis, jalur, dan durasi paparan. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menilai kemampuan peserta didik menghubungkan fenomena, model partikel, simbol, grafik, dan matematika, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Kerangka operasional

Aspek	Bukti yang dicari	Risiko interpretasi	Tindak lanjut
Target konsep	Penjelasan tentang kimia lingkungan	Target terlalu luas	Batasi subkonsep dan prasyarat
Respons	Stabilitas setelah jeda waktu	Mencampur kesalahan matematika dengan kesalahan konsep	Mencatat derajat keyakinan
Penalaran	Respons terhadap probing	Alasan diperlakukan sekadar tambahan	Membandingkan jawaban dengan alasan
Konfirmasi	Analisis grafik atau tugas isomorfik	Kesimpulan dibuat dari satu konteks	Gunakan bukti kedua dan rediagnosis

Implikasinya, model molekul dan keterbatasannya harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.9.17 Animasi dan simulasi dalam diagnosis

Animasi dan simulasi dalam diagnosis tidak berdiri sebagai prosedur administratif. Ia menjadi bagian dari keputusan tentang apa yang perlu dijelaskan ulang, representasi apa yang perlu dibandingkan, dan siapa yang memerlukan dukungan berbeda. Oleh sebab itu, bukti diagnosis harus langsung terhubung dengan tindakan pembelajaran, bukan berhenti pada daftar kesalahan (Gilbert & Treagust, 2009; Ainsworth, 2006).

Secara operasional, guru dapat menggunakan tes empat tingkat untuk menelusuri urutan langkah penyelesaian. Data tersebut perlu dibaca bersama ketepatan pilihan, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menggunakan ambang kategori tanpa bukti. Untuk mengurangnya, guru dapat menelusuri langkah penyelesaian dan menggunakan pertanyaan tindak lanjut. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai animasi dan simulasi dalam diagnosis, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan animasi dan simulasi dalam diagnosis melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan animasi dan simulasi dalam diagnosis;
- Bandingkan bukti animasi dan simulasi dalam diagnosis pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek animasi dan simulasi dalam diagnosis dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian animasi dan simulasi dalam diagnosis, topik gaya antarmolekul dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa pendidihan dianggap memutus ikatan kovalen dalam molekul. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta peserta didik membedakan perubahan intramolekul dan antarmolekul. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa perubahan fase terutama mengubah organisasi dan interaksi antarmolekul. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menilai kemampuan peserta didik menghubungkan fenomena, model partikel, simbol, grafik, dan

matematika, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, animasi dan simulasi dalam diagnosis harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.9.18 Tugas multiple external representations

Tugas multiple external representations perlu ditempatkan dalam kerangka diagnosis yang menilai kemampuan peserta didik menghubungkan fenomena, model partikel, simbol, grafik, dan matematika. Fokusnya bukan sekadar menghasilkan skor, melainkan memperoleh bukti yang cukup untuk menjelaskan bagaimana peserta didik membangun, menghubungkan, dan menggunakan konsep. Karena konsepsi tidak dapat diamati secara langsung, guru harus menarik inferensi dari respons verbal, tulisan, gambar, tindakan, dan perubahan strategi yang tampak selama penyelesaian tugas (Ainsworth, 2006; Chandrasegaran et al., 2007).

Secara operasional, guru dapat menggunakan think-aloud untuk menelusuri kualitas model mental yang dieksternalisasi. Data tersebut perlu dibaca bersama koherensi alasan, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah memperlakukan konsepsi sebagai sifat permanen. Untuk menguranginya, guru dapat menggunakan triangulasi bukti dan menyajikan tugas paralel. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai tugas multiple external representations, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan tugas multiple external representations melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan tugas multiple external representations;
- Bandingkan bukti tugas multiple external representations pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek tugas multiple external representations dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian tugas multiple external representations, topik larutan dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa zat terlarut dianggap menghilang ketika larut. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi

konseptual yang tepat. Guru dapat meminta representasi makro, partikel, dan simbolik proses pelarutan. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa pelarutan mempertahankan materi sambil mengubah interaksi dan penyebaran partikel. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menilai kemampuan peserta didik menghubungkan fenomena, model partikel, simbol, grafik, dan matematika, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, tugas multiple external representations harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.9.19 Representational competence dan representational fluency

Pembahasan tentang representational competence dan representational fluency berangkat dari prinsip bahwa data asesmen selalu merupakan jejak tidak langsung dari proses berpikir. Satu pilihan jawaban hanya menunjukkan produk akhir, sedangkan struktur pengetahuan yang melahirkannya dapat berbeda. Diagnosis yang kuat karena itu menggabungkan respons terhadap probing, tingkat keyakinan, dan bukti kontekstual sehingga keputusan pedagogis tidak bergantung pada satu indikator saja (Chandrasegaran et al., 2007; Johnstone, 2000).

Secara operasional, guru dapat menggunakan predict-observe-explain untuk menelusuri respons terhadap probing. Data tersebut perlu dibaca bersama tingkat keyakinan, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mengabaikan ketidakpastian pengukuran. Untuk mengurangnya, guru dapat menyusun tindak lanjut berbeda untuk kategori respons yang berbeda dan memeriksa konsep prasyarat. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai representational competence dan representational fluency, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan representational competence dan representational fluency melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan representational competence dan representational fluency;
- Bandingkan bukti representational competence dan representational fluency pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek representational competence dan representational fluency dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian representational competence dan representational fluency, topik asas Le Chatelier dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsep yang sering muncul ialah bahwa setiap gangguan dianggap otomatis menggeser sistem ke sisi yang mengurangi gangguan tanpa analisis. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta penggunaan kuosien reaksi dan tetapan kesetimbangan. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa prediksi pergeseran harus terkait perubahan Q terhadap K dan kondisi sistem. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menilai kemampuan peserta didik menghubungkan fenomena, model partikel, simbol, grafik, dan matematika, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, representational competence dan representational fluency harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.9.20 Rubrik penilaian representasi

Dalam pembelajaran kimia, rubrik penilaian representasi memiliki arti strategis karena banyak jawaban dapat diperoleh melalui hafalan prosedur atau pengenalan pola. Peserta didik mungkin tampak berhasil tetapi tidak mampu menerjemahkan makna simbol, model, atau besaran yang digunakan. Prinsip diagnosis menuntut bukti tentang cara konsep dipakai, bukan hanya apakah hasil akhirnya sesuai kunci (Johnstone, 2000; Gilbert & Treagust, 2009).

Secara operasional, guru dapat menggunakan tugas isomorfik untuk menelusuri ketepatan pilihan. Data tersebut perlu dibaca bersama konsistensi lintas butir, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menyimpulkan miskonsepsi hanya dari satu jawaban. Untuk menguranginya, guru dapat membandingkan jawaban dengan alasan dan meminta translasi antarrepresentasi. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai rubrik penilaian representasi, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan rubrik penilaian representasi melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan rubrik penilaian representasi;

- Bandingkan bukti rubrik penilaian representasi pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek rubrik penilaian representasi dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian rubrik penilaian representasi, topik kalor dan suhu dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa kalor dipandang sebagai zat yang tersimpan. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta penjelasan perpindahan energi dan perubahan suhu. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa kalor adalah energi dalam transit akibat perbedaan suhu. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menilai kemampuan peserta didik menghubungkan fenomena, model partikel, simbol, grafik, dan matematika, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Kerangka operasional

Aspek	Bukti yang dicari	Risiko interpretasi	Tindak lanjut
Target konsep	Penjelasan tentang kalor dan suhu	Target terlalu luas	Batasi subkonsep dan prasyarat
Respons	Ketepatan pilihan	Menyimpulkan miskonsepsi hanya dari satu jawaban	Membandingkan jawaban dengan alasan
Penalaran	Konsistensi lintas butir	Alasan diperlakukan sekadar tambahan	Meminta translasi antarrepresentasi
Konfirmasi	Tugas isomorfik atau tugas isomorfik	Kesimpulan dibuat dari satu konteks	Gunakan bukti kedua dan rediagnosis

Implikasinya, rubrik penilaian representasi harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.9.21 Validitas representasi diagnostik

Aspek validitas representasi diagnostik menegaskan bahwa diagnosis adalah kegiatan ilmiah berskala kelas: guru merumuskan pertanyaan, mengumpulkan bukti, mempertimbangkan penjelasan alternatif, lalu menentukan kesimpulan yang paling masuk akal. Inferensi harus

proporsional terhadap kualitas data. Semakin spesifik label yang diberikan, semakin kuat pula bukti yang diperlukan (Gilbert & Treagust, 2009; Ainsworth, 2006).

Secara operasional, guru dapat menggunakan kuis formatif digital untuk menelusuri koherensi alasan. Data tersebut perlu dibaca bersama konsistensi lintas representasi, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mencampur kesalahan matematika dengan kesalahan konsep. Untuk mengurangnya, guru dapat menggunakan pertanyaan tindak lanjut dan mengulang diagnosis setelah remediasi. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai validitas representasi diagnostik, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan validitas representasi diagnostik melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan validitas representasi diagnostik;
- Bandingkan bukti validitas representasi diagnostik pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek validitas representasi diagnostik dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian validitas representasi diagnostik, topik polimer dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa polimer dipandang sekadar molekul besar tanpa kaitan struktur-sifat. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta analisis monomer, ikatan silang, kristalinitas, dan sifat mekanik. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa sifat polimer muncul dari struktur pada beberapa skala. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menilai kemampuan peserta didik menghubungkan fenomena, model partikel, simbol, grafik, dan matematika, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, validitas representasi diagnostik harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.9.22 Aksesibilitas dan beban visual

Penerapan aksesibilitas dan beban visual harus menghubungkan teori pengukuran dengan pengetahuan pedagogis kimia. Instrumen yang teknis tetapi tidak memahami struktur konsep akan menghasilkan kategori dangkal; sebaliknya, pengetahuan konten tanpa prosedur pengumpulan bukti dapat menghasilkan kesimpulan intuitif yang sulit diuji. Keduanya perlu dipadukan dalam satu alur yang transparan (Ainsworth, 2006; Chandrasegaran et al., 2007).

Secara operasional, guru dapat menggunakan tes tiga tingkat untuk menelusuri tingkat keyakinan. Data tersebut perlu dibaca bersama kemampuan transfer, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menggunakan ambang kategori tanpa bukti. Untuk menguranginya, guru dapat menyajikan tugas paralel dan mencatat derajat keyakinan. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai aksesibilitas dan beban visual, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan aksesibilitas dan beban visual melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan aksesibilitas dan beban visual;
- Bandingkan bukti aksesibilitas dan beban visual pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek aksesibilitas dan beban visual dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian aksesibilitas dan beban visual, topik kimia lingkungan dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa zat alami dianggap selalu aman dan zat sintesis selalu berbahaya. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta penilaian berdasarkan dosis, paparan, sifat, dan konteks. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa risiko kimia bergantung pada bahaya, dosis, jalur, dan durasi paparan. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menilai kemampuan peserta didik menghubungkan fenomena, model partikel, simbol, grafik, dan matematika, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, aksesibilitas dan beban visual harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola

respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.9.23 Contoh diagnosis representasional pada reaksi kimia

Contoh diagnosis representasional pada reaksi kimia tidak berdiri sebagai prosedur administratif. Ia menjadi bagian dari keputusan tentang apa yang perlu dijelaskan ulang, representasi apa yang perlu dibandingkan, dan siapa yang memerlukan dukungan berbeda. Oleh sebab itu, bukti diagnosis harus langsung terhubung dengan tindakan pembelajaran, bukan berhenti pada daftar kesalahan (Chandrasegaran et al., 2007; Johnstone, 2000).

Secara operasional, guru dapat menggunakan wawancara semi-terstruktur untuk menelusuri konsistensi lintas butir. Data tersebut perlu dibaca bersama stabilitas setelah jeda waktu, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah memperlakukan konsepsi sebagai sifat permanen. Untuk menguranginya, guru dapat memeriksa konsep prasyarat dan menelusuri langkah penyelesaian. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai contoh diagnosis representasional pada reaksi kimia, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan contoh diagnosis representasional pada reaksi kimia melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan contoh diagnosis representasional pada reaksi kimia;
- Bandingkan bukti contoh diagnosis representasional pada reaksi kimia pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek contoh diagnosis representasional pada reaksi kimia dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian contoh diagnosis representasional pada reaksi kimia, topik gaya antarmolekul dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa pendidihan dianggap memutus ikatan kovalen dalam molekul. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta peserta didik membedakan perubahan intramolekul dan antarmolekul. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa perubahan fase terutama mengubah organisasi dan interaksi antarmolekul. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menilai kemampuan peserta didik menghubungkan fenomena, model partikel, simbol, grafik, dan matematika, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, contoh diagnosis representasional pada reaksi kimia harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.9.24 Contoh diagnosis representasional pada kesetimbangan

Contoh diagnosis representasional pada kesetimbangan perlu ditempatkan dalam kerangka diagnosis yang menilai kemampuan peserta didik menghubungkan fenomena, model partikel, simbol, grafik, dan matematika. Fokusnya bukan sekadar menghasilkan skor, melainkan memperoleh bukti yang cukup untuk menjelaskan bagaimana peserta didik membangun, menghubungkan, dan menggunakan konsep. Karena konsepsi tidak dapat diamati secara langsung, guru harus menarik inferensi dari respons verbal, tulisan, gambar, tindakan, dan perubahan strategi yang tampak selama penyelesaian tugas (Johnstone, 2000; Gilbert & Treagust, 2009).

Secara operasional, guru dapat menggunakan gambar partikel untuk menelusuri konsistensi lintas representasi. Data tersebut perlu dibaca bersama urutan langkah penyelesaian, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mengabaikan ketidakpastian pengukuran. Untuk menguranginya, guru dapat meminta translasi antarrepresentasi dan menggunakan triangulasi bukti. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai contoh diagnosis representasional pada kesetimbangan, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan contoh diagnosis representasional pada kesetimbangan melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan contoh diagnosis representasional pada kesetimbangan;
- Bandingkan bukti contoh diagnosis representasional pada kesetimbangan pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek contoh diagnosis representasional pada kesetimbangan dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian contoh diagnosis representasional pada kesetimbangan, topik larutan dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa zat terlarut dianggap menghilang ketika larut. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta representasi makro, partikel, dan simbolik proses pelarutan. Bukti yang diharapkan adalah kemampuan peserta didik

menjelaskan bahwa pelarutan mempertahankan materi sambil mengubah interaksi dan penyebaran partikel. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menilai kemampuan peserta didik menghubungkan fenomena, model partikel, simbol, grafik, dan matematika, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Kerangka operasional

Aspek	Bukti yang dicari	Risiko interpretasi	Tindak lanjut
Target konsep	Penjelasan tentang larutan	Target terlalu luas	Batasi subkonsep dan prasyarat
Respons	Konsistensi lintas representasi	Mengabaikan ketidakpastian pengukuran	Meminta translasi antarrepresentasi
Penalaran	Urutan langkah penyelesaian	Alasan diperlakukan sekadar tambahan	Menggunakan triangulasi bukti
Konfirmasi	Gambar partikel atau tugas isomorfik	Kesimpulan dibuat dari satu konteks	Gunakan bukti kedua dan rediagnosis

Implikasinya, contoh diagnosis representasional pada kesetimbangan harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.9.25 Sintesis diagnosis berbasis representasi

Pembahasan tentang sintesis diagnosis berbasis representasi berangkat dari prinsip bahwa data asesmen selalu merupakan jejak tidak langsung dari proses berpikir. Satu pilihan jawaban hanya menunjukkan produk akhir, sedangkan struktur pengetahuan yang melahirkannya dapat berbeda. Diagnosis yang kuat karena itu menggabungkan kemampuan transfer, kualitas model mental yang dieksternalisasi, dan bukti kontekstual sehingga keputusan pedagogis tidak bergantung pada satu indikator saja (Gilbert & Treagust, 2009; Ainsworth, 2006).

Secara operasional, guru dapat menggunakan observasi praktikum untuk menelusuri kemampuan transfer. Data tersebut perlu dibaca bersama kualitas model mental yang dieksternalisasi, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menyimpulkan miskonsepsi hanya dari satu jawaban. Untuk mengurangnya, guru dapat mengulang diagnosis setelah remediasi dan menyusun tindak lanjut berbeda untuk kategori respons yang berbeda. Prosedur tersebut membantu

memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai sintesis diagnosis berbasis representasi, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan sintesis diagnosis berbasis representasi melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan sintesis diagnosis berbasis representasi;
- Bandingkan bukti sintesis diagnosis berbasis representasi pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek sintesis diagnosis berbasis representasi dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian sintesis diagnosis berbasis representasi, topik asas Le Chatelier dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa setiap gangguan dianggap otomatis menggeser sistem ke sisi yang mengurangi gangguan tanpa analisis. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta penggunaan kuosien reaksi dan tetapan kesetimbangan. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa prediksi pergeseran harus terkait perubahan Q terhadap K dan kondisi sistem. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menilai kemampuan peserta didik menghubungkan fenomena, model partikel, simbol, grafik, dan matematika, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, sintesis diagnosis berbasis representasi harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.9.26 Implikasi praktis bagi guru kimia

Implikasi praktis subbab 5.9 dapat dirangkum sebagai alur: tentukan konsep target, pilih bukti yang relevan, kumpulkan respons dalam lebih dari satu bentuk, analisis pola, nyatakan derajat kepastian, susun tindak lanjut, dan lakukan rediagnosis. Alur ini menjaga agar diagnosis berbasis representasi tidak berubah menjadi administrasi skor, tetapi berfungsi sebagai mekanisme perbaikan pembelajaran yang berulang.

57. Identifikasi konsep inti dan konsep prasyarat sebelum menyusun instrumen.

58. Gunakan pertanyaan yang meminta alasan, representasi, atau langkah kerja.
59. Bandingkan respons pada beberapa konteks dan waktu.
60. Pisahkan bukti faktual dari interpretasi guru.
61. Tentukan tindak lanjut yang berbeda untuk kategori kebutuhan yang berbeda.
62. Simpan data secukupnya dan lindungi kerahasiaan peserta didik.
63. Periksa kembali perubahan melalui tugas transfer dan retensi.

5.10 Etika dalam Diagnosis Peserta Didik

Subbab ini menempatkan martabat, keadilan, privasi, dan kemanfaatan peserta didik sebagai batas penggunaan diagnosis. Uraian disusun dari prinsip teoretis menuju prosedur kelas, contoh kimia, risiko interpretasi, dan tindak lanjut. Setiap bagian menekankan bahwa kesimpulan diagnostik harus didasarkan pada pola bukti, bukan pada satu respons yang terpisah.

5.10.1 Diagnosis sebagai praktik pendidikan yang berorientasi kemaslahatan

Pembahasan tentang diagnosis sebagai praktik pendidikan yang berorientasi kemaslahatan berangkat dari prinsip bahwa data asesmen selalu merupakan jejak tidak langsung dari proses berpikir. Satu pilihan jawaban hanya menunjukkan produk akhir, sedangkan struktur pengetahuan yang melahirkannya dapat berbeda. Diagnosis yang kuat karena itu menggabungkan koherensi alasan, konsistensi lintas representasi, dan bukti kontekstual sehingga keputusan pedagogis tidak bergantung pada satu indikator saja (UNESCO IITE, 2020; Indonesia, 2022).

Secara operasional, guru dapat menggunakan peta konsep untuk menelusuri koherensi alasan. Data tersebut perlu dibaca bersama konsistensi lintas representasi, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mencampur kesalahan matematika dengan kesalahan konsep. Untuk mengurangnya, guru dapat menggunakan pertanyaan tindak lanjut dan mengulang diagnosis setelah remediasi. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai diagnosis sebagai praktik pendidikan yang berorientasi kemaslahatan, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan diagnosis sebagai praktik pendidikan yang berorientasi kemaslahatan melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan diagnosis sebagai praktik pendidikan yang berorientasi kemaslahatan;
- Bandingkan bukti diagnosis sebagai praktik pendidikan yang berorientasi kemaslahatan pada lebih dari satu konteks, format tugas, atau bentuk representasi;

- Hubungkan kategori hasil pada aspek diagnosis sebagai praktik pendidikan yang berorientasi kemaslahatan dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian diagnosis sebagai praktik pendidikan yang berorientasi kemaslahatan, topik kelarutan dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa larutan jenuh dianggap tidak mengandung proses pelarutan lagi. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta model partikel pelarutan dan pengendapan pada keadaan jenuh. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa kejenuhan adalah keadaan kesetimbangan dinamis. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menempatkan martabat, keadilan, privasi, dan kemanfaatan peserta didik sebagai batas penggunaan diagnosis, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, diagnosis sebagai praktik pendidikan yang berorientasi kemaslahatan harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.10.2 Prinsip tidak merugikan peserta didik

Dalam pembelajaran kimia, prinsip tidak merugikan peserta didik memiliki arti strategis karena banyak jawaban dapat diperoleh melalui hafalan prosedur atau pengenalan pola. Peserta didik mungkin tampak berhasil tetapi tidak mampu menerjemahkan makna simbol, model, atau besaran yang digunakan. Prinsip diagnosis menuntut bukti tentang cara konsep dipakai, bukan hanya apakah hasil akhirnya sesuai kunci (Indonesia, 2022; UNESCO, 2023).

Secara operasional, guru dapat menggunakan analisis respons tertulis untuk menelusuri tingkat keyakinan. Data tersebut perlu dibaca bersama kemampuan transfer, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menggunakan ambang kategori tanpa bukti. Untuk mengurangnya, guru dapat menyajikan tugas paralel dan mencatat derajat keyakinan. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai prinsip tidak merugikan peserta didik, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan prinsip tidak merugikan peserta didik melalui penggunaan konsep, bukan sekadar pengenalan istilah;

- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan prinsip tidak merugikan peserta didik;
- Bandingkan bukti prinsip tidak merugikan peserta didik pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek prinsip tidak merugikan peserta didik dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian prinsip tidak merugikan peserta didik, topik redoks dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa oksidasi selalu dianggap melibatkan oksigen. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta analisis transfer elektron dan perubahan bilangan oksidasi. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa pengikatan oksigen merupakan kasus khusus dari kerangka redoks yang lebih luas. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menempatkan martabat, keadilan, privasi, dan kemanfaatan peserta didik sebagai batas penggunaan diagnosis, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, prinsip tidak merugikan peserta didik harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.10.3 Menghindari pelabelan permanen

Aspek menghindari pelabelan permanen menegaskan bahwa diagnosis adalah kegiatan ilmiah berskala kelas: guru merumuskan pertanyaan, mengumpulkan bukti, mempertimbangkan penjelasan alternatif, lalu menentukan kesimpulan yang paling masuk akal. Inferensi harus proporsional terhadap kualitas data. Semakin spesifik label yang diberikan, semakin kuat pula bukti yang diperlukan (UNESCO, 2023; AERA et al., 2014).

Secara operasional, guru dapat menggunakan diskusi argumentatif untuk menelusuri konsistensi lintas butir. Data tersebut perlu dibaca bersama stabilitas setelah jeda waktu, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah memperlakukan konsepsi sebagai sifat permanen. Untuk mengurangnya, guru dapat memeriksa konsep prasyarat dan menelusuri langkah penyelesaian. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai menghindari pelabelan permanen, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan menghindari pelabelan permanen melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan menghindari pelabelan permanen;
- Bandingkan bukti menghindari pelabelan permanen pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek menghindari pelabelan permanen dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian menghindari pelabelan permanen, topik kimia organik dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa reaksi dipahami hanya sebagai penggantian simbol tanpa mekanisme. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta penelusuran pusat elektron dan perubahan ikatan. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa mekanisme menjelaskan urutan perpindahan elektron dan perubahan struktur. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menempatkan martabat, keadilan, privasi, dan kemanfaatan peserta didik sebagai batas penggunaan diagnosis, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, menghindari pelabelan permanen harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.10.4 Menghindari bahasa defisit dan stigmatisasi

Penerapan menghindari bahasa defisit dan stigmatisasi harus menghubungkan teori pengukuran dengan pengetahuan pedagogis kimia. Instrumen yang teknis tetapi tidak memahami struktur konsep akan menghasilkan kategori dangkal; sebaliknya, pengetahuan konten tanpa prosedur pengumpulan bukti dapat menghasilkan kesimpulan intuitif yang sulit diuji. Keduanya perlu dipadukan dalam satu alur yang transparan (AERA et al., 2014; UNESCO IITE, 2020).

Secara operasional, guru dapat menggunakan analisis grafik untuk menelusuri konsistensi lintas representasi. Data tersebut perlu dibaca bersama urutan langkah penyelesaian, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada

tahap ini adalah mengabaikan ketidakpastian pengukuran. Untuk menguranginya, guru dapat meminta translasi antarrepresentasi dan menggunakan triangulasi bukti. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai menghindari bahasa defisit dan stigmatisasi, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan menghindari bahasa defisit dan stigmatisasi melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan menghindari bahasa defisit dan stigmatisasi;
- Bandingkan bukti menghindari bahasa defisit dan stigmatisasi pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek menghindari bahasa defisit dan stigmatisasi dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian menghindari bahasa defisit dan stigmatisasi, topik struktur atom dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa elektron dipandang bergerak pada lintasan planet yang tetap. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta peserta didik membandingkan model Bohr dan model orbital serta menjelaskan batas setiap model. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa model atom adalah representasi teoretis dengan domain penggunaan tertentu. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menempatkan martabat, keadilan, privasi, dan kemanfaatan peserta didik sebagai batas penggunaan diagnosis, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Kerangka operasional

Aspek	Bukti yang dicari	Risiko interpretasi	Tindak lanjut
Target konsep	Penjelasan tentang struktur atom	Target terlalu luas	Batasi subkonsep dan prasyarat
Respons	Konsistensi lintas representasi	Mengabaikan ketidakpastian pengukuran	Meminta translasi antarrepresentasi
Penalaran	Urutan langkah penyelesaian	Alasan diperlakukan sekadar tambahan	Menggunakan triangulasi bukti
Konfirmasi	Analisis grafik atau tugas isomorfik	Kesimpulan dibuat dari satu konteks	Gunakan bukti kedua dan rediagnosis

Implikasinya, menghindari bahasa defisit dan stigmatisasi harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.10.5 Keadilan dan kesempatan yang setara

Keadilan dan kesempatan yang setara tidak berdiri sebagai prosedur administratif. Ia menjadi bagian dari keputusan tentang apa yang perlu dijelaskan ulang, representasi apa yang perlu dibandingkan, dan siapa yang memerlukan dukungan berbeda. Oleh sebab itu, bukti diagnosis harus langsung terhubung dengan tindakan pembelajaran, bukan berhenti pada daftar kesalahan (UNESCO IITE, 2020; Indonesia, 2022).

Secara operasional, guru dapat menggunakan tes empat tingkat untuk menelusuri kemampuan transfer. Data tersebut perlu dibaca bersama kualitas model mental yang dieksternalisasi, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menyimpulkan miskonsepsi hanya dari satu jawaban. Untuk mengurangnya, guru dapat mengulang diagnosis setelah remediasi dan menyusun tindak lanjut berbeda untuk kategori respons yang berbeda. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai keadilan dan kesempatan yang setara, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan keadilan dan kesempatan yang setara melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan keadilan dan kesempatan yang setara;
- Bandingkan bukti keadilan dan kesempatan yang setara pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek keadilan dan kesempatan yang setara dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian keadilan dan kesempatan yang setara, topik mol dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa satu mol semua zat dianggap memiliki massa yang sama. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta perbandingan jumlah partikel dan massa satu mol beberapa zat. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa mol menyatakan jumlah entitas, sedangkan massa bergantung pada massa molar.

Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menempatkan martabat, keadilan, privasi, dan kemanfaatan peserta didik sebagai batas penggunaan diagnosis, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, keadilan dan kesempatan yang setara harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.10.6 Persetujuan, pemberitahuan, dan transparansi

Persetujuan, pemberitahuan, dan transparansi perlu ditempatkan dalam kerangka diagnosis yang menempatkan martabat, keadilan, privasi, dan kemanfaatan peserta didik sebagai batas penggunaan diagnosis. Fokusnya bukan sekadar menghasilkan skor, melainkan memperoleh bukti yang cukup untuk menjelaskan bagaimana peserta didik membangun, menghubungkan, dan menggunakan konsep. Karena konsepsi tidak dapat diamati secara langsung, guru harus menarik inferensi dari respons verbal, tulisan, gambar, tindakan, dan perubahan strategi yang tampak selama penyelesaian tugas (Indonesia, 2022; UNESCO, 2023).

Secara operasional, guru dapat menggunakan think-aloud untuk menelusuri stabilitas setelah jeda waktu. Data tersebut perlu dibaca bersama respons terhadap probing, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mencampur kesalahan matematika dengan kesalahan konsep. Untuk mengurangnya, guru dapat mencatat derajat keyakinan dan membandingkan jawaban dengan alasan. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai persetujuan, pemberitahuan, dan transparansi, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan persetujuan, pemberitahuan, dan transparansi melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan persetujuan, pemberitahuan, dan transparansi;
- Bandingkan bukti persetujuan, pemberitahuan, dan transparansi pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek persetujuan, pemberitahuan, dan transparansi dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian persetujuan, pemberitahuan, dan transparansi, topik konsentrasi dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering

muncul ialah bahwa konsentrasi disamakan dengan jumlah zat terlarut tanpa memperhatikan volume. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta perbandingan beberapa larutan dengan jumlah mol dan volume berbeda. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa konsentrasi adalah besaran rasio yang bergantung pada pembilang dan penyebut. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menempatkan martabat, keadilan, privasi, dan kemanfaatan peserta didik sebagai batas penggunaan diagnosis, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, persetujuan, pemberitahuan, dan transparansi harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.10.7 Kerahasiaan hasil diagnosis

Pembahasan tentang kerahasiaan hasil diagnosis berangkat dari prinsip bahwa data asesmen selalu merupakan jejak tidak langsung dari proses berpikir. Satu pilihan jawaban hanya menunjukkan produk akhir, sedangkan struktur pengetahuan yang melahirkannya dapat berbeda. Diagnosis yang kuat karena itu menggabungkan urutan langkah penyelesaian, ketepatan pilihan, dan bukti kontekstual sehingga keputusan pedagogis tidak bergantung pada satu indikator saja (UNESCO, 2023; AERA et al., 2014).

Secara operasional, guru dapat menggunakan predict-observe-explain untuk menelusuri urutan langkah penyelesaian. Data tersebut perlu dibaca bersama ketepatan pilihan, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menggunakan ambang kategori tanpa bukti. Untuk mengurangnya, guru dapat menelusuri langkah penyelesaian dan menggunakan pertanyaan tindak lanjut. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai kerahasiaan hasil diagnosis, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan kerahasiaan hasil diagnosis melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan kerahasiaan hasil diagnosis;

- Bandingkan bukti kerahasiaan hasil diagnosis pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek kerahasiaan hasil diagnosis dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian kerahasiaan hasil diagnosis, topik kelarutan dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa larutan jenuh dianggap tidak mengandung proses pelarutan lagi. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta model partikel pelarutan dan pengendapan pada keadaan jenuh. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa kejenuhan adalah keadaan kesetimbangan dinamis. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menempatkan martabat, keadilan, privasi, dan kemanfaatan peserta didik sebagai batas penggunaan diagnosis, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, kerahasiaan hasil diagnosis harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.10.8 Pelindungan data pribadi peserta didik

Dalam pembelajaran kimia, pelindungan data pribadi peserta didik memiliki arti strategis karena banyak jawaban dapat diperoleh melalui hafalan prosedur atau pengenalan pola. Peserta didik mungkin tampak berhasil tetapi tidak mampu menerjemahkan makna simbol, model, atau besaran yang digunakan. Prinsip diagnosis menuntut bukti tentang cara konsep dipakai, bukan hanya apakah hasil akhirnya sesuai kunci (AERA et al., 2014; UNESCO IITE, 2020).

Secara operasional, guru dapat menggunakan tugas isomorfik untuk menelusuri kualitas model mental yang dieksternalisasi. Data tersebut perlu dibaca bersama koherensi alasan, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah memperlakukan konsepsi sebagai sifat permanen. Untuk mengurangnya, guru dapat menggunakan triangulasi bukti dan menyajikan tugas paralel. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai pelindungan data pribadi peserta didik, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;

- Pilih tugas yang merepresentasikan perlindungan data pribadi peserta didik melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan perlindungan data pribadi peserta didik;
- Bandingkan bukti perlindungan data pribadi peserta didik pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek perlindungan data pribadi peserta didik dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian perlindungan data pribadi peserta didik, topik redoks dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa oksidasi selalu dianggap melibatkan oksigen. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta analisis transfer elektron dan perubahan bilangan oksidasi. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa pengikatan oksigen merupakan kasus khusus dari kerangka redoks yang lebih luas. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menempatkan martabat, keadilan, privasi, dan kemanfaatan peserta didik sebagai batas penggunaan diagnosis, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Kerangka operasional

Aspek	Bukti yang dicari	Risiko interpretasi	Tindak lanjut
Target konsep	Penjelasan tentang redoks	Target terlalu luas	Batasi subkonsep dan prasyarat
Respons	Kualitas model mental yang dieksternalisasi	Memperlakukan konsepsi sebagai sifat permanen	Menggunakan triangulasi bukti
Penalaran	Koherensi alasan	Alasan diperlakukan sekadar tambahan	Menyajikan tugas paralel
Konfirmasi	Tugas isomorfik atau tugas isomorfik	Kesimpulan dibuat dari satu konteks	Gunakan bukti kedua dan rediagnosis

Implikasinya, perlindungan data pribadi peserta didik harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.10.9 Data anak sebagai data yang memerlukan perlindungan khusus

Aspek data anak sebagai data yang memerlukan perlindungan khusus menegaskan bahwa diagnosis adalah kegiatan ilmiah berskala kelas: guru merumuskan pertanyaan, mengumpulkan bukti, mempertimbangkan penjelasan alternatif, lalu menentukan kesimpulan yang paling masuk akal. Inferensi harus proporsional terhadap kualitas data. Semakin spesifik label yang diberikan, semakin kuat pula bukti yang diperlukan (UNESCO IITE, 2020; Indonesia, 2022).

Secara operasional, guru dapat menggunakan kuis formatif digital untuk menelusuri respons terhadap probing. Data tersebut perlu dibaca bersama tingkat keyakinan, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mengabaikan ketidakpastian pengukuran. Untuk mengurangnya, guru dapat menyusun tindak lanjut berbeda untuk kategori respons yang berbeda dan memeriksa konsep prasyarat. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai data anak sebagai data yang memerlukan perlindungan khusus, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan data anak sebagai data yang memerlukan perlindungan khusus melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan data anak sebagai data yang memerlukan perlindungan khusus;
- Bandingkan bukti data anak sebagai data yang memerlukan perlindungan khusus pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek data anak sebagai data yang memerlukan perlindungan khusus dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian data anak sebagai data yang memerlukan perlindungan khusus, topik kimia organik dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa reaksi dipahami hanya sebagai penggantian simbol tanpa mekanisme. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta penelusuran pusat elektron dan perubahan ikatan. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa mekanisme menjelaskan urutan perpindahan elektron dan perubahan struktur. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menempatkan martabat, keadilan, privasi, dan kemanfaatan peserta didik sebagai batas penggunaan diagnosis, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, data anak sebagai data yang memerlukan perlindungan khusus harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat

kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.10.10 Minimalisasi data dan pembatasan tujuan

Penerapan minimalisasi data dan pembatasan tujuan harus menghubungkan teori pengukuran dengan pengetahuan pedagogis kimia. Instrumen yang teknis tetapi tidak memahami struktur konsep akan menghasilkan kategori dangkal; sebaliknya, pengetahuan konten tanpa prosedur pengumpulan bukti dapat menghasilkan kesimpulan intuitif yang sulit diuji. Keduanya perlu dipadukan dalam satu alur yang transparan (Indonesia, 2022; UNESCO, 2023).

Secara operasional, guru dapat menggunakan tes tiga tingkat untuk menelusuri ketepatan pilihan. Data tersebut perlu dibaca bersama konsistensi lintas butir, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menyimpulkan miskonsepsi hanya dari satu jawaban. Untuk mengurangnya, guru dapat membandingkan jawaban dengan alasan dan meminta translasi antarrepresentasi. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai minimalisasi data dan pembatasan tujuan, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan minimalisasi data dan pembatasan tujuan melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan minimalisasi data dan pembatasan tujuan;
- Bandingkan bukti minimalisasi data dan pembatasan tujuan pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek minimalisasi data dan pembatasan tujuan dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian minimalisasi data dan pembatasan tujuan, topik struktur atom dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa elektron dipandang bergerak pada lintasan planet yang tetap. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta peserta didik membandingkan model Bohr dan model orbital serta menjelaskan batas setiap model. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa model atom adalah representasi teoretis dengan domain penggunaan tertentu. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menempatkan martabat, keadilan, privasi, dan kemanfaatan peserta didik sebagai batas penggunaan diagnosis, bukan diperlakukan sebagai skor yang berdiri

sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, minimalisasi data dan pembatasan tujuan harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.10.11 Keamanan penyimpanan data diagnosis

Keamanan penyimpanan data diagnosis tidak berdiri sebagai prosedur administratif. Ia menjadi bagian dari keputusan tentang apa yang perlu dijelaskan ulang, representasi apa yang perlu dibandingkan, dan siapa yang memerlukan dukungan berbeda. Oleh sebab itu, bukti diagnosis harus langsung terhubung dengan tindakan pembelajaran, bukan berhenti pada daftar kesalahan (UNESCO, 2023; AERA et al., 2014).

Secara operasional, guru dapat menggunakan wawancara semi-terstruktur untuk menelusuri koherensi alasan. Data tersebut perlu dibaca bersama konsistensi lintas representasi, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mencampur kesalahan matematika dengan kesalahan konsep. Untuk mengurangnya, guru dapat menggunakan pertanyaan tindak lanjut dan mengulang diagnosis setelah remediasi. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai keamanan penyimpanan data diagnosis, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan keamanan penyimpanan data diagnosis melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan keamanan penyimpanan data diagnosis;
- Bandingkan bukti keamanan penyimpanan data diagnosis pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek keamanan penyimpanan data diagnosis dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian keamanan penyimpanan data diagnosis, topik mol dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa satu mol semua zat dianggap memiliki massa yang sama. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta perbandingan jumlah partikel dan massa satu mol beberapa zat. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan

bahwa mol menyatakan jumlah entitas, sedangkan massa bergantung pada massa molar. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menempatkan martabat, keadilan, privasi, dan kemanfaatan peserta didik sebagai batas penggunaan diagnosis, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, keamanan penyimpanan data diagnosis harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.10.12 Akses terhadap data dan akuntabilitas

Akses terhadap data dan akuntabilitas perlu ditempatkan dalam kerangka diagnosis yang menempatkan martabat, keadilan, privasi, dan kemanfaatan peserta didik sebagai batas penggunaan diagnosis. Fokusnya bukan sekadar menghasilkan skor, melainkan memperoleh bukti yang cukup untuk menjelaskan bagaimana peserta didik membangun, menghubungkan, dan menggunakan konsep. Karena konsepsi tidak dapat diamati secara langsung, guru harus menarik inferensi dari respons verbal, tulisan, gambar, tindakan, dan perubahan strategi yang tampak selama penyelesaian tugas (AERA et al., 2014; UNESCO IITE, 2020).

Secara operasional, guru dapat menggunakan gambar partikel untuk menelusuri tingkat keyakinan. Data tersebut perlu dibaca bersama kemampuan transfer, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menggunakan ambang kategori tanpa bukti. Untuk mengurangnya, guru dapat menyajikan tugas paralel dan mencatat derajat keyakinan. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai akses terhadap data dan akuntabilitas, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan akses terhadap data dan akuntabilitas melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan akses terhadap data dan akuntabilitas;
- Bandingkan bukti akses terhadap data dan akuntabilitas pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek akses terhadap data dan akuntabilitas dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian akses terhadap data dan akuntabilitas, topik konsentrasi dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa konsentrasi disamakan dengan jumlah zat terlarut tanpa memperhatikan volume. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta perbandingan beberapa larutan dengan jumlah mol dan volume berbeda. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa konsentrasi adalah besaran rasio yang bergantung pada pembilang dan penyebut. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menempatkan martabat, keadilan, privasi, dan kemanfaatan peserta didik sebagai batas penggunaan diagnosis, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Kerangka operasional

Aspek	Bukti yang dicari	Risiko interpretasi	Tindak lanjut
Target konsep	Penjelasan tentang konsentrasi	Target terlalu luas	Batasi subkonsep dan prasyarat
Respons	Tingkat keyakinan	Menggunakan ambang kategori tanpa bukti	Menyajikan tugas paralel
Penalaran	Kemampuan transfer	Alasan diperlakukan sekadar tambahan	Mencatat derajat keyakinan
Konfirmasi	Gambar partikel atau tugas isomorfik	Kesimpulan dibuat dari satu konteks	Gunakan bukti kedua dan rediagnosis

Implikasinya, akses terhadap data dan akuntabilitas harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.10.13 Etika penggunaan platform digital

Pembahasan tentang etika penggunaan platform digital berangkat dari prinsip bahwa data asesmen selalu merupakan jejak tidak langsung dari proses berpikir. Satu pilihan jawaban hanya menunjukkan produk akhir, sedangkan struktur pengetahuan yang melahirkannya dapat berbeda. Diagnosis yang kuat karena itu menggabungkan konsistensi lintas butir, stabilitas setelah jeda waktu, dan bukti kontekstual sehingga keputusan pedagogis tidak bergantung pada satu indikator saja (UNESCO IITE, 2020; Indonesia, 2022).

Secara operasional, guru dapat menggunakan observasi praktikum untuk menelusuri konsistensi lintas butir. Data tersebut perlu dibaca bersama stabilitas setelah jeda waktu, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah memperlakukan konsepsi sebagai sifat permanen. Untuk mengurangnya, guru dapat memeriksa konsep prasyarat dan menelusuri langkah penyelesaian. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai etika penggunaan platform digital, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan etika penggunaan platform digital melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan etika penggunaan platform digital;
- Bandingkan bukti etika penggunaan platform digital pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek etika penggunaan platform digital dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian etika penggunaan platform digital, topik kelarutan dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa larutan jenuh dianggap tidak mengandung proses pelarutan lagi. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta model partikel pelarutan dan pengendapan pada keadaan jenuh. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa kejenuhan adalah keadaan kesetimbangan dinamis. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menempatkan martabat, keadilan, privasi, dan kemanfaatan peserta didik sebagai batas penggunaan diagnosis, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, etika penggunaan platform digital harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.10.14 Etika penggunaan kecerdasan artifisial

Dalam pembelajaran kimia, etika penggunaan kecerdasan artifisial memiliki arti strategis karena banyak jawaban dapat diperoleh melalui hafalan prosedur atau pengenalan pola.

Peserta didik mungkin tampak berhasil tetapi tidak mampu menerjemahkan makna simbol, model, atau besaran yang digunakan. Prinsip diagnosis menuntut bukti tentang cara konsep dipakai, bukan hanya apakah hasil akhirnya sesuai kunci (Indonesia, 2022; UNESCO, 2023).

Secara operasional, guru dapat menggunakan exit ticket untuk menelusuri konsistensi lintas representasi. Data tersebut perlu dibaca bersama urutan langkah penyelesaian, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mengabaikan ketidakpastian pengukuran. Untuk menguranginya, guru dapat meminta translasi antarrepresentasi dan menggunakan triangulasi bukti. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai etika penggunaan kecerdasan artifisial, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan etika penggunaan kecerdasan artifisial melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan etika penggunaan kecerdasan artifisial;
- Bandingkan bukti etika penggunaan kecerdasan artifisial pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek etika penggunaan kecerdasan artifisial dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian etika penggunaan kecerdasan artifisial, topik redoks dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa oksidasi selalu dianggap melibatkan oksigen. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta analisis transfer elektron dan perubahan bilangan oksidasi. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa pengikatan oksigen merupakan kasus khusus dari kerangka redoks yang lebih luas. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menempatkan martabat, keadilan, privasi, dan kemanfaatan peserta didik sebagai batas penggunaan diagnosis, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, etika penggunaan kecerdasan artifisial harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.10.15 Bias algoritmik dan keputusan otomatis

Aspek bias algoritmik dan keputusan otomatis menegaskan bahwa diagnosis adalah kegiatan ilmiah berskala kelas: guru merumuskan pertanyaan, mengumpulkan bukti, mempertimbangkan penjelasan alternatif, lalu menentukan kesimpulan yang paling masuk akal. Inferensi harus proporsional terhadap kualitas data. Semakin spesifik label yang diberikan, semakin kuat pula bukti yang diperlukan (UNESCO, 2023; AERA et al., 2014).

Secara operasional, guru dapat menggunakan tes pilihan ganda dua tingkat untuk menelusuri kemampuan transfer. Data tersebut perlu dibaca bersama kualitas model mental yang dieksternalisasi, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menyimpulkan miskonsepsi hanya dari satu jawaban. Untuk mengurangnya, guru dapat mengulang diagnosis setelah remediasi dan menyusun tindak lanjut berbeda untuk kategori respons yang berbeda. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai bias algoritmik dan keputusan otomatis, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan bias algoritmik dan keputusan otomatis melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan bias algoritmik dan keputusan otomatis;
- Bandingkan bukti bias algoritmik dan keputusan otomatis pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek bias algoritmik dan keputusan otomatis dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian bias algoritmik dan keputusan otomatis, topik kimia organik dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa reaksi dipahami hanya sebagai penggantian simbol tanpa mekanisme. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta penelusuran pusat elektron dan perubahan ikatan. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa mekanisme menjelaskan urutan perpindahan elektron dan perubahan struktur. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menempatkan martabat, keadilan, privasi, dan kemanfaatan peserta didik sebagai batas penggunaan diagnosis, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, bias algoritmik dan keputusan otomatis harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta

rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.10.16 Hak peserta didik untuk memperoleh penjelasan

Penerapan hak peserta didik untuk memperoleh penjelasan harus menghubungkan teori pengukuran dengan pengetahuan pedagogis kimia. Instrumen yang teknis tetapi tidak memahami struktur konsep akan menghasilkan kategori dangkal; sebaliknya, pengetahuan konten tanpa prosedur pengumpulan bukti dapat menghasilkan kesimpulan intuitif yang sulit diuji. Keduanya perlu dipadukan dalam satu alur yang transparan (AERA et al., 2014; UNESCO IITE, 2020).

Secara operasional, guru dapat menggunakan alasan terbuka untuk menelusuri stabilitas setelah jeda waktu. Data tersebut perlu dibaca bersama respons terhadap probing, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mencampur kesalahan matematika dengan kesalahan konsep. Untuk mengurangnya, guru dapat mencatat derajat keyakinan dan membandingkan jawaban dengan alasan. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai hak peserta didik untuk memperoleh penjelasan, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan hak peserta didik untuk memperoleh penjelasan melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan hak peserta didik untuk memperoleh penjelasan;
- Bandingkan bukti hak peserta didik untuk memperoleh penjelasan pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek hak peserta didik untuk memperoleh penjelasan dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian hak peserta didik untuk memperoleh penjelasan, topik struktur atom dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa elektron dipandang bergerak pada lintasan planet yang tetap. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta peserta didik membandingkan model Bohr dan model orbital serta menjelaskan batas setiap model. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa model atom adalah representasi teoretis dengan domain penggunaan tertentu. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menempatkan martabat, keadilan, privasi, dan kemanfaatan peserta didik sebagai batas penggunaan diagnosis, bukan diperlakukan sebagai skor yang berdiri

sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Kerangka operasional

Aspek	Bukti yang dicari	Risiko interpretasi	Tindak lanjut
Target konsep	Penjelasan tentang struktur atom	Target terlalu luas	Batasi subkonsep dan prasyarat
Respons	Stabilitas setelah jeda waktu	Mencampur kesalahan matematika dengan kesalahan konsep	Mencatat derajat keyakinan
Penalaran	Respons terhadap probing	Alasan diperlakukan sekadar tambahan	Membandingkan jawaban dengan alasan
Konfirmasi	Alasan terbuka atau tugas isomorfik	Kesimpulan dibuat dari satu konteks	Gunakan bukti kedua dan rediagnosis

Implikasinya, hak peserta didik untuk memperoleh penjelasan harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.10.17 Komunikasi hasil kepada orang tua atau wali

Komunikasi hasil kepada orang tua atau wali tidak berdiri sebagai prosedur administratif. Ia menjadi bagian dari keputusan tentang apa yang perlu dijelaskan ulang, representasi apa yang perlu dibandingkan, dan siapa yang memerlukan dukungan berbeda. Oleh sebab itu, bukti diagnosis harus langsung terhubung dengan tindakan pembelajaran, bukan berhenti pada daftar kesalahan (UNESCO IITE, 2020; Indonesia, 2022).

Secara operasional, guru dapat menggunakan peta konsep untuk menelusuri urutan langkah penyelesaian. Data tersebut perlu dibaca bersama ketepatan pilihan, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menggunakan ambang kategori tanpa bukti. Untuk mengurangnya, guru dapat menelusuri langkah penyelesaian dan menggunakan pertanyaan tindak lanjut. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai komunikasi hasil kepada orang tua atau wali, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;

- Pilih tugas yang merepresentasikan komunikasi hasil kepada orang tua atau wali melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan komunikasi hasil kepada orang tua atau wali;
- Bandingkan bukti komunikasi hasil kepada orang tua atau wali pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek komunikasi hasil kepada orang tua atau wali dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian komunikasi hasil kepada orang tua atau wali, topik mol dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa satu mol semua zat dianggap memiliki massa yang sama. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta perbandingan jumlah partikel dan massa satu mol beberapa zat. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa mol menyatakan jumlah entitas, sedangkan massa bergantung pada massa molar. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menempatkan martabat, keadilan, privasi, dan kemanfaatan peserta didik sebagai batas penggunaan diagnosis, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, komunikasi hasil kepada orang tua atau wali harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.10.18 Penggunaan hasil untuk remediasi, bukan hukuman

Penggunaan hasil untuk remediasi, bukan hukuman perlu ditempatkan dalam kerangka diagnosis yang menempatkan martabat, keadilan, privasi, dan kemanfaatan peserta didik sebagai batas penggunaan diagnosis. Fokusnya bukan sekadar menghasilkan skor, melainkan memperoleh bukti yang cukup untuk menjelaskan bagaimana peserta didik membangun, menghubungkan, dan menggunakan konsep. Karena konsepsi tidak dapat diamati secara langsung, guru harus menarik inferensi dari respons verbal, tulisan, gambar, tindakan, dan perubahan strategi yang tampak selama penyelesaian tugas (Indonesia, 2022; UNESCO, 2023).

Secara operasional, guru dapat menggunakan analisis respons tertulis untuk menelusuri kualitas model mental yang dieksternalisasi. Data tersebut perlu dibaca bersama koherensi alasan, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah memperlakukan konsepsi sebagai sifat permanen. Untuk

mengurangnya, guru dapat menggunakan triangulasi bukti dan menyajikan tugas paralel. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai penggunaan hasil untuk remediasi, bukan hukuman, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan penggunaan hasil untuk remediasi, bukan hukuman melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan penggunaan hasil untuk remediasi, bukan hukuman;
- Bandingkan bukti penggunaan hasil untuk remediasi, bukan hukuman pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek penggunaan hasil untuk remediasi, bukan hukuman dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian penggunaan hasil untuk remediasi, bukan hukuman, topik konsentrasi dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa konsentrasi disamakan dengan jumlah zat terlarut tanpa memperhatikan volume. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta perbandingan beberapa larutan dengan jumlah mol dan volume berbeda. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa konsentrasi adalah besaran rasio yang bergantung pada pembilang dan penyebut. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menempatkan martabat, keadilan, privasi, dan kemanfaatan peserta didik sebagai batas penggunaan diagnosis, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, penggunaan hasil untuk remediasi, bukan hukuman harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.10.19 Etika publikasi dan penelitian kelas

Pembahasan tentang etika publikasi dan penelitian kelas berangkat dari prinsip bahwa data asesmen selalu merupakan jejak tidak langsung dari proses berpikir. Satu pilihan jawaban hanya menunjukkan produk akhir, sedangkan struktur pengetahuan yang melahirkannya dapat berbeda. Diagnosis yang kuat karena itu menggabungkan respons terhadap probing,

tingkat keyakinan, dan bukti kontekstual sehingga keputusan pedagogis tidak bergantung pada satu indikator saja (UNESCO, 2023; AERA et al., 2014).

Secara operasional, guru dapat menggunakan diskusi argumentatif untuk menelusuri respons terhadap probing. Data tersebut perlu dibaca bersama tingkat keyakinan, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mengabaikan ketidakpastian pengukuran. Untuk mengurangnya, guru dapat menyusun tindak lanjut berbeda untuk kategori respons yang berbeda dan memeriksa konsep prasyarat. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai etika publikasi dan penelitian kelas, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan etika publikasi dan penelitian kelas melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan etika publikasi dan penelitian kelas;
- Bandingkan bukti etika publikasi dan penelitian kelas pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek etika publikasi dan penelitian kelas dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian etika publikasi dan penelitian kelas, topik kelarutan dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa larutan jenuh dianggap tidak mengandung proses pelarutan lagi. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta model partikel pelarutan dan pengendapan pada keadaan jenuh. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa kejenuhan adalah keadaan kesetimbangan dinamis. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menempatkan martabat, keadilan, privasi, dan kemanfaatan peserta didik sebagai batas penggunaan diagnosis, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, etika publikasi dan penelitian kelas harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.10.20 Anonimisasi dan pseudonimisasi data

Dalam pembelajaran kimia, anonimisasi dan pseudonimisasi data memiliki arti strategis karena banyak jawaban dapat diperoleh melalui hafalan prosedur atau pengenalan pola. Peserta didik mungkin tampak berhasil tetapi tidak mampu menerjemahkan makna simbol, model, atau besaran yang digunakan. Prinsip diagnosis menuntut bukti tentang cara konsep dipakai, bukan hanya apakah hasil akhirnya sesuai kunci (AERA et al., 2014; UNESCO IITE, 2020).

Secara operasional, guru dapat menggunakan analisis grafik untuk menelusuri ketepatan pilihan. Data tersebut perlu dibaca bersama konsistensi lintas butir, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menyimpulkan miskonsepsi hanya dari satu jawaban. Untuk mengurangnya, guru dapat membandingkan jawaban dengan alasan dan meminta translasi antarrepresentasi. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai anonimisasi dan pseudonimisasi data, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan anonimisasi dan pseudonimisasi data melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan anonimisasi dan pseudonimisasi data;
- Bandingkan bukti anonimisasi dan pseudonimisasi data pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek anonimisasi dan pseudonimisasi data dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian anonimisasi dan pseudonimisasi data, topik redoks dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa oksidasi selalu dianggap melibatkan oksigen. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta analisis transfer elektron dan perubahan bilangan oksidasi. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa pengikatan oksigen merupakan kasus khusus dari kerangka redoks yang lebih luas. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menempatkan martabat, keadilan, privasi, dan kemanfaatan peserta didik sebagai batas penggunaan diagnosis, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Kerangka operasional

Aspek	Bukti yang dicari	Risiko interpretasi	Tindak lanjut
Target konsep	Penjelasan tentang redoks	Target terlalu luas	Batasi subkonsep dan prasyarat
Respons	Ketepatan pilihan	Menyimpulkan miskonsepsi hanya dari satu jawaban	Membandingkan jawaban dengan alasan
Penalaran	Konsistensi lintas butir	Alasan diperlakukan sekadar tambahan	Meminta translasi antarrepresentasi
Konfirmasi	Analisis grafik atau tugas isomorfik	Kesimpulan dibuat dari satu konteks	Gunakan bukti kedua dan rediagnosis

Implikasinya, anonimisasi dan pseudonimisasi data harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.10.21 Keadilan bagi peserta didik berkebutuhan khusus

Aspek keadilan bagi peserta didik berkebutuhan khusus menegaskan bahwa diagnosis adalah kegiatan ilmiah berskala kelas: guru merumuskan pertanyaan, mengumpulkan bukti, mempertimbangkan penjelasan alternatif, lalu menentukan kesimpulan yang paling masuk akal. Inferensi harus proporsional terhadap kualitas data. Semakin spesifik label yang diberikan, semakin kuat pula bukti yang diperlukan (UNESCO IITE, 2020; Indonesia, 2022).

Secara operasional, guru dapat menggunakan tes empat tingkat untuk menelusuri koherensi alasan. Data tersebut perlu dibaca bersama konsistensi lintas representasi, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mencampur kesalahan matematika dengan kesalahan konsep. Untuk mengurangnya, guru dapat menggunakan pertanyaan tindak lanjut dan mengulang diagnosis setelah remediasi. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai keadilan bagi peserta didik berkebutuhan khusus, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan keadilan bagi peserta didik berkebutuhan khusus melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan keadilan bagi peserta didik berkebutuhan khusus;

- Bandingkan bukti keadilan bagi peserta didik berkebutuhan khusus pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek keadilan bagi peserta didik berkebutuhan khusus dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian keadilan bagi peserta didik berkebutuhan khusus, topik kimia organik dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa reaksi dipahami hanya sebagai penggantian simbol tanpa mekanisme. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta penelusuran pusat elektron dan perubahan ikatan. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa mekanisme menjelaskan urutan perpindahan elektron dan perubahan struktur. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menempatkan martabat, keadilan, privasi, dan kemanfaatan peserta didik sebagai batas penggunaan diagnosis, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, keadilan bagi peserta didik berkebutuhan khusus harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.10.22 Sensitivitas bahasa dan budaya

Penerapan sensitivitas bahasa dan budaya harus menghubungkan teori pengukuran dengan pengetahuan pedagogis kimia. Instrumen yang teknis tetapi tidak memahami struktur konsep akan menghasilkan kategori dangkal; sebaliknya, pengetahuan konten tanpa prosedur pengumpulan bukti dapat menghasilkan kesimpulan intuitif yang sulit diuji. Keduanya perlu dipadukan dalam satu alur yang transparan (Indonesia, 2022; UNESCO, 2023).

Secara operasional, guru dapat menggunakan think-aloud untuk menelusuri tingkat keyakinan. Data tersebut perlu dibaca bersama kemampuan transfer, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menggunakan ambang kategori tanpa bukti. Untuk menguranginya, guru dapat menyajikan tugas paralel dan mencatat derajat keyakinan. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai sensitivitas bahasa dan budaya, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;

- Pilih tugas yang merepresentasikan sensitivitas bahasa dan budaya melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan sensitivitas bahasa dan budaya;
- Bandingkan bukti sensitivitas bahasa dan budaya pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek sensitivitas bahasa dan budaya dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian sensitivitas bahasa dan budaya, topik struktur atom dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa elektron dipandang bergerak pada lintasan planet yang tetap. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta peserta didik membandingkan model Bohr dan model orbital serta menjelaskan batas setiap model. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa model atom adalah representasi teoretis dengan domain penggunaan tertentu. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menempatkan martabat, keadilan, privasi, dan kemanfaatan peserta didik sebagai batas penggunaan diagnosis, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, sensitivitas bahasa dan budaya harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.10.23 Konflik peran guru sebagai pengajar dan penilai

Konflik peran guru sebagai pengajar dan penilai tidak berdiri sebagai prosedur administratif. Ia menjadi bagian dari keputusan tentang apa yang perlu dijelaskan ulang, representasi apa yang perlu dibandingkan, dan siapa yang memerlukan dukungan berbeda. Oleh sebab itu, bukti diagnosis harus langsung terhubung dengan tindakan pembelajaran, bukan berhenti pada daftar kesalahan (UNESCO, 2023; AERA et al., 2014).

Secara operasional, guru dapat menggunakan predict-observe-explain untuk menelusuri konsistensi lintas butir. Data tersebut perlu dibaca bersama stabilitas setelah jeda waktu, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah memperlakukan konsepsi sebagai sifat permanen. Untuk menguranginya, guru dapat memeriksa konsep prasyarat dan menelusuri langkah penyelesaian. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai konflik peran guru sebagai pengajar dan penilai, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan konflik peran guru sebagai pengajar dan penilai melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan konflik peran guru sebagai pengajar dan penilai;
- Bandingkan bukti konflik peran guru sebagai pengajar dan penilai pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek konflik peran guru sebagai pengajar dan penilai dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian konflik peran guru sebagai pengajar dan penilai, topik mol dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa satu mol semua zat dianggap memiliki massa yang sama. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta perbandingan jumlah partikel dan massa satu mol beberapa zat. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa mol menyatakan jumlah entitas, sedangkan massa bergantung pada massa molar. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menempatkan martabat, keadilan, privasi, dan kemanfaatan peserta didik sebagai batas penggunaan diagnosis, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, konflik peran guru sebagai pengajar dan penilai harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.10.24 Prosedur penanganan temuan sensitif

Prosedur penanganan temuan sensitif perlu ditempatkan dalam kerangka diagnosis yang menempatkan martabat, keadilan, privasi, dan kemanfaatan peserta didik sebagai batas penggunaan diagnosis. Fokusnya bukan sekadar menghasilkan skor, melainkan memperoleh bukti yang cukup untuk menjelaskan bagaimana peserta didik membangun, menghubungkan, dan menggunakan konsep. Karena konsepsi tidak dapat diamati secara langsung, guru harus menarik inferensi dari respons verbal, tulisan, gambar, tindakan, dan perubahan strategi yang tampak selama penyelesaian tugas (AERA et al., 2014; UNESCO IITE, 2020).

Secara operasional, guru dapat menggunakan tugas isomorfik untuk menelusuri konsistensi lintas representasi. Data tersebut perlu dibaca bersama urutan langkah penyelesaian, karena

pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mengabaikan ketidakpastian pengukuran. Untuk mengurangnya, guru dapat meminta translasi antarrepresentasi dan menggunakan triangulasi bukti. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai prosedur penanganan temuan sensitif, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan prosedur penanganan temuan sensitif melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan prosedur penanganan temuan sensitif;
- Bandingkan bukti prosedur penanganan temuan sensitif pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek prosedur penanganan temuan sensitif dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian prosedur penanganan temuan sensitif, topik konsentrasi dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa konsentrasi disamakan dengan jumlah zat terlarut tanpa memperhatikan volume. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta perbandingan beberapa larutan dengan jumlah mol dan volume berbeda. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa konsentrasi adalah besaran rasio yang bergantung pada pembilang dan penyebut. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menempatkan martabat, keadilan, privasi, dan kemanfaatan peserta didik sebagai batas penggunaan diagnosis, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Kerangka operasional

Aspek	Bukti yang dicari	Risiko interpretasi	Tindak lanjut
Target konsep	Penjelasan tentang konsentrasi	Target terlalu luas	Batasi subkonsep dan prasyarat
Respons	Konsistensi lintas representasi	Mengabaikan ketidakpastian pengukuran	Meminta translasi antarrepresentasi
Penalaran	Urutan langkah penyelesaian	Alasan diperlakukan sekadar tambahan	Menggunakan triangulasi bukti

Aspek	Bukti yang dicari	Risiko interpretasi	Tindak lanjut
Konfirmasi	Tugas isomorfik atau tugas isomorfik	Kesimpulan dibuat dari satu konteks	Gunakan bukti kedua dan rediagnosis

Implikasinya, prosedur penanganan temuan sensitif harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.10.25 Audit etika diagnosis

Pembahasan tentang audit etika diagnosis berangkat dari prinsip bahwa data asesmen selalu merupakan jejak tidak langsung dari proses berpikir. Satu pilihan jawaban hanya menunjukkan produk akhir, sedangkan struktur pengetahuan yang melahirkannya dapat berbeda. Diagnosis yang kuat karena itu menggabungkan kemampuan transfer, kualitas model mental yang dieksternalisasi, dan bukti kontekstual sehingga keputusan pedagogis tidak bergantung pada satu indikator saja (UNESCO IITE, 2020; Indonesia, 2022).

Secara operasional, guru dapat menggunakan kuis formatif digital untuk menelusuri kemampuan transfer. Data tersebut perlu dibaca bersama kualitas model mental yang dieksternalisasi, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menyimpulkan miskonsepsi hanya dari satu jawaban. Untuk menguranginya, guru dapat mengulang diagnosis setelah remediasi dan menyusun tindak lanjut berbeda untuk kategori respons yang berbeda. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai audit etika diagnosis, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan audit etika diagnosis melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan audit etika diagnosis;
- Bandingkan bukti audit etika diagnosis pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek audit etika diagnosis dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian audit etika diagnosis, topik kelarutan dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa larutan jenuh

dianggap tidak mengandung proses pelarutan lagi. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta model partikel pelarutan dan pengendapan pada keadaan jenuh. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa kejenuhan adalah keadaan kesetimbangan dinamis. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menempatkan martabat, keadilan, privasi, dan kemanfaatan peserta didik sebagai batas penggunaan diagnosis, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, audit etika diagnosis harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.10.26 Sintesis etika diagnosis peserta didik

Dalam pembelajaran kimia, sintesis etika diagnosis peserta didik memiliki arti strategis karena banyak jawaban dapat diperoleh melalui hafalan prosedur atau pengenalan pola. Peserta didik mungkin tampak berhasil tetapi tidak mampu menerjemahkan makna simbol, model, atau besaran yang digunakan. Prinsip diagnosis menuntut bukti tentang cara konsep dipakai, bukan hanya apakah hasil akhirnya sesuai kunci (Indonesia, 2022; UNESCO, 2023).

Secara operasional, guru dapat menggunakan tes tiga tingkat untuk menelusuri stabilitas setelah jeda waktu. Data tersebut perlu dibaca bersama respons terhadap probing, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mencampur kesalahan matematika dengan kesalahan konsep. Untuk menguranginya, guru dapat mencatat derajat keyakinan dan membandingkan jawaban dengan alasan. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai sintesis etika diagnosis peserta didik, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan sintesis etika diagnosis peserta didik melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan sintesis etika diagnosis peserta didik;
- Bandingkan bukti sintesis etika diagnosis peserta didik pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek sintesis etika diagnosis peserta didik dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian sintesis etika diagnosis peserta didik, topik redoks dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa oksidasi selalu dianggap melibatkan oksigen. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta analisis transfer elektron dan perubahan bilangan oksidasi. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa pengikatan oksigen merupakan kasus khusus dari kerangka redoks yang lebih luas. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus menempatkan martabat, keadilan, privasi, dan kemanfaatan peserta didik sebagai batas penggunaan diagnosis, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, sintesis etika diagnosis peserta didik harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.10.27 Implikasi praktis bagi guru kimia

Implikasi praktis subbab 5.10 dapat dirangkum sebagai alur: tentukan konsep target, pilih bukti yang relevan, kumpulkan respons dalam lebih dari satu bentuk, analisis pola, nyatakan derajat kepastian, susun tindak lanjut, dan lakukan rediagnosis. Alur ini menjaga agar etika dalam diagnosis peserta didik tidak berubah menjadi administrasi skor, tetapi berfungsi sebagai mekanisme perbaikan pembelajaran yang berulang.

64. Identifikasi konsep inti dan konsep prasyarat sebelum menyusun instrumen.
65. Gunakan pertanyaan yang meminta alasan, representasi, atau langkah kerja.
66. Bandingkan respons pada beberapa konteks dan waktu.
67. Pisahkan bukti faktual dari interpretasi guru.
68. Tentukan tindak lanjut yang berbeda untuk kategori kebutuhan yang berbeda.
69. Simpan data secukupnya dan lindungi kerahasiaan peserta didik.
70. Periksa kembali perubahan melalui tugas transfer dan retensi.

5.11 Interpretasi Hasil Diagnosis

Subbab ini mengubah data menjadi kesimpulan yang terukur, transparan, dan langsung berhubungan dengan tindak lanjut. Uraian disusun dari prinsip teoretis menuju prosedur kelas, contoh kimia, risiko interpretasi, dan tindak lanjut. Setiap bagian menekankan bahwa kesimpulan diagnostik harus didasarkan pada pola bukti, bukan pada satu respons yang terpisah.

5.11.1 Interpretasi sebagai proses argumentatif

Pembahasan tentang interpretasi sebagai proses argumentatif berangkat dari prinsip bahwa data asesmen selalu merupakan jejak tidak langsung dari proses berpikir. Satu pilihan jawaban hanya menunjukkan produk akhir, sedangkan struktur pengetahuan yang melahirkannya dapat berbeda. Diagnosis yang kuat karena itu menggabungkan koherensi alasan, konsistensi lintas representasi, dan bukti kontekstual sehingga keputusan pedagogis tidak bergantung pada satu indikator saja (AERA et al., 2014; Treagust, 1988).

Secara operasional, guru dapat menggunakan analisis grafik untuk menelusuri koherensi alasan. Data tersebut perlu dibaca bersama konsistensi lintas representasi, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mencampur kesalahan matematika dengan kesalahan konsep. Untuk mengurangnya, guru dapat menggunakan pertanyaan tindak lanjut dan mengulang diagnosis setelah remediasi. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai interpretasi sebagai proses argumentatif, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan interpretasi sebagai proses argumentatif melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan interpretasi sebagai proses argumentatif;
- Bandingkan bukti interpretasi sebagai proses argumentatif pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek interpretasi sebagai proses argumentatif dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian interpretasi sebagai proses argumentatif, topik kinetika dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa reaksi eksoterm dianggap selalu cepat. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta perbandingan profil energi dan faktor kinetik. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa termodinamika dan kinetika menjawab pertanyaan berbeda. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengubah data menjadi kesimpulan yang terukur, transparan, dan langsung berhubungan dengan tindak lanjut, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, interpretasi sebagai proses argumentatif harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik

'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.11.2 Memulai interpretasi dari tujuan diagnosis

Dalam pembelajaran kimia, memulai interpretasi dari tujuan diagnosis memiliki arti strategis karena banyak jawaban dapat diperoleh melalui hafalan prosedur atau pengenalan pola. Peserta didik mungkin tampak berhasil tetapi tidak mampu menerjemahkan makna simbol, model, atau besaran yang digunakan. Prinsip diagnosis menuntut bukti tentang cara konsep dipakai, bukan hanya apakah hasil akhirnya sesuai kunci (Treagust, 1988; Black & Wiliam, 2009).

Secara operasional, guru dapat menggunakan tes empat tingkat untuk menelusuri tingkat keyakinan. Data tersebut perlu dibaca bersama kemampuan transfer, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menggunakan ambang kategori tanpa bukti. Untuk menguranginya, guru dapat menyajikan tugas paralel dan mencatat derajat keyakinan. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai memulai interpretasi dari tujuan diagnosis, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan memulai interpretasi dari tujuan diagnosis melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan memulai interpretasi dari tujuan diagnosis;
- Bandingkan bukti memulai interpretasi dari tujuan diagnosis pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek memulai interpretasi dari tujuan diagnosis dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian memulai interpretasi dari tujuan diagnosis, topik elektrokimia dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa elektron dianggap bergerak melalui jembatan garam. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta penelusuran jalur elektron dan ion pada sel galvanik. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa elektron mengalir pada rangkaian luar, sedangkan ion menjaga kenetralan muatan. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengubah data menjadi kesimpulan yang terukur, transparan, dan langsung berhubungan dengan tindak lanjut, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu

pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, memulai interpretasi dari tujuan diagnosis harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.11.3 Membaca skor total secara hati-hati

Aspek membaca skor total secara hati-hati menegaskan bahwa diagnosis adalah kegiatan ilmiah berskala kelas: guru merumuskan pertanyaan, mengumpulkan bukti, mempertimbangkan penjelasan alternatif, lalu menentukan kesimpulan yang paling masuk akal. Inferensi harus proporsional terhadap kualitas data. Semakin spesifik label yang diberikan, semakin kuat pula bukti yang diperlukan (Black & Wiliam, 2009; Kane, 2013).

Secara operasional, guru dapat menggunakan think-aloud untuk menelusuri konsistensi lintas butir. Data tersebut perlu dibaca bersama stabilitas setelah jeda waktu, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah memperlakukan konsepsi sebagai sifat permanen. Untuk mengurangnya, guru dapat memeriksa konsep prasyarat dan menelusuri langkah penyelesaian. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai membaca skor total secara hati-hati, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan membaca skor total secara hati-hati melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan membaca skor total secara hati-hati;
- Bandingkan bukti membaca skor total secara hati-hati pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek membaca skor total secara hati-hati dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian membaca skor total secara hati-hati, topik praktikum titrasi dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa titik akhir dianggap selalu identik dengan titik ekuivalen. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta analisis indikator, kurva titrasi, dan ketidakpastian. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa titik akhir adalah observasi operasional yang mendekati titik ekuivalen.

Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengubah data menjadi kesimpulan yang terukur, transparan, dan langsung berhubungan dengan tindak lanjut, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, membaca skor total secara hati-hati harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.11.4 Membaca pola per konsep dan subkonsep

Penerapan membaca pola per konsep dan subkonsep harus menghubungkan teori pengukuran dengan pengetahuan pedagogis kimia. Instrumen yang teknis tetapi tidak memahami struktur konsep akan menghasilkan kategori dangkal; sebaliknya, pengetahuan konten tanpa prosedur pengumpulan bukti dapat menghasilkan kesimpulan intuitif yang sulit diuji. Keduanya perlu dipadukan dalam satu alur yang transparan (Kane, 2013; AERA et al., 2014).

Secara operasional, guru dapat menggunakan predict-observe-explain untuk menelusuri konsistensi lintas representasi. Data tersebut perlu dibaca bersama urutan langkah penyelesaian, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mengabaikan ketidakpastian pengukuran. Untuk menguranginya, guru dapat meminta translasi antarrepresentasi dan menggunakan triangulasi bukti. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai membaca pola per konsep dan subkonsep, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan membaca pola per konsep dan subkonsep melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan membaca pola per konsep dan subkonsep;
- Bandingkan bukti membaca pola per konsep dan subkonsep pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek membaca pola per konsep dan subkonsep dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian membaca pola per konsep dan subkonsep, topik sistem periodik dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa jari-jari atom dianggap selalu bertambah bersama nomor atom. Diagnosis

tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta prediksi tren disertai penjelasan muatan inti efektif dan kulit elektron. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa tren periodik perlu dijelaskan melalui struktur dan interaksi, bukan hafalan arah panah. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengubah data menjadi kesimpulan yang terukur, transparan, dan langsung berhubungan dengan tindak lanjut, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Kerangka operasional

Aspek	Bukti yang dicari	Risiko interpretasi	Tindak lanjut
Target konsep	Penjelasan tentang sistem periodik	Target terlalu luas	Batasi subkonsep dan prasyarat
Respons	Konsistensi lintas representasi	Mengabaikan ketidakpastian pengukuran	Meminta translasi antarrepresentasi
Penalaran	Urutan langkah penyelesaian	Alasan diperlakukan sekadar tambahan	Menggunakan triangulasi bukti
Konfirmasi	Predict-observe-explain atau tugas isomorfik	Kesimpulan dibuat dari satu konteks	Gunakan bukti kedua dan rediagnosis

Implikasinya, membaca pola per konsep dan subkonsep harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.11.5 Membaca pola jawaban-alasan-keyakinan

Membaca pola jawaban-alasan-keyakinan tidak berdiri sebagai prosedur administratif. Ia menjadi bagian dari keputusan tentang apa yang perlu dijelaskan ulang, representasi apa yang perlu dibandingkan, dan siapa yang memerlukan dukungan berbeda. Oleh sebab itu, bukti diagnosis harus langsung terhubung dengan tindakan pembelajaran, bukan berhenti pada daftar kesalahan (AERA et al., 2014; Treagust, 1988).

Secara operasional, guru dapat menggunakan tugas isomorfik untuk menelusuri kemampuan transfer. Data tersebut perlu dibaca bersama kualitas model mental yang dieksternalisasi,

karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menyimpulkan miskonsepsi hanya dari satu jawaban. Untuk mengurangnya, guru dapat mengulang diagnosis setelah remediasi dan menyusun tindak lanjut berbeda untuk kategori respons yang berbeda. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai membaca pola jawaban-alasan-keyakinan, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan membaca pola jawaban-alasan-keyakinan melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan membaca pola jawaban-alasan-keyakinan;
- Bandingkan bukti membaca pola jawaban-alasan-keyakinan pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek membaca pola jawaban-alasan-keyakinan dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian membaca pola jawaban-alasan-keyakinan, topik stoikiometri dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa koefisien reaksi diperlakukan sebagai perbandingan massa langsung. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta translasi persamaan dari partikel ke mol dan massa. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa koefisien terutama menyatakan rasio jumlah entitas atau mol. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengubah data menjadi kesimpulan yang terukur, transparan, dan langsung berhubungan dengan tindak lanjut, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, membaca pola jawaban-alasan-keyakinan harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.11.6 Membedakan frekuensi dan tingkat keparahan

Membedakan frekuensi dan tingkat keparahan perlu ditempatkan dalam kerangka diagnosis yang mengubah data menjadi kesimpulan yang terukur, transparan, dan langsung berhubungan dengan tindak lanjut. Fokusnya bukan sekadar menghasilkan skor, melainkan

memperoleh bukti yang cukup untuk menjelaskan bagaimana peserta didik membangun, menghubungkan, dan menggunakan konsep. Karena konsepsi tidak dapat diamati secara langsung, guru harus menarik inferensi dari respons verbal, tulisan, gambar, tindakan, dan perubahan strategi yang tampak selama penyelesaian tugas (Treagust, 1988; Black & Wiliam, 2009).

Secara operasional, guru dapat menggunakan kuis formatif digital untuk menelusuri stabilitas setelah jeda waktu. Data tersebut perlu dibaca bersama respons terhadap probing, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mencampur kesalahan matematika dengan kesalahan konsep. Untuk mengurangnya, guru dapat mencatat derajat keyakinan dan membandingkan jawaban dengan alasan. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai membedakan frekuensi dan tingkat keparahan, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan membedakan frekuensi dan tingkat keparahan melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan membedakan frekuensi dan tingkat keparahan;
- Bandingkan bukti membedakan frekuensi dan tingkat keparahan pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek membedakan frekuensi dan tingkat keparahan dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian membedakan frekuensi dan tingkat keparahan, topik asam-basa dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa asam kuat dianggap selalu berkonsentrasi tinggi. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta pemisahan konsep derajat ionisasi, K_a , konsentrasi, dan pH. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa kekuatan asam dan konsentrasi adalah dimensi berbeda. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengubah data menjadi kesimpulan yang terukur, transparan, dan langsung berhubungan dengan tindak lanjut, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, membedakan frekuensi dan tingkat keparahan harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan

identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.11.7 Menafsirkan konsistensi lintas butir

Pembahasan tentang menafsirkan konsistensi lintas butir berangkat dari prinsip bahwa data asesmen selalu merupakan jejak tidak langsung dari proses berpikir. Satu pilihan jawaban hanya menunjukkan produk akhir, sedangkan struktur pengetahuan yang melahirkannya dapat berbeda. Diagnosis yang kuat karena itu menggabungkan urutan langkah penyelesaian, ketepatan pilihan, dan bukti kontekstual sehingga keputusan pedagogis tidak bergantung pada satu indikator saja (Black & Wiliam, 2009; Kane, 2013).

Secara operasional, guru dapat menggunakan tes tiga tingkat untuk menelusuri urutan langkah penyelesaian. Data tersebut perlu dibaca bersama ketepatan pilihan, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menggunakan ambang kategori tanpa bukti. Untuk mengurangnya, guru dapat menelusuri langkah penyelesaian dan menggunakan pertanyaan tindak lanjut. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai menafsirkan konsistensi lintas butir, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan menafsirkan konsistensi lintas butir melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan menafsirkan konsistensi lintas butir;
- Bandingkan bukti menafsirkan konsistensi lintas butir pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek menafsirkan konsistensi lintas butir dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian menafsirkan konsistensi lintas butir, topik kinetika dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa reaksi eksoterm dianggap selalu cepat. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta perbandingan profil energi dan faktor kinetik. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa termodinamika dan kinetika menjawab pertanyaan berbeda. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengubah data menjadi kesimpulan yang terukur, transparan, dan langsung berhubungan dengan tindak lanjut, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, menafsirkan konsistensi lintas butir harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.11.8 Menafsirkan konsistensi lintas representasi

Dalam pembelajaran kimia, menafsirkan konsistensi lintas representasi memiliki arti strategis karena banyak jawaban dapat diperoleh melalui hafalan prosedur atau pengenalan pola. Peserta didik mungkin tampak berhasil tetapi tidak mampu menerjemahkan makna simbol, model, atau besaran yang digunakan. Prinsip diagnosis menuntut bukti tentang cara konsep dipakai, bukan hanya apakah hasil akhirnya sesuai kunci (Kane, 2013; AERA et al., 2014).

Secara operasional, guru dapat menggunakan wawancara semi-terstruktur untuk menelusuri kualitas model mental yang dieksternalisasi. Data tersebut perlu dibaca bersama koherensi alasan, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah memperlakukan konsepsi sebagai sifat permanen. Untuk mengurangnya, guru dapat menggunakan triangulasi bukti dan menyajikan tugas paralel. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai menafsirkan konsistensi lintas representasi, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan menafsirkan konsistensi lintas representasi melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan menafsirkan konsistensi lintas representasi;
- Bandingkan bukti menafsirkan konsistensi lintas representasi pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek menafsirkan konsistensi lintas representasi dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian menafsirkan konsistensi lintas representasi, topik elektrokimia dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa elektron dianggap bergerak melalui jembatan garam. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta penelusuran jalur elektron dan ion pada sel galvanik. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa elektron mengalir pada rangkaian luar, sedangkan ion menjaga kenetralan muatan. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengubah data menjadi kesimpulan yang terukur, transparan, dan langsung berhubungan dengan tindak lanjut, bukan diperlakukan

sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Kerangka operasional

Aspek	Bukti yang dicari	Risiko interpretasi	Tindak lanjut
Target konsep	Penjelasan tentang elektrokimia	Target terlalu luas	Batasi subkonsep dan prasyarat
Respons	Kualitas model mental yang dieksternalisasi	Memperlakukan konsepsi sebagai sifat permanen	Menggunakan triangulasi bukti
Penalaran	Koherensi alasan	Alasan diperlakukan sekadar tambahan	Menyajikan tugas paralel
Konfirmasi	Wawancara semi-terstruktur atau tugas isomorfik	Kesimpulan dibuat dari satu konteks	Gunakan bukti kedua dan rediagnosis

Implikasinya, menafsirkan konsistensi lintas representasi harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.11.9 Menafsirkan perubahan dari waktu ke waktu

Aspek menafsirkan perubahan dari waktu ke waktu menegaskan bahwa diagnosis adalah kegiatan ilmiah berskala kelas: guru merumuskan pertanyaan, mengumpulkan bukti, mempertimbangkan penjelasan alternatif, lalu menentukan kesimpulan yang paling masuk akal. Inferensi harus proporsional terhadap kualitas data. Semakin spesifik label yang diberikan, semakin kuat pula bukti yang diperlukan (AERA et al., 2014; Treagust, 1988).

Secara operasional, guru dapat menggunakan gambar partikel untuk menelusuri respons terhadap probing. Data tersebut perlu dibaca bersama tingkat keyakinan, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mengabaikan ketidakpastian pengukuran. Untuk menguranginya, guru dapat menyusun tindak lanjut berbeda untuk kategori respons yang berbeda dan memeriksa konsep prasyarat. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai menafsirkan perubahan dari waktu ke waktu, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan menafsirkan perubahan dari waktu ke waktu melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan menafsirkan perubahan dari waktu ke waktu;
- Bandingkan bukti menafsirkan perubahan dari waktu ke waktu pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek menafsirkan perubahan dari waktu ke waktu dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian menafsirkan perubahan dari waktu ke waktu, topik praktikum titrasi dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa titik akhir dianggap selalu identik dengan titik ekuivalen. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta analisis indikator, kurva titrasi, dan ketidakpastian. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa titik akhir adalah observasi operasional yang mendekati titik ekuivalen. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengubah data menjadi kesimpulan yang terukur, transparan, dan langsung berhubungan dengan tindak lanjut, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, menafsirkan perubahan dari waktu ke waktu harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.11.10 Menafsirkan hasil pretest dan posttest

Penerapan menafsirkan hasil pretest dan posttest harus menghubungkan teori pengukuran dengan pengetahuan pedagogis kimia. Instrumen yang teknis tetapi tidak memahami struktur konsep akan menghasilkan kategori dangkal; sebaliknya, pengetahuan konten tanpa prosedur pengumpulan bukti dapat menghasilkan kesimpulan intuitif yang sulit diuji. Keduanya perlu dipadukan dalam satu alur yang transparan (Treagust, 1988; Black & Wiliam, 2009).

Secara operasional, guru dapat menggunakan observasi praktikum untuk menelusuri ketepatan pilihan. Data tersebut perlu dibaca bersama konsistensi lintas butir, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menyimpulkan miskonsepsi hanya dari satu jawaban. Untuk menguranginya, guru

dapat membandingkan jawaban dengan alasan dan meminta translasi antarrepresentasi. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai menafsirkan hasil pretest dan posttest, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan menafsirkan hasil pretest dan posttest melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan menafsirkan hasil pretest dan posttest;
- Bandingkan bukti menafsirkan hasil pretest dan posttest pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek menafsirkan hasil pretest dan posttest dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian menafsirkan hasil pretest dan posttest, topik sistem periodik dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa jari-jari atom dianggap selalu bertambah bersama nomor atom. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta prediksi tren disertai penjelasan muatan inti efektif dan kulit elektron. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa tren periodik perlu dijelaskan melalui struktur dan interaksi, bukan hafalan arah panah. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengubah data menjadi kesimpulan yang terukur, transparan, dan langsung berhubungan dengan tindak lanjut, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, menafsirkan hasil pretest dan posttest harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.11.11 Menafsirkan retensi dan transfer

Menafsirkan retensi dan transfer tidak berdiri sebagai prosedur administratif. Ia menjadi bagian dari keputusan tentang apa yang perlu dijelaskan ulang, representasi apa yang perlu dibandingkan, dan siapa yang memerlukan dukungan berbeda. Oleh sebab itu, bukti diagnosis harus langsung terhubung dengan tindakan pembelajaran, bukan berhenti pada daftar kesalahan (Black & Wiliam, 2009; Kane, 2013).

Secara operasional, guru dapat menggunakan exit ticket untuk menelusuri koherensi alasan. Data tersebut perlu dibaca bersama konsistensi lintas representasi, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mencampur kesalahan matematika dengan kesalahan konsep. Untuk mengurangnya, guru dapat menggunakan pertanyaan tindak lanjut dan mengulang diagnosis setelah remediasi. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai menafsirkan retensi dan transfer, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan menafsirkan retensi dan transfer melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan menafsirkan retensi dan transfer;
- Bandingkan bukti menafsirkan retensi dan transfer pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek menafsirkan retensi dan transfer dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian menafsirkan retensi dan transfer, topik stoikiometri dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa koefisien reaksi diperlakukan sebagai perbandingan massa langsung. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta translasi persamaan dari partikel ke mol dan massa. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa koefisien terutama menyatakan rasio jumlah entitas atau mol. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengubah data menjadi kesimpulan yang terukur, transparan, dan langsung berhubungan dengan tindak lanjut, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, menafsirkan retensi dan transfer harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.11.12 Menafsirkan respons tidak lengkap

Menafsirkan respons tidak lengkap perlu ditempatkan dalam kerangka diagnosis yang mengubah data menjadi kesimpulan yang terukur, transparan, dan langsung berhubungan dengan tindak lanjut. Fokusnya bukan sekadar menghasilkan skor, melainkan memperoleh

bukti yang cukup untuk menjelaskan bagaimana peserta didik membangun, menghubungkan, dan menggunakan konsep. Karena konsepsi tidak dapat diamati secara langsung, guru harus menarik inferensi dari respons verbal, tulisan, gambar, tindakan, dan perubahan strategi yang tampak selama penyelesaian tugas (Kane, 2013; AERA et al., 2014).

Secara operasional, guru dapat menggunakan tes pilihan ganda dua tingkat untuk menelusuri tingkat keyakinan. Data tersebut perlu dibaca bersama kemampuan transfer, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menggunakan ambang kategori tanpa bukti. Untuk mengurangnya, guru dapat menyajikan tugas paralel dan mencatat derajat keyakinan. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai menafsirkan respons tidak lengkap, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan menafsirkan respons tidak lengkap melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan menafsirkan respons tidak lengkap;
- Bandingkan bukti menafsirkan respons tidak lengkap pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek menafsirkan respons tidak lengkap dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian menafsirkan respons tidak lengkap, topik asam-basa dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa asam kuat dianggap selalu berkonsentrasi tinggi. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta pemisahan konsep derajat ionisasi, K_a , konsentrasi, dan pH. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa kekuatan asam dan konsentrasi adalah dimensi berbeda. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengubah data menjadi kesimpulan yang terukur, transparan, dan langsung berhubungan dengan tindak lanjut, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Kerangka operasional

Aspek	Bukti yang dicari	Risiko interpretasi	Tindak lanjut
Target konsep	Penjelasan tentang asam-basa	Target terlalu luas	Batasi subkonsep dan prasyarat

Aspek	Bukti yang dicari	Risiko interpretasi	Tindak lanjut
Respons	Tingkat keyakinan	Menggunakan ambang kategori tanpa bukti	Menyajikan tugas paralel
Penalaran	Kemampuan transfer	Alasan diperlakukan sekadar tambahan	Mencatat derajat keyakinan
Konfirmasi	Tes pilihan ganda dua tingkat atau tugas isomorfik	Kesimpulan dibuat dari satu konteks	Gunakan bukti kedua dan rediagnosis

Implikasinya, menafsirkan respons tidak lengkap harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.11.13 Menafsirkan data wawancara

Pembahasan tentang menafsirkan data wawancara berangkat dari prinsip bahwa data asesmen selalu merupakan jejak tidak langsung dari proses berpikir. Satu pilihan jawaban hanya menunjukkan produk akhir, sedangkan struktur pengetahuan yang melahirkannya dapat berbeda. Diagnosis yang kuat karena itu menggabungkan konsistensi lintas butir, stabilitas setelah jeda waktu, dan bukti kontekstual sehingga keputusan pedagogis tidak bergantung pada satu indikator saja (AERA et al., 2014; Treagust, 1988).

Secara operasional, guru dapat menggunakan alasan terbuka untuk menelusuri konsistensi lintas butir. Data tersebut perlu dibaca bersama stabilitas setelah jeda waktu, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah memperlakukan konsepsi sebagai sifat permanen. Untuk mengurangnya, guru dapat memeriksa konsep prasyarat dan menelusuri langkah penyelesaian. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai menafsirkan data wawancara, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan menafsirkan data wawancara melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan menafsirkan data wawancara;
- Bandingkan bukti menafsirkan data wawancara pada lebih dari satu konteks, format tugas, atau bentuk representasi;

- Hubungkan kategori hasil pada aspek menafsirkan data wawancara dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian menafsirkan data wawancara, topik kinetika dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa reaksi eksoterm dianggap selalu cepat. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta perbandingan profil energi dan faktor kinetik. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa termodinamika dan kinetika menjawab pertanyaan berbeda. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengubah data menjadi kesimpulan yang terukur, transparan, dan langsung berhubungan dengan tindak lanjut, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, menafsirkan data wawancara harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.11.14 Menafsirkan peta konsep dan gambar

Dalam pembelajaran kimia, menafsirkan peta konsep dan gambar memiliki arti strategis karena banyak jawaban dapat diperoleh melalui hafalan prosedur atau pengenalan pola. Peserta didik mungkin tampak berhasil tetapi tidak mampu menerjemahkan makna simbol, model, atau besaran yang digunakan. Prinsip diagnosis menuntut bukti tentang cara konsep dipakai, bukan hanya apakah hasil akhirnya sesuai kunci (Treagust, 1988; Black & Wiliam, 2009).

Secara operasional, guru dapat menggunakan peta konsep untuk menelusuri konsistensi lintas representasi. Data tersebut perlu dibaca bersama urutan langkah penyelesaian, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mengabaikan ketidakpastian pengukuran. Untuk menguranginya, guru dapat meminta translasi antarrepresentasi dan menggunakan triangulasi bukti. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai menafsirkan peta konsep dan gambar, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan menafsirkan peta konsep dan gambar melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan menafsirkan peta konsep dan gambar;

- Bandingkan bukti menafsirkan peta konsep dan gambar pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek menafsirkan peta konsep dan gambar dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian menafsirkan peta konsep dan gambar, topik elektrokimia dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa elektron dianggap bergerak melalui jembatan garam. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta penelusuran jalur elektron dan ion pada sel galvanik. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa elektron mengalir pada rangkaian luar, sedangkan ion menjaga kenetralan muatan. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengubah data menjadi kesimpulan yang terukur, transparan, dan langsung berhubungan dengan tindak lanjut, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, menafsirkan peta konsep dan gambar harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.11.15 Menafsirkan data observasi

Aspek menafsirkan data observasi menegaskan bahwa diagnosis adalah kegiatan ilmiah berskala kelas: guru merumuskan pertanyaan, mengumpulkan bukti, mempertimbangkan penjelasan alternatif, lalu menentukan kesimpulan yang paling masuk akal. Inferensi harus proporsional terhadap kualitas data. Semakin spesifik label yang diberikan, semakin kuat pula bukti yang diperlukan (Black & Wiliam, 2009; Kane, 2013).

Secara operasional, guru dapat menggunakan analisis respons tertulis untuk menelusuri kemampuan transfer. Data tersebut perlu dibaca bersama kualitas model mental yang dieksternalisasi, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menyimpulkan miskonsepsi hanya dari satu jawaban. Untuk mengurangnya, guru dapat mengulang diagnosis setelah remediasi dan menyusun tindak lanjut berbeda untuk kategori respons yang berbeda. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai menafsirkan data observasi, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;

- Pilih tugas yang merepresentasikan menafsirkan data observasi melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan menafsirkan data observasi;
- Bandingkan bukti menafsirkan data observasi pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek menafsirkan data observasi dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian menafsirkan data observasi, topik praktikum titrasi dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa titik akhir dianggap selalu identik dengan titik ekuivalen. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta analisis indikator, kurva titrasi, dan ketidakpastian. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa titik akhir adalah observasi operasional yang mendekati titik ekuivalen. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengubah data menjadi kesimpulan yang terukur, transparan, dan langsung berhubungan dengan tindak lanjut, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, menafsirkan data observasi harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.11.16 Menggabungkan data kuantitatif dan kualitatif

Penerapan menggabungkan data kuantitatif dan kualitatif harus menghubungkan teori pengukuran dengan pengetahuan pedagogis kimia. Instrumen yang teknis tetapi tidak memahami struktur konsep akan menghasilkan kategori dangkal; sebaliknya, pengetahuan konten tanpa prosedur pengumpulan bukti dapat menghasilkan kesimpulan intuitif yang sulit diuji. Keduanya perlu dipadukan dalam satu alur yang transparan (Kane, 2013; AERA et al., 2014).

Secara operasional, guru dapat menggunakan diskusi argumentatif untuk menelusuri stabilitas setelah jeda waktu. Data tersebut perlu dibaca bersama respons terhadap probing, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mencampur kesalahan matematika dengan kesalahan konsep. Untuk menguranginya, guru dapat mencatat derajat keyakinan dan membandingkan jawaban

dengan alasan. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai menggabungkan data kuantitatif dan kualitatif, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan menggabungkan data kuantitatif dan kualitatif melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan menggabungkan data kuantitatif dan kualitatif;
- Bandingkan bukti menggabungkan data kuantitatif dan kualitatif pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek menggabungkan data kuantitatif dan kualitatif dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian menggabungkan data kuantitatif dan kualitatif, topik sistem periodik dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa jari-jari atom dianggap selalu bertambah bersama nomor atom. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta prediksi tren disertai penjelasan muatan inti efektif dan kulit elektron. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa tren periodik perlu dijelaskan melalui struktur dan interaksi, bukan hafalan arah panah. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengubah data menjadi kesimpulan yang terukur, transparan, dan langsung berhubungan dengan tindak lanjut, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Kerangka operasional

Aspek	Bukti yang dicari	Risiko interpretasi	Tindak lanjut
Target konsep	Penjelasan tentang sistem periodik	Target terlalu luas	Batasi subkonsep dan prasyarat
Respons	Stabilitas setelah jeda waktu	Mencampur kesalahan matematika dengan kesalahan konsep	Mencatat derajat keyakinan
Penalaran	Respons terhadap probing	Alasan diperlakukan sekadar tambahan	Membandingkan jawaban dengan alasan
Konfirmasi	Diskusi argumentatif atau tugas isomorfik	Kesimpulan dibuat dari satu konteks	Gunakan bukti kedua dan rediagnosis

Implikasinya, menggabungkan data kuantitatif dan kualitatif harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.11.17 Menentukan ambang klasifikasi

Menentukan ambang klasifikasi tidak berdiri sebagai prosedur administratif. Ia menjadi bagian dari keputusan tentang apa yang perlu dijelaskan ulang, representasi apa yang perlu dibandingkan, dan siapa yang memerlukan dukungan berbeda. Oleh sebab itu, bukti diagnosis harus langsung terhubung dengan tindakan pembelajaran, bukan berhenti pada daftar kesalahan (AERA et al., 2014; Treagust, 1988).

Secara operasional, guru dapat menggunakan analisis grafik untuk menelusuri urutan langkah penyelesaian. Data tersebut perlu dibaca bersama ketepatan pilihan, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menggunakan ambang kategori tanpa bukti. Untuk mengurangnya, guru dapat menelusuri langkah penyelesaian dan menggunakan pertanyaan tindak lanjut. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai menentukan ambang klasifikasi, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan menentukan ambang klasifikasi melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan menentukan ambang klasifikasi;
- Bandingkan bukti menentukan ambang klasifikasi pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek menentukan ambang klasifikasi dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian menentukan ambang klasifikasi, topik stoikiometri dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa koefisien reaksi diperlakukan sebagai perbandingan massa langsung. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta translasi persamaan dari partikel ke mol dan massa. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa koefisien terutama menyatakan rasio jumlah entitas atau mol. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengubah data menjadi kesimpulan yang terukur, transparan, dan langsung berhubungan dengan tindak lanjut, bukan diperlakukan

sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, menentukan ambang klasifikasi harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.11.18 Ketidakpastian dan interval keputusan

Ketidakpastian dan interval keputusan perlu ditempatkan dalam kerangka diagnosis yang mengubah data menjadi kesimpulan yang terukur, transparan, dan langsung berhubungan dengan tindak lanjut. Fokusnya bukan sekadar menghasilkan skor, melainkan memperoleh bukti yang cukup untuk menjelaskan bagaimana peserta didik membangun, menghubungkan, dan menggunakan konsep. Karena konsepsi tidak dapat diamati secara langsung, guru harus menarik inferensi dari respons verbal, tulisan, gambar, tindakan, dan perubahan strategi yang tampak selama penyelesaian tugas (Treagust, 1988; Black & Wiliam, 2009).

Secara operasional, guru dapat menggunakan tes empat tingkat untuk menelusuri kualitas model mental yang dieksternalisasi. Data tersebut perlu dibaca bersama koherensi alasan, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah memperlakukan konsepsi sebagai sifat permanen. Untuk menguranginya, guru dapat menggunakan triangulasi bukti dan menyajikan tugas paralel. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai ketidakpastian dan interval keputusan, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan ketidakpastian dan interval keputusan melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan ketidakpastian dan interval keputusan;
- Bandingkan bukti ketidakpastian dan interval keputusan pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek ketidakpastian dan interval keputusan dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian ketidakpastian dan interval keputusan, topik asam-basa dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa asam kuat dianggap selalu berkonsentrasi tinggi. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta pemisahan konsep derajat ionisasi, K_a ,

konsentrasi, dan pH. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa kekuatan asam dan konsentrasi adalah dimensi berbeda. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengubah data menjadi kesimpulan yang terukur, transparan, dan langsung berhubungan dengan tindak lanjut, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, ketidakpastian dan interval keputusan harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.11.19 Menghindari overinterpretation

Pembahasan tentang menghindari overinterpretation berangkat dari prinsip bahwa data asesmen selalu merupakan jejak tidak langsung dari proses berpikir. Satu pilihan jawaban hanya menunjukkan produk akhir, sedangkan struktur pengetahuan yang melahirkannya dapat berbeda. Diagnosis yang kuat karena itu menggabungkan respons terhadap probing, tingkat keyakinan, dan bukti kontekstual sehingga keputusan pedagogis tidak bergantung pada satu indikator saja (Black & Wiliam, 2009; Kane, 2013).

Secara operasional, guru dapat menggunakan think-aloud untuk menelusuri respons terhadap probing. Data tersebut perlu dibaca bersama tingkat keyakinan, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mengabaikan ketidakpastian pengukuran. Untuk menguranginya, guru dapat menyusun tindak lanjut berbeda untuk kategori respons yang berbeda dan memeriksa konsep prasyarat. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai menghindari overinterpretation, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan menghindari overinterpretation melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan menghindari overinterpretation;
- Bandingkan bukti menghindari overinterpretation pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek menghindari overinterpretation dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian menghindari overinterpretation, topik kinetika dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa reaksi eksoterm dianggap selalu cepat. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta perbandingan profil energi dan faktor kinetik. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa termodinamika dan kinetika menjawab pertanyaan berbeda. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengubah data menjadi kesimpulan yang terukur, transparan, dan langsung berhubungan dengan tindak lanjut, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, menghindari overinterpretation harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.11.20 Menghindari causal fallacy

Dalam pembelajaran kimia, menghindari causal fallacy memiliki arti strategis karena banyak jawaban dapat diperoleh melalui hafalan prosedur atau pengenalan pola. Peserta didik mungkin tampak berhasil tetapi tidak mampu menerjemahkan makna simbol, model, atau besaran yang digunakan. Prinsip diagnosis menuntut bukti tentang cara konsep dipakai, bukan hanya apakah hasil akhirnya sesuai kunci (Kane, 2013; AERA et al., 2014).

Secara operasional, guru dapat menggunakan predict-observe-explain untuk menelusuri ketepatan pilihan. Data tersebut perlu dibaca bersama konsistensi lintas butir, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menyimpulkan miskonsepsi hanya dari satu jawaban. Untuk menguranginya, guru dapat membandingkan jawaban dengan alasan dan meminta translasi antarrepresentasi. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai menghindari causal fallacy, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan menghindari causal fallacy melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan menghindari causal fallacy;
- Bandingkan bukti menghindari causal fallacy pada lebih dari satu konteks, format tugas, atau bentuk representasi;

- Hubungkan kategori hasil pada aspek menghindari causal fallacy dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian menghindari causal fallacy, topik elektrokimia dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa elektron dianggap bergerak melalui jembatan garam. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta penelusuran jalur elektron dan ion pada sel galvanik. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa elektron mengalir pada rangkaian luar, sedangkan ion menjaga kenetralan muatan. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengubah data menjadi kesimpulan yang terukur, transparan, dan langsung berhubungan dengan tindak lanjut, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Kerangka operasional

Aspek	Bukti yang dicari	Risiko interpretasi	Tindak lanjut
Target konsep	Penjelasan tentang elektrokimia	Target terlalu luas	Batasi subkonsep dan prasyarat
Respons	Ketepatan pilihan	Menyimpulkan miskonsepsi hanya dari satu jawaban	Membandingkan jawaban dengan alasan
Penalaran	Konsistensi lintas butir	Alasan diperlakukan sekadar tambahan	Meminta translasi antarrepresentasi
Konfirmasi	Predict-observe-explain atau tugas isomorfik	Kesimpulan dibuat dari satu konteks	Gunakan bukti kedua dan rediagnosis

Implikasinya, menghindari causal fallacy harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.11.21 Menghubungkan interpretasi dengan tindak lanjut

Aspek menghubungkan interpretasi dengan tindak lanjut menegaskan bahwa diagnosis adalah kegiatan ilmiah berskala kelas: guru merumuskan pertanyaan, mengumpulkan bukti, mempertimbangkan penjelasan alternatif, lalu menentukan kesimpulan yang paling masuk

akal. Inferensi harus proporsional terhadap kualitas data. Semakin spesifik label yang diberikan, semakin kuat pula bukti yang diperlukan (AERA et al., 2014; Treagust, 1988).

Secara operasional, guru dapat menggunakan tugas isomorfik untuk menelusuri koherensi alasan. Data tersebut perlu dibaca bersama konsistensi lintas representasi, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mencampur kesalahan matematika dengan kesalahan konsep. Untuk mengurangnya, guru dapat menggunakan pertanyaan tindak lanjut dan mengulang diagnosis setelah remediasi. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai menghubungkan interpretasi dengan tindak lanjut, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan menghubungkan interpretasi dengan tindak lanjut melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan menghubungkan interpretasi dengan tindak lanjut;
- Bandingkan bukti menghubungkan interpretasi dengan tindak lanjut pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek menghubungkan interpretasi dengan tindak lanjut dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian menghubungkan interpretasi dengan tindak lanjut, topik praktikum titrasi dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa titik akhir dianggap selalu identik dengan titik ekuivalen. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta analisis indikator, kurva titrasi, dan ketidakpastian. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa titik akhir adalah observasi operasional yang mendekati titik ekuivalen. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengubah data menjadi kesimpulan yang terukur, transparan, dan langsung berhubungan dengan tindak lanjut, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, menghubungkan interpretasi dengan tindak lanjut harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.11.22 Menyusun laporan diagnosis yang komunikatif

Penerapan menyusun laporan diagnosis yang komunikatif harus menghubungkan teori pengukuran dengan pengetahuan pedagogis kimia. Instrumen yang teknis tetapi tidak memahami struktur konsep akan menghasilkan kategori dangkal; sebaliknya, pengetahuan konten tanpa prosedur pengumpulan bukti dapat menghasilkan kesimpulan intuitif yang sulit diuji. Keduanya perlu dipadukan dalam satu alur yang transparan (Treagust, 1988; Black & Wiliam, 2009).

Secara operasional, guru dapat menggunakan kuis formatif digital untuk menelusuri tingkat keyakinan. Data tersebut perlu dibaca bersama kemampuan transfer, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menggunakan ambang kategori tanpa bukti. Untuk mengurangnya, guru dapat menyajikan tugas paralel dan mencatat derajat keyakinan. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai menyusun laporan diagnosis yang komunikatif, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan menyusun laporan diagnosis yang komunikatif melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan menyusun laporan diagnosis yang komunikatif;
- Bandingkan bukti menyusun laporan diagnosis yang komunikatif pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek menyusun laporan diagnosis yang komunikatif dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian menyusun laporan diagnosis yang komunikatif, topik sistem periodik dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa jari-jari atom dianggap selalu bertambah bersama nomor atom. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta prediksi tren disertai penjelasan muatan inti efektif dan kulit elektron. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa tren periodik perlu dijelaskan melalui struktur dan interaksi, bukan hafalan arah panah. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengubah data menjadi kesimpulan yang terukur, transparan, dan langsung berhubungan dengan tindak lanjut, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, menyusun laporan diagnosis yang komunikatif harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.11.23 Visualisasi hasil yang tidak menyesatkan

Visualisasi hasil yang tidak menyesatkan tidak berdiri sebagai prosedur administratif. Ia menjadi bagian dari keputusan tentang apa yang perlu dijelaskan ulang, representasi apa yang perlu dibandingkan, dan siapa yang memerlukan dukungan berbeda. Oleh sebab itu, bukti diagnosis harus langsung terhubung dengan tindakan pembelajaran, bukan berhenti pada daftar kesalahan (Black & Wiliam, 2009; Kane, 2013).

Secara operasional, guru dapat menggunakan tes tiga tingkat untuk menelusuri konsistensi lintas butir. Data tersebut perlu dibaca bersama stabilitas setelah jeda waktu, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah memperlakukan konsepsi sebagai sifat permanen. Untuk mengurangnya, guru dapat memeriksa konsep prasyarat dan menelusuri langkah penyelesaian. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai visualisasi hasil yang tidak menyesatkan, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan visualisasi hasil yang tidak menyesatkan melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan visualisasi hasil yang tidak menyesatkan;
- Bandingkan bukti visualisasi hasil yang tidak menyesatkan pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek visualisasi hasil yang tidak menyesatkan dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian visualisasi hasil yang tidak menyesatkan, topik stoikiometri dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa koefisien reaksi diperlakukan sebagai perbandingan massa langsung. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta translasi persamaan dari partikel ke mol dan massa. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa koefisien terutama menyatakan rasio jumlah entitas atau mol. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengubah data menjadi kesimpulan yang terukur, transparan, dan langsung berhubungan dengan tindak lanjut, bukan diperlakukan

sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, visualisasi hasil yang tidak menyesatkan harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.11.24 Contoh interpretasi satu butir diagnostik

Contoh interpretasi satu butir diagnostik perlu ditempatkan dalam kerangka diagnosis yang mengubah data menjadi kesimpulan yang terukur, transparan, dan langsung berhubungan dengan tindak lanjut. Fokusnya bukan sekadar menghasilkan skor, melainkan memperoleh bukti yang cukup untuk menjelaskan bagaimana peserta didik membangun, menghubungkan, dan menggunakan konsep. Karena konsepsi tidak dapat diamati secara langsung, guru harus menarik inferensi dari respons verbal, tulisan, gambar, tindakan, dan perubahan strategi yang tampak selama penyelesaian tugas (Kane, 2013; AERA et al., 2014).

Secara operasional, guru dapat menggunakan wawancara semi-terstruktur untuk menelusuri konsistensi lintas representasi. Data tersebut perlu dibaca bersama urutan langkah penyelesaian, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mengabaikan ketidakpastian pengukuran. Untuk menguranginya, guru dapat meminta translasi antarrepresentasi dan menggunakan triangulasi bukti. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai contoh interpretasi satu butir diagnostik, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan contoh interpretasi satu butir diagnostik melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan contoh interpretasi satu butir diagnostik;
- Bandingkan bukti contoh interpretasi satu butir diagnostik pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek contoh interpretasi satu butir diagnostik dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian contoh interpretasi satu butir diagnostik, topik asam-basa dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa asam kuat dianggap selalu berkonsentrasi tinggi. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan

relasi konseptual yang tepat. Guru dapat meminta pemisahan konsep derajat ionisasi, K_a , konsentrasi, dan pH. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa kekuatan asam dan konsentrasi adalah dimensi berbeda. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengubah data menjadi kesimpulan yang terukur, transparan, dan langsung berhubungan dengan tindak lanjut, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Kerangka operasional

Aspek	Bukti yang dicari	Risiko interpretasi	Tindak lanjut
Target konsep	Penjelasan tentang asam-basa	Target terlalu luas	Batasi subkonsep dan prasyarat
Respons	Konsistensi lintas representasi	Mengabaikan ketidakpastian pengukuran	Meminta translasi antarrepresentasi
Penalaran	Urutan langkah penyelesaian	Alasan diperlakukan sekadar tambahan	Menggunakan triangulasi bukti
Konfirmasi	Wawancara semi-terstruktur atau tugas isomorfik	Kesimpulan dibuat dari satu konteks	Gunakan bukti kedua dan rediagnosis

Implikasinya, contoh interpretasi satu butir diagnostik harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.11.25 Contoh interpretasi satu kelas

Pembahasan tentang contoh interpretasi satu kelas berangkat dari prinsip bahwa data asesmen selalu merupakan jejak tidak langsung dari proses berpikir. Satu pilihan jawaban hanya menunjukkan produk akhir, sedangkan struktur pengetahuan yang melahirkannya dapat berbeda. Diagnosis yang kuat karena itu menggabungkan kemampuan transfer, kualitas model mental yang dieksternalisasi, dan bukti kontekstual sehingga keputusan pedagogis tidak bergantung pada satu indikator saja (AERA et al., 2014; Treagust, 1988).

Secara operasional, guru dapat menggunakan gambar partikel untuk menelusuri kemampuan transfer. Data tersebut perlu dibaca bersama kualitas model mental yang dieksternalisasi, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama

pada tahap ini adalah menyimpulkan miskonsepsi hanya dari satu jawaban. Untuk mengurangnya, guru dapat mengulang diagnosis setelah remediasi dan menyusun tindak lanjut berbeda untuk kategori respons yang berbeda. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai contoh interpretasi satu kelas, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan contoh interpretasi satu kelas melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan contoh interpretasi satu kelas;
- Bandingkan bukti contoh interpretasi satu kelas pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek contoh interpretasi satu kelas dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian contoh interpretasi satu kelas, topik kinetika dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa reaksi eksoterm dianggap selalu cepat. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta perbandingan profil energi dan faktor kinetik. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa termodinamika dan kinetika menjawab pertanyaan berbeda. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengubah data menjadi kesimpulan yang terukur, transparan, dan langsung berhubungan dengan tindak lanjut, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, contoh interpretasi satu kelas harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.11.26 Sintesis interpretasi hasil diagnosis

Dalam pembelajaran kimia, sintesis interpretasi hasil diagnosis memiliki arti strategis karena banyak jawaban dapat diperoleh melalui hafalan prosedur atau pengenalan pola. Peserta didik mungkin tampak berhasil tetapi tidak mampu menerjemahkan makna simbol, model, atau besaran yang digunakan. Prinsip diagnosis menuntut bukti tentang cara konsep dipakai, bukan hanya apakah hasil akhirnya sesuai kunci (Treagust, 1988; Black & Wiliam, 2009).

Secara operasional, guru dapat menggunakan observasi praktikum untuk menelusuri stabilitas setelah jeda waktu. Data tersebut perlu dibaca bersama respons terhadap probing, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mencampur kesalahan matematika dengan kesalahan konsep. Untuk mengurangnya, guru dapat mencatat derajat keyakinan dan membandingkan jawaban dengan alasan. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai sintesis interpretasi hasil diagnosis, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan sintesis interpretasi hasil diagnosis melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan sintesis interpretasi hasil diagnosis;
- Bandingkan bukti sintesis interpretasi hasil diagnosis pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek sintesis interpretasi hasil diagnosis dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian sintesis interpretasi hasil diagnosis, topik elektrokimia dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa elektron dianggap bergerak melalui jembatan garam. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta penelusuran jalur elektron dan ion pada sel galvanik. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa elektron mengalir pada rangkaian luar, sedangkan ion menjaga kenetralan muatan. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengubah data menjadi kesimpulan yang terukur, transparan, dan langsung berhubungan dengan tindak lanjut, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, sintesis interpretasi hasil diagnosis harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.11.27 Implikasi praktis bagi guru kimia

Implikasi praktis subbab 5.11 dapat dirangkum sebagai alur: tentukan konsep target, pilih bukti yang relevan, kumpulkan respons dalam lebih dari satu bentuk, analisis pola, nyatakan

derajat kepastian, susun tindak lanjut, dan lakukan rediagnosis. Alur ini menjaga agar interpretasi hasil diagnosis tidak berubah menjadi administrasi skor, tetapi berfungsi sebagai mekanisme perbaikan pembelajaran yang berulang.

71. Identifikasi konsep inti dan konsep prasyarat sebelum menyusun instrumen.
72. Gunakan pertanyaan yang meminta alasan, representasi, atau langkah kerja.
73. Bandingkan respons pada beberapa konteks dan waktu.
74. Pisahkan bukti faktual dari interpretasi guru.
75. Tentukan tindak lanjut yang berbeda untuk kategori kebutuhan yang berbeda.
76. Simpan data secukupnya dan lindungi kerahasiaan peserta didik.
77. Periksa kembali perubahan melalui tugas transfer dan retensi.

5.12 Penyusunan Profil Konseptual Peserta Didik

Subbab ini mengorganisasikan bukti multidimensi menjadi profil perkembangan konsep yang dinamis dan tidak stigmatis. Uraian disusun dari prinsip teoretis menuju prosedur kelas, contoh kimia, risiko interpretasi, dan tindak lanjut. Setiap bagian menekankan bahwa kesimpulan diagnostik harus didasarkan pada pola bukti, bukan pada satu respons yang terpisah.

5.12.1 Pengertian profil konseptual

Pembahasan tentang pengertian profil konseptual berangkat dari prinsip bahwa data asesmen selalu merupakan jejak tidak langsung dari proses berpikir. Satu pilihan jawaban hanya menunjukkan produk akhir, sedangkan struktur pengetahuan yang melahirkannya dapat berbeda. Diagnosis yang kuat karena itu menggabungkan koherensi alasan, konsistensi lintas representasi, dan bukti kontekstual sehingga keputusan pedagogis tidak bergantung pada satu indikator saja (Mavridi et al., 2026; Taber, 2002a).

Secara operasional, guru dapat menggunakan exit ticket untuk menelusuri koherensi alasan. Data tersebut perlu dibaca bersama konsistensi lintas representasi, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mencampur kesalahan matematika dengan kesalahan konsep. Untuk mengurangnya, guru dapat menggunakan pertanyaan tindak lanjut dan mengulang diagnosis setelah remediasi. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai pengertian profil konseptual, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan pengertian profil konseptual melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan pengertian profil konseptual;

- Bandingkan bukti pengertian profil konseptual pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek pengertian profil konseptual dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian pengertian profil konseptual, topik katalis dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa katalis dianggap menambah jumlah produk kesetimbangan. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta analisis jalur energi dan laju maju-balik. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa katalis mempercepat tercapainya kesetimbangan tanpa mengubah K . Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengorganisasikan bukti multidimensi menjadi profil perkembangan konsep yang dinamis dan tidak stigmatis, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, pengertian profil konseptual harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.12.2 Tujuan penyusunan profil konseptual

Dalam pembelajaran kimia, tujuan penyusunan profil konseptual memiliki arti strategis karena banyak jawaban dapat diperoleh melalui hafalan prosedur atau pengenalan pola. Peserta didik mungkin tampak berhasil tetapi tidak mampu menerjemahkan makna simbol, model, atau besaran yang digunakan. Prinsip diagnosis menuntut bukti tentang cara konsep dipakai, bukan hanya apakah hasil akhirnya sesuai kunci (Taber, 2002a; Vosniadou, 1994).

Secara operasional, guru dapat menggunakan tes pilihan ganda dua tingkat untuk menelusuri tingkat keyakinan. Data tersebut perlu dibaca bersama kemampuan transfer, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menggunakan ambang kategori tanpa bukti. Untuk mengurangnya, guru dapat menyajikan tugas paralel dan mencatat derajat keyakinan. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai tujuan penyusunan profil konseptual, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan tujuan penyusunan profil konseptual melalui penggunaan konsep, bukan sekadar pengenalan istilah;

- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan tujuan penyusunan profil konseptual;
- Bandingkan bukti tujuan penyusunan profil konseptual pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek tujuan penyusunan profil konseptual dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian tujuan penyusunan profil konseptual, topik elektrolisis dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa anoda selalu dianggap bermuatan negatif. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta pembedaan definisi anoda-katoda berdasarkan proses oksidasi-reduksi. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa tanda elektroda bergantung jenis sel, tetapi anoda tetap lokasi oksidasi. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengorganisasikan bukti multidimensi menjadi profil perkembangan konsep yang dinamis dan tidak stigmatis, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, tujuan penyusunan profil konseptual harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.12.3 Unit konsep dan subkonsep dalam profil

Aspek unit konsep dan subkonsep dalam profil menegaskan bahwa diagnosis adalah kegiatan ilmiah berskala kelas: guru merumuskan pertanyaan, mengumpulkan bukti, mempertimbangkan penjelasan alternatif, lalu menentukan kesimpulan yang paling masuk akal. Inferensi harus proporsional terhadap kualitas data. Semakin spesifik label yang diberikan, semakin kuat pula bukti yang diperlukan (Vosniadou, 1994; Mortimer, 1995).

Secara operasional, guru dapat menggunakan alasan terbuka untuk menelusuri konsistensi lintas butir. Data tersebut perlu dibaca bersama stabilitas setelah jeda waktu, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah memperlakukan konsepsi sebagai sifat permanen. Untuk menguranginya, guru dapat memeriksa konsep prasyarat dan menelusuri langkah penyelesaian. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai unit konsep dan subkonsep dalam profil, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan unit konsep dan subkonsep dalam profil melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan unit konsep dan subkonsep dalam profil;
- Bandingkan bukti unit konsep dan subkonsep dalam profil pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek unit konsep dan subkonsep dalam profil dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian unit konsep dan subkonsep dalam profil, topik pengukuran laboratorium dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa angka lebih banyak selalu dianggap lebih akurat. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta pembedaan akurasi, presisi, resolusi, dan angka penting. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa kualitas data ditentukan oleh desain, kalibrasi, dan ketidakpastian. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengorganisasikan bukti multidimensi menjadi profil perkembangan konsep yang dinamis dan tidak stigmatis, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, unit konsep dan subkonsep dalam profil harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.12.4 Dimensi ketepatan jawaban

Penerapan dimensi ketepatan jawaban harus menghubungkan teori pengukuran dengan pengetahuan pedagogis kimia. Instrumen yang teknis tetapi tidak memahami struktur konsep akan menghasilkan kategori dangkal; sebaliknya, pengetahuan konten tanpa prosedur pengumpulan bukti dapat menghasilkan kesimpulan intuitif yang sulit diuji. Keduanya perlu dipadukan dalam satu alur yang transparan (Mortimer, 1995; Mavridi et al., 2026).

Secara operasional, guru dapat menggunakan peta konsep untuk menelusuri konsistensi lintas representasi. Data tersebut perlu dibaca bersama urutan langkah penyelesaian, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mengabaikan ketidakpastian pengukuran. Untuk menguranginya, guru dapat

meminta translasi antarrepresentasi dan menggunakan triangulasi bukti. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai dimensi ketepatan jawaban, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan dimensi ketepatan jawaban melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan dimensi ketepatan jawaban;
- Bandingkan bukti dimensi ketepatan jawaban pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek dimensi ketepatan jawaban dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian dimensi ketepatan jawaban, topik ikatan ion dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa NaCl dipandang sebagai molekul diskrit. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta gambar partikel padatan NaCl dan penjelasan hubungan struktur-kisi dengan titik leleh. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa senyawa ionik membentuk kisi dan rumusnya menyatakan rasio ion. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengorganisasikan bukti multidimensi menjadi profil perkembangan konsep yang dinamis dan tidak stigmatis, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Kerangka operasional

Aspek	Bukti yang dicari	Risiko interpretasi	Tindak lanjut
Target konsep	Penjelasan tentang ikatan ion	Target terlalu luas	Batasi subkonsep dan prasyarat
Respons	Konsistensi lintas representasi	Mengabaikan ketidakpastian pengukuran	Meminta translasi antarrepresentasi
Penalaran	Urutan langkah penyelesaian	Alasan diperlakukan sekadar tambahan	Menggunakan triangulasi bukti
Konfirmasi	Peta konsep atau tugas isomorfik	Kesimpulan dibuat dari satu konteks	Gunakan bukti kedua dan rediagnosis

Implikasinya, dimensi ketepatan jawaban harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.12.5 Dimensi kualitas alasan

Dimensi kualitas alasan tidak berdiri sebagai prosedur administratif. Ia menjadi bagian dari keputusan tentang apa yang perlu dijelaskan ulang, representasi apa yang perlu dibandingkan, dan siapa yang memerlukan dukungan berbeda. Oleh sebab itu, bukti diagnosis harus langsung terhubung dengan tindakan pembelajaran, bukan berhenti pada daftar kesalahan (Mavridi et al., 2026; Taber, 2002a).

Secara operasional, guru dapat menggunakan analisis respons tertulis untuk menelusuri kemampuan transfer. Data tersebut perlu dibaca bersama kualitas model mental yang dieksternalisasi, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menyimpulkan miskonsepsi hanya dari satu jawaban. Untuk mengurangnya, guru dapat mengulang diagnosis setelah remediasi dan menyusun tindak lanjut berbeda untuk kategori respons yang berbeda. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai dimensi kualitas alasan, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan dimensi kualitas alasan melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan dimensi kualitas alasan;
- Bandingkan bukti dimensi kualitas alasan pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek dimensi kualitas alasan dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian dimensi kualitas alasan, topik pereaksi pembatas dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa pereaksi dengan massa terkecil selalu dianggap pembatas. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta analisis rasio mol terhadap koefisien. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa status pembatas ditentukan oleh hubungan jumlah zat dan stoikiometri. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengorganisasikan bukti multidimensi menjadi profil perkembangan konsep yang dinamis dan tidak stigmatis, bukan diperlakukan sebagai skor yang berdiri sendiri.

Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, dimensi kualitas alasan harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.12.6 Dimensi keyakinan

Dimensi keyakinan perlu ditempatkan dalam kerangka diagnosis yang mengorganisasikan bukti multidimensi menjadi profil perkembangan konsep yang dinamis dan tidak stigmatis. Fokusnya bukan sekadar menghasilkan skor, melainkan memperoleh bukti yang cukup untuk menjelaskan bagaimana peserta didik membangun, menghubungkan, dan menggunakan konsep. Karena konsepsi tidak dapat diamati secara langsung, guru harus menarik inferensi dari respons verbal, tulisan, gambar, tindakan, dan perubahan strategi yang tampak selama penyelesaian tugas (Taber, 2002a; Vosniadou, 1994).

Secara operasional, guru dapat menggunakan diskusi argumentatif untuk menelusuri stabilitas setelah jeda waktu. Data tersebut perlu dibaca bersama respons terhadap probing, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mencampur kesalahan matematika dengan kesalahan konsep. Untuk menguranginya, guru dapat mencatat derajat keyakinan dan membandingkan jawaban dengan alasan. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai dimensi keyakinan, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan dimensi keyakinan melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan dimensi keyakinan;
- Bandingkan bukti dimensi keyakinan pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek dimensi keyakinan dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian dimensi keyakinan, topik larutan penyangga dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa penyangga dianggap mempertahankan pH secara mutlak. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta analisis perubahan komposisi setelah

penambahan asam atau basa. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa penyangga membatasi perubahan dalam kapasitas tertentu, bukan meniadakannya. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengorganisasikan bukti multidimensi menjadi profil perkembangan konsep yang dinamis dan tidak stigmatis, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, dimensi keyakinan harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.12.7 Dimensi konsistensi

Pembahasan tentang dimensi konsistensi berangkat dari prinsip bahwa data asesmen selalu merupakan jejak tidak langsung dari proses berpikir. Satu pilihan jawaban hanya menunjukkan produk akhir, sedangkan struktur pengetahuan yang melahirkannya dapat berbeda. Diagnosis yang kuat karena itu menggabungkan urutan langkah penyelesaian, ketepatan pilihan, dan bukti kontekstual sehingga keputusan pedagogis tidak bergantung pada satu indikator saja (Vosniadou, 1994; Mortimer, 1995).

Secara operasional, guru dapat menggunakan analisis grafik untuk menelusuri urutan langkah penyelesaian. Data tersebut perlu dibaca bersama ketepatan pilihan, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menggunakan ambang kategori tanpa bukti. Untuk mengurangnya, guru dapat menelusuri langkah penyelesaian dan menggunakan pertanyaan tindak lanjut. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai dimensi konsistensi, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan dimensi konsistensi melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan dimensi konsistensi;
- Bandingkan bukti dimensi konsistensi pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek dimensi konsistensi dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian dimensi konsistensi, topik katalis dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa katalis dianggap

menambah jumlah produk kesetimbangan. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta analisis jalur energi dan laju maju-balik. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa katalis mempercepat tercapainya kesetimbangan tanpa mengubah K. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengorganisasikan bukti multidimensi menjadi profil perkembangan konsep yang dinamis dan tidak stigmatis, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, dimensi konsistensi harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.12.8 Dimensi representasi

Dalam pembelajaran kimia, dimensi representasi memiliki arti strategis karena banyak jawaban dapat diperoleh melalui hafalan prosedur atau pengenalan pola. Peserta didik mungkin tampak berhasil tetapi tidak mampu menerjemahkan makna simbol, model, atau besaran yang digunakan. Prinsip diagnosis menuntut bukti tentang cara konsep dipakai, bukan hanya apakah hasil akhirnya sesuai kunci (Mortimer, 1995; Mavridi et al., 2026).

Secara operasional, guru dapat menggunakan tes empat tingkat untuk menelusuri kualitas model mental yang dieksternalisasi. Data tersebut perlu dibaca bersama koherensi alasan, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah memperlakukan konsepsi sebagai sifat permanen. Untuk menguranginya, guru dapat menggunakan triangulasi bukti dan menyajikan tugas paralel. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai dimensi representasi, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan dimensi representasi melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan dimensi representasi;
- Bandingkan bukti dimensi representasi pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek dimensi representasi dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian dimensi representasi, topik elektrolisis dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa anoda selalu dianggap bermuatan negatif. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta pembedaan definisi anoda-katoda berdasarkan proses oksidasi-reduksi. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa tanda elektroda bergantung jenis sel, tetapi anoda tetap lokasi oksidasi. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengorganisasikan bukti multidimensi menjadi profil perkembangan konsep yang dinamis dan tidak stigmatis, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Kerangka operasional

Aspek	Bukti yang dicari	Risiko interpretasi	Tindak lanjut
Target konsep	Penjelasan tentang elektrolisis	Target terlalu luas	Batasi subkonsep dan prasyarat
Respons	Kualitas model mental yang dieksternalisasi	Memperlakukan konsepsi sebagai sifat permanen	Menggunakan triangulasi bukti
Penalaran	Koherensi alasan	Alasan diperlakukan sekadar tambahan	Menyajikan tugas paralel
Konfirmasi	Tes empat tingkat atau tugas isomorfik	Kesimpulan dibuat dari satu konteks	Gunakan bukti kedua dan rediagnosis

Implikasinya, dimensi representasi harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.12.9 Dimensi transfer

Aspek dimensi transfer menegaskan bahwa diagnosis adalah kegiatan ilmiah berskala kelas: guru merumuskan pertanyaan, mengumpulkan bukti, mempertimbangkan penjelasan alternatif, lalu menentukan kesimpulan yang paling masuk akal. Inferensi harus proporsional terhadap kualitas data. Semakin spesifik label yang diberikan, semakin kuat pula bukti yang diperlukan (Mavridi et al., 2026; Taber, 2002a).

Secara operasional, guru dapat menggunakan think-aloud untuk menelusuri respons terhadap probing. Data tersebut perlu dibaca bersama tingkat keyakinan, karena pola yang sama pada

permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mengabaikan ketidakpastian pengukuran. Untuk menguranginya, guru dapat menyusun tindak lanjut berbeda untuk kategori respons yang berbeda dan memeriksa konsep prasyarat. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai dimensi transfer, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan dimensi transfer melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan dimensi transfer;
- Bandingkan bukti dimensi transfer pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek dimensi transfer dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian dimensi transfer, topik pengukuran laboratorium dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa angka lebih banyak selalu dianggap lebih akurat. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta pembedaan akurasi, presisi, resolusi, dan angka penting. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa kualitas data ditentukan oleh desain, kalibrasi, dan ketidakpastian. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengorganisasikan bukti multidimensi menjadi profil perkembangan konsep yang dinamis dan tidak stigmatis, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, dimensi transfer harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.12.10 Dimensi retensi

Penerapan dimensi retensi harus menghubungkan teori pengukuran dengan pengetahuan pedagogis kimia. Instrumen yang teknis tetapi tidak memahami struktur konsep akan menghasilkan kategori dangkal; sebaliknya, pengetahuan konten tanpa prosedur pengumpulan bukti dapat menghasilkan kesimpulan intuitif yang sulit diuji. Keduanya perlu dipadukan dalam satu alur yang transparan (Taber, 2002a; Vosniadou, 1994).

Secara operasional, guru dapat menggunakan predict-observe-explain untuk menelusuri ketepatan pilihan. Data tersebut perlu dibaca bersama konsistensi lintas butir, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menyimpulkan miskonsepsi hanya dari satu jawaban. Untuk mengurangnya, guru dapat membandingkan jawaban dengan alasan dan meminta translasi antarrepresentasi. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai dimensi retensi, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan dimensi retensi melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan dimensi retensi;
- Bandingkan bukti dimensi retensi pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek dimensi retensi dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian dimensi retensi, topik ikatan ion dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa NaCl dipandang sebagai molekul diskrit. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta gambar partikel padatan NaCl dan penjelasan hubungan struktur-kisi dengan titik leleh. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa senyawa ionik membentuk kisi dan rumusnya menyatakan rasio ion. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengorganisasikan bukti multidimensi menjadi profil perkembangan konsep yang dinamis dan tidak stigmatis, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, dimensi retensi harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.12.11 Dimensi sumber miskonsepsi

Dimensi sumber miskonsepsi tidak berdiri sebagai prosedur administratif. Ia menjadi bagian dari keputusan tentang apa yang perlu dijelaskan ulang, representasi apa yang perlu dibandingkan, dan siapa yang memerlukan dukungan berbeda. Oleh sebab itu, bukti diagnosis

harus langsung terhubung dengan tindakan pembelajaran, bukan berhenti pada daftar kesalahan (Vosniadou, 1994; Mortimer, 1995).

Secara operasional, guru dapat menggunakan tugas isomorfik untuk menelusuri koherensi alasan. Data tersebut perlu dibaca bersama konsistensi lintas representasi, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mencampur kesalahan matematika dengan kesalahan konsep. Untuk mengurangnya, guru dapat menggunakan pertanyaan tindak lanjut dan mengulang diagnosis setelah remediasi. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai dimensi sumber miskonsepsi, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan dimensi sumber miskonsepsi melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan dimensi sumber miskonsepsi;
- Bandingkan bukti dimensi sumber miskonsepsi pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek dimensi sumber miskonsepsi dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian dimensi sumber miskonsepsi, topik pereaksi pembatas dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa pereaksi dengan massa terkecil selalu dianggap pembatas. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta analisis rasio mol terhadap koefisien. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa status pembatas ditentukan oleh hubungan jumlah zat dan stoikiometri. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengorganisasikan bukti multidimensi menjadi profil perkembangan konsep yang dinamis dan tidak stigmatis, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, dimensi sumber miskonsepsi harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.12.12 Dimensi konsep prasyarat

Dimensi konsep prasyarat perlu ditempatkan dalam kerangka diagnosis yang mengorganisasikan bukti multidimensi menjadi profil perkembangan konsep yang dinamis dan tidak stigmatis. Fokusnya bukan sekadar menghasilkan skor, melainkan memperoleh bukti yang cukup untuk menjelaskan bagaimana peserta didik membangun, menghubungkan, dan menggunakan konsep. Karena konsepsi tidak dapat diamati secara langsung, guru harus menarik inferensi dari respons verbal, tulisan, gambar, tindakan, dan perubahan strategi yang tampak selama penyelesaian tugas (Mortimer, 1995; Mavridi et al., 2026).

Secara operasional, guru dapat menggunakan kuis formatif digital untuk menelusuri tingkat keyakinan. Data tersebut perlu dibaca bersama kemampuan transfer, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menggunakan ambang kategori tanpa bukti. Untuk menguranginya, guru dapat menyajikan tugas paralel dan mencatat derajat keyakinan. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai dimensi konsep prasyarat, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan dimensi konsep prasyarat melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan dimensi konsep prasyarat;
- Bandingkan bukti dimensi konsep prasyarat pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek dimensi konsep prasyarat dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian dimensi konsep prasyarat, topik larutan penyangga dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa penyangga dianggap mempertahankan pH secara mutlak. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta analisis perubahan komposisi setelah penambahan asam atau basa. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa penyangga membatasi perubahan dalam kapasitas tertentu, bukan meniadakannya. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengorganisasikan bukti multidimensi menjadi profil perkembangan konsep yang dinamis dan tidak stigmatis, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Kerangka operasional

Aspek	Bukti yang dicari	Risiko interpretasi	Tindak lanjut
Target konsep	Penjelasan tentang larutan penyangga	Target terlalu luas	Batasi subkonsep dan prasyarat
Respons	Tingkat keyakinan	Menggunakan ambang kategori tanpa bukti	Menyajikan tugas paralel
Penalaran	Kemampuan transfer	Alasan diperlakukan sekadar tambahan	Mencatat derajat keyakinan
Konfirmasi	Kuis formatif digital atau tugas isomorfik	Kesimpulan dibuat dari satu konteks	Gunakan bukti kedua dan rediagnosis

Implikasinya, dimensi konsep prasyarat harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.12.13 Kategori paham, parsial, tidak tahu, dan miskonsepsi

Pembahasan tentang kategori paham, parsial, tidak tahu, dan miskonsepsi berangkat dari prinsip bahwa data asesmen selalu merupakan jejak tidak langsung dari proses berpikir. Satu pilihan jawaban hanya menunjukkan produk akhir, sedangkan struktur pengetahuan yang melahirkannya dapat berbeda. Diagnosis yang kuat karena itu menggabungkan konsistensi lintas butir, stabilitas setelah jeda waktu, dan bukti kontekstual sehingga keputusan pedagogis tidak bergantung pada satu indikator saja (Mavridi et al., 2026; Taber, 2002a).

Secara operasional, guru dapat menggunakan tes tiga tingkat untuk menelusuri konsistensi lintas butir. Data tersebut perlu dibaca bersama stabilitas setelah jeda waktu, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah memperlakukan konsepsi sebagai sifat permanen. Untuk mengurangnya, guru dapat memeriksa konsep prasyarat dan menelusuri langkah penyelesaian. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai kategori paham, parsial, tidak tahu, dan miskonsepsi, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan kategori paham, parsial, tidak tahu, dan miskonsepsi melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan kategori paham, parsial, tidak tahu, dan miskonsepsi;

- Bandingkan bukti kategori paham, parsial, tidak tahu, dan miskonsepsi pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek kategori paham, parsial, tidak tahu, dan miskonsepsi dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian kategori paham, parsial, tidak tahu, dan miskonsepsi, topik katalis dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa katalis dianggap menambah jumlah produk kesetimbangan. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta analisis jalur energi dan laju maju-balik. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa katalis mempercepat tercapainya kesetimbangan tanpa mengubah K. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengorganisasikan bukti multidimensi menjadi profil perkembangan konsep yang dinamis dan tidak stigmatis, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, kategori paham, parsial, tidak tahu, dan miskonsepsi harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.12.14 Profil konseptual berbasis matriks

Dalam pembelajaran kimia, profil konseptual berbasis matriks memiliki arti strategis karena banyak jawaban dapat diperoleh melalui hafalan prosedur atau pengenalan pola. Peserta didik mungkin tampak berhasil tetapi tidak mampu menerjemahkan makna simbol, model, atau besaran yang digunakan. Prinsip diagnosis menuntut bukti tentang cara konsep dipakai, bukan hanya apakah hasil akhirnya sesuai kunci (Taber, 2002a; Vosniadou, 1994).

Secara operasional, guru dapat menggunakan wawancara semi-terstruktur untuk menelusuri konsistensi lintas representasi. Data tersebut perlu dibaca bersama urutan langkah penyelesaian, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mengabaikan ketidakpastian pengukuran. Untuk menguranginya, guru dapat meminta translasi antarrepresentasi dan menggunakan triangulasi bukti. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai profil konseptual berbasis matriks, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;

- Pilih tugas yang merepresentasikan profil konseptual berbasis matriks melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan profil konseptual berbasis matriks;
- Bandingkan bukti profil konseptual berbasis matriks pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek profil konseptual berbasis matriks dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian profil konseptual berbasis matriks, topik elektrolisis dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa anoda selalu dianggap bermuatan negatif. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta pembedaan definisi anoda-katoda berdasarkan proses oksidasi-reduksi. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa tanda elektroda bergantung jenis sel, tetapi anoda tetap lokasi oksidasi. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengorganisasikan bukti multidimensi menjadi profil perkembangan konsep yang dinamis dan tidak stigmatis, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, profil konseptual berbasis matriks harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.12.15 Profil konseptual berbasis narasi

Aspek profil konseptual berbasis narasi menegaskan bahwa diagnosis adalah kegiatan ilmiah berskala kelas: guru merumuskan pertanyaan, mengumpulkan bukti, mempertimbangkan penjelasan alternatif, lalu menentukan kesimpulan yang paling masuk akal. Inferensi harus proporsional terhadap kualitas data. Semakin spesifik label yang diberikan, semakin kuat pula bukti yang diperlukan (Vosniadou, 1994; Mortimer, 1995).

Secara operasional, guru dapat menggunakan gambar partikel untuk menelusuri kemampuan transfer. Data tersebut perlu dibaca bersama kualitas model mental yang dieksternalisasi, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menyimpulkan miskonsepsi hanya dari satu jawaban. Untuk menguranginya, guru dapat mengulang diagnosis setelah remediasi dan menyusun tindak lanjut berbeda untuk kategori respons yang berbeda. Prosedur tersebut membantu

memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai profil konseptual berbasis narasi, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan profil konseptual berbasis narasi melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan profil konseptual berbasis narasi;
- Bandingkan bukti profil konseptual berbasis narasi pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek profil konseptual berbasis narasi dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian profil konseptual berbasis narasi, topik pengukuran laboratorium dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa angka lebih banyak selalu dianggap lebih akurat. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta pembedaan akurasi, presisi, resolusi, dan angka penting. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa kualitas data ditentukan oleh desain, kalibrasi, dan ketidakpastian. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengorganisasikan bukti multidimensi menjadi profil perkembangan konsep yang dinamis dan tidak stigmatis, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, profil konseptual berbasis narasi harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.12.16 Profil konseptual berbasis peta jaringan

Penerapan profil konseptual berbasis peta jaringan harus menghubungkan teori pengukuran dengan pengetahuan pedagogis kimia. Instrumen yang teknis tetapi tidak memahami struktur konsep akan menghasilkan kategori dangkal; sebaliknya, pengetahuan konten tanpa prosedur pengumpulan bukti dapat menghasilkan kesimpulan intuitif yang sulit diuji. Keduanya perlu dipadukan dalam satu alur yang transparan (Mortimer, 1995; Mavridi et al., 2026).

Secara operasional, guru dapat menggunakan observasi praktikum untuk menelusuri stabilitas setelah jeda waktu. Data tersebut perlu dibaca bersama respons terhadap probing, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada

tahap ini adalah mencampur kesalahan matematika dengan kesalahan konsep. Untuk mengurangnya, guru dapat mencatat derajat keyakinan dan membandingkan jawaban dengan alasan. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai profil konseptual berbasis peta jaringan, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan profil konseptual berbasis peta jaringan melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan profil konseptual berbasis peta jaringan;
- Bandingkan bukti profil konseptual berbasis peta jaringan pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek profil konseptual berbasis peta jaringan dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian profil konseptual berbasis peta jaringan, topik ikatan ion dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa NaCl dipandang sebagai molekul diskrit. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta gambar partikel padatan NaCl dan penjelasan hubungan struktur-kisi dengan titik leleh. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa senyawa ionik membentuk kisi dan rumusnya menyatakan rasio ion. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengorganisasikan bukti multidimensi menjadi profil perkembangan konsep yang dinamis dan tidak stigmatis, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Kerangka operasional

Aspek	Bukti yang dicari	Risiko interpretasi	Tindak lanjut
Target konsep	Penjelasan tentang ikatan ion	Target terlalu luas	Batasi subkonsep dan prasyarat
Respons	Stabilitas setelah jeda waktu	Mencampur kesalahan matematika dengan kesalahan konsep	Mencatat derajat keyakinan
Penalaran	Respons terhadap probing	Alasan diperlakukan sekadar tambahan	Membandingkan jawaban dengan alasan
Konfirmasi	Observasi praktikum atau tugas isomorfik	Kesimpulan dibuat dari satu konteks	Gunakan bukti kedua dan rediagnosis

Implikasinya, profil konseptual berbasis peta jaringan harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.12.17 Profil konseptual longitudinal

Profil konseptual longitudinal tidak berdiri sebagai prosedur administratif. Ia menjadi bagian dari keputusan tentang apa yang perlu dijelaskan ulang, representasi apa yang perlu dibandingkan, dan siapa yang memerlukan dukungan berbeda. Oleh sebab itu, bukti diagnosis harus langsung terhubung dengan tindakan pembelajaran, bukan berhenti pada daftar kesalahan (Mavridi et al., 2026; Taber, 2002a).

Secara operasional, guru dapat menggunakan exit ticket untuk menelusuri urutan langkah penyelesaian. Data tersebut perlu dibaca bersama ketepatan pilihan, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menggunakan ambang kategori tanpa bukti. Untuk mengurangnya, guru dapat menelusuri langkah penyelesaian dan menggunakan pertanyaan tindak lanjut. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai profil konseptual longitudinal, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan profil konseptual longitudinal melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan profil konseptual longitudinal;
- Bandingkan bukti profil konseptual longitudinal pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek profil konseptual longitudinal dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian profil konseptual longitudinal, topik pereaksi pembatas dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa pereaksi dengan massa terkecil selalu dianggap pembatas. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta analisis rasio mol terhadap koefisien. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa status pembatas ditentukan oleh hubungan jumlah zat dan stoikiometri. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengorganisasikan bukti multidimensi menjadi profil perkembangan konsep yang dinamis dan tidak stigmatis, bukan diperlakukan sebagai skor yang berdiri sendiri.

Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, profil konseptual longitudinal harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.12.18 Profil individual dan profil kelompok

Profil individual dan profil kelompok perlu ditempatkan dalam kerangka diagnosis yang mengorganisasikan bukti multidimensi menjadi profil perkembangan konsep yang dinamis dan tidak stigmatis. Fokusnya bukan sekadar menghasilkan skor, melainkan memperoleh bukti yang cukup untuk menjelaskan bagaimana peserta didik membangun, menghubungkan, dan menggunakan konsep. Karena konsepsi tidak dapat diamati secara langsung, guru harus menarik inferensi dari respons verbal, tulisan, gambar, tindakan, dan perubahan strategi yang tampak selama penyelesaian tugas (Taber, 2002a; Vosniadou, 1994).

Secara operasional, guru dapat menggunakan tes pilihan ganda dua tingkat untuk menelusuri kualitas model mental yang dieksternalisasi. Data tersebut perlu dibaca bersama koherensi alasan, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah memperlakukan konsepsi sebagai sifat permanen. Untuk mengurangnya, guru dapat menggunakan triangulasi bukti dan menyajikan tugas paralel. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai profil individual dan profil kelompok, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan profil individual dan profil kelompok melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan profil individual dan profil kelompok;
- Bandingkan bukti profil individual dan profil kelompok pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek profil individual dan profil kelompok dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian profil individual dan profil kelompok, topik larutan penyangga dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa penyangga dianggap mempertahankan pH secara mutlak. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta analisis perubahan

komposisi setelah penambahan asam atau basa. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa penyangga membatasi perubahan dalam kapasitas tertentu, bukan meniadakannya. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengorganisasikan bukti multidimensi menjadi profil perkembangan konsep yang dinamis dan tidak stigmatis, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, profil individual dan profil kelompok harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.12.19 Pemutakhiran profil setelah remediasi

Pembahasan tentang pemutakhiran profil setelah remediasi berangkat dari prinsip bahwa data asesmen selalu merupakan jejak tidak langsung dari proses berpikir. Satu pilihan jawaban hanya menunjukkan produk akhir, sedangkan struktur pengetahuan yang melahirkannya dapat berbeda. Diagnosis yang kuat karena itu menggabungkan respons terhadap probing, tingkat keyakinan, dan bukti kontekstual sehingga keputusan pedagogis tidak bergantung pada satu indikator saja (Vosniadou, 1994; Mortimer, 1995).

Secara operasional, guru dapat menggunakan alasan terbuka untuk menelusuri respons terhadap probing. Data tersebut perlu dibaca bersama tingkat keyakinan, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mengabaikan ketidakpastian pengukuran. Untuk mengurangnya, guru dapat menyusun tindak lanjut berbeda untuk kategori respons yang berbeda dan memeriksa konsep prasyarat. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai pemutakhiran profil setelah remediasi, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan pemutakhiran profil setelah remediasi melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan pemutakhiran profil setelah remediasi;
- Bandingkan bukti pemutakhiran profil setelah remediasi pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek pemutakhiran profil setelah remediasi dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian pemutakhiran profil setelah remediasi, topik katalis dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa katalis dianggap menambah jumlah produk kesetimbangan. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta analisis jalur energi dan laju maju-balik. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa katalis mempercepat tercapainya kesetimbangan tanpa mengubah K . Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengorganisasikan bukti multidimensi menjadi profil perkembangan konsep yang dinamis dan tidak stigmatis, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, pemutakhiran profil setelah remediasi harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.12.20 Keterhubungan profil dengan diferensiasi pembelajaran

Dalam pembelajaran kimia, keterhubungan profil dengan diferensiasi pembelajaran memiliki arti strategis karena banyak jawaban dapat diperoleh melalui hafalan prosedur atau pengenalan pola. Peserta didik mungkin tampak berhasil tetapi tidak mampu menerjemahkan makna simbol, model, atau besaran yang digunakan. Prinsip diagnosis menuntut bukti tentang cara konsep dipakai, bukan hanya apakah hasil akhirnya sesuai kunci (Mortimer, 1995; Mavridi et al., 2026).

Secara operasional, guru dapat menggunakan peta konsep untuk menelusuri ketepatan pilihan. Data tersebut perlu dibaca bersama konsistensi lintas butir, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menyimpulkan miskonsepsi hanya dari satu jawaban. Untuk menguranginya, guru dapat membandingkan jawaban dengan alasan dan meminta translasi antarrepresentasi. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai keterhubungan profil dengan diferensiasi pembelajaran, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan keterhubungan profil dengan diferensiasi pembelajaran melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan keterhubungan profil dengan diferensiasi pembelajaran;

- Bandingkan bukti keterhubungan profil dengan diferensiasi pembelajaran pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek keterhubungan profil dengan diferensiasi pembelajaran dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian keterhubungan profil dengan diferensiasi pembelajaran, topik elektrolisis dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa anoda selalu dianggap bermuatan negatif. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta pembedaan definisi anoda-katoda berdasarkan proses oksidasi-reduksi. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa tanda elektroda bergantung jenis sel, tetapi anoda tetap lokasi oksidasi. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengorganisasikan bukti multidimensi menjadi profil perkembangan konsep yang dinamis dan tidak stigmatis, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Kerangka operasional

Aspek	Bukti yang dicari	Risiko interpretasi	Tindak lanjut
Target konsep	Penjelasan tentang elektrolisis	Target terlalu luas	Batasi subkonsep dan prasyarat
Respons	Ketepatan pilihan	Menyimpulkan miskonsepsi hanya dari satu jawaban	Membandingkan jawaban dengan alasan
Penalaran	Konsistensi lintas butir	Alasan diperlakukan sekadar tambahan	Meminta translasi antarrepresentasi
Konfirmasi	Peta konsep atau tugas isomorfik	Kesimpulan dibuat dari satu konteks	Gunakan bukti kedua dan rediagnosis

Implikasinya, keterhubungan profil dengan diferensiasi pembelajaran harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.12.21 Keterhubungan profil dengan umpan balik

Aspek keterhubungan profil dengan umpan balik menegaskan bahwa diagnosis adalah kegiatan ilmiah berskala kelas: guru merumuskan pertanyaan, mengumpulkan bukti,

mempertimbangkan penjelasan alternatif, lalu menentukan kesimpulan yang paling masuk akal. Inferensi harus proporsional terhadap kualitas data. Semakin spesifik label yang diberikan, semakin kuat pula bukti yang diperlukan (Mavridi et al., 2026; Taber, 2002a).

Secara operasional, guru dapat menggunakan analisis respons tertulis untuk menelusuri koherensi alasan. Data tersebut perlu dibaca bersama konsistensi lintas representasi, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mencampur kesalahan matematika dengan kesalahan konsep. Untuk mengurangnya, guru dapat menggunakan pertanyaan tindak lanjut dan mengulang diagnosis setelah remediasi. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai keterhubungan profil dengan umpan balik, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan keterhubungan profil dengan umpan balik melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan keterhubungan profil dengan umpan balik;
- Bandingkan bukti keterhubungan profil dengan umpan balik pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek keterhubungan profil dengan umpan balik dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian keterhubungan profil dengan umpan balik, topik pengukuran laboratorium dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa angka lebih banyak selalu dianggap lebih akurat. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta pembedaan akurasi, presisi, resolusi, dan angka penting. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa kualitas data ditentukan oleh desain, kalibrasi, dan ketidakpastian. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengorganisasikan bukti multidimensi menjadi profil perkembangan konsep yang dinamis dan tidak stigmatis, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, keterhubungan profil dengan umpan balik harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.12.22 Keterhubungan profil dengan perencanaan unit berikutnya

Penerapan keterhubungan profil dengan perencanaan unit berikutnya harus menghubungkan teori pengukuran dengan pengetahuan pedagogis kimia. Instrumen yang teknis tetapi tidak memahami struktur konsep akan menghasilkan kategori dangkal; sebaliknya, pengetahuan konten tanpa prosedur pengumpulan bukti dapat menghasilkan kesimpulan intuitif yang sulit diuji. Keduanya perlu dipadukan dalam satu alur yang transparan (Taber, 2002a; Vosniadou, 1994).

Secara operasional, guru dapat menggunakan diskusi argumentatif untuk menelusuri tingkat keyakinan. Data tersebut perlu dibaca bersama kemampuan transfer, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menggunakan ambang kategori tanpa bukti. Untuk mengurangnya, guru dapat menyajikan tugas paralel dan mencatat derajat keyakinan. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai keterhubungan profil dengan perencanaan unit berikutnya, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan keterhubungan profil dengan perencanaan unit berikutnya melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan keterhubungan profil dengan perencanaan unit berikutnya;
- Bandingkan bukti keterhubungan profil dengan perencanaan unit berikutnya pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek keterhubungan profil dengan perencanaan unit berikutnya dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian keterhubungan profil dengan perencanaan unit berikutnya, topik ikatan ion dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa NaCl dipandang sebagai molekul diskrit. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta gambar partikel padatan NaCl dan penjelasan hubungan struktur-kisi dengan titik leleh. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa senyawa ionik membentuk kisi dan rumusnya menyatakan rasio ion. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengorganisasikan bukti multidimensi menjadi profil perkembangan konsep yang dinamis dan tidak stigmatis, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, keterhubungan profil dengan perencanaan unit berikutnya harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian

klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.12.23 Validasi profil melalui triangulasi

Validasi profil melalui triangulasi tidak berdiri sebagai prosedur administratif. Ia menjadi bagian dari keputusan tentang apa yang perlu dijelaskan ulang, representasi apa yang perlu dibandingkan, dan siapa yang memerlukan dukungan berbeda. Oleh sebab itu, bukti diagnosis harus langsung terhubung dengan tindakan pembelajaran, bukan berhenti pada daftar kesalahan (Vosniadou, 1994; Mortimer, 1995).

Secara operasional, guru dapat menggunakan analisis grafik untuk menelusuri konsistensi lintas butir. Data tersebut perlu dibaca bersama stabilitas setelah jeda waktu, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah memperlakukan konsepsi sebagai sifat permanen. Untuk mengurangnya, guru dapat memeriksa konsep prasyarat dan menelusuri langkah penyelesaian. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai validasi profil melalui triangulasi, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan validasi profil melalui triangulasi melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan validasi profil melalui triangulasi;
- Bandingkan bukti validasi profil melalui triangulasi pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek validasi profil melalui triangulasi dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian validasi profil melalui triangulasi, topik pereaksi pembatas dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa pereaksi dengan massa terkecil selalu dianggap pembatas. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta analisis rasio mol terhadap koefisien. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa status pembatas ditentukan oleh hubungan jumlah zat dan stoikiometri. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengorganisasikan bukti multidimensi menjadi profil perkembangan konsep yang dinamis dan tidak stigmatis, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan

klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, validasi profil melalui triangulasi harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.12.24 Reliabilitas pengodean profil

Reliabilitas pengodean profil perlu ditempatkan dalam kerangka diagnosis yang mengorganisasikan bukti multidimensi menjadi profil perkembangan konsep yang dinamis dan tidak stigmatis. Fokusnya bukan sekadar menghasilkan skor, melainkan memperoleh bukti yang cukup untuk menjelaskan bagaimana peserta didik membangun, menghubungkan, dan menggunakan konsep. Karena konsepsi tidak dapat diamati secara langsung, guru harus menarik inferensi dari respons verbal, tulisan, gambar, tindakan, dan perubahan strategi yang tampak selama penyelesaian tugas (Mortimer, 1995; Mavridi et al., 2026).

Secara operasional, guru dapat menggunakan tes empat tingkat untuk menelusuri konsistensi lintas representasi. Data tersebut perlu dibaca bersama urutan langkah penyelesaian, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mengabaikan ketidakpastian pengukuran. Untuk mengurangnya, guru dapat meminta translasi antarrepresentasi dan menggunakan triangulasi bukti. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai reliabilitas pengodean profil, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan reliabilitas pengodean profil melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan reliabilitas pengodean profil;
- Bandingkan bukti reliabilitas pengodean profil pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek reliabilitas pengodean profil dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian reliabilitas pengodean profil, topik larutan penyangga dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa penyangga dianggap mempertahankan pH secara mutlak. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta analisis perubahan komposisi setelah penambahan asam atau basa. Bukti yang diharapkan adalah kemampuan peserta didik

menjelaskan bahwa penyangga membatasi perubahan dalam kapasitas tertentu, bukan meniadakannya. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengorganisasikan bukti multidimensi menjadi profil perkembangan konsep yang dinamis dan tidak stigmatis, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Kerangka operasional

Aspek	Bukti yang dicari	Risiko interpretasi	Tindak lanjut
Target konsep	Penjelasan tentang larutan penyangga	Target terlalu luas	Batasi subkonsep dan prasyarat
Respons	Konsistensi lintas representasi	Mengabaikan ketidakpastian pengukuran	Meminta translasi antarrepresentasi
Penalaran	Urutan langkah penyelesaian	Alasan diperlakukan sekadar tambahan	Menggunakan triangulasi bukti
Konfirmasi	Tes empat tingkat atau tugas isomorfik	Kesimpulan dibuat dari satu konteks	Gunakan bukti kedua dan rediagnosis

Implikasinya, reliabilitas pengodean profil harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.12.25 Etika penyimpanan dan penggunaan profil

Pembahasan tentang etika penyimpanan dan penggunaan profil berangkat dari prinsip bahwa data asesmen selalu merupakan jejak tidak langsung dari proses berpikir. Satu pilihan jawaban hanya menunjukkan produk akhir, sedangkan struktur pengetahuan yang melahirkannya dapat berbeda. Diagnosis yang kuat karena itu menggabungkan kemampuan transfer, kualitas model mental yang dieksternalisasi, dan bukti kontekstual sehingga keputusan pedagogis tidak bergantung pada satu indikator saja (Mavridi et al., 2026; Taber, 2002a).

Secara operasional, guru dapat menggunakan think-aloud untuk menelusuri kemampuan transfer. Data tersebut perlu dibaca bersama kualitas model mental yang dieksternalisasi, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menyimpulkan miskonsepsi hanya dari satu jawaban. Untuk mengurangnya, guru dapat mengulang diagnosis setelah remediasi dan menyusun tindak lanjut berbeda untuk kategori respons yang berbeda. Prosedur tersebut membantu

memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai etika penyimpanan dan penggunaan profil, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan etika penyimpanan dan penggunaan profil melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan etika penyimpanan dan penggunaan profil;
- Bandingkan bukti etika penyimpanan dan penggunaan profil pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek etika penyimpanan dan penggunaan profil dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian etika penyimpanan dan penggunaan profil, topik katalis dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa katalis dianggap menambah jumlah produk kesetimbangan. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta analisis jalur energi dan laju maju-balik. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa katalis mempercepat tercapainya kesetimbangan tanpa mengubah K . Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengorganisasikan bukti multidimensi menjadi profil perkembangan konsep yang dinamis dan tidak stigmatis, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, etika penyimpanan dan penggunaan profil harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.12.26 Contoh profil konseptual topik asam-basa

Dalam pembelajaran kimia, contoh profil konseptual topik asam-basa memiliki arti strategis karena banyak jawaban dapat diperoleh melalui hafalan prosedur atau pengenalan pola. Peserta didik mungkin tampak berhasil tetapi tidak mampu menerjemahkan makna simbol, model, atau besaran yang digunakan. Prinsip diagnosis menuntut bukti tentang cara konsep dipakai, bukan hanya apakah hasil akhirnya sesuai kunci (Taber, 2002a; Vosniadou, 1994).

Secara operasional, guru dapat menggunakan predict-observe-explain untuk menelusuri stabilitas setelah jeda waktu. Data tersebut perlu dibaca bersama respons terhadap probing,

karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mencampur kesalahan matematika dengan kesalahan konsep. Untuk mengurangnya, guru dapat mencatat derajat keyakinan dan membandingkan jawaban dengan alasan. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai contoh profil konseptual topik asam-basa, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan contoh profil konseptual topik asam-basa melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan contoh profil konseptual topik asam-basa;
- Bandingkan bukti contoh profil konseptual topik asam-basa pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek contoh profil konseptual topik asam-basa dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian contoh profil konseptual topik asam-basa, topik elektrolisis dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa anoda selalu dianggap bermuatan negatif. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta pembedaan definisi anoda-katoda berdasarkan proses oksidasi-reduksi. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa tanda elektroda bergantung jenis sel, tetapi anoda tetap lokasi oksidasi. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengorganisasikan bukti multidimensi menjadi profil perkembangan konsep yang dinamis dan tidak stigmatis, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, contoh profil konseptual topik asam-basa harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.12.27 Contoh profil konseptual topik ikatan kimia

Aspek contoh profil konseptual topik ikatan kimia menegaskan bahwa diagnosis adalah kegiatan ilmiah berskala kelas: guru merumuskan pertanyaan, mengumpulkan bukti, mempertimbangkan penjelasan alternatif, lalu menentukan kesimpulan yang paling masuk

akal. Inferensi harus proporsional terhadap kualitas data. Semakin spesifik label yang diberikan, semakin kuat pula bukti yang diperlukan (Vosniadou, 1994; Mortimer, 1995).

Secara operasional, guru dapat menggunakan tugas isomorfik untuk menelusuri urutan langkah penyelesaian. Data tersebut perlu dibaca bersama ketepatan pilihan, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menggunakan ambang kategori tanpa bukti. Untuk mengurangnya, guru dapat menelusuri langkah penyelesaian dan menggunakan pertanyaan tindak lanjut. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai contoh profil konseptual topik ikatan kimia, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan contoh profil konseptual topik ikatan kimia melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan contoh profil konseptual topik ikatan kimia;
- Bandingkan bukti contoh profil konseptual topik ikatan kimia pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek contoh profil konseptual topik ikatan kimia dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian contoh profil konseptual topik ikatan kimia, topik pengukuran laboratorium dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa angka lebih banyak selalu dianggap lebih akurat. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta pembedaan akurasi, presisi, resolusi, dan angka penting. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa kualitas data ditentukan oleh desain, kalibrasi, dan ketidakpastian. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengorganisasikan bukti multidimensi menjadi profil perkembangan konsep yang dinamis dan tidak stigmatis, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, contoh profil konseptual topik ikatan kimia harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.12.28 Template operasional profil konseptual

Penerapan template operasional profil konseptual harus menghubungkan teori pengukuran dengan pengetahuan pedagogis kimia. Instrumen yang teknis tetapi tidak memahami struktur konsep akan menghasilkan kategori dangkal; sebaliknya, pengetahuan konten tanpa prosedur pengumpulan bukti dapat menghasilkan kesimpulan intuitif yang sulit diuji. Keduanya perlu dipadukan dalam satu alur yang transparan (Mortimer, 1995; Mavridi et al., 2026).

Secara operasional, guru dapat menggunakan kuis formatif digital untuk menelusuri kualitas model mental yang dieksternalisasi. Data tersebut perlu dibaca bersama koherensi alasan, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah memperlakukan konsepsi sebagai sifat permanen. Untuk menguranginya, guru dapat menggunakan triangulasi bukti dan menyajikan tugas paralel. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai template operasional profil konseptual, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan template operasional profil konseptual melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan template operasional profil konseptual;
- Bandingkan bukti template operasional profil konseptual pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek template operasional profil konseptual dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian template operasional profil konseptual, topik ikatan ion dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa NaCl dipandang sebagai molekul diskrit. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta gambar partikel padatan NaCl dan penjelasan hubungan struktur-kisi dengan titik leleh. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa senyawa ionik membentuk kisi dan rumusnya menyatakan rasio ion. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengorganisasikan bukti multidimensi menjadi profil perkembangan konsep yang dinamis dan tidak stigmatis, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Kerangka operasional

Aspek	Bukti yang dicari	Risiko interpretasi	Tindak lanjut
Target konsep	Penjelasan tentang ikatan ion	Target terlalu luas	Batasi subkonsep dan prasyarat
Respons	Kualitas model mental yang dieksternalisasi	Memperlakukan konsepsi sebagai sifat permanen	Menggunakan triangulasi bukti
Penalaran	Koherensi alasan	Alasan diperlakukan sekadar tambahan	Menyajikan tugas paralel
Konfirmasi	Kuis formatif digital atau tugas isomorfik	Kesimpulan dibuat dari satu konteks	Gunakan bukti kedua dan rediagnosis

Implikasinya, template operasional profil konseptual harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.12.29 Analisis kritis terhadap reduksi peserta didik menjadi kategori

Analisis kritis terhadap reduksi peserta didik menjadi kategori tidak berdiri sebagai prosedur administratif. Ia menjadi bagian dari keputusan tentang apa yang perlu dijelaskan ulang, representasi apa yang perlu dibandingkan, dan siapa yang memerlukan dukungan berbeda. Oleh sebab itu, bukti diagnosis harus langsung terhubung dengan tindakan pembelajaran, bukan berhenti pada daftar kesalahan (Mavridi et al., 2026; Taber, 2002a).

Secara operasional, guru dapat menggunakan tes tiga tingkat untuk menelusuri respons terhadap probing. Data tersebut perlu dibaca bersama tingkat keyakinan, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah mengabaikan ketidakpastian pengukuran. Untuk mengurangnya, guru dapat menyusun tindak lanjut berbeda untuk kategori respons yang berbeda dan memeriksa konsep prasyarat. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai analisis kritis terhadap reduksi peserta didik menjadi kategori, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan analisis kritis terhadap reduksi peserta didik menjadi kategori melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan analisis kritis terhadap reduksi peserta didik menjadi kategori;

- Bandingkan bukti analisis kritis terhadap reduksi peserta didik menjadi kategori pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek analisis kritis terhadap reduksi peserta didik menjadi kategori dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian analisis kritis terhadap reduksi peserta didik menjadi kategori, topik pereaksi pembatas dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa pereaksi dengan massa terkecil selalu dianggap pembatas. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta analisis rasio mol terhadap koefisien. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa status pembatas ditentukan oleh hubungan jumlah zat dan stoikiometri. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengorganisasikan bukti multidimensi menjadi profil perkembangan konsep yang dinamis dan tidak stigmatis, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, analisis kritis terhadap reduksi peserta didik menjadi kategori harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.12.30 Sintesis penyusunan profil konseptual

Sintesis penyusunan profil konseptual perlu ditempatkan dalam kerangka diagnosis yang mengorganisasikan bukti multidimensi menjadi profil perkembangan konsep yang dinamis dan tidak stigmatis. Fokusnya bukan sekadar menghasilkan skor, melainkan memperoleh bukti yang cukup untuk menjelaskan bagaimana peserta didik membangun, menghubungkan, dan menggunakan konsep. Karena konsepsi tidak dapat diamati secara langsung, guru harus menarik inferensi dari respons verbal, tulisan, gambar, tindakan, dan perubahan strategi yang tampak selama penyelesaian tugas (Taber, 2002a; Vosniadou, 1994).

Secara operasional, guru dapat menggunakan wawancara semi-terstruktur untuk menelusuri ketepatan pilihan. Data tersebut perlu dibaca bersama konsistensi lintas butir, karena pola yang sama pada permukaan dapat berasal dari sumber berbeda. Bahaya utama pada tahap ini adalah menyimpulkan miskonsepsi hanya dari satu jawaban. Untuk menguranginya, guru dapat membandingkan jawaban dengan alasan dan meminta translasi antarrepresentasi. Prosedur tersebut membantu memisahkan pengetahuan yang belum lengkap, kesalahan prosedural, respons spekulatif, dan konsepsi alternatif yang relatif stabil.

- Untuk menilai sintesis penyusunan profil konseptual, rumuskan konsep target, indikator, dan batas interpretasinya secara eksplisit;
- Pilih tugas yang merepresentasikan sintesis penyusunan profil konseptual melalui penggunaan konsep, bukan sekadar pengenalan istilah;
- Rekam alasan, representasi, atau langkah kerja yang secara langsung berkaitan dengan sintesis penyusunan profil konseptual;
- Bandingkan bukti sintesis penyusunan profil konseptual pada lebih dari satu konteks, format tugas, atau bentuk representasi;
- Hubungkan kategori hasil pada aspek sintesis penyusunan profil konseptual dengan tindak lanjut pembelajaran yang spesifik.

Dalam kajian sintesis penyusunan profil konseptual, topik larutan penyangga dapat digunakan untuk memperlihatkan kebutuhan akan bukti berlapis. Konsepsi yang sering muncul ialah bahwa penyangga dianggap mempertahankan pH secara mutlak. Diagnosis tidak cukup dilakukan dengan menanyakan definisi, sebab peserta didik dapat mengingat istilah tanpa menggunakan relasi konseptual yang tepat. Guru dapat meminta analisis perubahan komposisi setelah penambahan asam atau basa. Bukti yang diharapkan adalah kemampuan peserta didik menjelaskan bahwa penyangga membatasi perubahan dalam kapasitas tertentu, bukan meniadakannya. Kualitas bukti tersebut kemudian ditafsirkan sesuai fokus mengorganisasikan bukti multidimensi menjadi profil perkembangan konsep yang dinamis dan tidak stigmatis, bukan diperlakukan sebagai skor yang berdiri sendiri. Apabila peserta didik mengubah jawaban setelah satu pertanyaan klarifikasi, kasusnya mungkin berbeda dari peserta didik yang mempertahankan alasan yang sama pada beberapa tugas isomorfik.

Implikasinya, sintesis penyusunan profil konseptual harus berakhir pada keputusan yang dapat diuji kembali. Guru perlu mencatat bukti pendukung, tingkat kepastian klasifikasi, serta rencana rediagnosis. Bahasa laporan sebaiknya menyatakan bahwa peserta didik 'menunjukkan pola respons yang konsisten dengan' suatu konsepsi, bukan menyatakan identitas permanen. Cara ini menjaga ketelitian ilmiah sekaligus membuka ruang bagi perubahan konseptual.

5.12.31 Implikasi praktis bagi guru kimia

Implikasi praktis subbab 5.12 dapat dirangkum sebagai alur: tentukan konsep target, pilih bukti yang relevan, kumpulkan respons dalam lebih dari satu bentuk, analisis pola, nyatakan derajat kepastian, susun tindak lanjut, dan lakukan rediagnosis. Alur ini menjaga agar penyusunan profil konseptual peserta didik tidak berubah menjadi administrasi skor, tetapi berfungsi sebagai mekanisme perbaikan pembelajaran yang berulang.

78. Identifikasi konsep inti dan konsep prasyarat sebelum menyusun instrumen.
79. Gunakan pertanyaan yang meminta alasan, representasi, atau langkah kerja.
80. Bandingkan respons pada beberapa konteks dan waktu.
81. Pisahkan bukti faktual dari interpretasi guru.

82. Tentukan tindak lanjut yang berbeda untuk kategori kebutuhan yang berbeda.
83. Simpan data secukupnya dan lindungi kerahasiaan peserta didik.
84. Periksa kembali perubahan melalui tugas transfer dan retensi.

5.13 Sintesis Bab

Prinsip diagnosis miskonsepsi dibangun atas pengakuan bahwa jawaban peserta didik merupakan bukti yang harus ditafsirkan, bukan cermin langsung dari struktur pikirannya. Diagnosis yang dapat dipertanggungjawabkan menghubungkan konsep target, tugas, proses respons, alasan, keyakinan, konsistensi, representasi, dan konteks. Tidak ada satu indikator yang secara otomatis membuktikan keberadaan miskonsepsi.

Pertama, asesmen diagnostik harus memiliki tujuan pedagogis yang jelas. Data yang dikumpulkan harus diperlukan untuk menentukan tindakan. Kedua, diagnosis perlu berlangsung sebelum, selama, dan setelah pembelajaran agar perubahan konsep dapat diamati sebagai proses. Ketiga, kesalahan perlu dibedakan dari miskonsepsi dengan memeriksa sistematisitas, konsistensi, keyakinan, transfer, dan respons terhadap probing.

Keempat, kombinasi jawaban, alasan, dan keyakinan memperkaya diagnosis tetapi tidak menghapus kebutuhan konfirmasi. Kelima, validitas dan reliabilitas bukan sekadar angka teknis. Validitas menuntut argumentasi bahwa interpretasi dan penggunaan hasil didukung bukti, sedangkan reliabilitas menuntut pemahaman mengenai ketidakpastian dan konsistensi keputusan. Keenam, hasil tingkat kelas tidak boleh menutupi variasi individual, dan hasil individual tidak boleh dilepaskan dari konteks pembelajaran.

Ketujuh, diagnosis kimia harus berbasis representasi karena penguasaan konsep menuntut kemampuan menghubungkan fenomena, model submikroskopik, simbol, grafik, dan matematika. Kedelapan, etika menjadi batas seluruh proses: diagnosis harus bermanfaat, adil, transparan, tidak stigmatis, dan menjaga data pribadi. Kesembilan, interpretasi harus menyatakan bukti dan ketidakpastian secara eksplisit. Kesepuluh, profil konseptual harus diperlakukan sebagai dokumen dinamis yang diperbarui setelah intervensi.

Keseluruhan prinsip tersebut mengarah pada satu kesimpulan: kualitas diagnosis tidak ditentukan oleh kerumitan instrumen, melainkan oleh keselarasan antara tujuan, bukti, inferensi, tindakan, dan evaluasi ulang. Instrumen sederhana dapat sangat bermakna apabila pertanyaannya tepat dan tindak lanjutnya jelas. Sebaliknya, instrumen berlapis dapat menghasilkan kesimpulan lemah apabila distraktornya tidak valid, keyakinan ditafsirkan secara mekanis, atau hasilnya tidak dihubungkan dengan pembelajaran.

Bab berikutnya dapat menggunakan prinsip-prinsip ini sebagai dasar untuk membahas teknik dan instrumen pemetaan miskonsepsi secara lebih spesifik. Setiap teknik harus dinilai melalui pertanyaan yang sama: konsep apa yang hendak diungkap, bukti apa yang dihasilkan, seberapa kuat inferensinya, apa keterbatasannya, bagaimana data dilindungi, dan tindakan pembelajaran apa yang dapat dirancang dari hasil tersebut.

Daftar Pustaka Bab 5

Ainsworth, S. (2006). DeFT: A conceptual framework for considering learning with multiple representations. *Learning and Instruction*, 16(3), 183-198. <https://doi.org/10.1016/j.learninstruc.2006.03.001>

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). Standards for educational and psychological testing. American Educational Research Association. <https://www.testingstandards.net/open-access-files.html>

Anggraena, Y., Ginanto, D. E., Kesuma, A. T., & Setiyowati, D. (2025). Panduan pembelajaran dan asesmen pendidikan anak usia dini, jenjang pendidikan dasar, dan jenjang pendidikan menengah (Edisi revisi 2025). Badan Standar, Kurikulum, dan Asesmen Pendidikan. <https://repositori.kemendikdasmen.go.id/33290/>

Barke, H.-D., Hazari, A., & Yitbarek, S. (2009). Misconceptions in chemistry: Addressing perceptions in chemical education. Springer. <https://doi.org/10.1007/978-3-540-70989-3>

Bennett, R. E. (2011). Formative assessment: A critical review. *Assessment in Education: Principles, Policy & Practice*, 18(1), 5-25. <https://doi.org/10.1080/0969594X.2010.513678>

Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability*, 21(1), 5-31. <https://doi.org/10.1007/s11092-008-9068-5>

Brandriet, A. R., & Bretz, S. L. (2014). Measuring meta-ignorance through the lens of confidence: Examining students' redox misconceptions about oxidation numbers, charge, and electron transfer. *Chemistry Education Research and Practice*, 15(4), 729-746. <https://doi.org/10.1039/C4RP00129J>

Cahyani, N., Auliah, A., & Majid, A. F. (2024). Application of two-tier multiple choice assessment equipped with Certainty of Response Index as a misconception diagnosis instrument in chemistry learning. *Jurnal Pendidikan Matematika dan Sains*, 12(1), 13-21. <https://doi.org/10.21831/jpms.v12i1.74390>

Chandrasegaran, A. L., Treagust, D. F., & Mocerino, M. (2007). The development of a two-tier multiple-choice diagnostic instrument for evaluating secondary school students' ability to describe and explain chemical reactions using multiple levels of representation. *Chemistry Education Research and Practice*, 8(3), 293-307. <https://doi.org/10.1039/B7RP90006F>

Chi, M. T. H. (2005). Commonsense conceptions of emergent processes: Why some misconceptions are robust. *The Journal of the Learning Sciences*, 14(2), 161-199. https://doi.org/10.1207/s15327809jls1402_1

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16, 297-334. <https://doi.org/10.1007/BF02310555>

DeVellis, R. F., & Thorpe, C. T. (2021). *Scale development: Theory and applications* (5th ed.). SAGE. <https://us.sagepub.com/en-us/nam/scale-development/book269114>

Downing, S. M. (2003). Validity: On meaningful interpretation of assessment data. *Medical Education*, 37(9), 830-837. <https://doi.org/10.1046/j.1365-2923.2003.01594.x>

Duit, R., & Treagust, D. F. (2003). Conceptual change: A powerful framework for improving science teaching and learning. *International Journal of Science Education*, 25(6), 671-688. <https://doi.org/10.1080/09500690305016>

Gilbert, J. K., & Treagust, D. F. (Eds.). (2009). *Multiple representations in chemical education*. Springer. <https://doi.org/10.1007/978-1-4020-8872-8>

Hasan, S., Bagayoko, D., & Kelley, E. L. (1999). Misconceptions and the certainty of response index (CRI). *Physics Education*, 34(5), 294-299. <https://doi.org/10.1088/0031-9120/34/5/304>

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81-112. <https://doi.org/10.3102/003465430298487>

Hoe, K. Y., & Subramaniam, R. (2016). On the prevalence of alternative conceptions on acid-base chemistry among secondary students: Insights from cognitive and confidence measures. *Chemistry Education Research and Practice*, 17(2), 263-282. <https://doi.org/10.1039/C5RP00146C>

Indonesia. (2022). Undang-Undang Republik Indonesia Nomor 27 Tahun 2022 tentang Pelindungan Data Pribadi. Lembaran Negara Republik Indonesia Tahun 2022 Nomor 196. <https://peraturan.bpk.go.id/Details/229798/uu-no-27-tahun-2022>

Johnstone, A. H. (2000). Teaching of chemistry-Logical or psychological? *Chemistry Education Research and Practice*, 1(1), 9-15. <https://doi.org/10.1039/A9RP90001B>

Kane, M. T. (2013). Validating the interpretations and uses of test scores. *Journal of Educational Measurement*, 50(1), 1-73. <https://doi.org/10.1111/jedm.12000>

Koriat, A. (1997). Monitoring one's own knowledge during study: A cue-utilization approach to judgments of learning. *Journal of Experimental Psychology: General*, 126(4), 349-370. <https://doi.org/10.1037/0096-3445.126.4.349>

Lia, R. M., Rusilowati, A., & Isnaeni, W. (2020). NGSS-oriented chemistry test instruments: Validity and reliability analysis with the Rasch model. *Research and Evaluation in Education*, 6(1), 41-50. <https://doi.org/10.21831/reid.v6i1.30112>

Mathabathe, K. C., & Potgieter, M. (2014). Metacognitive monitoring and learning gain in foundation chemistry. *Chemistry Education Research and Practice*, 15(1), 94-104. <https://doi.org/10.1039/C3RP00119A>

Mavridi, M., Salta, K., & Koulougliotis, D. (2026). Students' thinking about chemical analysis: An investigation within the conceptual profile theoretical framework. *Chemistry Education Research and Practice*, 27, 1000-1023. <https://doi.org/10.1039/D5RP00441A>

Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist*, 50(9), 741-749. <https://doi.org/10.1037/0003-066X.50.9.741>

Mortimer, E. F. (1995). Conceptual change or conceptual profile change? *Science & Education*, 4(3), 267-285. <https://doi.org/10.1007/BF00486624>

Mulford, D. R., & Robinson, W. R. (2002). An inventory for alternate conceptions among first-semester general chemistry students. *Journal of Chemical Education*, 79(6), 739-744. <https://doi.org/10.1021/ed079p739>

Nakhleh, M. B. (1992). Why some students don't learn chemistry: Chemical misconceptions. *Journal of Chemical Education*, 69(3), 191-196. <https://doi.org/10.1021/ed069p191>

Odom, A. L., & Barrow, L. H. (1995). Development and application of a two-tier diagnostic test measuring college biology students' understanding of diffusion and osmosis after a course of instruction. *Journal of Research in Science Teaching*, 32(1), 45-61. <https://doi.org/10.1002/tea.3660320106>

Posner, G. J., Strike, K. A., Hewson, P. W., & Gertzog, W. A. (1982). Accommodation of a scientific conception: Toward a theory of conceptual change. *Science Education*, 66(2), 211-227. <https://doi.org/10.1002/sce.3730660207>

Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, 18, 119-144. <https://doi.org/10.1007/BF00117714>

Shiddiqi, M. H. A., Arthamena, V. D., Ayyubi, M., Manarisip, A. J., & Aznam, N. (2024). Systematic literature review: Analysis of misconception problems and diagnostic instruments for learning chemistry. *Jurnal Penelitian Pendidikan IPA*, 10(4), 168-179. <https://doi.org/10.29303/jppipa.v10i4.5189>

Shute, V. J. (2008). Focus on formative feedback. *Review of Educational Research*, 78(1), 153-189. <https://doi.org/10.3102/0034654307313795>

Smith, J. P., III, diSessa, A. A., & Roschelle, J. (1994). Misconceptions reconceived: A constructivist analysis of knowledge in transition. *The Journal of the Learning Sciences*, 3(2), 115-163. https://doi.org/10.1207/s15327809jls0302_1

Taber, K. S. (2001a). Building the structural concepts of chemistry: Some considerations from educational research. *Chemistry Education Research and Practice*, 2(2), 123-158. <https://doi.org/10.1039/B1RP90014E>

Taber, K. S. (2001b). Constructing chemical concepts in the classroom? Using research to inform practice. *Chemistry Education Research and Practice*, 2(1), 43-51. <https://doi.org/10.1039/B0RP90029J>

Taber, K. S. (2002a). Chemical misconceptions: Prevention, diagnosis and cure: Volume I- Theoretical background. Royal Society of Chemistry. <https://edu.rsc.org/resources/chemical-misconceptions-prevention-diagnosis-and-cure/4010255.article>

Taber, K. S. (2002b). Chemical misconceptions: Prevention, diagnosis and cure: Volume II- Classroom resources. Royal Society of Chemistry. <https://edu.rsc.org/resources/chemical-misconceptions-prevention-diagnosis-and-cure/4010255.article>

Talanquer, V. (2006). Commonsense chemistry: A model for understanding students' alternative conceptions. *Journal of Chemical Education*, 83(5), 811-816. <https://doi.org/10.1021/ed083p811>

Talanquer, V. (2013). When atoms want. *Journal of Chemical Education*, 90(11), 1419-1424. <https://doi.org/10.1021/ed400311x>

Talanquer, V. (2015). Threshold concepts in chemistry: The critical role of implicit schemas. *Journal of Chemical Education*, 92(1), 3-9. <https://doi.org/10.1021/ed500679k>

Talanquer, V. (2026). The importance of fostering conceptual reframing in chemistry education. *Chemistry Education Research and Practice*, 27(2), 693-700. <https://doi.org/10.1039/D5RP00421G>

Treagust, D. F. (1988). Development and use of diagnostic tests to evaluate students' misconceptions in science. *International Journal of Science Education*, 10(2), 159-169. <https://doi.org/10.1080/0950069880100204>

Tümay, H. (2016). Reconsidering learning difficulties and misconceptions in chemistry: Emergence in chemistry and its implications for chemical education. *Chemistry Education Research and Practice*, 17(2), 229-245. <https://doi.org/10.1039/C6RP00008H>

UNESCO Institute for Information Technologies in Education. (2020). Personal data and privacy protection in online learning: Guidance for students, teachers and parents. UNESCO IITE. <https://iite.unesco.org/publications/personal-data-and-privacy-protection-in-online-learning/>

UNESCO. (2023). Guidance for generative AI in education and research. UNESCO. <https://www.unesco.org/en/articles/guidance-generative-ai-education-and-research>

Vosniadou, S. (1994). Capturing and modeling the process of conceptual change. *Learning and Instruction*, 4(1), 45-69. [https://doi.org/10.1016/0959-4752\(94\)90018-3](https://doi.org/10.1016/0959-4752(94)90018-3)

Wren, D., & Barbera, J. (2014). Psychometric analysis of the thermochemistry concept inventory. *Chemistry Education Research and Practice*, 15(3), 380-390. <https://doi.org/10.1039/C3RP00170A>

Yan, Y. K., & Subramaniam, R. (2018). Using a multi-tier diagnostic test to explore the nature of students' alternative conceptions on reaction kinetics. *Chemistry Education Research and Practice*, 19(1), 213-226. <https://doi.org/10.1039/C7RP00143F>

BAB 6

INSTRUMEN DIAGNOSIS MISKONSEPSI

Pengantar Bab

Diagnosis miskonsepsi bukan kegiatan menghitung jawaban salah, melainkan proses mengumpulkan dan menafsirkan bukti tentang cara peserta didik memahami konsep. Jawaban salah dapat berasal dari tidak tahu, kelalaian, salah membaca, salah hitung, atau penggunaan prosedur yang tidak tepat. Sebaliknya, jawaban benar dapat diperoleh melalui tebakan, petunjuk opsi, atau algoritma yang tidak dipahami. Karena itu, instrumen diagnosis harus dirancang untuk mengungkap pola penalaran, representasi, alasan, keyakinan, dan konsistensi respons. (Gurel et al., 2015; Treagust, 1988; AERA, APA, & NCME, 2014)

Dalam pembelajaran kimia, kebutuhan tersebut menjadi sangat penting. Peserta didik harus berpindah antara fenomena makroskopik, model submikroskopik, simbol, grafik, dan hubungan matematis. Satu konsepsi dapat tampak benar pada pertanyaan verbal tetapi gagal ketika dinyatakan melalui gambar partikel atau perhitungan. Instrumen yang baik harus mempertimbangkan keragaman representasi dan memastikan bahwa kesalahan yang terdeteksi benar-benar berkaitan dengan konsep sasaran.

Bab ini membahas dua puluh dua komponen penting dalam diagnosis miskonsepsi. Pembahasan dimulai dari pilihan ganda konvensional, pilihan ganda beralasan, tes dua sampai lima tingkat, dan Certainty of Response Index. Selanjutnya dibahas wawancara, peta konsep, tes terbuka, Predict–Observe–Explain, gambar, analisis kesalahan, think-aloud, percakapan kelas, dan log aktivitas digital. Bagian akhir menguraikan pengembangan distraktor, validasi isi, uji coba, analisis butir, penetapan kategori, dan pelaporan.

Setiap instrumen dikaji melalui landasan konsep, sejarah, konstruk, struktur, desain, prosedur pengembangan, pelaksanaan, penskoran, interpretasi, validitas, reliabilitas, contoh kimia, kelebihan, keterbatasan, integrasi digital, etika, dan pedoman implementasi. Kerangka yang konsisten ini memungkinkan pembaca membandingkan teknik sekaligus memahami bahwa setiap teknik menghasilkan jenis bukti yang berbeda.

Prinsip utama bab ini adalah keselarasan. Tujuan diagnosis harus selaras dengan konstruk, tugas, bentuk respons, aturan skor, kategori, dan keputusan pembelajaran. Semakin besar konsekuensi keputusan, semakin kuat bukti yang diperlukan. Untuk kepentingan kelas, suatu instrumen dapat berfungsi sebagai penyaring awal. Untuk penelitian atau keputusan individual, hasil perlu dikonfirmasi melalui beberapa butir, representasi, atau metode lain. (Gurel et al., 2015; Treagust, 1988; AERA, APA, & NCME, 2014)

Tujuan bab. Setelah mempelajari bab ini, pembaca diharapkan mampu memilih, mengembangkan, memvalidasi, menganalisis, dan melaporkan instrumen diagnosis miskonsepsi secara etis dan berbasis bukti.

- membedakan kekuatan inferensi dari setiap instrumen;
- menyusun komponen instrumen berdasarkan konsepsi alternatif yang nyata;
- menggunakan validasi isi, uji proses respons, dan analisis empiris secara terpadu;
- menetapkan kategori pemahaman secara transparan dan tidak stigmatis;
- mengubah hasil diagnosis menjadi tindakan pembelajaran yang terarah.

6.1 Tes Pilihan Ganda Konvensional

6.1.1 Landasan Konseptual dan Kedudukan Instrumen

Pemilihan Tes Pilihan Ganda Konvensional harus berangkat dari pertanyaan diagnosis yang jelas, bukan dari kebiasaan menggunakan instrumen yang paling mudah tersedia. Dalam bab ini, instrumen tersebut didefinisikan sebagai instrumen yang meminta peserta didik memilih satu jawaban paling tepat dari beberapa opsi tanpa meminta alasan atau tingkat keyakinan secara eksplisit. Definisi operasional ini menegaskan bentuk bukti yang dikumpulkan sekaligus membatasi jenis kesimpulan yang layak dibuat. (Haladyna et al., 2002; Rodriguez, 2005; Tarrant et al., 2009)

Klaim utamanya adalah bahwa instrumen ini mengestimasi penguasaan hasil belajar dan memetakan pola pilihan jawaban, tetapi hanya memberikan bukti diagnosis miskonsepsi yang terbatas apabila digunakan sendirian. Klaim tersebut harus dipisahkan dari anggapan bahwa setiap jawaban salah merupakan miskonsepsi. Kesalahan dapat berasal dari kurang pengetahuan, ketidakcermatan, kesulitan bahasa, salah hitung, atau ketidakfamiliaran dengan representasi.

Data pokok yang dihasilkan berupa respons benar-salah dan frekuensi pemilihan setiap opsi. Data itu baru memperoleh makna diagnostik setelah dikaitkan dengan peta konsep sasaran, kemungkinan konsepsi alternatif, dan aturan interpretasi yang telah ditetapkan sebelum analisis.

Konsekuensinya, peneliti dan guru perlu membedakan data mentah, skor, kategori, dan klaim konseptual yang dibuat dari hasil tersebut. Posisi terbaik Tes Pilihan Ganda Konvensional adalah sebagai bagian dari sistem asesmen berlapis yang dapat dipadukan dengan bukti lain ketika konsekuensi keputusan semakin besar.

- menentukan konsep dan proposisi yang hendak didiagnosis;
- menjelaskan perilaku respons yang dianggap sebagai bukti;
- membatasi klaim sesuai kekuatan instrumen;
- menautkan hasil dengan tindakan pembelajaran yang realistis.

Dalam praktik kelas, Tes Pilihan Ganda Konvensional dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.1.2 Perkembangan Historis dan Rasional Penggunaan

Perkembangan Tes Pilihan Ganda Konvensional tidak dapat dipisahkan dari perubahan pandangan tentang belajar sains. Ketika peserta didik tidak lagi dipandang sebagai penerima informasi yang kosong, instrumen mulai dirancang untuk mengungkap gagasan awal, pola penalaran, dan keyakinan yang dibawa ke kelas. (Haladyna et al., 2002; Rodriguez, 2005; Tarrant et al., 2009; Treagust, 1988)

Pada tahap awal, penilaian banyak mengandalkan benar-salah atau skor total. Perkembangan penelitian konsepsi peserta didik kemudian mendorong penambahan alasan, keyakinan, tugas kinerja, rekaman proses, dan bukti sosial. Tes Pilihan Ganda Konvensional merupakan salah satu jawaban metodologis terhadap kebutuhan tersebut.

Dalam pendidikan kimia, kebutuhan diagnosis semakin kuat karena peserta didik harus mengoordinasikan fenomena makroskopik, model submikroskopik, simbol, grafik, dan perhitungan. Instrumen yang hanya menguji hafalan definisi sering gagal memperlihatkan cara representasi itu dihubungkan.

Sejarah penggunaannya juga menunjukkan bahwa format yang lebih kompleks tidak otomatis lebih baik. Setiap tambahan tier, rubrik, atau kanal data harus mempunyai fungsi yang jelas dan didukung bukti bahwa informasi tambahan benar-benar memperbaiki keputusan.

- pergeseran dari pengukuran skor menuju diagnosis pola pemahaman;
- penggunaan respons terbuka dan wawancara untuk menemukan konsepsi alternatif;
- pengembangan teknik kuantitatif agar diagnosis dapat dilakukan pada kelompok besar;
- integrasi data digital dan metakognitif pada perkembangan mutakhir.

Dalam praktik kelas, Tes Pilihan Ganda Konvensional dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.1.3 Konstruk, Klaim, dan Bukti Diagnosis

Konstruk yang dinilai melalui Tes Pilihan Ganda Konvensional bukan objek yang terlihat langsung. Yang diamati adalah respons benar-salah dan frekuensi pemilihan setiap opsi; dari sana dibuat inferensi tentang pemahaman konsep, miskonsepsi, pengetahuan parsial, atau ketidakpastian. (Sadler, 1998; AERA, APA, & NCME, 2014; Messick, 1995; Kane, 2013)

Klaim diagnosis perlu dirumuskan sebagai argumen. Pertama, tugas harus memerlukan konsep sasaran. Kedua, respons harus diberi skor atau kode secara konsisten. Ketiga, pola skor harus mempunyai hubungan yang masuk akal dengan struktur pengetahuan. Keempat, penggunaan hasil harus menghasilkan keputusan yang proporsional.

Sumber ancaman utama adalah ketidakjelasan konstruk. Misalnya, peserta didik dapat gagal karena istilah yang ambigu, beban membaca, representasi yang tidak dikenal, atau prosedur matematika. Kegagalan semacam itu tidak boleh langsung diperlakukan sebagai miskonsepsi kimia.

Oleh sebab itu, setiap pengembang perlu menulis pernyataan klaim-bukti: instrumen mengukur apa, perilaku apa yang menjadi indikator, alternatif penjelasan apa yang mungkin, dan bukti tambahan apa yang diperlukan untuk memperkuat kesimpulan.

- konstruk utama: pemahaman konseptual melalui tes objektif satu tingkat;
- bukti observabel: respons, alasan, tindakan, atau representasi;
- inferensi: profil konsepsi pada konten dan konteks tertentu;
- penggunaan: perencanaan pembelajaran, remediasi, atau penelitian.

Dalam praktik kelas, Tes Pilihan Ganda Konvensional dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.1.4 Struktur, Komponen, dan Alur Respons

Struktur Tes Pilihan Ganda Konvensional dibangun dari komponen yang saling bergantung. Komponen inti meliputi stimulus atau stem, satu kunci jawaban, dua sampai empat distraktor, aturan penskoran dikotomis, dan peta indikator dan konten. Hilangnya satu komponen dapat mengubah makna data yang diperoleh. (Haladyna et al., 2002; Rodriguez, 2005; Tarrant et al., 2009; Treagust, 1988)

Urutan penyajian harus ditetapkan secara sadar. Komponen yang muncul lebih awal dapat memberi petunjuk pada komponen berikutnya, sedangkan permintaan alasan atau keyakinan dapat mengubah cara peserta didik memeriksa jawabannya. Penguncian respons dan uji urutan diperlukan pada format digital.

Setiap komponen membutuhkan spesifikasi. Stimulus harus memuat informasi cukup, opsi atau prompt harus sejajar dengan indikator, skala harus memiliki jangkar, dan rubrik harus menjelaskan perbedaan kategori yang berdekatan.

Struktur yang baik memungkinkan peneliti menelusuri kembali setiap kategori hasil ke respons mentah. Transparansi ini penting agar guru tidak hanya menerima label otomatis tanpa memahami alasan di baliknya.

- stimulus atau stem;
- satu kunci jawaban;
- dua sampai empat distraktor;
- aturan penskoran dikotomis;
- peta indikator dan konten;

Dalam praktik kelas, Tes Pilihan Ganda Konvensional dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

No.	Komponen	Fungsi Diagnostik
1	stimulus atau stem	Bukti yang harus ditelusuri
2	satu kunci jawaban	Bukti yang harus ditelusuri
3	dua sampai empat distraktor	Bukti yang harus ditelusuri
4	aturan penskoran dikotomis	Bukti yang harus ditelusuri

6.1.5 Desain Stimulus, Tugas, atau Pertanyaan

Desain stimulus pada Tes Pilihan Ganda Konvensional harus memunculkan penggunaan konsep, bukan sekadar pengenalan istilah. Situasi dapat berbentuk fenomena, gambar partikel, grafik, data eksperimen, persamaan, atau masalah kontekstual, tetapi seluruh informasi yang tidak relevan perlu dikendalikan. (Haladyna et al., 2002; Rodriguez, 2005; Tarrant et al., 2009)

Dalam kimia, satu stimulus sering menggabungkan beberapa tingkat representasi. Pengembang harus memutuskan apakah tujuan butir adalah menerjemahkan representasi, menjelaskan hubungan sebab-akibat, membedakan contoh dan noncontoh, atau menerapkan prinsip pada konteks baru.

Bahasa perlu ringkas tanpa menghilangkan syarat ilmiah. Kata seperti kuat, stabil, jenuh, spontan, atau setimbang mempunyai arti teknis dan sehari-hari yang berbeda. Ketidakjelasan istilah dapat menghasilkan kesalahan yang berasal dari bahasa, bukan struktur konsep.

Desain yang baik juga menghindari petunjuk format, pilihan terlalu panjang, negasi bertumpuk, gambar dekoratif, dan data yang tidak diperlukan. Setiap elemen harus mempunyai alasan diagnostik.

- gunakan konteks yang cukup dikenal tetapi tidak menghafalkan contoh;
- pastikan hanya satu interpretasi ilmiah yang menjadi target;
- uji semua simbol, satuan, grafik, dan gambar;
- sediakan ruang bagi respons yang tidak diprediksi bila tujuan masih eksploratif.

Dalam praktik kelas, Tes Pilihan Ganda Konvensional dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.1.6 Prosedur Pengembangan Secara Sistematis

Pengembangan Tes Pilihan Ganda Konvensional harus mengikuti alur yang terdokumentasi. Langkah yang lazim ialah menetapkan batas konten, memetakan indikator konseptual, mengumpulkan konsepsi alternatif, menulis stem yang bebas petunjuk, kemudian menyusun kunci dan distraktor homogen, melakukan telaah ahli, melaksanakan uji coba, menganalisis butir. Urutan dapat berulang karena temuan uji coba sering memaksa revisi konstruk atau redaksi. (Haladyna et al., 2002; Rodriguez, 2005; Tarrant et al., 2009; Adams & Wieman, 2011; AERA, APA, & NCME, 2014)

Tahap konseptual dimulai dengan analisis kurikulum, literatur miskonsepsi, dan peta konsep. Tahap empiris awal menggunakan wawancara, respons terbuka, atau observasi untuk memastikan bahwa kategori yang dirancang benar-benar muncul pada populasi sasaran.

Setiap versi instrumen perlu diberi nomor, tanggal, dan catatan perubahan. Revisi harus menyebutkan masalah yang ditemukan, bukti yang digunakan, perubahan yang dilakukan, dan alasan mempertahankan atau menghapus butir.

Pengembangan dianggap belum selesai ketika hanya menghasilkan butir yang tampak baik. Instrumen perlu diuji pada peserta didik yang serupa dengan pengguna sasaran, lalu dianalisis secara konseptual dan empiris sebelum dipakai untuk keputusan penting.

- 1. menetapkan batas konten.
- 2. memetakan indikator konseptual.
- 3. mengumpulkan konsepsi alternatif.
- 4. menulis stem yang bebas petunjuk.
- 5. menyusun kunci dan distraktor homogen.

Dalam praktik kelas, Tes Pilihan Ganda Konvensional dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.1.7 Pelaksanaan dan Pengumpulan Data

Pelaksanaan Tes Pilihan Ganda Konvensional harus distandardisasi secukupnya agar perbedaan hasil lebih mungkin berasal dari pemahaman, bukan variasi instruksi, waktu, alat, atau bantuan. Petunjuk perlu menjelaskan tujuan umum tanpa mengajarkan jawaban. (Haladyna et al., 2002; Rodriguez, 2005; AERA, APA, & NCME, 2014)

Kondisi administrasi mencakup waktu, urutan tugas, kesempatan bertanya, penggunaan kalkulator, akses tabel periodik, kualitas gambar, dan perangkat digital. Pada instrumen kualitatif, standardisasi berarti konsistensi fokus dan prinsip probing, bukan membaca naskah secara kaku.

Data yang dikumpulkan ialah respons benar-salah dan frekuensi pemilihan setiap opsi. Data mentah harus disimpan bersama metadata yang diperlukan, seperti kode butir, versi instrumen, waktu administrasi, dan kondisi khusus, tetapi hanya data yang relevan yang boleh dikumpulkan.

Sebelum administrasi luas, lakukan uji teknis. Pastikan tautan, urutan, penguncian respons, rekaman, penyimpanan, serta tampilan rumus dan subskrip bekerja. Gangguan teknis perlu dicatat sebagai bagian dari interpretasi.

- berikan instruksi yang sama dan mudah dipahami;
- catat kondisi yang dapat memengaruhi respons;
- hindari umpan balik sebelum seluruh respons diagnostik selesai;
- sediakan prosedur penanganan data tidak lengkap.

Dalam praktik kelas, Tes Pilihan Ganda Konvensional dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.1.8 Penskoran, Pengodean, dan Pengolahan Data

Penskoran Tes Pilihan Ganda Konvensional dapat diringkas sebagai skor 1 untuk jawaban benar dan 0 untuk jawaban salah, dilengkapi analisis distraktor per butir. Aturan ini harus ditulis sebelum melihat hasil utama agar keputusan tidak disesuaikan secara oportunistik dengan pola data. (Haladyna et al., 2002; Rodriguez, 2005; Tarrant et al., 2009; Chi, 1997; AERA, APA, & NCME, 2014)

Skor total jarang cukup untuk diagnosis. Analisis perlu mempertahankan informasi tingkat butir, konsep, opsi, alasan, keyakinan, proses, atau representasi. Penggabungan data terlalu dini dapat menyembunyikan pola yang secara pedagogis penting.

Untuk respons terbuka atau kualitatif, buku kode harus mencantumkan definisi, contoh positif, contoh batas, noncontoh, dan aturan ketika satu respons memuat beberapa gagasan. Sebagian data perlu dinilai oleh lebih dari satu penilai.

Setiap transformasi data—misalnya dari respons ke kode, dari kode ke kategori, dan dari kategori ke persentase—harus dapat diaudit. Simpan respons asli dan keputusan pengodean agar kasus ambigu dapat ditinjau kembali.

- pisahkan skor ketepatan dari skor keyakinan atau kualitas alasan;
- laporkan kasus tidak lengkap dan aturan penanganannya;
- uji konsistensi penilai untuk data berkode;
- hindari mengubah data ordinal menjadi ukuran interval tanpa alasan.

Dalam praktik kelas, Tes Pilihan Ganda Konvensional dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.1.9 Interpretasi dan Penetapan Kategori

Interpretasi hasil Tes Pilihan Ganda Konvensional harus bergerak dari respons menuju pola, bukan langsung dari satu jawaban menuju label. Status konseptual menjadi lebih kuat apabila gagasan yang sama muncul pada beberapa butir, representasi, atau konteks. (Haladyna et al., 2002; Rodriguez, 2005; Tarrant et al., 2009; Treagust, 1988; Kane, 2013)

Kategori dapat mencakup pemahaman ilmiah, pemahaman parsial, kurang pengetahuan, respons tidak konsisten, false positive, false negative, dan miskonsepsi. Tidak semua instrumen mampu membedakan seluruh kategori tersebut; skema harus mengikuti informasi yang benar-benar tersedia.

Ambang, seperti batas keyakinan atau persentase pemilihan, perlu memiliki rasional dan diuji melalui analisis sensitivitas. Perubahan kecil pada ambang dapat mengubah prevalensi kategori secara berarti.

Interpretasi yang bertanggung jawab menyatakan bahwa kategori berlaku pada konsep, bentuk tugas, waktu, dan kondisi tertentu. Hasil merupakan hipotesis kerja untuk pembelajaran, bukan diagnosis psikologis permanen.

- gunakan lebih dari satu butir untuk konsepsi penting;
- periksa konsistensi antara jawaban dan alasan;
- tandai kasus ambigu untuk konfirmasi;
- hubungkan kategori dengan rekomendasi yang berbeda.

Dalam praktik kelas, Tes Pilihan Ganda Konvensional dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.1.10 Validitas dan Kualitas Inferensi

Validitas Tes Pilihan Ganda Konvensional berkaitan dengan kualitas interpretasi dan penggunaan hasil, bukan sifat tetap yang dimiliki instrumen. Fokus khususnya adalah kesesuaian butir dengan domain konsep, keterbacaan, representasi kimia, dan inferensi yang diizinkan oleh skor. (AERA, APA, & NCME, 2014; Messick, 1995; Kane, 2013; Haynes et al., 1995)

Bukti berdasarkan isi menilai keterwakilan domain dan akurasi konsep. Bukti proses respons memeriksa bagaimana peserta didik memahami tugas. Bukti struktur internal menilai kesesuaian pola respons. Bukti hubungan dengan variabel lain dan konsekuensi penggunaan melengkapi argumen.

Ancaman validitas dapat berasal dari konten yang kurang terwakili, tuntutan bahasa berlebihan, petunjuk opsi, rubrik yang tidak stabil, keyakinan yang tidak terkalibrasi, atau algoritma yang belajar dari label keliru.

Dokumen validasi perlu menghubungkan setiap klaim dengan bukti dan alternatif penjelasan. Menampilkan satu indeks tidak cukup; validasi adalah program argumentatif yang diperbarui setiap kali instrumen dipakai pada populasi atau tujuan baru.

- bukti isi dan kesesuaian kurikulum;
- bukti proses respons melalui wawancara kognitif;
- bukti struktur dan fungsi butir;
- bukti kegunaan serta konsekuensi keputusan.

Dalam praktik kelas, Tes Pilihan Ganda Konvensional dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.1.11 Reliabilitas, Keajegan, dan Analisis Empiris

Keajegan Tes Pilihan Ganda Konvensional harus dinilai sesuai jenis data. Pada respons objektif dapat digunakan konsistensi internal, stabilitas waktu, atau model pengukuran. Pada data kualitatif, perhatian diarahkan pada konsistensi pengodean, audit trail, dan ketepatan penerapan definisi kode. (Kuder & Richardson, 1937; Cronbach, 1951; Ding et al., 2006; Wren & Barbera, 2014; AERA, APA, & NCME, 2014)

Reliabilitas tinggi tidak otomatis menunjukkan diagnosis yang benar. Butir yang berulang-ulang mengukur hafalan dapat menghasilkan koefisien tinggi tetapi tidak mewakili pemahaman konseptual. Sebaliknya, konstruk yang multidimensi tidak selalu layak dipaksa menjadi satu skala.

Analisis empiris perlu memeriksa kesukaran, daya beda, fungsi opsi, pola data hilang, dan keterhubungan komponen. Untuk instrumen bertingkat, reliabilitas dapat dihitung pada setiap tier serta kombinasi yang mempunyai makna teoritis.

Ukuran sampel, distribusi kemampuan, dan homogenitas konten memengaruhi statistik. Karena itu, laporan harus menyatakan kondisi estimasi dan tidak menyalin batas angka sebagai aturan universal.

- pilih statistik yang sesuai dengan jenis skor;
- laporkan interval atau ketidakpastian bila tersedia;
- periksa butir secara substantif dan statistik;
- uji kembali setelah revisi atau perubahan platform.

Dalam praktik kelas, Tes Pilihan Ganda Konvensional dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.1.12 Contoh Penerapan dalam Pembelajaran Kimia

Penerapan Tes Pilihan Ganda Konvensional dalam kimia dapat dimulai dari konsep membedakan asam kuat dari larutan berkonsentrasi tinggi. Pengembang perlu menuliskan konsep ilmiah target, miskonsepsi yang mungkin, bukti respons yang diharapkan, dan keputusan pembelajaran setelah hasil diperoleh. (Haladyna et al., 2002; Rodriguez, 2005; Tarrant et al., 2009; Treagust, 1988; Nakhleh, 1992)

Contoh kedua adalah menafsirkan arah pergeseran kesetimbangan. Stimulus sebaiknya menghubungkan fenomena dengan model partikel atau persamaan, lalu meminta peserta didik menunjukkan hubungan sebab-akibat. Jawaban yang hanya menyebut istilah belum cukup menjadi bukti pemahaman.

Pada mengenali hubungan energi aktivasi dan katalis, diagnosis dapat memeriksa apakah peserta didik membedakan perubahan laju, keadaan setimbang, energi, dan komposisi. Kesalahan sering muncul ketika satu aturan dipindahkan ke konteks yang tidak sesuai.

Kasus membedakan molekul, ion, dan unit formula menunjukkan pentingnya representasi. Respons verbal, gambar, simbol, dan hitungan dapat tidak konsisten. Hasil instrumen sebaiknya digunakan untuk memilih demonstrasi, diskusi, contoh-noncontoh, atau tugas remediasi yang secara langsung menantang hubungan salah.

- tentukan proposisi ilmiah dan konsepsi alternatif;
- gunakan minimal dua bentuk representasi;
- minta alasan atau bukti ketika format memungkinkan;
- rancang tindak lanjut berdasarkan pola, bukan hanya skor.

Dalam praktik kelas, Tes Pilihan Ganda Konvensional dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.1.13 Kelebihan, Keterbatasan, dan Potensi Bias

Kekuatan utama Tes Pilihan Ganda Konvensional meliputi efisien untuk kelompok besar, mudah dikoreksi secara otomatis, cakupan materi dapat luas, dan hasil antarwaktu mudah dibandingkan. Kekuatan tersebut menjelaskan situasi ketika instrumen layak dipilih dibanding teknik lain. (Haladyna et al., 2002; Rodriguez, 2005; Tarrant et al., 2009; Gurel et al., 2015; AERA, APA, & NCME, 2014)

Keterbatasannya mencakup jawaban benar dapat berasal dari tebakan, jawaban salah tidak otomatis merupakan miskonsepsi, alasan peserta didik tidak terekam, dan opsi yang lemah menurunkan daya diagnosis. Keterbatasan harus diperlakukan sebagai batas inferensi, bukan sekadar paragraf formal yang tidak memengaruhi analisis.

Potensi bias dapat berasal dari bahasa, pengalaman konteks, akses teknologi, kecemasan, gaya menggunakan skala, keterampilan menggambar, atau kemampuan menulis. Instrumen perlu diuji pada kelompok yang beragam dan ditinjau untuk fungsi berbeda antarkelompok bila data mencukupi.

Keputusan memilih instrumen sebaiknya mempertimbangkan kedalaman data, ukuran kelompok, waktu, keterampilan analisis, tujuan diagnosis, dan konsekuensi kesalahan klasifikasi. Tidak ada satu metode yang unggul untuk semua tujuan.

- Kekuatan: efisien untuk kelompok besar.
- Kekuatan: mudah dikoreksi secara otomatis.
- Batas: jawaban benar dapat berasal dari tebakan.
- Batas: jawaban salah tidak otomatis merupakan miskonsepsi.

Dalam praktik kelas, Tes Pilihan Ganda Konvensional dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

Aspek	Temuan	Respons Metodologis
Kekuatan	efisien untuk kelompok besar	Pilih ketika sesuai tujuan
Keterbatasan	jawaban benar dapat berasal dari tebakan	Tambahkan bukti atau kontrol
Bias potensial	Bahasa, konteks, perangkat, gaya respons	Uji proses respons

6.1.14 Integrasi Digital, Etika, dan Tata Kelola Data

Integrasi digital Tes Pilihan Ganda Konvensional dapat berupa bank soal adaptif, analisis opsi waktu nyata, dan umpan balik tertunda yang tidak membocorkan kunci sebelum diagnosis selesai. Teknologi seharusnya memperluas kualitas bukti, bukan hanya memindahkan lembar kertas ke layar. (Baker & Inventado, 2014; Romero & Ventura, 2020; AERA, APA, & NCME, 2014; Haladyna et al., 2002)

Platform perlu menyimpan versi instrumen, urutan respons, timestamp yang relevan, aturan skor, dan log perubahan. Namun, hanya data yang dibutuhkan untuk tujuan diagnosis yang boleh dikumpulkan; pengumpulan berlebihan meningkatkan risiko privasi tanpa manfaat ilmiah.

Aspek etika utama adalah bahwa hasil tidak boleh diberi label miskonsepsi secara individual tanpa bukti alasan atau konsistensi lintasbutir. Peserta didik dan guru perlu mengetahui data apa yang dikumpulkan, untuk tujuan apa, siapa yang dapat mengakses, berapa lama disimpan, dan bagaimana kesalahan data diperbaiki.

Algoritma otomatis harus dapat diaudit. Kategori yang dihasilkan sistem perlu ditautkan ke respons mentah dan aturan keputusan. Kasus berisiko atau ambigu harus ditinjau manusia sebelum tindak lanjut penting dilakukan.

- gunakan identitas pseudonim dan akses berbasis peran;
- uji tampilan rumus, subskrip, grafik, dan gambar;
- sediakan opsi aksesibilitas dan alternatif perangkat;
- dokumentasikan algoritma serta perubahan versi.

Dalam praktik kelas, Tes Pilihan Ganda Konvensional dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.1.15 Sintesis dan Pedoman Implementasi

Tes Pilihan Ganda Konvensional bermanfaat apabila dipilih berdasarkan tujuan dan digunakan sesuai batasnya. Instrumen ini menghasilkan respons benar-salah dan frekuensi pemilihan setiap opsi, lalu data diolah melalui skor 1 untuk jawaban benar dan 0 untuk jawaban salah, dilengkapi analisis distraktor per butir. (Haladyna et al., 2002; Rodriguez, 2005; Tarrant et al., 2009; AERA, APA, & NCME, 2014; Kane, 2013)

Rangkaian kerja yang disarankan dimulai dari analisis konsep, identifikasi konsepsi alternatif, desain tugas, telaah ahli, uji proses respons, uji lapangan, analisis, penetapan kategori, dan penyusunan tindak lanjut. Setiap tahap harus meninggalkan dokumentasi yang dapat diperiksa.

Hasil diagnosis tidak berhenti pada tabel prevalensi. Guru perlu memilih konsepsi prioritas berdasarkan frekuensi, keyakinan, posisi sebagai konsep prasyarat, dampak terhadap topik lanjutan, dan kemungkinan diperbaiki melalui pembelajaran.

Prinsip akhirnya ialah kehati-hatian inferensial: sebutkan apa yang benar-benar ditunjukkan data, bedakan bukti kuat dari indikasi, gunakan triangulasi ketika diperlukan, dan hindari label permanen pada peserta didik.

- selaraskan tujuan, konstruk, tugas, skor, kategori, dan tindakan;
- utamakan konsepsi yang berdampak pada pembelajaran lanjutan;
- gabungkan bukti bila keputusan memiliki konsekuensi besar;
- evaluasi kembali efektivitas tindak lanjut melalui diagnosis ulang.

Dalam praktik kelas, Tes Pilihan Ganda Konvensional dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.2 Tes Pilihan Ganda Beralasan

6.2.1 Landasan Konseptual dan Kedudukan Instrumen

Secara metodologis, Tes Pilihan Ganda Beralasan menghubungkan perilaku yang dapat diamati dengan konstruksi pemahaman yang tidak dapat diamati secara langsung. Dalam bab ini, instrumen tersebut didefinisikan sebagai pilihan ganda yang diikuti permintaan alasan singkat, uraian pembenaran, atau pemilihan alasan dari daftar terpisah yang belum dibakukan sebagai format dua tingkat. Definisi operasional ini menegaskan bentuk bukti yang dikumpulkan sekaligus membatasi jenis kesimpulan yang layak dibuat. (Treagust, 1988; Chandrasegaran et al., 2007; Chi, 1997)

Klaim utamanya adalah bahwa instrumen ini membedakan hasil jawaban dari kualitas penalaran sehingga diagnosis tidak berhenti pada benar atau salahnya opsi utama. Klaim tersebut harus dipisahkan dari anggapan bahwa setiap jawaban salah merupakan miskonsepsi. Kesalahan dapat berasal dari kurang pengetahuan, ketidakcermatan, kesulitan bahasa, salah hitung, atau ketidakfamiliaran dengan representasi.

Data pokok yang dihasilkan berupa kombinasi opsi terpilih dan kualitas justifikasi verbal atau tertulis. Data itu baru memperoleh makna diagnostik setelah dikaitkan dengan peta konsep sasaran, kemungkinan konsepsi alternatif, dan aturan interpretasi yang telah ditetapkan sebelum analisis.

Dengan demikian, mutu diagnosis ditentukan oleh keselarasan antara tujuan, desain, pelaksanaan, analisis, dan keputusan yang diambil. Posisi terbaik Tes Pilihan Ganda Beralasan adalah sebagai bagian dari sistem asesmen berlapis yang dapat dipadukan dengan bukti lain ketika konsekuensi keputusan semakin besar.

- menentukan konsep dan proposisi yang hendak didiagnosis;
- menjelaskan perilaku respons yang dianggap sebagai bukti;
- membatasi klaim sesuai kekuatan instrumen;
- menautkan hasil dengan tindakan pembelajaran yang realistis.

Dalam praktik kelas, Tes Pilihan Ganda Beralasan dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.2.2 Perkembangan Historis dan Rasional Penggunaan

Perkembangan Tes Pilihan Ganda Beralasan tidak dapat dipisahkan dari perubahan pandangan tentang belajar sains. Ketika peserta didik tidak lagi dipandang sebagai penerima informasi yang kosong, instrumen mulai dirancang untuk mengungkap gagasan awal, pola penalaran, dan keyakinan yang dibawa ke kelas. (Treagust, 1988; Chandrasegaran et al., 2007; Chi, 1997; AERA, APA, & NCME, 2014)

Pada tahap awal, penilaian banyak mengandalkan benar-salah atau skor total. Perkembangan penelitian konsepsi peserta didik kemudian mendorong penambahan alasan, keyakinan, tugas kinerja, rekaman proses, dan bukti sosial. Tes Pilihan Ganda Beralasan merupakan salah satu jawaban metodologis terhadap kebutuhan tersebut.

Dalam pendidikan kimia, kebutuhan diagnosis semakin kuat karena peserta didik harus mengoordinasikan fenomena makroskopik, model submikroskopik, simbol, grafik, dan perhitungan. Instrumen yang hanya menguji hafalan definisi sering gagal memperlihatkan cara representasi itu dihubungkan.

Sejarah penggunaannya juga menunjukkan bahwa format yang lebih kompleks tidak otomatis lebih baik. Setiap tambahan tier, rubrik, atau kanal data harus mempunyai fungsi yang jelas dan didukung bukti bahwa informasi tambahan benar-benar memperbaiki keputusan.

- pergeseran dari pengukuran skor menuju diagnosis pola pemahaman;
- penggunaan respons terbuka dan wawancara untuk menemukan konsepsi alternatif;
- pengembangan teknik kuantitatif agar diagnosis dapat dilakukan pada kelompok besar;
- integrasi data digital dan metakognitif pada perkembangan mutakhir.

Dalam praktik kelas, Tes Pilihan Ganda Beralasan dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.2.3 Konstruk, Klaim, dan Bukti Diagnosis

Konstruk yang dinilai melalui Tes Pilihan Ganda Beralasan bukan objek yang terlihat langsung. Yang diamati adalah kombinasi opsi terpilih dan kualitas justifikasi verbal atau tertulis; dari sana dibuat inferensi tentang pemahaman konsep, miskonsepsi, pengetahuan parsial, atau ketidakpastian. (AERA, APA, & NCME, 2014; Taber, 2002; Messick, 1995; Kane, 2013)

Klaim diagnosis perlu dirumuskan sebagai argumen. Pertama, tugas harus memerlukan konsep sasaran. Kedua, respons harus diberi skor atau kode secara konsisten. Ketiga, pola skor harus mempunyai hubungan yang masuk akal dengan struktur pengetahuan. Keempat, penggunaan hasil harus menghasilkan keputusan yang proporsional.

Sumber ancaman utama adalah ketidakjelasan konstruk. Misalnya, peserta didik dapat gagal karena istilah yang ambigu, beban membaca, representasi yang tidak dikenal, atau prosedur matematika. Kegagalan semacam itu tidak boleh langsung diperlakukan sebagai miskonsepsi kimia.

Oleh sebab itu, setiap pengembang perlu menulis pernyataan klaim-bukti: instrumen mengukur apa, perilaku apa yang menjadi indikator, alternatif penjelasan apa yang mungkin, dan bukti tambahan apa yang diperlukan untuk memperkuat kesimpulan.

- konstruk utama: pemahaman konseptual melalui tes objektif dengan justifikasi;
- bukti observabel: respons, alasan, tindakan, atau representasi;
- inferensi: profil konsepsi pada konten dan konteks tertentu;
- penggunaan: perencanaan pembelajaran, remediasi, atau penelitian.

Dalam praktik kelas, Tes Pilihan Ganda Beralasan dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.2.4 Struktur, Komponen, dan Alur Respons

Struktur Tes Pilihan Ganda Beralasan dibangun dari komponen yang saling bergantung. Komponen inti meliputi pertanyaan utama, opsi jawaban, ruang alasan atau daftar alasan, rubrik kualitas alasan, dan kode hubungan jawaban-alasan. Hilangnya satu komponen dapat mengubah makna data yang diperoleh. (Treagust, 1988; Chandrasegaran et al., 2007; Chi, 1997; AERA, APA, & NCME, 2014)

Urutan penyajian harus ditetapkan secara sadar. Komponen yang muncul lebih awal dapat memberi petunjuk pada komponen berikutnya, sedangkan permintaan alasan atau keyakinan dapat mengubah cara peserta didik memeriksa jawabannya. Penguncian respons dan uji urutan diperlukan pada format digital.

Setiap komponen membutuhkan spesifikasi. Stimulus harus memuat informasi cukup, opsi atau prompt harus sejajar dengan indikator, skala harus memiliki jangkar, dan rubrik harus menjelaskan perbedaan kategori yang berdekatan.

Struktur yang baik memungkinkan peneliti menelusuri kembali setiap kategori hasil ke respons mentah. Transparansi ini penting agar guru tidak hanya menerima label otomatis tanpa memahami alasan di baliknya.

- pertanyaan utama;
- opsi jawaban;
- ruang alasan atau daftar alasan;
- rubrik kualitas alasan;
- kode hubungan jawaban-alasan;

Dalam praktik kelas, Tes Pilihan Ganda Beralasan dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

No.	Komponen	Fungsi Diagnostik
1	pertanyaan utama	Bukti yang harus ditelusuri
2	opsi jawaban	Bukti yang harus ditelusuri
3	ruang alasan atau daftar alasan	Bukti yang harus ditelusuri
4	rubrik kualitas alasan	Bukti yang harus ditelusuri

6.2.5 Desain Stimulus, Tugas, atau Pertanyaan

Desain stimulus pada Tes Pilihan Ganda Beralasan harus memunculkan penggunaan konsep, bukan sekadar pengenalan istilah. Situasi dapat berbentuk fenomena, gambar partikel, grafik, data eksperimen, persamaan, atau masalah kontekstual, tetapi seluruh informasi yang tidak relevan perlu dikendalikan. (Haladyna et al., 2002; Treagust, 1988; Chandrasegaran et al., 2007; Chi, 1997)

Dalam kimia, satu stimulus sering menggabungkan beberapa tingkat representasi. Pengembang harus memutuskan apakah tujuan butir adalah menerjemahkan representasi, menjelaskan hubungan sebab-akibat, membedakan contoh dan noncontoh, atau menerapkan prinsip pada konteks baru.

Bahasa perlu ringkas tanpa menghilangkan syarat ilmiah. Kata seperti kuat, stabil, jenuh, spontan, atau setimbang mempunyai arti teknis dan sehari-hari yang berbeda. Ketidakjelasan istilah dapat menghasilkan kesalahan yang berasal dari bahasa, bukan struktur konsep.

Desain yang baik juga menghindari petunjuk format, pilihan terlalu panjang, negasi bertumpuk, gambar dekoratif, dan data yang tidak diperlukan. Setiap elemen harus mempunyai alasan diagnostik.

- gunakan konteks yang cukup dikenal tetapi tidak menghafalkan contoh;
- pastikan hanya satu interpretasi ilmiah yang menjadi target;
- uji semua simbol, satuan, grafik, dan gambar;
- sediakan ruang bagi respons yang tidak diprediksi bila tujuan masih eksploratif.

Dalam praktik kelas, Tes Pilihan Ganda Beralasan dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.2.6 Prosedur Pengembangan Secara Sistematis

Pengembangan Tes Pilihan Ganda Beralasan harus mengikuti alur yang terdokumentasi. Langkah yang lazim ialah menentukan klaim konseptual, menulis item utama, mengembangkan prompt alasan, menguji rubrik pada respons nyata, kemudian melatih penilai, menelaah konsistensi jawaban-alasan, merevisi bahasa, mengintegrasikan hasil. Urutan dapat berulang karena temuan uji coba sering memaksa revisi konstruk atau redaksi. (Treagust, 1988; Chandrasegaran et al., 2007; Chi, 1997; Adams & Wieman, 2011; AERA, APA, & NCME, 2014)

Tahap konseptual dimulai dengan analisis kurikulum, literatur miskonsepsi, dan peta konsep. Tahap empiris awal menggunakan wawancara, respons terbuka, atau observasi untuk memastikan bahwa kategori yang dirancang benar-benar muncul pada populasi sasaran.

Setiap versi instrumen perlu diberi nomor, tanggal, dan catatan perubahan. Revisi harus menyebutkan masalah yang ditemukan, bukti yang digunakan, perubahan yang dilakukan, dan alasan mempertahankan atau menghapus butir.

Pengembangan dianggap belum selesai ketika hanya menghasilkan butir yang tampak baik. Instrumen perlu diuji pada peserta didik yang serupa dengan pengguna sasaran, lalu dianalisis secara konseptual dan empiris sebelum dipakai untuk keputusan penting.

- 1. menentukan klaim konseptual.
- 2. menulis item utama.
- 3. mengembangkan prompt alasan.
- 4. menguji rubrik pada respons nyata.
- 5. melatih penilai.

Dalam praktik kelas, Tes Pilihan Ganda Beralasan dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.2.7 Pelaksanaan dan Pengumpulan Data

Pelaksanaan Tes Pilihan Ganda Beralasan harus distandardisasi secukupnya agar perbedaan hasil lebih mungkin berasal dari pemahaman, bukan variasi instruksi, waktu, alat, atau bantuan. Petunjuk perlu menjelaskan tujuan umum tanpa mengajarkan jawaban. (Treagust, 1988; Chandrasegaran et al., 2007; AERA, APA, & NCME, 2014)

Kondisi administrasi mencakup waktu, urutan tugas, kesempatan bertanya, penggunaan kalkulator, akses tabel periodik, kualitas gambar, dan perangkat digital. Pada instrumen kualitatif, standardisasi berarti konsistensi fokus dan prinsip probing, bukan membaca naskah secara kaku.

Data yang dikumpulkan ialah kombinasi opsi terpilih dan kualitas justifikasi verbal atau tertulis. Data mentah harus disimpan bersama metadata yang diperlukan, seperti kode butir, versi instrumen, waktu administrasi, dan kondisi khusus, tetapi hanya data yang relevan yang boleh dikumpulkan.

Sebelum administrasi luas, lakukan uji teknis. Pastikan tautan, urutan, penguncian respons, rekaman, penyimpanan, serta tampilan rumus dan subskrip bekerja. Gangguan teknis perlu dicatat sebagai bagian dari interpretasi.

- berikan instruksi yang sama dan mudah dipahami;
- catat kondisi yang dapat memengaruhi respons;
- hindari umpan balik sebelum seluruh respons diagnostik selesai;
- sediakan prosedur penanganan data tidak lengkap.

Dalam praktik kelas, Tes Pilihan Ganda Beralasan dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.2.8 Penskoran, Pengodean, dan Pengolahan Data

Penskoran Tes Pilihan Ganda Beralasan dapat diringkas sebagai penskoran analitik yang memisahkan ketepatan jawaban, relevansi alasan, penggunaan bukti, dan koherensi penjelasan. Aturan ini harus ditulis sebelum melihat hasil utama agar keputusan tidak disesuaikan secara oportunistik dengan pola data. (Treagust, 1988; Chandrasegaran et al., 2007; Chi, 1997; AERA, APA, & NCME, 2014)

Skor total jarang cukup untuk diagnosis. Analisis perlu mempertahankan informasi tingkat butir, konsep, opsi, alasan, keyakinan, proses, atau representasi. Penggabungan data terlalu dini dapat menyembunyikan pola yang secara pedagogis penting.

Untuk respons terbuka atau kualitatif, buku kode harus mencantumkan definisi, contoh positif, contoh batas, noncontoh, dan aturan ketika satu respons memuat beberapa gagasan. Sebagian data perlu dinilai oleh lebih dari satu penilai.

Setiap transformasi data—misalnya dari respons ke kode, dari kode ke kategori, dan dari kategori ke persentase—harus dapat diaudit. Simpan respons asli dan keputusan pengodean agar kasus ambigu dapat ditinjau kembali.

- pisahkan skor ketepatan dari skor keyakinan atau kualitas alasan;
- laporkan kasus tidak lengkap dan aturan penanganannya;
- uji konsistensi penilai untuk data berkode;
- hindari mengubah data ordinal menjadi ukuran interval tanpa alasan.

Dalam praktik kelas, Tes Pilihan Ganda Beralasan dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.2.9 Interpretasi dan Penetapan Kategori

Interpretasi hasil Tes Pilihan Ganda Beralasan harus bergerak dari respons menuju pola, bukan langsung dari satu jawaban menuju label. Status konseptual menjadi lebih kuat apabila gagasan yang sama muncul pada beberapa butir, representasi, atau konteks. (Treagust, 1988; Chandrasegaran et al., 2007; Chi, 1997; AERA, APA, & NCME, 2014; Kane, 2013)

Kategori dapat mencakup pemahaman ilmiah, pemahaman parsial, kurang pengetahuan, respons tidak konsisten, false positive, false negative, dan miskonsepsi. Tidak semua instrumen mampu membedakan seluruh kategori tersebut; skema harus mengikuti informasi yang benar-benar tersedia.

Ambang, seperti batas keyakinan atau persentase pemilihan, perlu memiliki rasional dan diuji melalui analisis sensitivitas. Perubahan kecil pada ambang dapat mengubah prevalensi kategori secara berarti.

Interpretasi yang bertanggung jawab menyatakan bahwa kategori berlaku pada konsep, bentuk tugas, waktu, dan kondisi tertentu. Hasil merupakan hipotesis kerja untuk pembelajaran, bukan diagnosis psikologis permanen.

- gunakan lebih dari satu butir untuk konsepsi penting;
- periksa konsistensi antara jawaban dan alasan;
- tandai kasus ambigu untuk konfirmasi;
- hubungkan kategori dengan rekomendasi yang berbeda.

Dalam praktik kelas, Tes Pilihan Ganda Beralasan dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.2.10 Validitas dan Kualitas Inferensi

Validitas Tes Pilihan Ganda Beralasan berkaitan dengan kualitas interpretasi dan penggunaan hasil, bukan sifat tetap yang dimiliki instrumen. Fokus khususnya adalah kecocokan rubrik dengan tingkat penjelasan yang ditargetkan dan bukti bahwa alasan benar-benar merepresentasikan pemahaman kimia. (AERA, APA, & NCME, 2014; Messick, 1995; Kane, 2013; Haynes et al., 1995)

Bukti berdasarkan isi menilai keterwakilan domain dan akurasi konsep. Bukti proses respons memeriksa bagaimana peserta didik memahami tugas. Bukti struktur internal menilai kesesuaian pola respons. Bukti hubungan dengan variabel lain dan konsekuensi penggunaan melengkapi argumen.

Ancaman validitas dapat berasal dari konten yang kurang terwakili, tuntutan bahasa berlebihan, petunjuk opsi, rubrik yang tidak stabil, keyakinan yang tidak terkalibrasi, atau algoritma yang belajar dari label keliru.

Dokumen validasi perlu menghubungkan setiap klaim dengan bukti dan alternatif penjelasan. Menampilkan satu indeks tidak cukup; validasi adalah program argumentatif yang diperbarui setiap kali instrumen dipakai pada populasi atau tujuan baru.

- bukti isi dan kesesuaian kurikulum;
- bukti proses respons melalui wawancara kognitif;
- bukti struktur dan fungsi butir;
- bukti kegunaan serta konsekuensi keputusan.

Dalam praktik kelas, Tes Pilihan Ganda Beralasan dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.2.11 Reliabilitas, Keajegan, dan Analisis Empiris

Keajegan Tes Pilihan Ganda Beralasan harus dinilai sesuai jenis data. Pada respons objektif dapat digunakan konsistensi internal, stabilitas waktu, atau model pengukuran. Pada data kualitatif, perhatian diarahkan pada konsistensi pengodean, audit trail, dan ketepatan penerapan definisi kode. (Kuder & Richardson, 1937; Cronbach, 1951; Ding et al., 2006; Wren & Barbera, 2014; AERA, APA, & NCME, 2014)

Reliabilitas tinggi tidak otomatis menunjukkan diagnosis yang benar. Butir yang berulang-ulang mengukur hafalan dapat menghasilkan koefisien tinggi tetapi tidak mewakili pemahaman konseptual. Sebaliknya, konstruk yang multidimensi tidak selalu layak dipaksa menjadi satu skala.

Analisis empiris perlu memeriksa kesukaran, daya beda, fungsi opsi, pola data hilang, dan keterhubungan komponen. Untuk instrumen bertingkat, reliabilitas dapat dihitung pada setiap tier serta kombinasi yang mempunyai makna teoritis.

Ukuran sampel, distribusi kemampuan, dan homogenitas konten memengaruhi statistik. Karena itu, laporan harus menyatakan kondisi estimasi dan tidak menyalin batas angka sebagai aturan universal.

- pilih statistik yang sesuai dengan jenis skor;
- laporkan interval atau ketidakpastian bila tersedia;
- periksa butir secara substantif dan statistik;
- uji kembali setelah revisi atau perubahan platform.

Dalam praktik kelas, Tes Pilihan Ganda Beralasan dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.2.12 Contoh Penerapan dalam Pembelajaran Kimia

Penerapan Tes Pilihan Ganda Beralasan dalam kimia dapat dimulai dari konsep menjelaskan mengapa NaCl padat tidak menghantarkan listrik. Pengembang perlu menuliskan konsep ilmiah target, miskonsepsi yang mungkin, bukti respons yang diharapkan, dan keputusan pembelajaran setelah hasil diperoleh. (Treagust, 1988; Chandrasegaran et al., 2007; Chi, 1997; AERA, APA, & NCME, 2014; Nakhleh, 1992)

Contoh kedua adalah menerangkan mengapa katalis tidak mengubah tetapan kesetimbangan. Stimulus sebaiknya menghubungkan fenomena dengan model partikel atau persamaan, lalu meminta peserta didik menunjukkan hubungan sebab-akibat. Jawaban yang hanya menyebut istilah belum cukup menjadi bukti pemahaman.

Pada membedakan ikatan intramolekul dan gaya antarmolekul, diagnosis dapat memeriksa apakah peserta didik membedakan perubahan laju, keadaan setimbang, energi, dan komposisi. Kesalahan sering muncul ketika satu aturan dipindahkan ke konteks yang tidak sesuai.

Kasus menafsirkan pH larutan asam lemah menunjukkan pentingnya representasi. Respons verbal, gambar, simbol, dan hitungan dapat tidak konsisten. Hasil instrumen sebaiknya digunakan untuk memilih demonstrasi, diskusi, contoh-noncontoh, atau tugas remediasi yang secara langsung menantang hubungan salah.

- tentukan proposisi ilmiah dan konsepsi alternatif;
- gunakan minimal dua bentuk representasi;
- minta alasan atau bukti ketika format memungkinkan;
- rancang tindak lanjut berdasarkan pola, bukan hanya skor.

Dalam praktik kelas, Tes Pilihan Ganda Beralasan dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.2.13 Kelebihan, Keterbatasan, dan Potensi Bias

Kekuatan utama Tes Pilihan Ganda Beralasan meliputi mengurangi interpretasi berlebihan atas jawaban salah, mengungkap penalaran parsial, dapat digunakan sebagai jembatan menuju tes dua tingkat, dan memberi bahan pengembangan distraktor. Kekuatan tersebut menjelaskan situasi ketika instrumen layak dipilih dibanding teknik lain. (Treagust, 1988; Chandrasegaran et al., 2007; Chi, 1997; Gurel et al., 2015; AERA, APA, & NCME, 2014)

Keterbatasannya mencakup koreksi alasan terbuka memerlukan waktu, rubrik dapat menghasilkan perbedaan penilai, peserta didik mungkin memiliki konsep tetapi kesulitan menulis, dan alasan yang disediakan peneliti dapat membatasi ekspresi. Keterbatasan harus diperlakukan sebagai batas inferensi, bukan sekadar paragraf formal yang tidak memengaruhi analisis.

Potensi bias dapat berasal dari bahasa, pengalaman konteks, akses teknologi, kecemasan, gaya menggunakan skala, keterampilan menggambar, atau kemampuan menulis. Instrumen perlu diuji pada kelompok yang beragam dan ditinjau untuk fungsi berbeda antarkelompok bila data mencukupi.

Keputusan memilih instrumen sebaiknya mempertimbangkan kedalaman data, ukuran kelompok, waktu, keterampilan analisis, tujuan diagnosis, dan konsekuensi kesalahan klasifikasi. Tidak ada satu metode yang unggul untuk semua tujuan.

- Kekuatan: mengurangi interpretasi berlebihan atas jawaban salah.
- Kekuatan: mengungkap penalaran parsial.
- Batas: koreksi alasan terbuka memerlukan waktu.
- Batas: rubrik dapat menghasilkan perbedaan penilai.

Dalam praktik kelas, Tes Pilihan Ganda Beralasan dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

Aspek	Temuan	Respons Metodologis
Kekuatan	mengurangi interpretasi berlebihan atas jawaban salah	Pilih ketika sesuai tujuan
Keterbatasan	koreksi alasan terbuka memerlukan waktu	Tambahkan bukti atau kontrol
Bias potensial	Bahasa, konteks, perangkat, gaya respons	Uji proses respons

6.2.14 Integrasi Digital, Etika, dan Tata Kelola Data

Integrasi digital Tes Pilihan Ganda Beralasan dapat berupa kolom alasan singkat, pengodean berbantuan teks, dan pemeriksaan manual untuk kasus ambigu. Teknologi seharusnya memperluas kualitas bukti, bukan hanya memindahkan lembar kertas ke layar. (Baker & Inventado, 2014; Romero & Ventura, 2020; AERA, APA, & NCME, 2014; Treagust, 1988)

Platform perlu menyimpan versi instrumen, urutan respons, timestamp yang relevan, aturan skor, dan log perubahan. Namun, hanya data yang dibutuhkan untuk tujuan diagnosis yang boleh dikumpulkan; pengumpulan berlebihan meningkatkan risiko privasi tanpa manfaat ilmiah.

Aspek etika utama adalah bahwa variasi kemampuan bahasa harus dipisahkan dari ketepatan konseptual agar peserta didik tidak dirugikan. Peserta didik dan guru perlu mengetahui data apa yang dikumpulkan, untuk tujuan apa, siapa yang dapat mengakses, berapa lama disimpan, dan bagaimana kesalahan data diperbaiki.

Algoritma otomatis harus dapat diaudit. Kategori yang dihasilkan sistem perlu ditautkan ke respons mentah dan aturan keputusan. Kasus berisiko atau ambigu harus ditinjau manusia sebelum tindak lanjut penting dilakukan.

- gunakan identitas pseudonim dan akses berbasis peran;
- uji tampilan rumus, subskrip, grafik, dan gambar;
- sediakan opsi aksesibilitas dan alternatif perangkat;
- dokumentasikan algoritma serta perubahan versi.

Dalam praktik kelas, Tes Pilihan Ganda Beralasan dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.2.15 Sintesis dan Pedoman Implementasi

Tes Pilihan Ganda Beralasan bermanfaat apabila dipilih berdasarkan tujuan dan digunakan sesuai batasnya. Instrumen ini menghasilkan kombinasi opsi terpilih dan kualitas justifikasi verbal atau tertulis, lalu data diolah melalui penskoran analitik yang memisahkan ketepatan jawaban, relevansi alasan, penggunaan bukti, dan koherensi penjelasan. (Treagust, 1988; Chandrasegaran et al., 2007; Chi, 1997; AERA, APA, & NCME, 2014; Kane, 2013)

Rangkaian kerja yang disarankan dimulai dari analisis konsep, identifikasi konsepsi alternatif, desain tugas, telaah ahli, uji proses respons, uji lapangan, analisis, penetapan kategori, dan penyusunan tindak lanjut. Setiap tahap harus meninggalkan dokumentasi yang dapat diperiksa.

Hasil diagnosis tidak berhenti pada tabel prevalensi. Guru perlu memilih konsepsi prioritas berdasarkan frekuensi, keyakinan, posisi sebagai konsep prasyarat, dampak terhadap topik lanjutan, dan kemungkinan diperbaiki melalui pembelajaran.

Prinsip akhirnya ialah kehati-hatian inferensial: sebutkan apa yang benar-benar ditunjukkan data, bedakan bukti kuat dari indikasi, gunakan triangulasi ketika diperlukan, dan hindari label permanen pada peserta didik.

- selaraskan tujuan, konstruk, tugas, skor, kategori, dan tindakan;
- utamakan konsepsi yang berdampak pada pembelajaran lanjutan;
- gabungkan bukti bila keputusan memiliki konsekuensi besar;
- evaluasi kembali efektivitas tindak lanjut melalui diagnosis ulang.

Dalam praktik kelas, Tes Pilihan Ganda Beralasan dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.3 Two-Tier Diagnostic Test

6.3.1 Landasan Konseptual dan Kedudukan Instrumen

Dalam diagnosis miskonsepsi, Two-Tier Diagnostic Test perlu dipahami bukan sekadar sebagai bentuk soal atau teknik pengumpulan data, melainkan sebagai perangkat inferensi. Dalam bab ini, instrumen tersebut didefinisikan sebagai instrumen yang memasangkan tingkat jawaban konten dengan tingkat alasan sehingga setiap butir menghasilkan bukti tentang keputusan dan dasar penalarannya. Definisi operasional ini menegaskan bentuk bukti yang dikumpulkan sekaligus membatasi jenis kesimpulan yang layak dibuat. (Treagust, 1988; Odom & Barrow, 1995; Chandrasegaran et al., 2007)

Klaim utamanya adalah bahwa instrumen ini mengidentifikasi pola konsepsi alternatif secara lebih spesifik daripada pilihan ganda satu tingkat, terutama ketika distraktor alasan berasal dari wawancara dan respons terbuka. Klaim tersebut harus dipisahkan dari anggapan bahwa setiap jawaban salah merupakan miskonsepsi. Kesalahan dapat berasal dari kurang pengetahuan, ketidakcermatan, kesulitan bahasa, salah hitung, atau ketidakfamiliaran dengan representasi.

Data pokok yang dihasilkan berupa kombinasi benar-salah pada tier jawaban dan tier alasan. Data itu baru memperoleh makna diagnostik setelah dikaitkan dengan peta konsep sasaran, kemungkinan konsepsi alternatif, dan aturan interpretasi yang telah ditetapkan sebelum analisis.

Karena miskonsepsi bersifat kontekstual dan kadang tersembunyi, bukti tunggal perlu ditafsirkan secara terbatas. Posisi terbaik Two-Tier Diagnostic Test adalah sebagai bagian dari sistem asesmen berlapis yang dapat dipadukan dengan bukti lain ketika konsekuensi keputusan semakin besar.

- menentukan konsep dan proposisi yang hendak didiagnosis;
- menjelaskan perilaku respons yang dianggap sebagai bukti;
- membatasi klaim sesuai kekuatan instrumen;
- menautkan hasil dengan tindakan pembelajaran yang realistis.

Dalam praktik kelas, Two-Tier Diagnostic Test dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.3.2 Perkembangan Historis dan Rasional Penggunaan

Perkembangan Two-Tier Diagnostic Test tidak dapat dipisahkan dari perubahan pandangan tentang belajar sains. Ketika peserta didik tidak lagi dipandang sebagai penerima informasi yang kosong, instrumen mulai dirancang untuk mengungkap gagasan awal, pola penalaran, dan keyakinan yang dibawa ke kelas. (Treagust, 1988; Odom & Barrow, 1995; Chandrasegaran et al., 2007; Çil, 2015)

Pada tahap awal, penilaian banyak mengandalkan benar-salah atau skor total. Perkembangan penelitian konsepsi peserta didik kemudian mendorong penambahan alasan, keyakinan, tugas kinerja, rekaman proses, dan bukti sosial. Two-Tier Diagnostic Test merupakan salah satu jawaban metodologis terhadap kebutuhan tersebut.

Dalam pendidikan kimia, kebutuhan diagnosis semakin kuat karena peserta didik harus mengoordinasikan fenomena makroskopik, model submikroskopik, simbol, grafik, dan perhitungan. Instrumen yang hanya menguji hafalan definisi sering gagal memperlihatkan cara representasi itu dihubungkan.

Sejarah penggunaannya juga menunjukkan bahwa format yang lebih kompleks tidak otomatis lebih baik. Setiap tambahan tier, rubrik, atau kanal data harus mempunyai fungsi yang jelas dan didukung bukti bahwa informasi tambahan benar-benar memperbaiki keputusan.

- pergeseran dari pengukuran skor menuju diagnosis pola pemahaman;
- penggunaan respons terbuka dan wawancara untuk menemukan konsepsi alternatif;
- pengembangan teknik kuantitatif agar diagnosis dapat dilakukan pada kelompok besar;
- integrasi data digital dan metakognitif pada perkembangan mutakhir.

Dalam praktik kelas, Two-Tier Diagnostic Test dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.3.3 Konstruk, Klaim, dan Bukti Diagnosis

Konstruk yang dinilai melalui Two-Tier Diagnostic Test bukan objek yang terlihat langsung. Yang diamati adalah kombinasi benar-salah pada tier jawaban dan tier alasan; dari sana dibuat inferensi tentang pemahaman konsep, miskonsepsi, pengetahuan parsial, atau ketidakpastian. (Gurel et al., 2015; AERA, APA, & NCME, 2014; Messick, 1995; Kane, 2013)

Klaim diagnosis perlu dirumuskan sebagai argumen. Pertama, tugas harus memerlukan konsep sasaran. Kedua, respons harus diberi skor atau kode secara konsisten. Ketiga, pola skor harus mempunyai hubungan yang masuk akal dengan struktur pengetahuan. Keempat, penggunaan hasil harus menghasilkan keputusan yang proporsional.

Sumber ancaman utama adalah ketidakjelasan konstruk. Misalnya, peserta didik dapat gagal karena istilah yang ambigu, beban membaca, representasi yang tidak dikenal, atau prosedur matematika. Kegagalan semacam itu tidak boleh langsung diperlakukan sebagai miskonsepsi kimia.

Oleh sebab itu, setiap pengembang perlu menulis pernyataan klaim-bukti: instrumen mengukur apa, perilaku apa yang menjadi indikator, alternatif penjelasan apa yang mungkin, dan bukti tambahan apa yang diperlukan untuk memperkuat kesimpulan.

- konstruk utama: pemahaman konseptual melalui tes diagnostik dua tingkat;
- bukti observabel: respons, alasan, tindakan, atau representasi;
- inferensi: profil konsepsi pada konten dan konteks tertentu;
- penggunaan: perencanaan pembelajaran, remediasi, atau penelitian.

Dalam praktik kelas, Two-Tier Diagnostic Test dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.3.4 Struktur, Komponen, dan Alur Respons

Struktur Two-Tier Diagnostic Test dibangun dari komponen yang saling bergantung. Komponen inti meliputi tier pertama berupa jawaban konten, tier kedua berupa alasan, kunci kombinasi jawaban-alasan, distraktor berbasis miskonsepsi, dan peta konsep dan proposisi. Hilangnya satu komponen dapat mengubah makna data yang diperoleh. (Treagust, 1988; Odom & Barrow, 1995; Chandrasegaran et al., 2007; Çil, 2015)

Urutan penyajian harus ditetapkan secara sadar. Komponen yang muncul lebih awal dapat memberi petunjuk pada komponen berikutnya, sedangkan permintaan alasan atau keyakinan dapat mengubah cara peserta didik memeriksa jawabannya. Penguncian respons dan uji urutan diperlukan pada format digital.

Setiap komponen membutuhkan spesifikasi. Stimulus harus memuat informasi cukup, opsi atau prompt harus sejajar dengan indikator, skala harus memiliki jangkar, dan rubrik harus menjelaskan perbedaan kategori yang berdekatan.

Struktur yang baik memungkinkan peneliti menelusuri kembali setiap kategori hasil ke respons mentah. Transparansi ini penting agar guru tidak hanya menerima label otomatis tanpa memahami alasan di baliknya.

- tier pertama berupa jawaban konten;
- tier kedua berupa alasan;
- kunci kombinasi jawaban-alasan;
- distraktor berbasis miskonsepsi;
- peta konsep dan proposisi;

Dalam praktik kelas, Two-Tier Diagnostic Test dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

No.	Komponen	Fungsi Diagnostik
1	tier pertama berupa jawaban konten	Bukti yang harus ditelusuri
2	tier kedua berupa alasan	Bukti yang harus ditelusuri
3	kunci kombinasi jawaban-alasan	Bukti yang harus ditelusuri
4	distraktor berbasis miskonsepsi	Bukti yang harus ditelusuri

6.3.5 Desain Stimulus, Tugas, atau Pertanyaan

Desain stimulus pada Two-Tier Diagnostic Test harus memunculkan penggunaan konsep, bukan sekadar pengenalan istilah. Situasi dapat berbentuk fenomena, gambar partikel, grafik, data eksperimen, persamaan, atau masalah kontekstual, tetapi seluruh informasi yang tidak relevan perlu dikendalikan. (Haladyna et al., 2002; Treagust, 1988; Odom & Barrow, 1995; Chandrasegaran et al., 2007)

Dalam kimia, satu stimulus sering menggabungkan beberapa tingkat representasi. Pengembang harus memutuskan apakah tujuan butir adalah menerjemahkan representasi, menjelaskan hubungan sebab-akibat, membedakan contoh dan noncontoh, atau menerapkan prinsip pada konteks baru.

Bahasa perlu ringkas tanpa menghilangkan syarat ilmiah. Kata seperti kuat, stabil, jenuh, spontan, atau setimbang mempunyai arti teknis dan sehari-hari yang berbeda. Ketidakjelasan istilah dapat menghasilkan kesalahan yang berasal dari bahasa, bukan struktur konsep.

Desain yang baik juga menghindari petunjuk format, pilihan terlalu panjang, negasi bertumpuk, gambar dekoratif, dan data yang tidak diperlukan. Setiap elemen harus mempunyai alasan diagnostik.

- gunakan konteks yang cukup dikenal tetapi tidak menghafalkan contoh;
- pastikan hanya satu interpretasi ilmiah yang menjadi target;
- uji semua simbol, satuan, grafik, dan gambar;
- sediakan ruang bagi respons yang tidak diprediksi bila tujuan masih eksploratif.

Dalam praktik kelas, Two-Tier Diagnostic Test dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.3.6 Prosedur Pengembangan Secara Sistematis

Pengembangan Two-Tier Diagnostic Test harus mengikuti alur yang terdokumentasi. Langkah yang lazim ialah mendefinisikan konten, mengidentifikasi konsepsi alternatif, menyusun pertanyaan terbuka, melakukan wawancara, kemudian membangun dua tier, validasi ahli, uji coba, analisis kombinasi respons. Urutan dapat berulang karena temuan uji coba sering memaksa revisi konstruk atau redaksi. (Treagust, 1988; Odom & Barrow, 1995; Chandrasegaran et al., 2007; Adams & Wieman, 2011; AERA, APA, & NCME, 2014)

Tahap konseptual dimulai dengan analisis kurikulum, literatur miskonsepsi, dan peta konsep. Tahap empiris awal menggunakan wawancara, respons terbuka, atau observasi untuk memastikan bahwa kategori yang dirancang benar-benar muncul pada populasi sasaran.

Setiap versi instrumen perlu diberi nomor, tanggal, dan catatan perubahan. Revisi harus menyebutkan masalah yang ditemukan, bukti yang digunakan, perubahan yang dilakukan, dan alasan mempertahankan atau menghapus butir.

Pengembangan dianggap belum selesai ketika hanya menghasilkan butir yang tampak baik. Instrumen perlu diuji pada peserta didik yang serupa dengan pengguna sasaran, lalu dianalisis secara konseptual dan empiris sebelum dipakai untuk keputusan penting.

- 1. mendefinisikan konten.
- 2. mengidentifikasi konsepsi alternatif.
- 3. menyusun pertanyaan terbuka.
- 4. melakukan wawancara.
- 5. membangun dua tier.

Dalam praktik kelas, Two-Tier Diagnostic Test dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.3.7 Pelaksanaan dan Pengumpulan Data

Pelaksanaan Two-Tier Diagnostic Test harus distandardisasi secukupnya agar perbedaan hasil lebih mungkin berasal dari pemahaman, bukan variasi instruksi, waktu, alat, atau bantuan. Petunjuk perlu menjelaskan tujuan umum tanpa mengajarkan jawaban. (Treagust, 1988; Odom & Barrow, 1995; AERA, APA, & NCME, 2014)

Kondisi administrasi mencakup waktu, urutan tugas, kesempatan bertanya, penggunaan kalkulator, akses tabel periodik, kualitas gambar, dan perangkat digital. Pada instrumen kualitatif, standardisasi berarti konsistensi fokus dan prinsip probing, bukan membaca naskah secara kaku.

Data yang dikumpulkan ialah kombinasi benar-salah pada tier jawaban dan tier alasan. Data mentah harus disimpan bersama metadata yang diperlukan, seperti kode butir, versi instrumen, waktu administrasi, dan kondisi khusus, tetapi hanya data yang relevan yang boleh dikumpulkan.

Sebelum administrasi luas, lakukan uji teknis. Pastikan tautan, urutan, penguncian respons, rekaman, penyimpanan, serta tampilan rumus dan subskrip bekerja. Gangguan teknis perlu dicatat sebagai bagian dari interpretasi.

- berikan instruksi yang sama dan mudah dipahami;
- catat kondisi yang dapat memengaruhi respons;
- hindari umpan balik sebelum seluruh respons diagnostik selesai;
- sediakan prosedur penanganan data tidak lengkap.

Dalam praktik kelas, Two-Tier Diagnostic Test dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.3.8 Penskoran, Pengodean, dan Pengolahan Data

Penskoran Two-Tier Diagnostic Test dapat diringkas sebagai kombinasi BB, BS, SB, dan SS yang dapat dipetakan menjadi pemahaman utuh, pemahaman parsial, false positive, false negative, atau ketidakpahaman. Aturan ini harus ditulis sebelum melihat hasil utama agar keputusan tidak disesuaikan secara oportunistik dengan pola data. (Treagust, 1988; Odom & Barrow, 1995; Chandrasegaran et al., 2007; Chi, 1997; AERA, APA, & NCME, 2014)

Skor total jarang cukup untuk diagnosis. Analisis perlu mempertahankan informasi tingkat butir, konsep, opsi, alasan, keyakinan, proses, atau representasi. Penggabungan data terlalu dini dapat menyembunyikan pola yang secara pedagogis penting.

Untuk respons terbuka atau kualitatif, buku kode harus mencantumkan definisi, contoh positif, contoh batas, noncontoh, dan aturan ketika satu respons memuat beberapa gagasan. Sebagian data perlu dinilai oleh lebih dari satu penilai.

Setiap transformasi data—misalnya dari respons ke kode, dari kode ke kategori, dan dari kategori ke persentase—harus dapat diaudit. Simpan respons asli dan keputusan pengodean agar kasus ambigu dapat ditinjau kembali.

- pisahkan skor ketepatan dari skor keyakinan atau kualitas alasan;
- laporkan kasus tidak lengkap dan aturan penanganannya;
- uji konsistensi penilai untuk data berkode;
- hindari mengubah data ordinal menjadi ukuran interval tanpa alasan.

Dalam praktik kelas, Two-Tier Diagnostic Test dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.3.9 Interpretasi dan Penetapan Kategori

Interpretasi hasil Two-Tier Diagnostic Test harus bergerak dari respons menuju pola, bukan langsung dari satu jawaban menuju label. Status konseptual menjadi lebih kuat apabila gagasan yang sama muncul pada beberapa butir, representasi, atau konteks. (Treagust, 1988; Odom & Barrow, 1995; Chandrasegaran et al., 2007; Çil, 2015; Kane, 2013)

Kategori dapat mencakup pemahaman ilmiah, pemahaman parsial, kurang pengetahuan, respons tidak konsisten, false positive, false negative, dan miskonsepsi. Tidak semua instrumen mampu membedakan seluruh kategori tersebut; skema harus mengikuti informasi yang benar-benar tersedia.

Ambang, seperti batas keyakinan atau persentase pemilihan, perlu memiliki rasional dan diuji melalui analisis sensitivitas. Perubahan kecil pada ambang dapat mengubah prevalensi kategori secara berarti.

Interpretasi yang bertanggung jawab menyatakan bahwa kategori berlaku pada konsep, bentuk tugas, waktu, dan kondisi tertentu. Hasil merupakan hipotesis kerja untuk pembelajaran, bukan diagnosis psikologis permanen.

- gunakan lebih dari satu butir untuk konsepsi penting;
- periksa konsistensi antara jawaban dan alasan;
- tandai kasus ambigu untuk konfirmasi;
- hubungkan kategori dengan rekomendasi yang berbeda.

Dalam praktik kelas, Two-Tier Diagnostic Test dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.3.10 Validitas dan Kualitas Inferensi

Validitas Two-Tier Diagnostic Test berkaitan dengan kualitas interpretasi dan penggunaan hasil, bukan sifat tetap yang dimiliki instrumen. Fokus khususnya adalah keterwakilan konsepsi alternatif, independensi relatif antartier, dan ketepatan interpretasi kombinasi respons. (AERA, APA, & NCME, 2014; Messick, 1995; Kane, 2013; Haynes et al., 1995)

Bukti berdasarkan isi menilai keterwakilan domain dan akurasi konsep. Bukti proses respons memeriksa bagaimana peserta didik memahami tugas. Bukti struktur internal menilai kesesuaian pola respons. Bukti hubungan dengan variabel lain dan konsekuensi penggunaan melengkapi argumen.

Ancaman validitas dapat berasal dari konten yang kurang terwakili, tuntutan bahasa berlebihan, petunjuk opsi, rubrik yang tidak stabil, keyakinan yang tidak terkalibrasi, atau algoritma yang belajar dari label keliru.

Dokumen validasi perlu menghubungkan setiap klaim dengan bukti dan alternatif penjelasan. Menampilkan satu indeks tidak cukup; validasi adalah program argumentatif yang diperbarui setiap kali instrumen dipakai pada populasi atau tujuan baru.

- bukti isi dan kesesuaian kurikulum;
- bukti proses respons melalui wawancara kognitif;
- bukti struktur dan fungsi butir;
- bukti kegunaan serta konsekuensi keputusan.

Dalam praktik kelas, Two-Tier Diagnostic Test dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.3.11 Reliabilitas, Keajegan, dan Analisis Empiris

Keajegan Two-Tier Diagnostic Test harus dinilai sesuai jenis data. Pada respons objektif dapat digunakan konsistensi internal, stabilitas waktu, atau model pengukuran. Pada data kualitatif, perhatian diarahkan pada konsistensi pengodean, audit trail, dan ketepatan penerapan definisi kode. (Kuder & Richardson, 1937; Cronbach, 1951; Ding et al., 2006; Wren & Barbera, 2014; AERA, APA, & NCME, 2014)

Reliabilitas tinggi tidak otomatis menunjukkan diagnosis yang benar. Butir yang berulang-ulang mengukur hafalan dapat menghasilkan koefisien tinggi tetapi tidak mewakili pemahaman konseptual. Sebaliknya, konstruk yang multidimensi tidak selalu layak dipaksa menjadi satu skala.

Analisis empiris perlu memeriksa kesukaran, daya beda, fungsi opsi, pola data hilang, dan keterhubungan komponen. Untuk instrumen bertingkat, reliabilitas dapat dihitung pada setiap tier serta kombinasi yang mempunyai makna teoritis.

Ukuran sampel, distribusi kemampuan, dan homogenitas konten memengaruhi statistik. Karena itu, laporan harus menyatakan kondisi estimasi dan tidak menyalin batas angka sebagai aturan universal.

- pilih statistik yang sesuai dengan jenis skor;
- laporkan interval atau ketidakpastian bila tersedia;
- periksa butir secara substantif dan statistik;
- uji kembali setelah revisi atau perubahan platform.

Dalam praktik kelas, Two-Tier Diagnostic Test dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.3.12 Contoh Penerapan dalam Pembelajaran Kimia

Penerapan Two-Tier Diagnostic Test dalam kimia dapat dimulai dari konsep reaksi kimia pada tingkat makroskopik dan submikroskopik. Pengembang perlu menuliskan konsep ilmiah target, miskonsepsi yang mungkin, bukti respons yang diharapkan, dan keputusan pembelajaran setelah hasil diperoleh. (Treagust, 1988; Odom & Barrow, 1995; Chandrasegaran et al., 2007; Çil, 2015; Nakhleh, 1992)

Contoh kedua adalah kesetimbangan dinamis. Stimulus sebaiknya menghubungkan fenomena dengan model partikel atau persamaan, lalu meminta peserta didik menunjukkan hubungan sebab-akibat. Jawaban yang hanya menyebut istilah belum cukup menjadi bukti pemahaman.

Pada ikatan kovalen dan struktur, diagnosis dapat memeriksa apakah peserta didik membedakan perubahan laju, keadaan setimbang, energi, dan komposisi. Kesalahan sering muncul ketika satu aturan dipindahkan ke konteks yang tidak sesuai.

Kasus difusi ion dalam larutan menunjukkan pentingnya representasi. Respons verbal, gambar, simbol, dan hitungan dapat tidak konsisten. Hasil instrumen sebaiknya digunakan untuk memilih demonstrasi, diskusi, contoh-noncontoh, atau tugas remediasi yang secara langsung menantang hubungan salah.

- tentukan proposisi ilmiah dan konsepsi alternatif;
- gunakan minimal dua bentuk representasi;
- minta alasan atau bukti ketika format memungkinkan;
- rancang tindak lanjut berdasarkan pola, bukan hanya skor.

Dalam praktik kelas, Two-Tier Diagnostic Test dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.3.13 Kelebihan, Keterbatasan, dan Potensi Bias

Kekuatan utama Two-Tier Diagnostic Test meliputi alasan meningkatkan spesifisitas diagnosis, dapat diterapkan pada kelompok besar, cocok untuk representasi kimia berlapis, dan distraktor menyimpan pengetahuan dari studi kualitatif. Kekuatan tersebut menjelaskan situasi ketika instrumen layak dipilih dibanding teknik lain. (Treagust, 1988; Odom & Barrow, 1995; Chandrasegaran et al., 2007; Gurel et al., 2015; AERA, APA, & NCME, 2014)

Keterbatasannya mencakup tidak membedakan respons yakin dan tidak yakin, kombinasi salah-salah dapat berasal dari ketidaktahuan, tier alasan dapat memberi petunjuk pada tier pertama, dan skema kategori harus dijelaskan secara transparan. Keterbatasan harus diperlakukan sebagai batas inferensi, bukan sekadar paragraf formal yang tidak memengaruhi analisis.

Potensi bias dapat berasal dari bahasa, pengalaman konteks, akses teknologi, kecemasan, gaya menggunakan skala, keterampilan menggambar, atau kemampuan menulis. Instrumen perlu diuji pada kelompok yang beragam dan ditinjau untuk fungsi berbeda antarkelompok bila data mencukupi.

Keputusan memilih instrumen sebaiknya mempertimbangkan kedalaman data, ukuran kelompok, waktu, keterampilan analisis, tujuan diagnosis, dan konsekuensi kesalahan klasifikasi. Tidak ada satu metode yang unggul untuk semua tujuan.

- Kekuatan: alasan meningkatkan spesifisitas diagnosis.
- Kekuatan: dapat diterapkan pada kelompok besar.
- Batas: tidak membedakan respons yakin dan tidak yakin.
- Batas: kombinasi salah-salah dapat berasal dari ketidaktahuan.

Dalam praktik kelas, Two-Tier Diagnostic Test dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

Aspek	Temuan	Respons Metodologis
Kekuatan	alasan meningkatkan spesifisitas diagnosis	Pilih ketika sesuai tujuan
Keterbatasan	tidak membedakan respons yakin dan tidak yakin	Tambahkan bukti atau kontrol
Bias potensial	Bahasa, konteks, perangkat, gaya respons	Uji proses respons

6.3.14 Integrasi Digital, Etika, dan Tata Kelola Data

Integrasi digital Two-Tier Diagnostic Test dapat berupa penyajian berurutan yang menyimpan tier pertama sebelum alasan ditampilkan serta analitik kombinasi respons otomatis. Teknologi seharusnya memperluas kualitas bukti, bukan hanya memindahkan lembar kertas ke layar. (Baker & Inventado, 2014; Romero & Ventura, 2020; AERA, APA, & NCME, 2014; Treagust, 1988)

Platform perlu menyimpan versi instrumen, urutan respons, timestamp yang relevan, aturan skor, dan log perubahan. Namun, hanya data yang dibutuhkan untuk tujuan diagnosis yang boleh dikumpulkan; pengumpulan berlebihan meningkatkan risiko privasi tanpa manfaat ilmiah.

Aspek etika utama adalah bahwa kategori false positive dan false negative harus diperlakukan sebagai sinyal pemeriksaan lanjut, bukan label kemampuan tetap. Peserta didik dan guru perlu mengetahui data apa yang dikumpulkan, untuk tujuan apa, siapa yang dapat mengakses, berapa lama disimpan, dan bagaimana kesalahan data diperbaiki.

Algoritma otomatis harus dapat diaudit. Kategori yang dihasilkan sistem perlu ditautkan ke respons mentah dan aturan keputusan. Kasus berisiko atau ambigu harus ditinjau manusia sebelum tindak lanjut penting dilakukan.

- gunakan identitas pseudonim dan akses berbasis peran;
- uji tampilan rumus, subskrip, grafik, dan gambar;
- sediakan opsi aksesibilitas dan alternatif perangkat;
- dokumentasikan algoritma serta perubahan versi.

Dalam praktik kelas, Two-Tier Diagnostic Test dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.3.15 Sintesis dan Pedoman Implementasi

Two-Tier Diagnostic Test bermanfaat apabila dipilih berdasarkan tujuan dan digunakan sesuai batasnya. Instrumen ini menghasilkan kombinasi benar-salah pada tier jawaban dan tier alasan, lalu data diolah melalui kombinasi BB, BS, SB, dan SS yang dapat dipetakan menjadi pemahaman utuh, pemahaman parsial, false positive, false negative, atau ketidakpahaman. (Treagust, 1988; Odom & Barrow, 1995; Chandrasegaran et al., 2007; AERA, APA, & NCME, 2014; Kane, 2013)

Rangkaian kerja yang disarankan dimulai dari analisis konsep, identifikasi konsepsi alternatif, desain tugas, telaah ahli, uji proses respons, uji lapangan, analisis, penetapan kategori, dan penyusunan tindak lanjut. Setiap tahap harus meninggalkan dokumentasi yang dapat diperiksa.

Hasil diagnosis tidak berhenti pada tabel prevalensi. Guru perlu memilih konsepsi prioritas berdasarkan frekuensi, keyakinan, posisi sebagai konsep prasyarat, dampak terhadap topik lanjutan, dan kemungkinan diperbaiki melalui pembelajaran.

Prinsip akhirnya ialah kehati-hatian inferensial: sebutkan apa yang benar-benar ditunjukkan data, bedakan bukti kuat dari indikasi, gunakan triangulasi ketika diperlukan, dan hindari label permanen pada peserta didik.

- selaraskan tujuan, konstruk, tugas, skor, kategori, dan tindakan;
- utamakan konsepsi yang berdampak pada pembelajaran lanjutan;
- gabungkan bukti bila keputusan memiliki konsekuensi besar;
- evaluasi kembali efektivitas tindak lanjut melalui diagnosis ulang.

Dalam praktik kelas, Two-Tier Diagnostic Test dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.4 Three-Tier Diagnostic Test

6.4.1 Landasan Konseptual dan Kedudukan Instrumen

Dalam diagnosis miskonsepsi, Three-Tier Diagnostic Test perlu dipahami bukan sekadar sebagai bentuk soal atau teknik pengumpulan data, melainkan sebagai perangkat inferensi. Dalam bab ini, instrumen tersebut didefinisikan sebagai tes dua tingkat yang ditambah pengukuran keyakinan terhadap respons, biasanya berupa pilihan yakin-tidak yakin atau skala keyakinan. Definisi operasional ini menegaskan bentuk bukti yang dikumpulkan sekaligus membatasi jenis kesimpulan yang layak dibuat. (Caleon & Subramaniam, 2010a; Peşman & Eryılmaz, 2010; Gurcay & Gulbas, 2015)

Klaim utamanya adalah bahwa instrumen ini membedakan kesalahan yang diyakini dari kesalahan yang disertai keraguan sehingga inferensi miskonsepsi lebih kuat daripada sekadar kombinasi jawaban dan alasan. Klaim tersebut harus dipisahkan dari anggapan bahwa setiap jawaban salah merupakan miskonsepsi. Kesalahan dapat berasal dari kurang pengetahuan, ketidakcermatan, kesulitan bahasa, salah hitung, atau ketidakfamiliaran dengan representasi.

Data pokok yang dihasilkan berupa ketepatan jawaban, ketepatan alasan, dan keyakinan yang dinyatakan peserta didik. Data itu baru memperoleh makna diagnostik setelah dikaitkan dengan peta konsep sasaran, kemungkinan konsepsi alternatif, dan aturan interpretasi yang telah ditetapkan sebelum analisis.

Dengan demikian, mutu diagnosis ditentukan oleh keselarasan antara tujuan, desain, pelaksanaan, analisis, dan keputusan yang diambil. Posisi terbaik Three-Tier Diagnostic Test adalah sebagai bagian dari sistem asesmen berlapis yang dapat dipadukan dengan bukti lain ketika konsekuensi keputusan semakin besar.

- menentukan konsep dan proposisi yang hendak didiagnosis;
- menjelaskan perilaku respons yang dianggap sebagai bukti;
- membatasi klaim sesuai kekuatan instrumen;
- menautkan hasil dengan tindakan pembelajaran yang realistis.

Dalam praktik kelas, Three-Tier Diagnostic Test dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.4.2 Perkembangan Historis dan Rasional Penggunaan

Perkembangan Three-Tier Diagnostic Test tidak dapat dipisahkan dari perubahan pandangan tentang belajar sains. Ketika peserta didik tidak lagi dipandang sebagai penerima informasi yang kosong, instrumen mulai dirancang untuk mengungkap gagasan awal, pola penalaran, dan keyakinan yang dibawa ke kelas. (Caleon & Subramaniam, 2010a; Peşman & Eryılmaz, 2010; Gurcay & Gulbas, 2015; Taslidere, 2016)

Pada tahap awal, penilaian banyak mengandalkan benar-salah atau skor total. Perkembangan penelitian konsepsi peserta didik kemudian mendorong penambahan alasan, keyakinan, tugas kinerja, rekaman proses, dan bukti sosial. Three-Tier Diagnostic Test merupakan salah satu jawaban metodologis terhadap kebutuhan tersebut.

Dalam pendidikan kimia, kebutuhan diagnosis semakin kuat karena peserta didik harus mengoordinasikan fenomena makroskopik, model submikroskopik, simbol, grafik, dan perhitungan. Instrumen yang hanya menguji hafalan definisi sering gagal memperlihatkan cara representasi itu dihubungkan.

Sejarah penggunaannya juga menunjukkan bahwa format yang lebih kompleks tidak otomatis lebih baik. Setiap tambahan tier, rubrik, atau kanal data harus mempunyai fungsi yang jelas dan didukung bukti bahwa informasi tambahan benar-benar memperbaiki keputusan.

- pergeseran dari pengukuran skor menuju diagnosis pola pemahaman;
- penggunaan respons terbuka dan wawancara untuk menemukan konsepsi alternatif;
- pengembangan teknik kuantitatif agar diagnosis dapat dilakukan pada kelompok besar;
- integrasi data digital dan metakognitif pada perkembangan mutakhir.

Dalam praktik kelas, Three-Tier Diagnostic Test dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.4.3 Konstruk, Klaim, dan Bukti Diagnosis

Konstruk yang dinilai melalui Three-Tier Diagnostic Test bukan objek yang terlihat langsung. Yang diamati adalah ketepatan jawaban, ketepatan alasan, dan keyakinan yang dinyatakan peserta didik; dari sana dibuat inferensi tentang pemahaman konsep, miskonsepsi, pengetahuan parsial, atau ketidakpastian. (Mathabathe & Potgieter, 2014; Gurel et al., 2015; Messick, 1995; Kane, 2013)

Klaim diagnosis perlu dirumuskan sebagai argumen. Pertama, tugas harus memerlukan konsep sasaran. Kedua, respons harus diberi skor atau kode secara konsisten. Ketiga, pola skor harus mempunyai hubungan yang masuk akal dengan struktur pengetahuan. Keempat, penggunaan hasil harus menghasilkan keputusan yang proporsional.

Sumber ancaman utama adalah ketidakjelasan konstruk. Misalnya, peserta didik dapat gagal karena istilah yang ambigu, beban membaca, representasi yang tidak dikenal, atau prosedur matematika. Kegagalan semacam itu tidak boleh langsung diperlakukan sebagai miskonsepsi kimia.

Oleh sebab itu, setiap pengembang perlu menulis pernyataan klaim-bukti: instrumen mengukur apa, perilaku apa yang menjadi indikator, alternatif penjelasan apa yang mungkin, dan bukti tambahan apa yang diperlukan untuk memperkuat kesimpulan.

- konstruk utama: pemahaman konseptual melalui tes diagnostik tiga tingkat;
- bukti observabel: respons, alasan, tindakan, atau representasi;
- inferensi: profil konsepsi pada konten dan konteks tertentu;
- penggunaan: perencanaan pembelajaran, remediasi, atau penelitian.

Dalam praktik kelas, Three-Tier Diagnostic Test dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.4.4 Struktur, Komponen, dan Alur Respons

Struktur Three-Tier Diagnostic Test dibangun dari komponen yang saling bergantung. Komponen inti meliputi jawaban konten, alasan, tingkat keyakinan, aturan ambang keyakinan, dan matriks kategori tiga tier. Hilangnya satu komponen dapat mengubah makna data yang diperoleh. (Caleon & Subramaniam, 2010a; Peşman & Eryılmaz, 2010; Gurcay & Gulbas, 2015; Taslidere, 2016)

Urutan penyajian harus ditetapkan secara sadar. Komponen yang muncul lebih awal dapat memberi petunjuk pada komponen berikutnya, sedangkan permintaan alasan atau keyakinan dapat mengubah cara peserta didik memeriksa jawabannya. Penguncian respons dan uji urutan diperlukan pada format digital.

Setiap komponen membutuhkan spesifikasi. Stimulus harus memuat informasi cukup, opsi atau prompt harus sejajar dengan indikator, skala harus memiliki jangkar, dan rubrik harus menjelaskan perbedaan kategori yang berdekatan.

Struktur yang baik memungkinkan peneliti menelusuri kembali setiap kategori hasil ke respons mentah. Transparansi ini penting agar guru tidak hanya menerima label otomatis tanpa memahami alasan di baliknya.

- jawaban konten;
- alasan;
- tingkat keyakinan;
- aturan ambang keyakinan;
- matriks kategori tiga tier;

Dalam praktik kelas, Three-Tier Diagnostic Test dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

No.	Komponen	Fungsi Diagnostik
1	jawaban konten	Bukti yang harus ditelusuri
2	alasan	Bukti yang harus ditelusuri
3	tingkat keyakinan	Bukti yang harus ditelusuri
4	aturan ambang keyakinan	Bukti yang harus ditelusuri

6.4.5 Desain Stimulus, Tugas, atau Pertanyaan

Desain stimulus pada Three-Tier Diagnostic Test harus memunculkan penggunaan konsep, bukan sekadar pengenalan istilah. Situasi dapat berbentuk fenomena, gambar partikel, grafik, data eksperimen, persamaan, atau masalah kontekstual, tetapi seluruh informasi yang tidak relevan perlu dikendalikan. (Haladyna et al., 2002; Caleon & Subramaniam, 2010a; Peşman & Eryilmaz, 2010; Gurcay & Gulbas, 2015)

Dalam kimia, satu stimulus sering menggabungkan beberapa tingkat representasi. Pengembang harus memutuskan apakah tujuan butir adalah menerjemahkan representasi, menjelaskan hubungan sebab-akibat, membedakan contoh dan noncontoh, atau menerapkan prinsip pada konteks baru.

Bahasa perlu ringkas tanpa menghilangkan syarat ilmiah. Kata seperti kuat, stabil, jenuh, spontan, atau setimbang mempunyai arti teknis dan sehari-hari yang berbeda. Ketidakjelasan istilah dapat menghasilkan kesalahan yang berasal dari bahasa, bukan struktur konsep.

Desain yang baik juga menghindari petunjuk format, pilihan terlalu panjang, negasi bertumpuk, gambar dekoratif, dan data yang tidak diperlukan. Setiap elemen harus mempunyai alasan diagnostik.

- gunakan konteks yang cukup dikenal tetapi tidak menghafalkan contoh;
- pastikan hanya satu interpretasi ilmiah yang menjadi target;
- uji semua simbol, satuan, grafik, dan gambar;
- sediakan ruang bagi respons yang tidak diprediksi bila tujuan masih eksploratif.

Dalam praktik kelas, Three-Tier Diagnostic Test dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.4.6 Prosedur Pengembangan Secara Sistematis

Pengembangan Three-Tier Diagnostic Test harus mengikuti alur yang terdokumentasi. Langkah yang lazim ialah mengembangkan dua tier, memilih bentuk skala keyakinan, mendefinisikan ambang, menguji keterbacaan skala, kemudian memeriksa hubungan keyakinan-kinerja, menyusun kategori, uji coba luas, validasi silang. Urutan dapat berulang karena temuan uji coba sering memaksa revisi konstruk atau redaksi. (Caleon & Subramaniam, 2010a; Peşman & Eryilmaz, 2010; Gurcay & Gulbas, 2015; Adams & Wieman, 2011; AERA, APA, & NCME, 2014)

Tahap konseptual dimulai dengan analisis kurikulum, literatur miskonsepsi, dan peta konsep. Tahap empiris awal menggunakan wawancara, respons terbuka, atau observasi untuk memastikan bahwa kategori yang dirancang benar-benar muncul pada populasi sasaran.

Setiap versi instrumen perlu diberi nomor, tanggal, dan catatan perubahan. Revisi harus menyebutkan masalah yang ditemukan, bukti yang digunakan, perubahan yang dilakukan, dan alasan mempertahankan atau menghapus butir.

Pengembangan dianggap belum selesai ketika hanya menghasilkan butir yang tampak baik. Instrumen perlu diuji pada peserta didik yang serupa dengan pengguna sasaran, lalu dianalisis secara konseptual dan empiris sebelum dipakai untuk keputusan penting.

- 1. mengembangkan dua tier.
- 2. memilih bentuk skala keyakinan.
- 3. mendefinisikan ambang.
- 4. menguji keterbacaan skala.
- 5. memeriksa hubungan keyakinan-kinerja.

Dalam praktik kelas, Three-Tier Diagnostic Test dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.4.7 Pelaksanaan dan Pengumpulan Data

Pelaksanaan Three-Tier Diagnostic Test harus distandardisasi secukupnya agar perbedaan hasil lebih mungkin berasal dari pemahaman, bukan variasi instruksi, waktu, alat, atau bantuan. Petunjuk perlu menjelaskan tujuan umum tanpa mengajarkan jawaban. (Caleon & Subramaniam, 2010a; Peşman & Eryılmaz, 2010; AERA, APA, & NCME, 2014)

Kondisi administrasi mencakup waktu, urutan tugas, kesempatan bertanya, penggunaan kalkulator, akses tabel periodik, kualitas gambar, dan perangkat digital. Pada instrumen kualitatif, standardisasi berarti konsistensi fokus dan prinsip probing, bukan membaca naskah secara kaku.

Data yang dikumpulkan ialah ketepatan jawaban, ketepatan alasan, dan keyakinan yang dinyatakan peserta didik. Data mentah harus disimpan bersama metadata yang diperlukan, seperti kode butir, versi instrumen, waktu administrasi, dan kondisi khusus, tetapi hanya data yang relevan yang boleh dikumpulkan.

Sebelum administrasi luas, lakukan uji teknis. Pastikan tautan, urutan, penguncian respons, rekaman, penyimpanan, serta tampilan rumus dan subskrip bekerja. Gangguan teknis perlu dicatat sebagai bagian dari interpretasi.

- berikan instruksi yang sama dan mudah dipahami;
- catat kondisi yang dapat memengaruhi respons;
- hindari umpan balik sebelum seluruh respons diagnostik selesai;
- sediakan prosedur penanganan data tidak lengkap.

Dalam praktik kelas, Three-Tier Diagnostic Test dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.4.8 Penskoran, Pengodean, dan Pengolahan Data

Penskoran Three-Tier Diagnostic Test dapat diringkas sebagai kombinasi tiga dimensi yang memisahkan pemahaman ilmiah, pengetahuan parsial, tidak tahu, menebak, dan miskonsepsi dengan keyakinan tinggi. Aturan ini harus ditulis sebelum melihat hasil utama agar keputusan tidak disesuaikan secara oportunistik dengan pola data. (Caleon & Subramaniam, 2010a; Peşman & Eryilmaz, 2010; Gurcay & Gulbas, 2015; Chi, 1997; AERA, APA, & NCME, 2014)

Skor total jarang cukup untuk diagnosis. Analisis perlu mempertahankan informasi tingkat butir, konsep, opsi, alasan, keyakinan, proses, atau representasi. Penggabungan data terlalu dini dapat menyembunyikan pola yang secara pedagogis penting.

Untuk respons terbuka atau kualitatif, buku kode harus mencantumkan definisi, contoh positif, contoh batas, noncontoh, dan aturan ketika satu respons memuat beberapa gagasan. Sebagian data perlu dinilai oleh lebih dari satu penilai.

Setiap transformasi data—misalnya dari respons ke kode, dari kode ke kategori, dan dari kategori ke persentase—harus dapat diaudit. Simpan respons asli dan keputusan pengodean agar kasus ambigu dapat ditinjau kembali.

- pisahkan skor ketepatan dari skor keyakinan atau kualitas alasan;
- laporkan kasus tidak lengkap dan aturan penanganannya;
- uji konsistensi penilai untuk data berkode;
- hindari mengubah data ordinal menjadi ukuran interval tanpa alasan.

Dalam praktik kelas, Three-Tier Diagnostic Test dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.4.9 Interpretasi dan Penetapan Kategori

Interpretasi hasil Three-Tier Diagnostic Test harus bergerak dari respons menuju pola, bukan langsung dari satu jawaban menuju label. Status konseptual menjadi lebih kuat apabila gagasan yang sama muncul pada beberapa butir, representasi, atau konteks. (Caleon & Subramaniam, 2010a; Peşman & Eryılmaz, 2010; Gurcay & Gulbas, 2015; Taslidere, 2016; Kane, 2013)

Kategori dapat mencakup pemahaman ilmiah, pemahaman parsial, kurang pengetahuan, respons tidak konsisten, false positive, false negative, dan miskonsepsi. Tidak semua instrumen mampu membedakan seluruh kategori tersebut; skema harus mengikuti informasi yang benar-benar tersedia.

Ambang, seperti batas keyakinan atau persentase pemilihan, perlu memiliki rasional dan diuji melalui analisis sensitivitas. Perubahan kecil pada ambang dapat mengubah prevalensi kategori secara berarti.

Interpretasi yang bertanggung jawab menyatakan bahwa kategori berlaku pada konsep, bentuk tugas, waktu, dan kondisi tertentu. Hasil merupakan hipotesis kerja untuk pembelajaran, bukan diagnosis psikologis permanen.

- gunakan lebih dari satu butir untuk konsepsi penting;
- periksa konsistensi antara jawaban dan alasan;
- tandai kasus ambigu untuk konfirmasi;
- hubungkan kategori dengan rekomendasi yang berbeda.

Dalam praktik kelas, Three-Tier Diagnostic Test dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.4.10 Validitas dan Kualitas Inferensi

Validitas Three-Tier Diagnostic Test berkaitan dengan kualitas interpretasi dan penggunaan hasil, bukan sifat tetap yang dimiliki instrumen. Fokus khususnya adalah bukti bahwa keyakinan berfungsi sebagai informasi metakognitif dan tidak sekadar merefleksikan keberanian atau kebiasaan menggunakan skala. (AERA, APA, & NCME, 2014; Messick, 1995; Kane, 2013; Haynes et al., 1995)

Bukti berdasarkan isi menilai keterwakilan domain dan akurasi konsep. Bukti proses respons memeriksa bagaimana peserta didik memahami tugas. Bukti struktur internal menilai kesesuaian pola respons. Bukti hubungan dengan variabel lain dan konsekuensi penggunaan melengkapi argumen.

Ancaman validitas dapat berasal dari konten yang kurang terwakili, tuntutan bahasa berlebihan, petunjuk opsi, rubrik yang tidak stabil, keyakinan yang tidak terkalibrasi, atau algoritma yang belajar dari label keliru.

Dokumen validasi perlu menghubungkan setiap klaim dengan bukti dan alternatif penjelasan. Menampilkan satu indeks tidak cukup; validasi adalah program argumentatif yang diperbarui setiap kali instrumen dipakai pada populasi atau tujuan baru.

- bukti isi dan kesesuaian kurikulum;
- bukti proses respons melalui wawancara kognitif;
- bukti struktur dan fungsi butir;
- bukti kegunaan serta konsekuensi keputusan.

Dalam praktik kelas, Three-Tier Diagnostic Test dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.4.11 Reliabilitas, Keajegan, dan Analisis Empiris

Keajegan Three-Tier Diagnostic Test harus dinilai sesuai jenis data. Pada respons objektif dapat digunakan konsistensi internal, stabilitas waktu, atau model pengukuran. Pada data kualitatif, perhatian diarahkan pada konsistensi pengodean, audit trail, dan ketepatan penerapan definisi kode. (Kuder & Richardson, 1937; Cronbach, 1951; Ding et al., 2006; Wren & Barbera, 2014; AERA, APA, & NCME, 2014)

Reliabilitas tinggi tidak otomatis menunjukkan diagnosis yang benar. Butir yang berulang-ulang mengukur hafalan dapat menghasilkan koefisien tinggi tetapi tidak mewakili pemahaman konseptual. Sebaliknya, konstruk yang multidimensi tidak selalu layak dipaksa menjadi satu skala.

Analisis empiris perlu memeriksa kesukaran, daya beda, fungsi opsi, pola data hilang, dan keterhubungan komponen. Untuk instrumen bertingkat, reliabilitas dapat dihitung pada setiap tier serta kombinasi yang mempunyai makna teoritis.

Ukuran sampel, distribusi kemampuan, dan homogenitas konten memengaruhi statistik. Karena itu, laporan harus menyatakan kondisi estimasi dan tidak menyalin batas angka sebagai aturan universal.

- pilih statistik yang sesuai dengan jenis skor;
- laporkan interval atau ketidakpastian bila tersedia;
- periksa butir secara substantif dan statistik;
- uji kembali setelah revisi atau perubahan platform.

Dalam praktik kelas, Three-Tier Diagnostic Test dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.4.12 Contoh Penerapan dalam Pembelajaran Kimia

Penerapan Three-Tier Diagnostic Test dalam kimia dapat dimulai dari konsep kalor dan suhu. Pengembang perlu menuliskan konsep ilmiah target, miskonsepsi yang mungkin, bukti respons yang diharapkan, dan keputusan pembelajaran setelah hasil diperoleh. (Caleon & Subramaniam, 2010a; Peşman & Eryılmaz, 2010; Gurcay & Gulbas, 2015; Taslidere, 2016; Nakhleh, 1992)

Contoh kedua adalah laju reaksi. Stimulus sebaiknya menghubungkan fenomena dengan model partikel atau persamaan, lalu meminta peserta didik menunjukkan hubungan sebab-akibat. Jawaban yang hanya menyebut istilah belum cukup menjadi bukti pemahaman.

Pada efek katalis, diagnosis dapat memeriksa apakah peserta didik membedakan perubahan laju, keadaan setimbang, energi, dan komposisi. Kesalahan sering muncul ketika satu aturan dipindahkan ke konteks yang tidak sesuai.

Kasus hubungan konsentrasi dengan kekuatan asam menunjukkan pentingnya representasi. Respons verbal, gambar, simbol, dan hitungan dapat tidak konsisten. Hasil instrumen sebaiknya digunakan untuk memilih demonstrasi, diskusi, contoh-noncontoh, atau tugas remediasi yang secara langsung menantang hubungan salah.

- tentukan proposisi ilmiah dan konsepsi alternatif;
- gunakan minimal dua bentuk representasi;
- minta alasan atau bukti ketika format memungkinkan;
- rancang tindak lanjut berdasarkan pola, bukan hanya skor.

Dalam praktik kelas, Three-Tier Diagnostic Test dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.4.13 Kelebihan, Keterbatasan, dan Potensi Bias

Kekuatan utama Three-Tier Diagnostic Test meliputi membedakan ketidaktahuan dari miskonsepsi, menyediakan indikator metakognitif, membantu prioritas remediasi, dan dapat dianalisis pada tingkat butir dan konsepsi. Kekuatan tersebut menjelaskan situasi ketika instrumen layak dipilih dibanding teknik lain. (Caleon & Subramaniam, 2010a; Peşman & Eryilmaz, 2010; Gurcay & Gulbas, 2015; Gurel et al., 2015; AERA, APA, & NCME, 2014)

Keterbatasannya mencakup satu keyakinan gabungan tidak menunjukkan apakah keraguan berada pada jawaban atau alasan, ambang keyakinan dapat arbitrer, skala keyakinan rentan gaya respons, dan waktu pengerjaan lebih panjang. Keterbatasan harus diperlakukan sebagai batas inferensi, bukan sekadar paragraf formal yang tidak memengaruhi analisis.

Potensi bias dapat berasal dari bahasa, pengalaman konteks, akses teknologi, kecemasan, gaya menggunakan skala, keterampilan menggambar, atau kemampuan menulis. Instrumen perlu diuji pada kelompok yang beragam dan ditinjau untuk fungsi berbeda antarkelompok bila data mencukupi.

Keputusan memilih instrumen sebaiknya mempertimbangkan kedalaman data, ukuran kelompok, waktu, keterampilan analisis, tujuan diagnosis, dan konsekuensi kesalahan klasifikasi. Tidak ada satu metode yang unggul untuk semua tujuan.

- Kekuatan: membedakan ketidaktahuan dari miskonsepsi.
- Kekuatan: menyediakan indikator metakognitif.
- Batas: satu keyakinan gabungan tidak menunjukkan apakah keraguan berada pada jawaban atau alasan.
- Batas: ambang keyakinan dapat arbitrer.

Dalam praktik kelas, Three-Tier Diagnostic Test dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

Aspek	Temuan	Respons Metodologis
Kekuatan	membedakan ketidaktahuan dari miskonsepsi	Pilih ketika sesuai tujuan
Keterbatasan	satu keyakinan gabungan tidak menunjukkan apakah keraguan berada pada jawaban atau alasan	Tambahkan bukti atau kontrol
Bias potensial	Bahasa, konteks, perangkat, gaya respons	Uji proses respons

6.4.14 Integrasi Digital, Etika, dan Tata Kelola Data

Integrasi digital Three-Tier Diagnostic Test dapat berupa slider atau kategori keyakinan dengan urutan tetap, pencatatan waktu, dan analisis kalibrasi keyakinan. Teknologi seharusnya memperluas kualitas bukti, bukan hanya memindahkan lembar kertas ke layar. (Baker & Inventado, 2014; Romero & Ventura, 2020; AERA, APA, & NCME, 2014; Caleon & Subramaniam, 2010a)

Platform perlu menyimpan versi instrumen, urutan respons, timestamp yang relevan, aturan skor, dan log perubahan. Namun, hanya data yang dibutuhkan untuk tujuan diagnosis yang boleh dikumpulkan; pengumpulan berlebihan meningkatkan risiko privasi tanpa manfaat ilmiah.

Aspek etika utama adalah bahwa pelaporan keyakinan perlu menjaga privasi karena peserta didik dapat merasa dipermalukan ketika yakin tetapi salah. Peserta didik dan guru perlu mengetahui data apa yang dikumpulkan, untuk tujuan apa, siapa yang dapat mengakses, berapa lama disimpan, dan bagaimana kesalahan data diperbaiki.

Algoritma otomatis harus dapat diaudit. Kategori yang dihasilkan sistem perlu ditautkan ke respons mentah dan aturan keputusan. Kasus berisiko atau ambigu harus ditinjau manusia sebelum tindak lanjut penting dilakukan.

- gunakan identitas pseudonim dan akses berbasis peran;
- uji tampilan rumus, subskrip, grafik, dan gambar;
- sediakan opsi aksesibilitas dan alternatif perangkat;
- dokumentasikan algoritma serta perubahan versi.

Dalam praktik kelas, Three-Tier Diagnostic Test dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.4.15 Sintesis dan Pedoman Implementasi

Three-Tier Diagnostic Test bermanfaat apabila dipilih berdasarkan tujuan dan digunakan sesuai batasnya. Instrumen ini menghasilkan ketepatan jawaban, ketepatan alasan, dan keyakinan yang dinyatakan peserta didik, lalu data diolah melalui kombinasi tiga dimensi yang memisahkan pemahaman ilmiah, pengetahuan parsial, tidak tahu, menebak, dan miskonsepsi dengan keyakinan tinggi. (Caleon & Subramaniam, 2010a; Peşman & Eryilmaz, 2010; Gurcay & Gulbas, 2015; AERA, APA, & NCME, 2014; Kane, 2013)

Rangkaian kerja yang disarankan dimulai dari analisis konsep, identifikasi konsepsi alternatif, desain tugas, telaah ahli, uji proses respons, uji lapangan, analisis, penetapan kategori, dan penyusunan tindak lanjut. Setiap tahap harus meninggalkan dokumentasi yang dapat diperiksa.

Hasil diagnosis tidak berhenti pada tabel prevalensi. Guru perlu memilih konsepsi prioritas berdasarkan frekuensi, keyakinan, posisi sebagai konsep prasyarat, dampak terhadap topik lanjutan, dan kemungkinan diperbaiki melalui pembelajaran.

Prinsip akhirnya ialah kehati-hatian inferensial: sebutkan apa yang benar-benar ditunjukkan data, bedakan bukti kuat dari indikasi, gunakan triangulasi ketika diperlukan, dan hindari label permanen pada peserta didik.

- selaraskan tujuan, konstruk, tugas, skor, kategori, dan tindakan;
- utamakan konsepsi yang berdampak pada pembelajaran lanjutan;
- gabungkan bukti bila keputusan memiliki konsekuensi besar;
- evaluasi kembali efektivitas tindak lanjut melalui diagnosis ulang.

Dalam praktik kelas, Three-Tier Diagnostic Test dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.5 Four-Tier Diagnostic Test

6.5.1 Landasan Konseptual dan Kedudukan Instrumen

Secara metodologis, Four-Tier Diagnostic Test menghubungkan perilaku yang dapat diamati dengan konstruksi pemahaman yang tidak dapat diamati secara langsung. Dalam bab ini, instrumen tersebut didefinisikan sebagai instrumen yang mengukur jawaban dan keyakinan jawaban secara terpisah, kemudian alasan dan keyakinan alasan secara terpisah. Definisi operasional ini menegaskan bentuk bukti yang dikumpulkan sekaligus membatasi jenis kesimpulan yang layak dibuat. (Caleon & Subramaniam, 2010b; Kaltakci-Gurel et al., 2017; Yan & Subramaniam, 2018)

Klaim utamanya adalah bahwa instrumen ini mengurangi kelemahan keyakinan tunggal karena peneliti dapat mengetahui apakah kepastian berasal dari keputusan konten, alasan, atau keduanya. Klaim tersebut harus dipisahkan dari anggapan bahwa setiap jawaban salah merupakan miskonsepsi. Kesalahan dapat berasal dari kurang pengetahuan, ketidakcermatan, kesulitan bahasa, salah hitung, atau ketidakfamiliaran dengan representasi.

Data pokok yang dihasilkan berupa empat respons yang memotret ketepatan dan kalibrasi keyakinan pada dua komponen kognitif. Data itu baru memperoleh makna diagnostik setelah dikaitkan dengan peta konsep sasaran, kemungkinan konsepsi alternatif, dan aturan interpretasi yang telah ditetapkan sebelum analisis.

Konsekuensinya, peneliti dan guru perlu membedakan data mentah, skor, kategori, dan klaim konseptual yang dibuat dari hasil tersebut. Posisi terbaik Four-Tier Diagnostic Test adalah sebagai bagian dari sistem asesmen berlapis yang dapat dipadukan dengan bukti lain ketika konsekuensi keputusan semakin besar.

- menentukan konsep dan proposisi yang hendak didiagnosis;
- menjelaskan perilaku respons yang dianggap sebagai bukti;
- membatasi klaim sesuai kekuatan instrumen;
- menautkan hasil dengan tindakan pembelajaran yang realistis.

Dalam praktik kelas, Four-Tier Diagnostic Test dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.5.2 Perkembangan Historis dan Rasional Penggunaan

Perkembangan Four-Tier Diagnostic Test tidak dapat dipisahkan dari perubahan pandangan tentang belajar sains. Ketika peserta didik tidak lagi dipandang sebagai penerima informasi yang kosong, instrumen mulai dirancang untuk mengungkap gagasan awal, pola penalaran, dan keyakinan yang dibawa ke kelas. (Caleon & Subramaniam, 2010b; Kaltakci-Gurel et al., 2017; Yan & Subramaniam, 2018; Yang & Lin, 2015)

Pada tahap awal, penilaian banyak mengandalkan benar-salah atau skor total. Perkembangan penelitian konsepsi peserta didik kemudian mendorong penambahan alasan, keyakinan, tugas kinerja, rekaman proses, dan bukti sosial. Four-Tier Diagnostic Test merupakan salah satu jawaban metodologis terhadap kebutuhan tersebut.

Dalam pendidikan kimia, kebutuhan diagnosis semakin kuat karena peserta didik harus mengoordinasikan fenomena makroskopik, model submikroskopik, simbol, grafik, dan perhitungan. Instrumen yang hanya menguji hafalan definisi sering gagal memperlihatkan cara representasi itu dihubungkan.

Sejarah penggunaannya juga menunjukkan bahwa format yang lebih kompleks tidak otomatis lebih baik. Setiap tambahan tier, rubrik, atau kanal data harus mempunyai fungsi yang jelas dan didukung bukti bahwa informasi tambahan benar-benar memperbaiki keputusan.

- pergeseran dari pengukuran skor menuju diagnosis pola pemahaman;
- penggunaan respons terbuka dan wawancara untuk menemukan konsepsi alternatif;
- pengembangan teknik kuantitatif agar diagnosis dapat dilakukan pada kelompok besar;
- integrasi data digital dan metakognitif pada perkembangan mutakhir.

Dalam praktik kelas, Four-Tier Diagnostic Test dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.5.3 Konstruk, Klaim, dan Bukti Diagnosis

Konstruk yang dinilai melalui Four-Tier Diagnostic Test bukan objek yang terlihat langsung. Yang diamati adalah empat respons yang memotret ketepatan dan kalibrasi keyakinan pada dua komponen kognitif; dari sana dibuat inferensi tentang pemahaman konsep, miskonsepsi, pengetahuan parsial, atau ketidakpastian. (Hoe & Subramaniam, 2016; Brandriet & Bretz, 2014; Messick, 1995; Kane, 2013)

Klaim diagnosis perlu dirumuskan sebagai argumen. Pertama, tugas harus memerlukan konsep sasaran. Kedua, respons harus diberi skor atau kode secara konsisten. Ketiga, pola skor harus mempunyai hubungan yang masuk akal dengan struktur pengetahuan. Keempat, penggunaan hasil harus menghasilkan keputusan yang proporsional.

Sumber ancaman utama adalah ketidakjelasan konstruk. Misalnya, peserta didik dapat gagal karena istilah yang ambigu, beban membaca, representasi yang tidak dikenal, atau prosedur matematika. Kegagalan semacam itu tidak boleh langsung diperlakukan sebagai miskonsepsi kimia.

Oleh sebab itu, setiap pengembang perlu menulis pernyataan klaim-bukti: instrumen mengukur apa, perilaku apa yang menjadi indikator, alternatif penjelasan apa yang mungkin, dan bukti tambahan apa yang diperlukan untuk memperkuat kesimpulan.

- konstruk utama: pemahaman konseptual melalui tes diagnostik empat tingkat;
- bukti observabel: respons, alasan, tindakan, atau representasi;
- inferensi: profil konsepsi pada konten dan konteks tertentu;
- penggunaan: perencanaan pembelajaran, remediasi, atau penelitian.

Dalam praktik kelas, Four-Tier Diagnostic Test dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.5.4 Struktur, Komponen, dan Alur Respons

Struktur Four-Tier Diagnostic Test dibangun dari komponen yang saling bergantung. Komponen inti meliputi jawaban konten, keyakinan jawaban, alasan, keyakinan alasan, dan matriks klasifikasi empat tier. Hilangnya satu komponen dapat mengubah makna data yang diperoleh. (Caleon & Subramaniam, 2010b; Kaltakci-Gurel et al., 2017; Yan & Subramaniam, 2018; Yang & Lin, 2015)

Urutan penyajian harus ditetapkan secara sadar. Komponen yang muncul lebih awal dapat memberi petunjuk pada komponen berikutnya, sedangkan permintaan alasan atau keyakinan dapat mengubah cara peserta didik memeriksa jawabannya. Penguncian respons dan uji urutan diperlukan pada format digital.

Setiap komponen membutuhkan spesifikasi. Stimulus harus memuat informasi cukup, opsi atau prompt harus sejajar dengan indikator, skala harus memiliki jangkar, dan rubrik harus menjelaskan perbedaan kategori yang berdekatan.

Struktur yang baik memungkinkan peneliti menelusuri kembali setiap kategori hasil ke respons mentah. Transparansi ini penting agar guru tidak hanya menerima label otomatis tanpa memahami alasan di baliknya.

- jawaban konten;
- keyakinan jawaban;
- alasan;
- keyakinan alasan;
- matriks klasifikasi empat tier;

Dalam praktik kelas, Four-Tier Diagnostic Test dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

No.	Komponen	Fungsi Diagnostik
1	jawaban konten	Bukti yang harus ditelusuri
2	keyakinan jawaban	Bukti yang harus ditelusuri
3	alasan	Bukti yang harus ditelusuri
4	keyakinan alasan	Bukti yang harus ditelusuri

6.5.5 Desain Stimulus, Tugas, atau Pertanyaan

Desain stimulus pada Four-Tier Diagnostic Test harus memunculkan penggunaan konsep, bukan sekadar pengenalan istilah. Situasi dapat berbentuk fenomena, gambar partikel, grafik, data eksperimen, persamaan, atau masalah kontekstual, tetapi seluruh informasi yang tidak relevan perlu dikendalikan. (Haladyna et al., 2002; Caleon & Subramaniam, 2010b; Kaltakci-Gurel et al., 2017; Yan & Subramaniam, 2018)

Dalam kimia, satu stimulus sering menggabungkan beberapa tingkat representasi. Pengembang harus memutuskan apakah tujuan butir adalah menerjemahkan representasi, menjelaskan hubungan sebab-akibat, membedakan contoh dan noncontoh, atau menerapkan prinsip pada konteks baru.

Bahasa perlu ringkas tanpa menghilangkan syarat ilmiah. Kata seperti kuat, stabil, jenuh, spontan, atau setimbang mempunyai arti teknis dan sehari-hari yang berbeda. Ketidakjelasan istilah dapat menghasilkan kesalahan yang berasal dari bahasa, bukan struktur konsep.

Desain yang baik juga menghindari petunjuk format, pilihan terlalu panjang, negasi bertumpuk, gambar dekoratif, dan data yang tidak diperlukan. Setiap elemen harus mempunyai alasan diagnostik.

- gunakan konteks yang cukup dikenal tetapi tidak menghafalkan contoh;
- pastikan hanya satu interpretasi ilmiah yang menjadi target;
- uji semua simbol, satuan, grafik, dan gambar;
- sediakan ruang bagi respons yang tidak diprediksi bila tujuan masih eksploratif.

Dalam praktik kelas, Four-Tier Diagnostic Test dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.5.6 Prosedur Pengembangan Secara Sistematis

Pengembangan Four-Tier Diagnostic Test harus mengikuti alur yang terdokumentasi. Langkah yang lazim ialah membangun item dua tier, menetapkan dua skala keyakinan, menguji urutan penyajian, menyusun matriks kategori, kemudian memeriksa konsistensi internal, melakukan wawancara konfirmasi, uji lapangan, menetapkan aturan pelaporan. Urutan dapat berulang karena temuan uji coba sering memaksa revisi konstruk atau redaksi. (Caleon & Subramaniam, 2010b; Kaltakci-Gurel et al., 2017; Yan & Subramaniam, 2018; Adams & Wieman, 2011; AERA, APA, & NCME, 2014)

Tahap konseptual dimulai dengan analisis kurikulum, literatur miskonsepsi, dan peta konsep. Tahap empiris awal menggunakan wawancara, respons terbuka, atau observasi untuk memastikan bahwa kategori yang dirancang benar-benar muncul pada populasi sasaran.

Setiap versi instrumen perlu diberi nomor, tanggal, dan catatan perubahan. Revisi harus menyebutkan masalah yang ditemukan, bukti yang digunakan, perubahan yang dilakukan, dan alasan mempertahankan atau menghapus butir.

Pengembangan dianggap belum selesai ketika hanya menghasilkan butir yang tampak baik. Instrumen perlu diuji pada peserta didik yang serupa dengan pengguna sasaran, lalu dianalisis secara konseptual dan empiris sebelum dipakai untuk keputusan penting.

- 1. membangun item dua tier.
- 2. menetapkan dua skala keyakinan.
- 3. menguji urutan penyajian.
- 4. menyusun matriks kategori.
- 5. memeriksa konsistensi internal.

Dalam praktik kelas, Four-Tier Diagnostic Test dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.5.7 Pelaksanaan dan Pengumpulan Data

Pelaksanaan Four-Tier Diagnostic Test harus distandardisasi secukupnya agar perbedaan hasil lebih mungkin berasal dari pemahaman, bukan variasi instruksi, waktu, alat, atau bantuan. Petunjuk perlu menjelaskan tujuan umum tanpa mengajarkan jawaban. (Caleon & Subramaniam, 2010b; Kaltakci-Gurel et al., 2017; AERA, APA, & NCME, 2014)

Kondisi administrasi mencakup waktu, urutan tugas, kesempatan bertanya, penggunaan kalkulator, akses tabel periodik, kualitas gambar, dan perangkat digital. Pada instrumen kualitatif, standardisasi berarti konsistensi fokus dan prinsip probing, bukan membaca naskah secara kaku.

Data yang dikumpulkan ialah empat respons yang memotret ketepatan dan kalibrasi keyakinan pada dua komponen kognitif. Data mentah harus disimpan bersama metadata yang diperlukan, seperti kode butir, versi instrumen, waktu administrasi, dan kondisi khusus, tetapi hanya data yang relevan yang boleh dikumpulkan.

Sebelum administrasi luas, lakukan uji teknis. Pastikan tautan, urutan, penguncian respons, rekaman, penyimpanan, serta tampilan rumus dan subskrip bekerja. Gangguan teknis perlu dicatat sebagai bagian dari interpretasi.

- berikan instruksi yang sama dan mudah dipahami;
- catat kondisi yang dapat memengaruhi respons;
- hindari umpan balik sebelum seluruh respons diagnostik selesai;
- sediakan prosedur penanganan data tidak lengkap.

Dalam praktik kelas, Four-Tier Diagnostic Test dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.5.8 Penskoran, Pengodean, dan Pengolahan Data

Penskoran Four-Tier Diagnostic Test dapat diringkas sebagai kombinasi jawaban-alasan dan dua keyakinan untuk mengidentifikasi pemahaman ilmiah, pengetahuan parsial, kurang pengetahuan, false positive, false negative, dan miskonsepsi. Aturan ini harus ditulis sebelum melihat hasil utama agar keputusan tidak disesuaikan secara oportunistik dengan pola data. (Caleon & Subramaniam, 2010b; Kaltakci-Gurel et al., 2017; Yan & Subramaniam, 2018; Chi, 1997; AERA, APA, & NCME, 2014)

Skor total jarang cukup untuk diagnosis. Analisis perlu mempertahankan informasi tingkat butir, konsep, opsi, alasan, keyakinan, proses, atau representasi. Penggabungan data terlalu dini dapat menyembunyikan pola yang secara pedagogis penting.

Untuk respons terbuka atau kualitatif, buku kode harus mencantumkan definisi, contoh positif, contoh batas, noncontoh, dan aturan ketika satu respons memuat beberapa gagasan. Sebagian data perlu dinilai oleh lebih dari satu penilai.

Setiap transformasi data—misalnya dari respons ke kode, dari kode ke kategori, dan dari kategori ke persentase—harus dapat diaudit. Simpan respons asli dan keputusan pengodean agar kasus ambigu dapat ditinjau kembali.

- pisahkan skor ketepatan dari skor keyakinan atau kualitas alasan;
- laporkan kasus tidak lengkap dan aturan penanganannya;
- uji konsistensi penilai untuk data berkode;
- hindari mengubah data ordinal menjadi ukuran interval tanpa alasan.

Dalam praktik kelas, Four-Tier Diagnostic Test dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.5.9 Interpretasi dan Penetapan Kategori

Interpretasi hasil Four-Tier Diagnostic Test harus bergerak dari respons menuju pola, bukan langsung dari satu jawaban menuju label. Status konseptual menjadi lebih kuat apabila gagasan yang sama muncul pada beberapa butir, representasi, atau konteks. (Caleon & Subramaniam, 2010b; Kaltakci-Gurel et al., 2017; Yan & Subramaniam, 2018; Yang & Lin, 2015; Kane, 2013)

Kategori dapat mencakup pemahaman ilmiah, pemahaman parsial, kurang pengetahuan, respons tidak konsisten, false positive, false negative, dan miskonsepsi. Tidak semua instrumen mampu membedakan seluruh kategori tersebut; skema harus mengikuti informasi yang benar-benar tersedia.

Ambang, seperti batas keyakinan atau persentase pemilihan, perlu memiliki rasional dan diuji melalui analisis sensitivitas. Perubahan kecil pada ambang dapat mengubah prevalensi kategori secara berarti.

Interpretasi yang bertanggung jawab menyatakan bahwa kategori berlaku pada konsep, bentuk tugas, waktu, dan kondisi tertentu. Hasil merupakan hipotesis kerja untuk pembelajaran, bukan diagnosis psikologis permanen.

- gunakan lebih dari satu butir untuk konsepsi penting;
- periksa konsistensi antara jawaban dan alasan;
- tandai kasus ambigu untuk konfirmasi;
- hubungkan kategori dengan rekomendasi yang berbeda.

Dalam praktik kelas, Four-Tier Diagnostic Test dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.5.10 Validitas dan Kualitas Inferensi

Validitas Four-Tier Diagnostic Test berkaitan dengan kualitas interpretasi dan penggunaan hasil, bukan sifat tetap yang dimiliki instrumen. Fokus khususnya adalah bukti hubungan antar-tier, independensi yang memadai, kalibrasi keyakinan, dan kesesuaian kategori dengan data wawancara. (AERA, APA, & NCME, 2014; Messick, 1995; Kane, 2013; Haynes et al., 1995)

Bukti berdasarkan isi menilai keterwakilan domain dan akurasi konsep. Bukti proses respons memeriksa bagaimana peserta didik memahami tugas. Bukti struktur internal menilai kesesuaian pola respons. Bukti hubungan dengan variabel lain dan konsekuensi penggunaan melengkapi argumen.

Ancaman validitas dapat berasal dari konten yang kurang terwakili, tuntutan bahasa berlebihan, petunjuk opsi, rubrik yang tidak stabil, keyakinan yang tidak terkalibrasi, atau algoritma yang belajar dari label keliru.

Dokumen validasi perlu menghubungkan setiap klaim dengan bukti dan alternatif penjelasan. Menampilkan satu indeks tidak cukup; validasi adalah program argumentatif yang diperbarui setiap kali instrumen dipakai pada populasi atau tujuan baru.

- bukti isi dan kesesuaian kurikulum;
- bukti proses respons melalui wawancara kognitif;
- bukti struktur dan fungsi butir;
- bukti kegunaan serta konsekuensi keputusan.

Dalam praktik kelas, Four-Tier Diagnostic Test dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.5.11 Reliabilitas, Keajegan, dan Analisis Empiris

Keajegan Four-Tier Diagnostic Test harus dinilai sesuai jenis data. Pada respons objektif dapat digunakan konsistensi internal, stabilitas waktu, atau model pengukuran. Pada data kualitatif, perhatian diarahkan pada konsistensi pengodean, audit trail, dan ketepatan penerapan definisi kode. (Kuder & Richardson, 1937; Cronbach, 1951; Ding et al., 2006; Wren & Barbera, 2014; AERA, APA, & NCME, 2014)

Reliabilitas tinggi tidak otomatis menunjukkan diagnosis yang benar. Butir yang berulang-ulang mengukur hafalan dapat menghasilkan koefisien tinggi tetapi tidak mewakili pemahaman konseptual. Sebaliknya, konstruk yang multidimensi tidak selalu layak dipaksa menjadi satu skala.

Analisis empiris perlu memeriksa kesukaran, daya beda, fungsi opsi, pola data hilang, dan keterhubungan komponen. Untuk instrumen bertingkat, reliabilitas dapat dihitung pada setiap tier serta kombinasi yang mempunyai makna teoritis.

Ukuran sampel, distribusi kemampuan, dan homogenitas konten memengaruhi statistik. Karena itu, laporan harus menyatakan kondisi estimasi dan tidak menyalin batas angka sebagai aturan universal.

- pilih statistik yang sesuai dengan jenis skor;
- laporkan interval atau ketidakpastian bila tersedia;
- periksa butir secara substantif dan statistik;
- uji kembali setelah revisi atau perubahan platform.

Dalam praktik kelas, Four-Tier Diagnostic Test dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.5.12 Contoh Penerapan dalam Pembelajaran Kimia

Penerapan Four-Tier Diagnostic Test dalam kimia dapat dimulai dari konsep kinetika reaksi berbasis grafik. Pengembang perlu menuliskan konsep ilmiah target, miskonsepsi yang mungkin, bukti respons yang diharapkan, dan keputusan pembelajaran setelah hasil diperoleh. (Caleon & Subramaniam, 2010b; Kaltakci-Gurel et al., 2017; Yan & Subramaniam, 2018; Yang & Lin, 2015; Nakhleh, 1992)

Contoh kedua adalah asam-basa dan ionisasi. Stimulus sebaiknya menghubungkan fenomena dengan model partikel atau persamaan, lalu meminta peserta didik menunjukkan hubungan sebab-akibat. Jawaban yang hanya menyebut istilah belum cukup menjadi bukti pemahaman.

Pada struktur-properti senyawa, diagnosis dapat memeriksa apakah peserta didik membedakan perubahan laju, keadaan setimbang, energi, dan komposisi. Kesalahan sering muncul ketika satu aturan dipindahkan ke konteks yang tidak sesuai.

Kasus termokimia dan energi ikatan menunjukkan pentingnya representasi. Respons verbal, gambar, simbol, dan hitungan dapat tidak konsisten. Hasil instrumen sebaiknya digunakan untuk memilih demonstrasi, diskusi, contoh-noncontoh, atau tugas remediasi yang secara langsung menantang hubungan salah.

- tentukan proposisi ilmiah dan konsepsi alternatif;
- gunakan minimal dua bentuk representasi;
- minta alasan atau bukti ketika format memungkinkan;
- rancang tindak lanjut berdasarkan pola, bukan hanya skor.

Dalam praktik kelas, Four-Tier Diagnostic Test dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.5.13 Kelebihan, Keterbatasan, dan Potensi Bias

Kekuatan utama Four-Tier Diagnostic Test meliputi lokasi ketidakpastian lebih jelas, miskonsepsi kuat dapat dibedakan dari kesalahan ragu-ragu, berguna untuk prioritas pembelajaran, dan memungkinkan analisis kalibrasi yang rinci. Kekuatan tersebut menjelaskan situasi ketika instrumen layak dipilih dibanding teknik lain. (Caleon & Subramaniam, 2010b; Kaltakci-Gurel et al., 2017; Yan & Subramaniam, 2018; Gurel et al., 2015; AERA, APA, & NCME, 2014)

Keterbatasannya mencakup beban respons lebih tinggi, matriks kategori mudah menjadi terlalu rumit, hasil dipengaruhi kualitas skala keyakinan, dan konstruk tidak otomatis unidimensional. Keterbatasan harus diperlakukan sebagai batas inferensi, bukan sekadar paragraf formal yang tidak memengaruhi analisis.

Potensi bias dapat berasal dari bahasa, pengalaman konteks, akses teknologi, kecemasan, gaya menggunakan skala, keterampilan menggambar, atau kemampuan menulis. Instrumen perlu diuji pada kelompok yang beragam dan ditinjau untuk fungsi berbeda antarkelompok bila data mencukupi.

Keputusan memilih instrumen sebaiknya mempertimbangkan kedalaman data, ukuran kelompok, waktu, keterampilan analisis, tujuan diagnosis, dan konsekuensi kesalahan klasifikasi. Tidak ada satu metode yang unggul untuk semua tujuan.

- Kekuatan: lokasi ketidakpastian lebih jelas.
- Kekuatan: miskonsepsi kuat dapat dibedakan dari kesalahan ragu-ragu.
- Batas: beban respons lebih tinggi.
- Batas: matriks kategori mudah menjadi terlalu rumit.

Dalam praktik kelas, Four-Tier Diagnostic Test dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

Aspek	Temuan	Respons Metodologis
Kekuatan	lokasi ketidakpastian lebih jelas	Pilih ketika sesuai tujuan
Keterbatasan	beban respons lebih tinggi	Tambahkan bukti atau kontrol
Bias potensial	Bahasa, konteks, perangkat, gaya respons	Uji proses respons

6.5.14 Integrasi Digital, Etika, dan Tata Kelola Data

Integrasi digital Four-Tier Diagnostic Test dapat berupa penyajian empat tahap dengan penguncian respons, waktu per tier, dan dashboard profil konsepsi. Teknologi seharusnya memperluas kualitas bukti, bukan hanya memindahkan lembar kertas ke layar. (Baker & Inventado, 2014; Romero & Ventura, 2020; AERA, APA, & NCME, 2014; Caleon & Subramaniam, 2010b)

Platform perlu menyimpan versi instrumen, urutan respons, timestamp yang relevan, aturan skor, dan log perubahan. Namun, hanya data yang dibutuhkan untuk tujuan diagnosis yang boleh dikumpulkan; pengumpulan berlebihan meningkatkan risiko privasi tanpa manfaat ilmiah.

Aspek etika utama adalah bahwa kategori miskonsepsi kuat harus dilaporkan sebagai pola respons pada konteks instrumen, bukan identitas tetap peserta didik. Peserta didik dan guru perlu mengetahui data apa yang dikumpulkan, untuk tujuan apa, siapa yang dapat mengakses, berapa lama disimpan, dan bagaimana kesalahan data diperbaiki.

Algoritma otomatis harus dapat diaudit. Kategori yang dihasilkan sistem perlu ditautkan ke respons mentah dan aturan keputusan. Kasus berisiko atau ambigu harus ditinjau manusia sebelum tindak lanjut penting dilakukan.

- gunakan identitas pseudonim dan akses berbasis peran;
- uji tampilan rumus, subskrip, grafik, dan gambar;
- sediakan opsi aksesibilitas dan alternatif perangkat;
- dokumentasikan algoritma serta perubahan versi.

Dalam praktik kelas, Four-Tier Diagnostic Test dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.5.15 Sintesis dan Pedoman Implementasi

Four-Tier Diagnostic Test bermanfaat apabila dipilih berdasarkan tujuan dan digunakan sesuai batasnya. Instrumen ini menghasilkan empat respons yang memotret ketepatan dan kalibrasi keyakinan pada dua komponen kognitif, lalu data diolah melalui kombinasi jawaban-alasan dan dua keyakinan untuk mengidentifikasi pemahaman ilmiah, pengetahuan parsial, kurang pengetahuan, false positive, false negative, dan miskonsepsi. (Caleon & Subramaniam, 2010b; Kaltakci-Gurel et al., 2017; Yan & Subramaniam, 2018; AERA, APA, & NCME, 2014; Kane, 2013)

Rangkaian kerja yang disarankan dimulai dari analisis konsep, identifikasi konsepsi alternatif, desain tugas, telaah ahli, uji proses respons, uji lapangan, analisis, penetapan kategori, dan penyusunan tindak lanjut. Setiap tahap harus meninggalkan dokumentasi yang dapat diperiksa.

Hasil diagnosis tidak berhenti pada tabel prevalensi. Guru perlu memilih konsepsi prioritas berdasarkan frekuensi, keyakinan, posisi sebagai konsep prasyarat, dampak terhadap topik lanjutan, dan kemungkinan diperbaiki melalui pembelajaran.

Prinsip akhirnya ialah kehati-hatian inferensial: sebutkan apa yang benar-benar ditunjukkan data, bedakan bukti kuat dari indikasi, gunakan triangulasi ketika diperlukan, dan hindari label permanen pada peserta didik.

- selaraskan tujuan, konstruk, tugas, skor, kategori, dan tindakan;
- utamakan konsepsi yang berdampak pada pembelajaran lanjutan;
- gabungkan bukti bila keputusan memiliki konsekuensi besar;
- evaluasi kembali efektivitas tindak lanjut melalui diagnosis ulang.

Dalam praktik kelas, Four-Tier Diagnostic Test dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.6 Five-Tier Diagnostic Test

6.6.1 Landasan Konseptual dan Kedudukan Instrumen

Secara metodologis, Five-Tier Diagnostic Test menghubungkan perilaku yang dapat diamati dengan konstruksi pemahaman yang tidak dapat diamati secara langsung. Dalam bab ini, instrumen tersebut didefinisikan sebagai pengembangan multi-tier yang menambahkan tier kelima untuk menelusuri sumber konsepsi, mekanisme pembentukan, jenis representasi, atau alasan metakognitif setelah jawaban, alasan, dan keyakinan. Definisi operasional ini menegaskan bentuk bukti yang dikumpulkan sekaligus membatasi jenis kesimpulan yang layak dibuat. (Ma et al., 2025; Putra et al., 2020; Caleon & Subramaniam, 2010b)

Klaim utamanya adalah bahwa instrumen ini tidak hanya mengidentifikasi apakah miskonsepsi muncul, tetapi juga mengumpulkan petunjuk tentang asal, konteks, atau proses yang mempertahankannya. Klaim tersebut harus dipisahkan dari anggapan bahwa setiap jawaban salah merupakan miskonsepsi. Kesalahan dapat berasal dari kurang pengetahuan, ketidakcermatan, kesulitan bahasa, salah hitung, atau ketidakfamiliaran dengan representasi.

Data pokok yang dihasilkan berupa pola kognitif, metakognitif, dan atribusi sumber yang terhubung pada satu butir. Data itu baru memperoleh makna diagnostik setelah dikaitkan dengan peta konsep sasaran, kemungkinan konsepsi alternatif, dan aturan interpretasi yang telah ditetapkan sebelum analisis.

Pembedaan ini penting karena satu pola respons dapat memiliki lebih dari satu penjelasan yang masuk akal. Posisi terbaik Five-Tier Diagnostic Test adalah sebagai bagian dari sistem asesmen berlapis yang dapat dipadukan dengan bukti lain ketika konsekuensi keputusan semakin besar.

- menentukan konsep dan proposisi yang hendak didiagnosis;
- menjelaskan perilaku respons yang dianggap sebagai bukti;
- membatasi klaim sesuai kekuatan instrumen;
- menautkan hasil dengan tindakan pembelajaran yang realistis.

Dalam praktik kelas, Five-Tier Diagnostic Test dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.6.2 Perkembangan Historis dan Rasional Penggunaan

Perkembangan Five-Tier Diagnostic Test tidak dapat dipisahkan dari perubahan pandangan tentang belajar sains. Ketika peserta didik tidak lagi dipandang sebagai penerima informasi yang kosong, instrumen mulai dirancang untuk mengungkap gagasan awal, pola penalaran, dan keyakinan yang dibawa ke kelas. (Ma et al., 2025; Putra et al., 2020; Caleon & Subramaniam, 2010b; Kaltakci-Gurel et al., 2017)

Pada tahap awal, penilaian banyak mengandalkan benar-salah atau skor total. Perkembangan penelitian konsepsi peserta didik kemudian mendorong penambahan alasan, keyakinan, tugas kinerja, rekaman proses, dan bukti sosial. Five-Tier Diagnostic Test merupakan salah satu jawaban metodologis terhadap kebutuhan tersebut.

Dalam pendidikan kimia, kebutuhan diagnosis semakin kuat karena peserta didik harus mengoordinasikan fenomena makroskopik, model submikroskopik, simbol, grafik, dan perhitungan. Instrumen yang hanya menguji hafalan definisi sering gagal memperlihatkan cara representasi itu dihubungkan.

Sejarah penggunaannya juga menunjukkan bahwa format yang lebih kompleks tidak otomatis lebih baik. Setiap tambahan tier, rubrik, atau kanal data harus mempunyai fungsi yang jelas dan didukung bukti bahwa informasi tambahan benar-benar memperbaiki keputusan.

- pergeseran dari pengukuran skor menuju diagnosis pola pemahaman;
- penggunaan respons terbuka dan wawancara untuk menemukan konsepsi alternatif;
- pengembangan teknik kuantitatif agar diagnosis dapat dilakukan pada kelompok besar;
- integrasi data digital dan metakognitif pada perkembangan mutakhir.

Dalam praktik kelas, Five-Tier Diagnostic Test dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.6.3 Konstruk, Klaim, dan Bukti Diagnosis

Konstruk yang dinilai melalui Five-Tier Diagnostic Test bukan objek yang terlihat langsung. Yang diamati adalah pola kognitif, metakognitif, dan atribusi sumber yang terhubung pada satu butir; dari sana dibuat inferensi tentang pemahaman konsep, miskonsepsi, pengetahuan parsial, atau ketidakpastian. (Gurel et al., 2015; AERA, APA, & NCME, 2014; Messick, 1995; Kane, 2013)

Klaim diagnosis perlu dirumuskan sebagai argumen. Pertama, tugas harus memerlukan konsep sasaran. Kedua, respons harus diberi skor atau kode secara konsisten. Ketiga, pola skor harus mempunyai hubungan yang masuk akal dengan struktur pengetahuan. Keempat, penggunaan hasil harus menghasilkan keputusan yang proporsional.

Sumber ancaman utama adalah ketidakjelasan konstruk. Misalnya, peserta didik dapat gagal karena istilah yang ambigu, beban membaca, representasi yang tidak dikenal, atau prosedur matematika. Kegagalan semacam itu tidak boleh langsung diperlakukan sebagai miskonsepsi kimia.

Oleh sebab itu, setiap pengembang perlu menulis pernyataan klaim-bukti: instrumen mengukur apa, perilaku apa yang menjadi indikator, alternatif penjelasan apa yang mungkin, dan bukti tambahan apa yang diperlukan untuk memperkuat kesimpulan.

- konstruk utama: pemahaman konseptual melalui tes diagnostik lima tingkat;
- bukti observabel: respons, alasan, tindakan, atau representasi;
- inferensi: profil konsepsi pada konten dan konteks tertentu;
- penggunaan: perencanaan pembelajaran, remediasi, atau penelitian.

Dalam praktik kelas, Five-Tier Diagnostic Test dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.6.4 Struktur, Komponen, dan Alur Respons

Struktur Five-Tier Diagnostic Test dibangun dari komponen yang saling bergantung. Komponen inti meliputi jawaban konten, keyakinan jawaban, alasan, keyakinan alasan, dan tier sumber atau mekanisme miskonsepsi. Hilangnya satu komponen dapat mengubah makna data yang diperoleh. (Ma et al., 2025; Putra et al., 2020; Caleon & Subramaniam, 2010b; Kaltakci-Gurel et al., 2017)

Urutan penyajian harus ditetapkan secara sadar. Komponen yang muncul lebih awal dapat memberi petunjuk pada komponen berikutnya, sedangkan permintaan alasan atau keyakinan dapat mengubah cara peserta didik memeriksa jawabannya. Penguncian respons dan uji urutan diperlukan pada format digital.

Setiap komponen membutuhkan spesifikasi. Stimulus harus memuat informasi cukup, opsi atau prompt harus sejajar dengan indikator, skala harus memiliki jangkar, dan rubrik harus menjelaskan perbedaan kategori yang berdekatan.

Struktur yang baik memungkinkan peneliti menelusuri kembali setiap kategori hasil ke respons mentah. Transparansi ini penting agar guru tidak hanya menerima label otomatis tanpa memahami alasan di baliknya.

- jawaban konten;
- keyakinan jawaban;
- alasan;
- keyakinan alasan;
- tier sumber atau mekanisme miskonsepsi;

Dalam praktik kelas, Five-Tier Diagnostic Test dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

No.	Komponen	Fungsi Diagnostik
1	jawaban konten	Bukti yang harus ditelusuri
2	keyakinan jawaban	Bukti yang harus ditelusuri
3	alasan	Bukti yang harus ditelusuri
4	keyakinan alasan	Bukti yang harus ditelusuri

6.6.5 Desain Stimulus, Tugas, atau Pertanyaan

Desain stimulus pada Five-Tier Diagnostic Test harus memunculkan penggunaan konsep, bukan sekadar pengenalan istilah. Situasi dapat berbentuk fenomena, gambar partikel, grafik, data eksperimen, persamaan, atau masalah kontekstual, tetapi seluruh informasi yang tidak relevan perlu dikendalikan. (Haladyna et al., 2002; Ma et al., 2025; Putra et al., 2020; Caleon & Subramaniam, 2010b)

Dalam kimia, satu stimulus sering menggabungkan beberapa tingkat representasi. Pengembang harus memutuskan apakah tujuan butir adalah menerjemahkan representasi, menjelaskan hubungan sebab-akibat, membedakan contoh dan noncontoh, atau menerapkan prinsip pada konteks baru.

Bahasa perlu ringkas tanpa menghilangkan syarat ilmiah. Kata seperti kuat, stabil, jenuh, spontan, atau setimbang mempunyai arti teknis dan sehari-hari yang berbeda. Ketidakjelasan istilah dapat menghasilkan kesalahan yang berasal dari bahasa, bukan struktur konsep.

Desain yang baik juga menghindari petunjuk format, pilihan terlalu panjang, negasi bertumpuk, gambar dekoratif, dan data yang tidak diperlukan. Setiap elemen harus mempunyai alasan diagnostik.

- gunakan konteks yang cukup dikenal tetapi tidak menghafalkan contoh;
- pastikan hanya satu interpretasi ilmiah yang menjadi target;
- uji semua simbol, satuan, grafik, dan gambar;
- sediakan ruang bagi respons yang tidak diprediksi bila tujuan masih eksploratif.

Dalam praktik kelas, Five-Tier Diagnostic Test dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.6.6 Prosedur Pengembangan Secara Sistematis

Pengembangan Five-Tier Diagnostic Test harus mengikuti alur yang terdokumentasi. Langkah yang lazim ialah menetapkan fungsi tier kelima, mengumpulkan sumber konsepsi melalui wawancara, membangun kategori sumber, menguji kelengkapan kategori, kemudian menilai beban respons, menguji model pengukuran, mengonfirmasi melalui wawancara, merevisi aturan interpretasi. Urutan dapat berulang karena temuan uji coba sering memaksa revisi konstruk atau redaksi. (Ma et al., 2025; Putra et al., 2020; Caleon & Subramaniam, 2010b; Adams & Wieman, 2011; AERA, APA, & NCME, 2014)

Tahap konseptual dimulai dengan analisis kurikulum, literatur miskonsepsi, dan peta konsep. Tahap empiris awal menggunakan wawancara, respons terbuka, atau observasi untuk memastikan bahwa kategori yang dirancang benar-benar muncul pada populasi sasaran.

Setiap versi instrumen perlu diberi nomor, tanggal, dan catatan perubahan. Revisi harus menyebutkan masalah yang ditemukan, bukti yang digunakan, perubahan yang dilakukan, dan alasan mempertahankan atau menghapus butir.

Pengembangan dianggap belum selesai ketika hanya menghasilkan butir yang tampak baik. Instrumen perlu diuji pada peserta didik yang serupa dengan pengguna sasaran, lalu dianalisis secara konseptual dan empiris sebelum dipakai untuk keputusan penting.

- 1. menetapkan fungsi tier kelima.
- 2. mengumpulkan sumber konsepsi melalui wawancara.
- 3. membangun kategori sumber.
- 4. menguji kelengkapan kategori.
- 5. menilai beban respons.

Dalam praktik kelas, Five-Tier Diagnostic Test dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.6.7 Pelaksanaan dan Pengumpulan Data

Pelaksanaan Five-Tier Diagnostic Test harus distandardisasi secukupnya agar perbedaan hasil lebih mungkin berasal dari pemahaman, bukan variasi instruksi, waktu, alat, atau bantuan. Petunjuk perlu menjelaskan tujuan umum tanpa mengajarkan jawaban. (Ma et al., 2025; Putra et al., 2020; AERA, APA, & NCME, 2014)

Kondisi administrasi mencakup waktu, urutan tugas, kesempatan bertanya, penggunaan kalkulator, akses tabel periodik, kualitas gambar, dan perangkat digital. Pada instrumen kualitatif, standardisasi berarti konsistensi fokus dan prinsip probing, bukan membaca naskah secara kaku.

Data yang dikumpulkan ialah pola kognitif, metakognitif, dan atribusi sumber yang terhubung pada satu butir. Data mentah harus disimpan bersama metadata yang diperlukan, seperti kode butir, versi instrumen, waktu administrasi, dan kondisi khusus, tetapi hanya data yang relevan yang boleh dikumpulkan.

Sebelum administrasi luas, lakukan uji teknis. Pastikan tautan, urutan, penguncian respons, rekaman, penyimpanan, serta tampilan rumus dan subskrip bekerja. Gangguan teknis perlu dicatat sebagai bagian dari interpretasi.

- berikan instruksi yang sama dan mudah dipahami;
- catat kondisi yang dapat memengaruhi respons;
- hindari umpan balik sebelum seluruh respons diagnostik selesai;
- sediakan prosedur penanganan data tidak lengkap.

Dalam praktik kelas, Five-Tier Diagnostic Test dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.6.8 Penskoran, Pengodean, dan Pengolahan Data

Penskoran Five-Tier Diagnostic Test dapat diringkas sebagai matriks empat tier untuk status konseptual yang dilengkapi kode sumber seperti pengalaman sehari-hari, buku, guru, analogi, representasi, atau penalaran intuitif. Aturan ini harus ditulis sebelum melihat hasil utama agar keputusan tidak disesuaikan secara oportunistik dengan pola data. (Ma et al., 2025; Putra et al., 2020; Caleon & Subramaniam, 2010b; Chi, 1997; AERA, APA, & NCME, 2014)

Skor total jarang cukup untuk diagnosis. Analisis perlu mempertahankan informasi tingkat butir, konsep, opsi, alasan, keyakinan, proses, atau representasi. Penggabungan data terlalu dini dapat menyembunyikan pola yang secara pedagogis penting.

Untuk respons terbuka atau kualitatif, buku kode harus mencantumkan definisi, contoh positif, contoh batas, noncontoh, dan aturan ketika satu respons memuat beberapa gagasan. Sebagian data perlu dinilai oleh lebih dari satu penilai.

Setiap transformasi data—misalnya dari respons ke kode, dari kode ke kategori, dan dari kategori ke persentase—harus dapat diaudit. Simpan respons asli dan keputusan pengodean agar kasus ambigu dapat ditinjau kembali.

- pisahkan skor ketepatan dari skor keyakinan atau kualitas alasan;
- laporkan kasus tidak lengkap dan aturan penanganannya;
- uji konsistensi penilai untuk data berkode;
- hindari mengubah data ordinal menjadi ukuran interval tanpa alasan.

Dalam praktik kelas, Five-Tier Diagnostic Test dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.6.9 Interpretasi dan Penetapan Kategori

Interpretasi hasil Five-Tier Diagnostic Test harus bergerak dari respons menuju pola, bukan langsung dari satu jawaban menuju label. Status konseptual menjadi lebih kuat apabila gagasan yang sama muncul pada beberapa butir, representasi, atau konteks. (Ma et al., 2025; Putra et al., 2020; Caleon & Subramaniam, 2010b; Kaltakci-Gurel et al., 2017; Kane, 2013)

Kategori dapat mencakup pemahaman ilmiah, pemahaman parsial, kurang pengetahuan, respons tidak konsisten, false positive, false negative, dan miskonsepsi. Tidak semua instrumen mampu membedakan seluruh kategori tersebut; skema harus mengikuti informasi yang benar-benar tersedia.

Ambang, seperti batas keyakinan atau persentase pemilihan, perlu memiliki rasional dan diuji melalui analisis sensitivitas. Perubahan kecil pada ambang dapat mengubah prevalensi kategori secara berarti.

Interpretasi yang bertanggung jawab menyatakan bahwa kategori berlaku pada konsep, bentuk tugas, waktu, dan kondisi tertentu. Hasil merupakan hipotesis kerja untuk pembelajaran, bukan diagnosis psikologis permanen.

- gunakan lebih dari satu butir untuk konsepsi penting;
- periksa konsistensi antara jawaban dan alasan;
- tandai kasus ambigu untuk konfirmasi;
- hubungkan kategori dengan rekomendasi yang berbeda.

Dalam praktik kelas, Five-Tier Diagnostic Test dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.6.10 Validitas dan Kualitas Inferensi

Validitas Five-Tier Diagnostic Test berkaitan dengan kualitas interpretasi dan penggunaan hasil, bukan sifat tetap yang dimiliki instrumen. Fokus khususnya adalah kejelasan fungsi tier kelima, kemampuan peserta didik mengatribusikan sumber secara wajar, dan konsistensi dengan bukti kualitatif. (AERA, APA, & NCME, 2014; Messick, 1995; Kane, 2013; Haynes et al., 1995)

Bukti berdasarkan isi menilai keterwakilan domain dan akurasi konsep. Bukti proses respons memeriksa bagaimana peserta didik memahami tugas. Bukti struktur internal menilai kesesuaian pola respons. Bukti hubungan dengan variabel lain dan konsekuensi penggunaan melengkapi argumen.

Ancaman validitas dapat berasal dari konten yang kurang terwakili, tuntutan bahasa berlebihan, petunjuk opsi, rubrik yang tidak stabil, keyakinan yang tidak terkalibrasi, atau algoritma yang belajar dari label keliru.

Dokumen validasi perlu menghubungkan setiap klaim dengan bukti dan alternatif penjelasan. Menampilkan satu indeks tidak cukup; validasi adalah program argumentatif yang diperbarui setiap kali instrumen dipakai pada populasi atau tujuan baru.

- bukti isi dan kesesuaian kurikulum;
- bukti proses respons melalui wawancara kognitif;
- bukti struktur dan fungsi butir;
- bukti kegunaan serta konsekuensi keputusan.

Dalam praktik kelas, Five-Tier Diagnostic Test dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.6.11 Reliabilitas, Keajegan, dan Analisis Empiris

Keajegan Five-Tier Diagnostic Test harus dinilai sesuai jenis data. Pada respons objektif dapat digunakan konsistensi internal, stabilitas waktu, atau model pengukuran. Pada data kualitatif, perhatian diarahkan pada konsistensi pengodean, audit trail, dan ketepatan penerapan definisi kode. (Kuder & Richardson, 1937; Cronbach, 1951; Ding et al., 2006; Wren & Barbera, 2014; AERA, APA, & NCME, 2014)

Reliabilitas tinggi tidak otomatis menunjukkan diagnosis yang benar. Butir yang berulang-ulang mengukur hafalan dapat menghasilkan koefisien tinggi tetapi tidak mewakili pemahaman konseptual. Sebaliknya, konstruk yang multidimensi tidak selalu layak dipaksa menjadi satu skala.

Analisis empiris perlu memeriksa kesukaran, daya beda, fungsi opsi, pola data hilang, dan keterhubungan komponen. Untuk instrumen bertingkat, reliabilitas dapat dihitung pada setiap tier serta kombinasi yang mempunyai makna teoritis.

Ukuran sampel, distribusi kemampuan, dan homogenitas konten memengaruhi statistik. Karena itu, laporan harus menyatakan kondisi estimasi dan tidak menyalin batas angka sebagai aturan universal.

- pilih statistik yang sesuai dengan jenis skor;
- laporkan interval atau ketidakpastian bila tersedia;
- periksa butir secara substantif dan statistik;
- uji kembali setelah revisi atau perubahan platform.

Dalam praktik kelas, Five-Tier Diagnostic Test dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.6.12 Contoh Penerapan dalam Pembelajaran Kimia

Penerapan Five-Tier Diagnostic Test dalam kimia dapat dimulai dari konsep sumber miskonsepsi respirasi dan fotosintesis. Pengembang perlu menuliskan konsep ilmiah target, miskonsepsi yang mungkin, bukti respons yang diharapkan, dan keputusan pembelajaran setelah hasil diperoleh. (Ma et al., 2025; Putra et al., 2020; Caleon & Subramaniam, 2010b; Kaltakci-Gurel et al., 2017; Nakhleh, 1992)

Contoh kedua adalah asal gagasan tentang gelombang dan optik. Stimulus sebaiknya menghubungkan fenomena dengan model partikel atau persamaan, lalu meminta peserta didik menunjukkan hubungan sebab-akibat. Jawaban yang hanya menyebut istilah belum cukup menjadi bukti pemahaman.

Pada sumber kekeliruan asam-basa, diagnosis dapat memeriksa apakah peserta didik membedakan perubahan laju, keadaan setimbang, energi, dan komposisi. Kesalahan sering muncul ketika satu aturan dipindahkan ke konteks yang tidak sesuai.

Kasus pengaruh diagram pada konsep partikel menunjukkan pentingnya representasi. Respons verbal, gambar, simbol, dan hitungan dapat tidak konsisten. Hasil instrumen sebaiknya digunakan untuk memilih demonstrasi, diskusi, contoh-noncontoh, atau tugas remediasi yang secara langsung menantang hubungan salah.

- tentukan proposisi ilmiah dan konsepsi alternatif;
- gunakan minimal dua bentuk representasi;
- minta alasan atau bukti ketika format memungkinkan;
- rancang tindak lanjut berdasarkan pola, bukan hanya skor.

Dalam praktik kelas, Five-Tier Diagnostic Test dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.6.13 Kelebihan, Keterbatasan, dan Potensi Bias

Kekuatan utama Five-Tier Diagnostic Test meliputi menghasilkan diagnosis yang lebih kaya, mendukung perancangan remediasi berbasis sumber, menghubungkan respons kuantitatif dengan penjelasan etiologis, dan cocok untuk penelitian pengembangan. Kekuatan tersebut menjelaskan situasi ketika instrumen layak dipilih dibanding teknik lain. (Ma et al., 2025; Putra et al., 2020; Caleon & Subramaniam, 2010b; Gurel et al., 2015; AERA, APA, & NCME, 2014)

Keterbatasannya mencakup tier kelima belum memiliki format tunggal, atribusi sumber dapat bersifat retrospektif, instrumen panjang dan melelahkan, dan validasi membutuhkan bukti multipel. Keterbatasan harus diperlakukan sebagai batas inferensi, bukan sekadar paragraf formal yang tidak memengaruhi analisis.

Potensi bias dapat berasal dari bahasa, pengalaman konteks, akses teknologi, kecemasan, gaya menggunakan skala, keterampilan menggambar, atau kemampuan menulis. Instrumen perlu diuji pada kelompok yang beragam dan ditinjau untuk fungsi berbeda antarkelompok bila data mencukupi.

Keputusan memilih instrumen sebaiknya mempertimbangkan kedalaman data, ukuran kelompok, waktu, keterampilan analisis, tujuan diagnosis, dan konsekuensi kesalahan klasifikasi. Tidak ada satu metode yang unggul untuk semua tujuan.

- Kekuatan: menghasilkan diagnosis yang lebih kaya.
- Kekuatan: mendukung perancangan remediasi berbasis sumber.
- Batas: tier kelima belum memiliki format tunggal.
- Batas: atribusi sumber dapat bersifat retrospektif.

Dalam praktik kelas, Five-Tier Diagnostic Test dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

Aspek	Temuan	Respons Metodologis
Kekuatan	menghasilkan diagnosis yang lebih kaya	Pilih ketika sesuai tujuan
Keterbatasan	tier kelima belum memiliki format tunggal	Tambahkan bukti atau kontrol
Bias potensial	Bahasa, konteks, perangkat, gaya respons	Uji proses respons

6.6.14 Integrasi Digital, Etika, dan Tata Kelola Data

Integrasi digital Five-Tier Diagnostic Test dapat berupa menu sumber berjenjang, opsi lain dengan isian, pelacakan navigasi, dan analitik hubungan sumber dengan jenis miskonsepsi. Teknologi seharusnya memperluas kualitas bukti, bukan hanya memindahkan lembar kertas ke layar. (Baker & Inventado, 2014; Romero & Ventura, 2020; AERA, APA, & NCME, 2014; Ma et al., 2025)

Platform perlu menyimpan versi instrumen, urutan respons, timestamp yang relevan, aturan skor, dan log perubahan. Namun, hanya data yang dibutuhkan untuk tujuan diagnosis yang boleh dikumpulkan; pengumpulan berlebihan meningkatkan risiko privasi tanpa manfaat ilmiah.

Aspek etika utama adalah bahwa pertanyaan sumber tidak boleh diarahkan untuk menyalahkan guru, keluarga, atau buku tertentu tanpa triangulasi. Peserta didik dan guru perlu mengetahui data apa yang dikumpulkan, untuk tujuan apa, siapa yang dapat mengakses, berapa lama disimpan, dan bagaimana kesalahan data diperbaiki.

Algoritma otomatis harus dapat diaudit. Kategori yang dihasilkan sistem perlu ditautkan ke respons mentah dan aturan keputusan. Kasus berisiko atau ambigu harus ditinjau manusia sebelum tindak lanjut penting dilakukan.

- gunakan identitas pseudonim dan akses berbasis peran;
- uji tampilan rumus, subskrip, grafik, dan gambar;
- sediakan opsi aksesibilitas dan alternatif perangkat;
- dokumentasikan algoritma serta perubahan versi.

Dalam praktik kelas, Five-Tier Diagnostic Test dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.6.15 Sintesis dan Pedoman Implementasi

Five-Tier Diagnostic Test bermanfaat apabila dipilih berdasarkan tujuan dan digunakan sesuai batasnya. Instrumen ini menghasilkan pola kognitif, metakognitif, dan atribusi sumber yang terhubung pada satu butir, lalu data diolah melalui matriks empat tier untuk status konseptual yang dilengkapi kode sumber seperti pengalaman sehari-hari, buku, guru, analogi, representasi, atau penalaran intuitif. (Ma et al., 2025; Putra et al., 2020; Caleon & Subramaniam, 2010b; AERA, APA, & NCME, 2014; Kane, 2013)

Rangkaian kerja yang disarankan dimulai dari analisis konsep, identifikasi konsepsi alternatif, desain tugas, telaah ahli, uji proses respons, uji lapangan, analisis, penetapan kategori, dan penyusunan tindak lanjut. Setiap tahap harus meninggalkan dokumentasi yang dapat diperiksa.

Hasil diagnosis tidak berhenti pada tabel prevalensi. Guru perlu memilih konsepsi prioritas berdasarkan frekuensi, keyakinan, posisi sebagai konsep prasyarat, dampak terhadap topik lanjutan, dan kemungkinan diperbaiki melalui pembelajaran.

Prinsip akhirnya ialah kehati-hatian inferensial: sebutkan apa yang benar-benar ditunjukkan data, bedakan bukti kuat dari indikasi, gunakan triangulasi ketika diperlukan, dan hindari label permanen pada peserta didik.

- selaraskan tujuan, konstruk, tugas, skor, kategori, dan tindakan;
- utamakan konsepsi yang berdampak pada pembelajaran lanjutan;
- gabungkan bukti bila keputusan memiliki konsekuensi besar;
- evaluasi kembali efektivitas tindak lanjut melalui diagnosis ulang.

Dalam praktik kelas, Five-Tier Diagnostic Test dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.7 Certainty of Response Index

6.7.1 Landasan Konseptual dan Kedudukan Instrumen

Certainty of Response Index menempati posisi tertentu dalam rangkaian bukti diagnosis, dari pengungkapan respons awal sampai penentuan tindak lanjut pembelajaran. Dalam bab ini, instrumen tersebut didefinisikan sebagai skala yang meminta peserta didik menyatakan derajat kepastian terhadap jawabannya sehingga jawaban salah yang yakin dapat dibedakan dari jawaban salah yang tidak yakin. Definisi operasional ini menegaskan bentuk bukti yang dikumpulkan sekaligus membatasi jenis kesimpulan yang layak dibuat. (Hasan et al., 1999; Brandriet & Bretz, 2014; Mathabathe & Potgieter, 2014)

Klaim utamanya adalah bahwa instrumen ini memberikan petunjuk praktis tentang perbedaan antara kurang pengetahuan dan miskonsepsi, terutama bila dipadukan dengan pola jawaban, alasan, dan nilai rata-rata kelompok. Klaim tersebut harus dipisahkan dari anggapan bahwa setiap jawaban salah merupakan miskonsepsi. Kesalahan dapat berasal dari kurang pengetahuan, ketidakcermatan, kesulitan bahasa, salah hitung, atau ketidakfamiliaran dengan representasi.

Data pokok yang dihasilkan berupa ketepatan respons dan rating kepastian pada skala ordinal atau numerik. Data itu baru memperoleh makna diagnostik setelah dikaitkan dengan peta konsep sasaran, kemungkinan konsepsi alternatif, dan aturan interpretasi yang telah ditetapkan sebelum analisis.

Konsekuensinya, peneliti dan guru perlu membedakan data mentah, skor, kategori, dan klaim konseptual yang dibuat dari hasil tersebut. Posisi terbaik Certainty of Response Index adalah sebagai bagian dari sistem asesmen berlapis yang dapat dipadukan dengan bukti lain ketika konsekuensi keputusan semakin besar.

- menentukan konsep dan proposisi yang hendak didiagnosis;
- menjelaskan perilaku respons yang dianggap sebagai bukti;
- membatasi klaim sesuai kekuatan instrumen;
- menautkan hasil dengan tindakan pembelajaran yang realistis.

Dalam praktik kelas, Certainty of Response Index dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.7.2 Perkembangan Historis dan Rasional Penggunaan

Perkembangan Certainty of Response Index tidak dapat dipisahkan dari perubahan pandangan tentang belajar sains. Ketika peserta didik tidak lagi dipandang sebagai penerima informasi yang kosong, instrumen mulai dirancang untuk mengungkap gagasan awal, pola penalaran, dan keyakinan yang dibawa ke kelas. (Hasan et al., 1999; Brandriet & Bretz, 2014; Mathabathe & Potgieter, 2014; Hoe & Subramaniam, 2016)

Pada tahap awal, penilaian banyak mengandalkan benar-salah atau skor total. Perkembangan penelitian konsepsi peserta didik kemudian mendorong penambahan alasan, keyakinan, tugas kinerja, rekaman proses, dan bukti sosial. Certainty of Response Index merupakan salah satu jawaban metodologis terhadap kebutuhan tersebut.

Dalam pendidikan kimia, kebutuhan diagnosis semakin kuat karena peserta didik harus mengoordinasikan fenomena makroskopik, model submikroskopik, simbol, grafik, dan perhitungan. Instrumen yang hanya menguji hafalan definisi sering gagal memperlihatkan cara representasi itu dihubungkan.

Sejarah penggunaannya juga menunjukkan bahwa format yang lebih kompleks tidak otomatis lebih baik. Setiap tambahan tier, rubrik, atau kanal data harus mempunyai fungsi yang jelas dan didukung bukti bahwa informasi tambahan benar-benar memperbaiki keputusan.

- pergeseran dari pengukuran skor menuju diagnosis pola pemahaman;
- penggunaan respons terbuka dan wawancara untuk menemukan konsepsi alternatif;
- pengembangan teknik kuantitatif agar diagnosis dapat dilakukan pada kelompok besar;
- integrasi data digital dan metakognitif pada perkembangan mutakhir.

Dalam praktik kelas, Certainty of Response Index dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.7.3 Konstruk, Klaim, dan Bukti Diagnosis

Konstruk yang dinilai melalui Certainty of Response Index bukan objek yang terlihat langsung. Yang diamati adalah ketepatan respons dan rating kepastian pada skala ordinal atau numerik; dari sana dibuat inferensi tentang pemahaman konsep, miskonsepsi, pengetahuan parsial, atau ketidakpastian. (Hoe & Subramaniam, 2016; Caleon & Subramaniam, 2010b; Messick, 1995; Kane, 2013)

Klaim diagnosis perlu dirumuskan sebagai argumen. Pertama, tugas harus memerlukan konsep sasaran. Kedua, respons harus diberi skor atau kode secara konsisten. Ketiga, pola skor harus mempunyai hubungan yang masuk akal dengan struktur pengetahuan. Keempat, penggunaan hasil harus menghasilkan keputusan yang proporsional.

Sumber ancaman utama adalah ketidakjelasan konstruk. Misalnya, peserta didik dapat gagal karena istilah yang ambigu, beban membaca, representasi yang tidak dikenal, atau prosedur matematika. Kegagalan semacam itu tidak boleh langsung diperlakukan sebagai miskonsepsi kimia.

Oleh sebab itu, setiap pengembang perlu menulis pernyataan klaim-bukti: instrumen mengukur apa, perilaku apa yang menjadi indikator, alternatif penjelasan apa yang mungkin, dan bukti tambahan apa yang diperlukan untuk memperkuat kesimpulan.

- konstruk utama: pemahaman konseptual melalui indeks keyakinan respons;
- bukti observabel: respons, alasan, tindakan, atau representasi;
- inferensi: profil konsepsi pada konten dan konteks tertentu;
- penggunaan: perencanaan pembelajaran, remediasi, atau penelitian.

Dalam praktik kelas, Certainty of Response Index dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.7.4 Struktur, Komponen, dan Alur Respons

Struktur Certainty of Response Index dibangun dari komponen yang saling bergantung. Komponen inti meliputi jawaban terhadap butir, skala CRI, ambang rendah-tinggi, matriks benar-salah dan yakin-ragu, dan agregasi individu dan kelompok. Hilangnya satu komponen dapat mengubah makna data yang diperoleh. (Hasan et al., 1999; Brandriet & Bretz, 2014; Mathabathe & Potgieter, 2014; Hoe & Subramaniam, 2016)

Urutan penyajian harus ditetapkan secara sadar. Komponen yang muncul lebih awal dapat memberi petunjuk pada komponen berikutnya, sedangkan permintaan alasan atau keyakinan dapat mengubah cara peserta didik memeriksa jawabannya. Penguncian respons dan uji urutan diperlukan pada format digital.

Setiap komponen membutuhkan spesifikasi. Stimulus harus memuat informasi cukup, opsi atau prompt harus sejajar dengan indikator, skala harus memiliki jangkar, dan rubrik harus menjelaskan perbedaan kategori yang berdekatan.

Struktur yang baik memungkinkan peneliti menelusuri kembali setiap kategori hasil ke respons mentah. Transparansi ini penting agar guru tidak hanya menerima label otomatis tanpa memahami alasan di baliknya.

- jawaban terhadap butir;
- skala CRI;
- ambang rendah-tinggi;
- matriks benar-salah dan yakin-ragu;
- agregasi individu dan kelompok;

Dalam praktik kelas, Certainty of Response Index dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

No.	Komponen	Fungsi Diagnostik
1	jawaban terhadap butir	Bukti yang harus ditelusuri
2	skala CRI	Bukti yang harus ditelusuri
3	ambang rendah-tinggi	Bukti yang harus ditelusuri
4	matriks benar-salah dan yakin-ragu	Bukti yang harus ditelusuri

6.7.5 Desain Stimulus, Tugas, atau Pertanyaan

Desain stimulus pada Certainty of Response Index harus memunculkan penggunaan konsep, bukan sekadar pengenalan istilah. Situasi dapat berbentuk fenomena, gambar partikel, grafik, data eksperimen, persamaan, atau masalah kontekstual, tetapi seluruh informasi yang tidak relevan perlu dikendalikan. (Haladyna et al., 2002; Hasan et al., 1999; Brandriet & Bretz, 2014; Mathabathe & Potgieter, 2014)

Dalam kimia, satu stimulus sering menggabungkan beberapa tingkat representasi. Pengembang harus memutuskan apakah tujuan butir adalah menerjemahkan representasi, menjelaskan hubungan sebab-akibat, membedakan contoh dan noncontoh, atau menerapkan prinsip pada konteks baru.

Bahasa perlu ringkas tanpa menghilangkan syarat ilmiah. Kata seperti kuat, stabil, jenuh, spontan, atau setimbang mempunyai arti teknis dan sehari-hari yang berbeda. Ketidakjelasan istilah dapat menghasilkan kesalahan yang berasal dari bahasa, bukan struktur konsep.

Desain yang baik juga menghindari petunjuk format, pilihan terlalu panjang, negasi bertumpuk, gambar dekoratif, dan data yang tidak diperlukan. Setiap elemen harus mempunyai alasan diagnostik.

- gunakan konteks yang cukup dikenal tetapi tidak menghafalkan contoh;
- pastikan hanya satu interpretasi ilmiah yang menjadi target;
- uji semua simbol, satuan, grafik, dan gambar;
- sediakan ruang bagi respons yang tidak diprediksi bila tujuan masih eksploratif.

Dalam praktik kelas, Certainty of Response Index dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.7.6 Prosedur Pengembangan Secara Sistematis

Pengembangan Certainty of Response Index harus mengikuti alur yang terdokumentasi. Langkah yang lazim ialah menentukan skala, menulis jangkar verbal, melatih peserta didik menggunakan skala, menetapkan ambang, kemudian menggabungkan dengan jawaban dan alasan, menganalisis kalibrasi, menguji stabilitas, menafsirkan secara hati-hati. Urutan dapat berulang karena temuan uji coba sering memaksa revisi konstruk atau redaksi. (Hasan et al., 1999; Brandriet & Bretz, 2014; Mathabathe & Potgieter, 2014; Adams & Wieman, 2011; AERA, APA, & NCME, 2014)

Tahap konseptual dimulai dengan analisis kurikulum, literatur miskonsepsi, dan peta konsep. Tahap empiris awal menggunakan wawancara, respons terbuka, atau observasi untuk memastikan bahwa kategori yang dirancang benar-benar muncul pada populasi sasaran.

Setiap versi instrumen perlu diberi nomor, tanggal, dan catatan perubahan. Revisi harus menyebutkan masalah yang ditemukan, bukti yang digunakan, perubahan yang dilakukan, dan alasan mempertahankan atau menghapus butir.

Pengembangan dianggap belum selesai ketika hanya menghasilkan butir yang tampak baik. Instrumen perlu diuji pada peserta didik yang serupa dengan pengguna sasaran, lalu dianalisis secara konseptual dan empiris sebelum dipakai untuk keputusan penting.

- 1. menentukan skala.
- 2. menulis jangkar verbal.
- 3. melatih peserta didik menggunakan skala.
- 4. menetapkan ambang.
- 5. menggabungkan dengan jawaban dan alasan.

Dalam praktik kelas, Certainty of Response Index dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.7.7 Pelaksanaan dan Pengumpulan Data

Pelaksanaan Certainty of Response Index harus distandardisasi secukupnya agar perbedaan hasil lebih mungkin berasal dari pemahaman, bukan variasi instruksi, waktu, alat, atau bantuan. Petunjuk perlu menjelaskan tujuan umum tanpa mengajarkan jawaban. (Hasan et al., 1999; Brandriet & Bretz, 2014; AERA, APA, & NCME, 2014)

Kondisi administrasi mencakup waktu, urutan tugas, kesempatan bertanya, penggunaan kalkulator, akses tabel periodik, kualitas gambar, dan perangkat digital. Pada instrumen kualitatif, standardisasi berarti konsistensi fokus dan prinsip probing, bukan membaca naskah secara kaku.

Data yang dikumpulkan ialah ketepatan respons dan rating kepastian pada skala ordinal atau numerik. Data mentah harus disimpan bersama metadata yang diperlukan, seperti kode butir, versi instrumen, waktu administrasi, dan kondisi khusus, tetapi hanya data yang relevan yang boleh dikumpulkan.

Sebelum administrasi luas, lakukan uji teknis. Pastikan tautan, urutan, penguncian respons, rekaman, penyimpanan, serta tampilan rumus dan subskrip bekerja. Gangguan teknis perlu dicatat sebagai bagian dari interpretasi.

- berikan instruksi yang sama dan mudah dipahami;
- catat kondisi yang dapat memengaruhi respons;
- hindari umpan balik sebelum seluruh respons diagnostik selesai;
- sediakan prosedur penanganan data tidak lengkap.

Dalam praktik kelas, Certainty of Response Index dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.7.8 Penskoran, Pengodean, dan Pengolahan Data

Penskoran Certainty of Response Index dapat diringkas sebagai tabulasi jawaban benar-salah dengan CRI rendah-tinggi serta analisis rata-rata CRI untuk opsi tertentu. Aturan ini harus ditulis sebelum melihat hasil utama agar keputusan tidak disesuaikan secara oportunistik dengan pola data. (Hasan et al., 1999; Brandriet & Bretz, 2014; Mathabathe & Potgieter, 2014; Chi, 1997; AERA, APA, & NCME, 2014)

Skor total jarang cukup untuk diagnosis. Analisis perlu mempertahankan informasi tingkat butir, konsep, opsi, alasan, keyakinan, proses, atau representasi. Penggabungan data terlalu dini dapat menyembunyikan pola yang secara pedagogis penting.

Untuk respons terbuka atau kualitatif, buku kode harus mencantumkan definisi, contoh positif, contoh batas, noncontoh, dan aturan ketika satu respons memuat beberapa gagasan. Sebagian data perlu dinilai oleh lebih dari satu penilai.

Setiap transformasi data—misalnya dari respons ke kode, dari kode ke kategori, dan dari kategori ke persentase—harus dapat diaudit. Simpan respons asli dan keputusan pengodean agar kasus ambigu dapat ditinjau kembali.

- pisahkan skor ketepatan dari skor keyakinan atau kualitas alasan;
- laporkan kasus tidak lengkap dan aturan penanganannya;
- uji konsistensi penilai untuk data berkode;
- hindari mengubah data ordinal menjadi ukuran interval tanpa alasan.

Dalam praktik kelas, Certainty of Response Index dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.7.9 Interpretasi dan Penetapan Kategori

Interpretasi hasil Certainty of Response Index harus bergerak dari respons menuju pola, bukan langsung dari satu jawaban menuju label. Status konseptual menjadi lebih kuat apabila gagasan yang sama muncul pada beberapa butir, representasi, atau konteks. (Hasan et al., 1999; Brandriet & Bretz, 2014; Mathabathe & Potgieter, 2014; Hoe & Subramaniam, 2016; Kane, 2013)

Kategori dapat mencakup pemahaman ilmiah, pemahaman parsial, kurang pengetahuan, respons tidak konsisten, false positive, false negative, dan miskonsepsi. Tidak semua instrumen mampu membedakan seluruh kategori tersebut; skema harus mengikuti informasi yang benar-benar tersedia.

Ambang, seperti batas keyakinan atau persentase pemilihan, perlu memiliki rasional dan diuji melalui analisis sensitivitas. Perubahan kecil pada ambang dapat mengubah prevalensi kategori secara berarti.

Interpretasi yang bertanggung jawab menyatakan bahwa kategori berlaku pada konsep, bentuk tugas, waktu, dan kondisi tertentu. Hasil merupakan hipotesis kerja untuk pembelajaran, bukan diagnosis psikologis permanen.

- gunakan lebih dari satu butir untuk konsepsi penting;
- periksa konsistensi antara jawaban dan alasan;
- tandai kasus ambigu untuk konfirmasi;
- hubungkan kategori dengan rekomendasi yang berbeda.

Dalam praktik kelas, Certainty of Response Index dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.7.10 Validitas dan Kualitas Inferensi

Validitas Certainty of Response Index berkaitan dengan kualitas interpretasi dan penggunaan hasil, bukan sifat tetap yang dimiliki instrumen. Fokus khususnya adalah hubungan CRI dengan kinerja, stabilitas rating, keterpahaman jangkar, dan bukti bahwa skala mengukur kepastian bukan tingkat kesulitan yang dirasakan. (AERA, APA, & NCME, 2014; Messick, 1995; Kane, 2013; Haynes et al., 1995)

Bukti berdasarkan isi menilai keterwakilan domain dan akurasi konsep. Bukti proses respons memeriksa bagaimana peserta didik memahami tugas. Bukti struktur internal menilai kesesuaian pola respons. Bukti hubungan dengan variabel lain dan konsekuensi penggunaan melengkapi argumen.

Ancaman validitas dapat berasal dari konten yang kurang terwakili, tuntutan bahasa berlebihan, petunjuk opsi, rubrik yang tidak stabil, keyakinan yang tidak terkalibrasi, atau algoritma yang belajar dari label keliru.

Dokumen validasi perlu menghubungkan setiap klaim dengan bukti dan alternatif penjelasan. Menampilkan satu indeks tidak cukup; validasi adalah program argumentatif yang diperbarui setiap kali instrumen dipakai pada populasi atau tujuan baru.

- bukti isi dan kesesuaian kurikulum;
- bukti proses respons melalui wawancara kognitif;
- bukti struktur dan fungsi butir;
- bukti kegunaan serta konsekuensi keputusan.

Dalam praktik kelas, Certainty of Response Index dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.7.11 Reliabilitas, Keajegan, dan Analisis Empiris

Keajegan Certainty of Response Index harus dinilai sesuai jenis data. Pada respons objektif dapat digunakan konsistensi internal, stabilitas waktu, atau model pengukuran. Pada data kualitatif, perhatian diarahkan pada konsistensi pengodean, audit trail, dan ketepatan penerapan definisi kode. (Kuder & Richardson, 1937; Cronbach, 1951; Ding et al., 2006; Wren & Barbera, 2014; AERA, APA, & NCME, 2014)

Reliabilitas tinggi tidak otomatis menunjukkan diagnosis yang benar. Butir yang berulang-ulang mengukur hafalan dapat menghasilkan koefisien tinggi tetapi tidak mewakili pemahaman konseptual. Sebaliknya, konstruk yang multidimensi tidak selalu layak dipaksa menjadi satu skala.

Analisis empiris perlu memeriksa kesukaran, daya beda, fungsi opsi, pola data hilang, dan keterhubungan komponen. Untuk instrumen bertingkat, reliabilitas dapat dihitung pada setiap tier serta kombinasi yang mempunyai makna teoritis.

Ukuran sampel, distribusi kemampuan, dan homogenitas konten memengaruhi statistik. Karena itu, laporan harus menyatakan kondisi estimasi dan tidak menyalin batas angka sebagai aturan universal.

- pilih statistik yang sesuai dengan jenis skor;
- laporkan interval atau ketidakpastian bila tersedia;
- periksa butir secara substantif dan statistik;
- uji kembali setelah revisi atau perubahan platform.

Dalam praktik kelas, Certainty of Response Index dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.7.12 Contoh Penerapan dalam Pembelajaran Kimia

Penerapan Certainty of Response Index dalam kimia dapat dimulai dari konsep bilangan oksidasi dan transfer elektron. Pengembang perlu menuliskan konsep ilmiah target, miskonsepsi yang mungkin, bukti respons yang diharapkan, dan keputusan pembelajaran setelah hasil diperoleh. (Hasan et al., 1999; Brandriet & Bretz, 2014; Mathabathe & Potgieter, 2014; Hoe & Subramaniam, 2016; Nakhleh, 1992)

Contoh kedua adalah konsep mol. Stimulus sebaiknya menghubungkan fenomena dengan model partikel atau persamaan, lalu meminta peserta didik menunjukkan hubungan sebab-akibat. Jawaban yang hanya menyebut istilah belum cukup menjadi bukti pemahaman.

Pada termokimia, diagnosis dapat memeriksa apakah peserta didik membedakan perubahan laju, keadaan setimbang, energi, dan komposisi. Kesalahan sering muncul ketika satu aturan dipindahkan ke konteks yang tidak sesuai.

Kasus kesetimbangan asam-basa menunjukkan pentingnya representasi. Respons verbal, gambar, simbol, dan hitungan dapat tidak konsisten. Hasil instrumen sebaiknya digunakan untuk memilih demonstrasi, diskusi, contoh-noncontoh, atau tugas remediasi yang secara langsung menantang hubungan salah.

- tentukan proposisi ilmiah dan konsepsi alternatif;
- gunakan minimal dua bentuk representasi;
- minta alasan atau bukti ketika format memungkinkan;
- rancang tindak lanjut berdasarkan pola, bukan hanya skor.

Dalam praktik kelas, Certainty of Response Index dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.7.13 Kelebihan, Keterbatasan, dan Potensi Bias

Kekuatan utama Certainty of Response Index meliputi mudah ditambahkan pada tes yang sudah ada, ringkas, mendukung prioritas miskonsepsi yakin, dan meningkatkan kesadaran metakognitif. Kekuatan tersebut menjelaskan situasi ketika instrumen layak dipilih dibanding teknik lain. (Hasan et al., 1999; Brandriet & Bretz, 2014; Mathabathe & Potgieter, 2014; Gurel et al., 2015; AERA, APA, & NCME, 2014)

Keterbatasannya mencakup keyakinan bukan bukti alasan, ambang dapat berubah menurut skala, budaya dan gaya respons memengaruhi rating, dan peserta didik dapat terlalu percaya diri atau terlalu hati-hati. Keterbatasan harus diperlakukan sebagai batas inferensi, bukan sekadar paragraf formal yang tidak memengaruhi analisis.

Potensi bias dapat berasal dari bahasa, pengalaman konteks, akses teknologi, kecemasan, gaya menggunakan skala, keterampilan menggambar, atau kemampuan menulis. Instrumen perlu diuji pada kelompok yang beragam dan ditinjau untuk fungsi berbeda antarkelompok bila data mencukupi.

Keputusan memilih instrumen sebaiknya mempertimbangkan kedalaman data, ukuran kelompok, waktu, keterampilan analisis, tujuan diagnosis, dan konsekuensi kesalahan klasifikasi. Tidak ada satu metode yang unggul untuk semua tujuan.

- Kekuatan: mudah ditambahkan pada tes yang sudah ada.
- Kekuatan: ringkas.
- Batas: keyakinan bukan bukti alasan.
- Batas: ambang dapat berubah menurut skala.

Dalam praktik kelas, Certainty of Response Index dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

Aspek	Temuan	Respons Metodologis
Kekuatan	mudah ditambahkan pada tes yang sudah ada	Pilih ketika sesuai tujuan
Keterbatasan	keyakinan bukan bukti alasan	Tambahkan bukti atau kontrol
Bias potensial	Bahasa, konteks, perangkat, gaya respons	Uji proses respons

6.7.14 Integrasi Digital, Etika, dan Tata Kelola Data

Integrasi digital Certainty of Response Index dapat berupa slider berjangkar, penguncian rating sebelum umpan balik, serta grafik kalibrasi keyakinan dan akurasi. Teknologi seharusnya memperluas kualitas bukti, bukan hanya memindahkan lembar kertas ke layar. (Baker & Inventado, 2014; Romero & Ventura, 2020; AERA, APA, & NCME, 2014; Hasan et al., 1999)

Platform perlu menyimpan versi instrumen, urutan respons, timestamp yang relevan, aturan skor, dan log perubahan. Namun, hanya data yang dibutuhkan untuk tujuan diagnosis yang boleh dikumpulkan; pengumpulan berlebihan meningkatkan risiko privasi tanpa manfaat ilmiah.

Aspek etika utama adalah bahwa hasil CRI tidak boleh digunakan untuk mempermalukan peserta didik yang terlalu yakin; fokusnya adalah dukungan metakognitif. Peserta didik dan guru perlu mengetahui data apa yang dikumpulkan, untuk tujuan apa, siapa yang dapat mengakses, berapa lama disimpan, dan bagaimana kesalahan data diperbaiki.

Algoritma otomatis harus dapat diaudit. Kategori yang dihasilkan sistem perlu ditautkan ke respons mentah dan aturan keputusan. Kasus berisiko atau ambigu harus ditinjau manusia sebelum tindak lanjut penting dilakukan.

- gunakan identitas pseudonim dan akses berbasis peran;
- uji tampilan rumus, subskrip, grafik, dan gambar;
- sediakan opsi aksesibilitas dan alternatif perangkat;
- dokumentasikan algoritma serta perubahan versi.

Dalam praktik kelas, Certainty of Response Index dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.7.15 Sintesis dan Pedoman Implementasi

Certainty of Response Index bermanfaat apabila dipilih berdasarkan tujuan dan digunakan sesuai batasnya. Instrumen ini menghasilkan ketepatan respons dan rating kepastian pada skala ordinal atau numerik, lalu data diolah melalui tabulasi jawaban benar-salah dengan CRI rendah-tinggi serta analisis rata-rata CRI untuk opsi tertentu. (Hasan et al., 1999; Brandriet & Bretz, 2014; Mathabathe & Potgieter, 2014; AERA, APA, & NCME, 2014; Kane, 2013)

Rangkaian kerja yang disarankan dimulai dari analisis konsep, identifikasi konsepsi alternatif, desain tugas, telaah ahli, uji proses respons, uji lapangan, analisis, penetapan kategori, dan penyusunan tindak lanjut. Setiap tahap harus meninggalkan dokumentasi yang dapat diperiksa.

Hasil diagnosis tidak berhenti pada tabel prevalensi. Guru perlu memilih konsepsi prioritas berdasarkan frekuensi, keyakinan, posisi sebagai konsep prasyarat, dampak terhadap topik lanjutan, dan kemungkinan diperbaiki melalui pembelajaran.

Prinsip akhirnya ialah kehati-hatian inferensial: sebutkan apa yang benar-benar ditunjukkan data, bedakan bukti kuat dari indikasi, gunakan triangulasi ketika diperlukan, dan hindari label permanen pada peserta didik.

- selaraskan tujuan, konstruk, tugas, skor, kategori, dan tindakan;
- utamakan konsepsi yang berdampak pada pembelajaran lanjutan;
- gabungkan bukti bila keputusan memiliki konsekuensi besar;
- evaluasi kembali efektivitas tindak lanjut melalui diagnosis ulang.

Dalam praktik kelas, Certainty of Response Index dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.8 Wawancara Klinis

6.8.1 Landasan Konseptual dan Kedudukan Instrumen

Dalam diagnosis miskonsepsi, Wawancara Klinis perlu dipahami bukan sekadar sebagai bentuk soal atau teknik pengumpulan data, melainkan sebagai perangkat inferensi. Dalam bab ini, instrumen tersebut didefinisikan sebagai percakapan terstruktur secara fleksibel yang menelusuri cara peserta didik memahami konsep melalui pertanyaan utama, probing, klarifikasi, dan pengujian konsistensi. Definisi operasional ini menegaskan bentuk bukti yang dikumpulkan sekaligus membatasi jenis kesimpulan yang layak dibuat. (Ginsburg, 1997; Posner & Gertzog, 1982; Chi, 1997)

Klaim utamanya adalah bahwa instrumen ini mengungkap model mental, asumsi implisit, perubahan penalaran, dan makna istilah yang tidak terlihat melalui skor tes tertulis. Klaim tersebut harus dipisahkan dari anggapan bahwa setiap jawaban salah merupakan miskonsepsi. Kesalahan dapat berasal dari kurang pengetahuan, ketidakcermatan, kesulitan bahasa, salah hitung, atau ketidakfamiliaran dengan representasi.

Data pokok yang dihasilkan berupa ucapan, jeda, perubahan jawaban, gambar, gestur, dan respons terhadap probe. Data itu baru memperoleh makna diagnostik setelah dikaitkan dengan peta konsep sasaran, kemungkinan konsepsi alternatif, dan aturan interpretasi yang telah ditetapkan sebelum analisis.

Konsekuensinya, peneliti dan guru perlu membedakan data mentah, skor, kategori, dan klaim konseptual yang dibuat dari hasil tersebut. Posisi terbaik Wawancara Klinis adalah sebagai bagian dari sistem asesmen berlapis yang dapat dipadukan dengan bukti lain ketika konsekuensi keputusan semakin besar.

- menentukan konsep dan proposisi yang hendak didiagnosis;
- menjelaskan perilaku respons yang dianggap sebagai bukti;
- membatasi klaim sesuai kekuatan instrumen;
- menautkan hasil dengan tindakan pembelajaran yang realistis.

Dalam praktik kelas, Wawancara Klinis dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.8.2 Perkembangan Historis dan Rasional Penggunaan

Perkembangan Wawancara Klinis tidak dapat dipisahkan dari perubahan pandangan tentang belajar sains. Ketika peserta didik tidak lagi dipandang sebagai penerima informasi yang kosong, instrumen mulai dirancang untuk mengungkap gagasan awal, pola penalaran, dan keyakinan yang dibawa ke kelas. (Ginsburg, 1997; Posner & Gertzog, 1982; Chi, 1997; Gurel et al., 2015)

Pada tahap awal, penilaian banyak mengandalkan benar-salah atau skor total. Perkembangan penelitian konsepsi peserta didik kemudian mendorong penambahan alasan, keyakinan, tugas kinerja, rekaman proses, dan bukti sosial. Wawancara Klinis merupakan salah satu jawaban metodologis terhadap kebutuhan tersebut.

Dalam pendidikan kimia, kebutuhan diagnosis semakin kuat karena peserta didik harus mengoordinasikan fenomena makroskopik, model submikroskopik, simbol, grafik, dan perhitungan. Instrumen yang hanya menguji hafalan definisi sering gagal memperlihatkan cara representasi itu dihubungkan.

Sejarah penggunaannya juga menunjukkan bahwa format yang lebih kompleks tidak otomatis lebih baik. Setiap tambahan tier, rubrik, atau kanal data harus mempunyai fungsi yang jelas dan didukung bukti bahwa informasi tambahan benar-benar memperbaiki keputusan.

- pergeseran dari pengukuran skor menuju diagnosis pola pemahaman;
- penggunaan respons terbuka dan wawancara untuk menemukan konsepsi alternatif;
- pengembangan teknik kuantitatif agar diagnosis dapat dilakukan pada kelompok besar;
- integrasi data digital dan metakognitif pada perkembangan mutakhir.

Dalam praktik kelas, Wawancara Klinis dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.8.3 Konstruk, Klaim, dan Bukti Diagnosis

Konstruk yang dinilai melalui Wawancara Klinis bukan objek yang terlihat langsung. Yang diamati adalah ucapan, jeda, perubahan jawaban, gambar, gestur, dan respons terhadap probe; dari sana dibuat inferensi tentang pemahaman konsep, miskonsepsi, pengetahuan parsial, atau ketidakpastian. (Gurel et al., 2015; AERA, APA, & NCME, 2014; Messick, 1995; Kane, 2013)

Klaim diagnosis perlu dirumuskan sebagai argumen. Pertama, tugas harus memerlukan konsep sasaran. Kedua, respons harus diberi skor atau kode secara konsisten. Ketiga, pola skor harus mempunyai hubungan yang masuk akal dengan struktur pengetahuan. Keempat, penggunaan hasil harus menghasilkan keputusan yang proporsional.

Sumber ancaman utama adalah ketidakjelasan konstruk. Misalnya, peserta didik dapat gagal karena istilah yang ambigu, beban membaca, representasi yang tidak dikenal, atau prosedur matematika. Kegagalan semacam itu tidak boleh langsung diperlakukan sebagai miskonsepsi kimia.

Oleh sebab itu, setiap pengembang perlu menulis pernyataan klaim-bukti: instrumen mengukur apa, perilaku apa yang menjadi indikator, alternatif penjelasan apa yang mungkin, dan bukti tambahan apa yang diperlukan untuk memperkuat kesimpulan.

- konstruk utama: pemahaman konseptual melalui diagnosis kualitatif mendalam;
- bukti observabel: respons, alasan, tindakan, atau representasi;
- inferensi: profil konsepsi pada konten dan konteks tertentu;
- penggunaan: perencanaan pembelajaran, remediasi, atau penelitian.

Dalam praktik kelas, Wawancara Klinis dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.8.4 Struktur, Komponen, dan Alur Respons

Struktur Wawancara Klinis dibangun dari komponen yang saling bergantung. Komponen inti meliputi panduan wawancara, pertanyaan pemicu, probe netral, tugas atau fenomena, rekaman dan transkripsi, dan kerangka pengodean. Hilangnya satu komponen dapat mengubah makna data yang diperoleh. (Ginsburg, 1997; Posner & Gertzog, 1982; Chi, 1997; Gurel et al., 2015)

Urutan penyajian harus ditetapkan secara sadar. Komponen yang muncul lebih awal dapat memberi petunjuk pada komponen berikutnya, sedangkan permintaan alasan atau keyakinan dapat mengubah cara peserta didik memeriksa jawabannya. Penguncian respons dan uji urutan diperlukan pada format digital.

Setiap komponen membutuhkan spesifikasi. Stimulus harus memuat informasi cukup, opsi atau prompt harus sejajar dengan indikator, skala harus memiliki jangkar, dan rubrik harus menjelaskan perbedaan kategori yang berdekatan.

Struktur yang baik memungkinkan peneliti menelusuri kembali setiap kategori hasil ke respons mentah. Transparansi ini penting agar guru tidak hanya menerima label otomatis tanpa memahami alasan di baliknya.

- panduan wawancara;
- pertanyaan pemicu;
- probe netral;
- tugas atau fenomena;
- rekaman dan transkripsi;
- kerangka pengodean;

Dalam praktik kelas, Wawancara Klinis dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

No.	Komponen	Fungsi Diagnostik
1	panduan wawancara	Bukti yang harus ditelusuri
2	pertanyaan pemicu	Bukti yang harus ditelusuri
3	probe netral	Bukti yang harus ditelusuri
4	tugas atau fenomena	Bukti yang harus ditelusuri

6.8.5 Desain Stimulus, Tugas, atau Pertanyaan

Desain stimulus pada Wawancara Klinis harus memunculkan penggunaan konsep, bukan sekadar pengenalan istilah. Situasi dapat berbentuk fenomena, gambar partikel, grafik, data eksperimen, persamaan, atau masalah kontekstual, tetapi seluruh informasi yang tidak relevan perlu dikendalikan. (Haladyna et al., 2002; Ginsburg, 1997; Posner & Gertzog, 1982; Chi, 1997)

Dalam kimia, satu stimulus sering menggabungkan beberapa tingkat representasi. Pengembang harus memutuskan apakah tujuan butir adalah menerjemahkan representasi, menjelaskan hubungan sebab-akibat, membedakan contoh dan noncontoh, atau menerapkan prinsip pada konteks baru.

Bahasa perlu ringkas tanpa menghilangkan syarat ilmiah. Kata seperti kuat, stabil, jenuh, spontan, atau setimbang mempunyai arti teknis dan sehari-hari yang berbeda. Ketidakjelasan istilah dapat menghasilkan kesalahan yang berasal dari bahasa, bukan struktur konsep.

Desain yang baik juga menghindari petunjuk format, pilihan terlalu panjang, negasi bertumpuk, gambar dekoratif, dan data yang tidak diperlukan. Setiap elemen harus mempunyai alasan diagnostik.

- gunakan konteks yang cukup dikenal tetapi tidak menghafalkan contoh;
- pastikan hanya satu interpretasi ilmiah yang menjadi target;
- uji semua simbol, satuan, grafik, dan gambar;
- sediakan ruang bagi respons yang tidak diprediksi bila tujuan masih eksploratif.

Dalam praktik kelas, Wawancara Klinis dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.8.6 Prosedur Pengembangan Secara Sistematis

Pengembangan Wawancara Klinis harus mengikuti alur yang terdokumentasi. Langkah yang lazim ialah merumuskan fokus, menyusun protokol, menguji pertanyaan, melatih probing netral, kemudian merekam data, mentranskripsi, mengodekan, melakukan pemeriksaan antarpeliteli. Urutan dapat berulang karena temuan uji coba sering memaksa revisi konstruk atau redaksi. (Ginsburg, 1997; Posner & Gertzog, 1982; Chi, 1997; Adams & Wieman, 2011; AERA, APA, & NCME, 2014)

Tahap konseptual dimulai dengan analisis kurikulum, literatur miskonsepsi, dan peta konsep. Tahap empiris awal menggunakan wawancara, respons terbuka, atau observasi untuk memastikan bahwa kategori yang dirancang benar-benar muncul pada populasi sasaran.

Setiap versi instrumen perlu diberi nomor, tanggal, dan catatan perubahan. Revisi harus menyebutkan masalah yang ditemukan, bukti yang digunakan, perubahan yang dilakukan, dan alasan mempertahankan atau menghapus butir.

Pengembangan dianggap belum selesai ketika hanya menghasilkan butir yang tampak baik. Instrumen perlu diuji pada peserta didik yang serupa dengan pengguna sasaran, lalu dianalisis secara konseptual dan empiris sebelum dipakai untuk keputusan penting.

- 1. merumuskan fokus.
- 2. menyusun protokol.
- 3. menguji pertanyaan.
- 4. melatih probing netral.
- 5. merekam data.

Dalam praktik kelas, Wawancara Klinis dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.8.7 Pelaksanaan dan Pengumpulan Data

Pelaksanaan Wawancara Klinis harus distandardisasi secukupnya agar perbedaan hasil lebih mungkin berasal dari pemahaman, bukan variasi instruksi, waktu, alat, atau bantuan. Petunjuk perlu menjelaskan tujuan umum tanpa mengajarkan jawaban. (Ginsburg, 1997; Posner & Gertzog, 1982; AERA, APA, & NCME, 2014)

Kondisi administrasi mencakup waktu, urutan tugas, kesempatan bertanya, penggunaan kalkulator, akses tabel periodik, kualitas gambar, dan perangkat digital. Pada instrumen kualitatif, standardisasi berarti konsistensi fokus dan prinsip probing, bukan membaca naskah secara kaku.

Data yang dikumpulkan ialah ucapan, jeda, perubahan jawaban, gambar, gestur, dan respons terhadap probe. Data mentah harus disimpan bersama metadata yang diperlukan, seperti kode butir, versi instrumen, waktu administrasi, dan kondisi khusus, tetapi hanya data yang relevan yang boleh dikumpulkan.

Sebelum administrasi luas, lakukan uji teknis. Pastikan tautan, urutan, penguncian respons, rekaman, penyimpanan, serta tampilan rumus dan subskrip bekerja. Gangguan teknis perlu dicatat sebagai bagian dari interpretasi.

- berikan instruksi yang sama dan mudah dipahami;
- catat kondisi yang dapat memengaruhi respons;
- hindari umpan balik sebelum seluruh respons diagnostik selesai;
- sediakan prosedur penanganan data tidak lengkap.

Dalam praktik kelas, Wawancara Klinis dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.8.8 Penskoran, Pengodean, dan Pengolahan Data

Penskoran Wawancara Klinis dapat diringkas sebagai pengodean tematik atau proposisional, rekonstruksi alur penalaran, dan penentuan profil konsepsi dengan bukti kutipan. Aturan ini harus ditulis sebelum melihat hasil utama agar keputusan tidak disesuaikan secara oportunistik dengan pola data. (Ginsburg, 1997; Posner & Gertzog, 1982; Chi, 1997; AERA, APA, & NCME, 2014)

Skor total jarang cukup untuk diagnosis. Analisis perlu mempertahankan informasi tingkat butir, konsep, opsi, alasan, keyakinan, proses, atau representasi. Penggabungan data terlalu dini dapat menyembunyikan pola yang secara pedagogis penting.

Untuk respons terbuka atau kualitatif, buku kode harus mencantumkan definisi, contoh positif, contoh batas, noncontoh, dan aturan ketika satu respons memuat beberapa gagasan. Sebagian data perlu dinilai oleh lebih dari satu penilai.

Setiap transformasi data—misalnya dari respons ke kode, dari kode ke kategori, dan dari kategori ke persentase—harus dapat diaudit. Simpan respons asli dan keputusan pengodean agar kasus ambigu dapat ditinjau kembali.

- pisahkan skor ketepatan dari skor keyakinan atau kualitas alasan;
- laporkan kasus tidak lengkap dan aturan penanganannya;
- uji konsistensi penilai untuk data berkode;
- hindari mengubah data ordinal menjadi ukuran interval tanpa alasan.

Dalam praktik kelas, Wawancara Klinis dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.8.9 Interpretasi dan Penetapan Kategori

Interpretasi hasil Wawancara Klinis harus bergerak dari respons menuju pola, bukan langsung dari satu jawaban menuju label. Status konseptual menjadi lebih kuat apabila gagasan yang sama muncul pada beberapa butir, representasi, atau konteks. (Ginsburg, 1997; Posner & Gertzog, 1982; Chi, 1997; Gurel et al., 2015; Kane, 2013)

Kategori dapat mencakup pemahaman ilmiah, pemahaman parsial, kurang pengetahuan, respons tidak konsisten, false positive, false negative, dan miskonsepsi. Tidak semua instrumen mampu membedakan seluruh kategori tersebut; skema harus mengikuti informasi yang benar-benar tersedia.

Ambang, seperti batas keyakinan atau persentase pemilihan, perlu memiliki rasional dan diuji melalui analisis sensitivitas. Perubahan kecil pada ambang dapat mengubah prevalensi kategori secara berarti.

Interpretasi yang bertanggung jawab menyatakan bahwa kategori berlaku pada konsep, bentuk tugas, waktu, dan kondisi tertentu. Hasil merupakan hipotesis kerja untuk pembelajaran, bukan diagnosis psikologis permanen.

- gunakan lebih dari satu butir untuk konsepsi penting;
- periksa konsistensi antara jawaban dan alasan;
- tandai kasus ambigu untuk konfirmasi;
- hubungkan kategori dengan rekomendasi yang berbeda.

Dalam praktik kelas, Wawancara Klinis dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.8.10 Validitas dan Kualitas Inferensi

Validitas Wawancara Klinis berkaitan dengan kualitas interpretasi dan penggunaan hasil, bukan sifat tetap yang dimiliki instrumen. Fokus khususnya adalah kredibilitas interpretasi, transparansi jejak analisis, triangulasi, dan bukti bahwa pertanyaan tidak menuntun jawaban. (AERA, APA, & NCME, 2014; Messick, 1995; Kane, 2013; Haynes et al., 1995)

Bukti berdasarkan isi menilai keterwakilan domain dan akurasi konsep. Bukti proses respons memeriksa bagaimana peserta didik memahami tugas. Bukti struktur internal menilai kesesuaian pola respons. Bukti hubungan dengan variabel lain dan konsekuensi penggunaan melengkapi argumen.

Ancaman validitas dapat berasal dari konten yang kurang terwakili, tuntutan bahasa berlebihan, petunjuk opsi, rubrik yang tidak stabil, keyakinan yang tidak terkalibrasi, atau algoritma yang belajar dari label keliru.

Dokumen validasi perlu menghubungkan setiap klaim dengan bukti dan alternatif penjelasan. Menampilkan satu indeks tidak cukup; validasi adalah program argumentatif yang diperbarui setiap kali instrumen dipakai pada populasi atau tujuan baru.

- bukti isi dan kesesuaian kurikulum;
- bukti proses respons melalui wawancara kognitif;
- bukti struktur dan fungsi butir;
- bukti kegunaan serta konsekuensi keputusan.

Dalam praktik kelas, Wawancara Klinis dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.8.11 Reliabilitas, Keajegan, dan Analisis Empiris

Keajegan Wawancara Klinis harus dinilai sesuai jenis data. Pada respons objektif dapat digunakan konsistensi internal, stabilitas waktu, atau model pengukuran. Pada data kualitatif, perhatian diarahkan pada konsistensi pengodean, audit trail, dan ketepatan penerapan definisi kode. (Kuder & Richardson, 1937; Cronbach, 1951; Ding et al., 2006; Wren & Barbera, 2014; AERA, APA, & NCME, 2014)

Reliabilitas tinggi tidak otomatis menunjukkan diagnosis yang benar. Butir yang berulang-ulang mengukur hafalan dapat menghasilkan koefisien tinggi tetapi tidak mewakili pemahaman konseptual. Sebaliknya, konstruk yang multidimensi tidak selalu layak dipaksa menjadi satu skala.

Analisis empiris perlu memeriksa kesukaran, daya beda, fungsi opsi, pola data hilang, dan keterhubungan komponen. Untuk instrumen bertingkat, reliabilitas dapat dihitung pada setiap tier serta kombinasi yang mempunyai makna teoritis.

Ukuran sampel, distribusi kemampuan, dan homogenitas konten memengaruhi statistik. Karena itu, laporan harus menyatakan kondisi estimasi dan tidak menyalin batas angka sebagai aturan universal.

- pilih statistik yang sesuai dengan jenis skor;
- laporkan interval atau ketidakpastian bila tersedia;
- periksa butir secara substantif dan statistik;
- uji kembali setelah revisi atau perubahan platform.

Dalam praktik kelas, Wawancara Klinis dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.8.12 Contoh Penerapan dalam Pembelajaran Kimia

Penerapan Wawancara Klinis dalam kimia dapat dimulai dari konsep makna kesetimbangan dinamis. Pengembang perlu menuliskan konsep ilmiah target, miskonsepsi yang mungkin, bukti respons yang diharapkan, dan keputusan pembelajaran setelah hasil diperoleh. (Ginsburg, 1997; Posner & Gertzog, 1982; Chi, 1997; Gurel et al., 2015; Nakhleh, 1992)

Contoh kedua adalah bayangan partikel saat pelarutan. Stimulus sebaiknya menghubungkan fenomena dengan model partikel atau persamaan, lalu meminta peserta didik menunjukkan hubungan sebab-akibat. Jawaban yang hanya menyebut istilah belum cukup menjadi bukti pemahaman.

Pada penalaran energi ikatan, diagnosis dapat memeriksa apakah peserta didik membedakan perubahan laju, keadaan setimbang, energi, dan komposisi. Kesalahan sering muncul ketika satu aturan dipindahkan ke konteks yang tidak sesuai.

Kasus konsepsi orbital dan lintasan elektron menunjukkan pentingnya representasi. Respons verbal, gambar, simbol, dan hitungan dapat tidak konsisten. Hasil instrumen sebaiknya digunakan untuk memilih demonstrasi, diskusi, contoh-noncontoh, atau tugas remediasi yang secara langsung menantang hubungan salah.

- tentukan proposisi ilmiah dan konsepsi alternatif;
- gunakan minimal dua bentuk representasi;
- minta alasan atau bukti ketika format memungkinkan;
- rancang tindak lanjut berdasarkan pola, bukan hanya skor.

Dalam praktik kelas, Wawancara Klinis dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.8.13 Kelebihan, Keterbatasan, dan Potensi Bias

Kekuatan utama Wawancara Klinis meliputi kedalaman data tinggi, dapat mengejar respons tak terduga, membedakan bahasa sederhana dari konsep salah, dan berguna untuk mengembangkan instrumen tertulis. Kekuatan tersebut menjelaskan situasi ketika instrumen layak dipilih dibanding teknik lain. (Ginsburg, 1997; Posner & Gertzog, 1982; Chi, 1997; Gurel et al., 2015; AERA, APA, & NCME, 2014)

Keterbatasannya mencakup memerlukan pewawancara terlatih, sampel biasanya kecil, probe dapat memengaruhi jawaban, dan analisis membutuhkan waktu dan audit trail. Keterbatasan harus diperlakukan sebagai batas inferensi, bukan sekadar paragraf formal yang tidak memengaruhi analisis.

Potensi bias dapat berasal dari bahasa, pengalaman konteks, akses teknologi, kecemasan, gaya menggunakan skala, keterampilan menggambar, atau kemampuan menulis. Instrumen perlu diuji pada kelompok yang beragam dan ditinjau untuk fungsi berbeda antarkelompok bila data mencukupi.

Keputusan memilih instrumen sebaiknya mempertimbangkan kedalaman data, ukuran kelompok, waktu, keterampilan analisis, tujuan diagnosis, dan konsekuensi kesalahan klasifikasi. Tidak ada satu metode yang unggul untuk semua tujuan.

- Kekuatan: kedalaman data tinggi.
- Kekuatan: dapat mengejar respons tak terduga.
- Batas: memerlukan pewawancara terlatih.
- Batas: sampel biasanya kecil.

Dalam praktik kelas, Wawancara Klinis dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

Aspek	Temuan	Respons Metodologis
Kekuatan	kedalaman data tinggi	Pilih ketika sesuai tujuan
Keterbatasan	memerlukan pewawancara terlatih	Tambahkan bukti atau kontrol
Bias potensial	Bahasa, konteks, perangkat, gaya respons	Uji proses respons

6.8.14 Integrasi Digital, Etika, dan Tata Kelola Data

Integrasi digital Wawancara Klinis dapat berupa wawancara video, papan tulis digital, penanda waktu, dan anotasi multimodal dengan persetujuan peserta. Teknologi seharusnya memperluas kualitas bukti, bukan hanya memindahkan lembar kertas ke layar. (Baker & Inventado, 2014; Romero & Ventura, 2020; AERA, APA, & NCME, 2014; Ginsburg, 1997)

Platform perlu menyimpan versi instrumen, urutan respons, timestamp yang relevan, aturan skor, dan log perubahan. Namun, hanya data yang dibutuhkan untuk tujuan diagnosis yang boleh dikumpulkan; pengumpulan berlebihan meningkatkan risiko privasi tanpa manfaat ilmiah.

Aspek etika utama adalah bahwa persetujuan, privasi rekaman, hak berhenti, dan perlindungan agar peserta didik tidak merasa diuji secara mengancam. Peserta didik dan guru perlu mengetahui data apa yang dikumpulkan, untuk tujuan apa, siapa yang dapat mengakses, berapa lama disimpan, dan bagaimana kesalahan data diperbaiki.

Algoritma otomatis harus dapat diaudit. Kategori yang dihasilkan sistem perlu ditautkan ke respons mentah dan aturan keputusan. Kasus berisiko atau ambigu harus ditinjau manusia sebelum tindak lanjut penting dilakukan.

- gunakan identitas pseudonim dan akses berbasis peran;
- uji tampilan rumus, subskrip, grafik, dan gambar;
- sediakan opsi aksesibilitas dan alternatif perangkat;
- dokumentasikan algoritma serta perubahan versi.

Dalam praktik kelas, Wawancara Klinis dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.8.15 Sintesis dan Pedoman Implementasi

Wawancara Klinis bermanfaat apabila dipilih berdasarkan tujuan dan digunakan sesuai batasnya. Instrumen ini menghasilkan ucapan, jeda, perubahan jawaban, gambar, gestur, dan respons terhadap probe, lalu data diolah melalui pengodean tematik atau proposisional, rekonstruksi alur penalaran, dan penentuan profil konsepsi dengan bukti kutipan. (Ginsburg, 1997; Posner & Gertzog, 1982; Chi, 1997; AERA, APA, & NCME, 2014; Kane, 2013)

Rangkaian kerja yang disarankan dimulai dari analisis konsep, identifikasi konsepsi alternatif, desain tugas, telaah ahli, uji proses respons, uji lapangan, analisis, penetapan kategori, dan penyusunan tindak lanjut. Setiap tahap harus meninggalkan dokumentasi yang dapat diperiksa.

Hasil diagnosis tidak berhenti pada tabel prevalensi. Guru perlu memilih konsepsi prioritas berdasarkan frekuensi, keyakinan, posisi sebagai konsep prasyarat, dampak terhadap topik lanjutan, dan kemungkinan diperbaiki melalui pembelajaran.

Prinsip akhirnya ialah kehati-hatian inferensial: sebutkan apa yang benar-benar ditunjukkan data, bedakan bukti kuat dari indikasi, gunakan triangulasi ketika diperlukan, dan hindari label permanen pada peserta didik.

- selaraskan tujuan, konstruk, tugas, skor, kategori, dan tindakan;
- utamakan konsepsi yang berdampak pada pembelajaran lanjutan;
- gabungkan bukti bila keputusan memiliki konsekuensi besar;
- evaluasi kembali efektivitas tindak lanjut melalui diagnosis ulang.

Dalam praktik kelas, Wawancara Klinis dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.9 Wawancara Berbasis Tugas

6.9.1 Landasan Konseptual dan Kedudukan Instrumen

Wawancara Berbasis Tugas menempati posisi tertentu dalam rangkaian bukti diagnosis, dari pengungkapan respons awal sampai penentuan tindak lanjut pembelajaran. Dalam bab ini, instrumen tersebut didefinisikan sebagai wawancara yang berpusat pada penyelesaian tugas konkret, eksperimen mini, gambar, grafik, atau masalah sehingga penalaran muncul saat peserta didik bertindak. Definisi operasional ini menegaskan bentuk bukti yang dikumpulkan sekaligus membatasi jenis kesimpulan yang layak dibuat. (Posner & Gertzog, 1982; Ericsson & Simon, 1993; van Someren et al., 1994)

Klaim utamanya adalah bahwa instrumen ini menghubungkan apa yang dikatakan peserta didik dengan keputusan yang dibuat selama mengerjakan tugas, sehingga diagnosis lebih kontekstual. Klaim tersebut harus dipisahkan dari anggapan bahwa setiap jawaban salah merupakan miskonsepsi. Kesalahan dapat berasal dari kurang pengetahuan, ketidakcermatan, kesulitan bahasa, salah hitung, atau ketidakfamiliaran dengan representasi.

Data pokok yang dihasilkan berupa langkah kerja, pilihan representasi, ucapan, revisi, kesalahan, dan produk tugas. Data itu baru memperoleh makna diagnostik setelah dikaitkan dengan peta konsep sasaran, kemungkinan konsepsi alternatif, dan aturan interpretasi yang telah ditetapkan sebelum analisis.

Pembedaan ini penting karena satu pola respons dapat memiliki lebih dari satu penjelasan yang masuk akal. Posisi terbaik Wawancara Berbasis Tugas adalah sebagai bagian dari sistem asesmen berlapis yang dapat dipadukan dengan bukti lain ketika konsekuensi keputusan semakin besar.

- menentukan konsep dan proposisi yang hendak didiagnosis;
- menjelaskan perilaku respons yang dianggap sebagai bukti;
- membatasi klaim sesuai kekuatan instrumen;
- menautkan hasil dengan tindakan pembelajaran yang realistis.

Dalam praktik kelas, Wawancara Berbasis Tugas dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.9.2 Perkembangan Historis dan Rasional Penggunaan

Perkembangan Wawancara Berbasis Tugas tidak dapat dipisahkan dari perubahan pandangan tentang belajar sains. Ketika peserta didik tidak lagi dipandang sebagai penerima informasi yang kosong, instrumen mulai dirancang untuk mengungkap gagasan awal, pola penalaran, dan keyakinan yang dibawa ke kelas. (Posner & Gertzog, 1982; Ericsson & Simon, 1993; van Someren et al., 1994; Chi, 1997)

Pada tahap awal, penilaian banyak mengandalkan benar-salah atau skor total. Perkembangan penelitian konsepsi peserta didik kemudian mendorong penambahan alasan, keyakinan, tugas kinerja, rekaman proses, dan bukti sosial. Wawancara Berbasis Tugas merupakan salah satu jawaban metodologis terhadap kebutuhan tersebut.

Dalam pendidikan kimia, kebutuhan diagnosis semakin kuat karena peserta didik harus mengoordinasikan fenomena makroskopik, model submikroskopik, simbol, grafik, dan perhitungan. Instrumen yang hanya menguji hafalan definisi sering gagal memperlihatkan cara representasi itu dihubungkan.

Sejarah penggunaannya juga menunjukkan bahwa format yang lebih kompleks tidak otomatis lebih baik. Setiap tambahan tier, rubrik, atau kanal data harus mempunyai fungsi yang jelas dan didukung bukti bahwa informasi tambahan benar-benar memperbaiki keputusan.

- pergeseran dari pengukuran skor menuju diagnosis pola pemahaman;
- penggunaan respons terbuka dan wawancara untuk menemukan konsepsi alternatif;
- pengembangan teknik kuantitatif agar diagnosis dapat dilakukan pada kelompok besar;
- integrasi data digital dan metakognitif pada perkembangan mutakhir.

Dalam praktik kelas, Wawancara Berbasis Tugas dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.9.3 Konstruk, Klaim, dan Bukti Diagnosis

Konstruk yang dinilai melalui Wawancara Berbasis Tugas bukan objek yang terlihat langsung. Yang diamati adalah langkah kerja, pilihan representasi, ucapan, revisi, kesalahan, dan produk tugas; dari sana dibuat inferensi tentang pemahaman konsep, miskonsepsi, pengetahuan parsial, atau ketidakpastian. (Chi, 1997; White & Gunstone, 1992; Messick, 1995; Kane, 2013)

Klaim diagnosis perlu dirumuskan sebagai argumen. Pertama, tugas harus memerlukan konsep sasaran. Kedua, respons harus diberi skor atau kode secara konsisten. Ketiga, pola skor harus mempunyai hubungan yang masuk akal dengan struktur pengetahuan. Keempat, penggunaan hasil harus menghasilkan keputusan yang proporsional.

Sumber ancaman utama adalah ketidakjelasan konstruk. Misalnya, peserta didik dapat gagal karena istilah yang ambigu, beban membaca, representasi yang tidak dikenal, atau prosedur matematika. Kegagalan semacam itu tidak boleh langsung diperlakukan sebagai miskonsepsi kimia.

Oleh sebab itu, setiap pengembang perlu menulis pernyataan klaim-bukti: instrumen mengukur apa, perilaku apa yang menjadi indikator, alternatif penjelasan apa yang mungkin, dan bukti tambahan apa yang diperlukan untuk memperkuat kesimpulan.

- konstruk utama: pemahaman konseptual melalui diagnosis kualitatif berorientasi kinerja;
- bukti observabel: respons, alasan, tindakan, atau representasi;
- inferensi: profil konsepsi pada konten dan konteks tertentu;
- penggunaan: perencanaan pembelajaran, remediasi, atau penelitian.

Dalam praktik kelas, Wawancara Berbasis Tugas dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.9.4 Struktur, Komponen, dan Alur Respons

Struktur Wawancara Berbasis Tugas dibangun dari komponen yang saling bergantung. Komponen inti meliputi tugas pemicu, artefak atau bahan, urutan prompt, pertanyaan prediksi dan alasan, rekaman tindakan, dan rubrik proses. Hilangnya satu komponen dapat mengubah makna data yang diperoleh. (Posner & Gertzog, 1982; Ericsson & Simon, 1993; van Someren et al., 1994; Chi, 1997)

Urutan penyajian harus ditetapkan secara sadar. Komponen yang muncul lebih awal dapat memberi petunjuk pada komponen berikutnya, sedangkan permintaan alasan atau keyakinan dapat mengubah cara peserta didik memeriksa jawabannya. Penguncian respons dan uji urutan diperlukan pada format digital.

Setiap komponen membutuhkan spesifikasi. Stimulus harus memuat informasi cukup, opsi atau prompt harus sejajar dengan indikator, skala harus memiliki jangkar, dan rubrik harus menjelaskan perbedaan kategori yang berdekatan.

Struktur yang baik memungkinkan peneliti menelusuri kembali setiap kategori hasil ke respons mentah. Transparansi ini penting agar guru tidak hanya menerima label otomatis tanpa memahami alasan di baliknya.

- tugas pemicu;
- artefak atau bahan;
- urutan prompt;
- pertanyaan prediksi dan alasan;
- rekaman tindakan;
- rubrik proses;

Dalam praktik kelas, Wawancara Berbasis Tugas dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

No.	Komponen	Fungsi Diagnostik
1	tugas pemicu	Bukti yang harus ditelusuri
2	artefak atau bahan	Bukti yang harus ditelusuri
3	urutan prompt	Bukti yang harus ditelusuri
4	pertanyaan prediksi dan alasan	Bukti yang harus ditelusuri

6.9.5 Desain Stimulus, Tugas, atau Pertanyaan

Desain stimulus pada Wawancara Berbasis Tugas harus memunculkan penggunaan konsep, bukan sekadar pengenalan istilah. Situasi dapat berbentuk fenomena, gambar partikel, grafik, data eksperimen, persamaan, atau masalah kontekstual, tetapi seluruh informasi yang tidak relevan perlu dikendalikan. (Haladyna et al., 2002; Posner & Gertzog, 1982; Ericsson & Simon, 1993; van Someren et al., 1994)

Dalam kimia, satu stimulus sering menggabungkan beberapa tingkat representasi. Pengembang harus memutuskan apakah tujuan butir adalah menerjemahkan representasi, menjelaskan hubungan sebab-akibat, membedakan contoh dan noncontoh, atau menerapkan prinsip pada konteks baru.

Bahasa perlu ringkas tanpa menghilangkan syarat ilmiah. Kata seperti kuat, stabil, jenuh, spontan, atau setimbang mempunyai arti teknis dan sehari-hari yang berbeda. Ketidakjelasan istilah dapat menghasilkan kesalahan yang berasal dari bahasa, bukan struktur konsep.

Desain yang baik juga menghindari petunjuk format, pilihan terlalu panjang, negasi bertumpuk, gambar dekoratif, dan data yang tidak diperlukan. Setiap elemen harus mempunyai alasan diagnostik.

- gunakan konteks yang cukup dikenal tetapi tidak menghafalkan contoh;
- pastikan hanya satu interpretasi ilmiah yang menjadi target;
- uji semua simbol, satuan, grafik, dan gambar;
- sediakan ruang bagi respons yang tidak diprediksi bila tujuan masih eksploratif.

Dalam praktik kelas, Wawancara Berbasis Tugas dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.9.6 Prosedur Pengembangan Secara Sistematis

Pengembangan Wawancara Berbasis Tugas harus mengikuti alur yang terdokumentasi. Langkah yang lazim ialah memilih tugas diagnostik, memetakan keputusan kunci, menyiapkan probe, menguji alat, kemudian menetapkan urutan, merekam proses, mengodekan episode, membandingkan lintaspartisipan. Urutan dapat berulang karena temuan uji coba sering memaksa revisi konstruk atau redaksi. (Posner & Gertzog, 1982; Ericsson & Simon, 1993; van Someren et al., 1994; Adams & Wieman, 2011; AERA, APA, & NCME, 2014)

Tahap konseptual dimulai dengan analisis kurikulum, literatur miskonsepsi, dan peta konsep. Tahap empiris awal menggunakan wawancara, respons terbuka, atau observasi untuk memastikan bahwa kategori yang dirancang benar-benar muncul pada populasi sasaran.

Setiap versi instrumen perlu diberi nomor, tanggal, dan catatan perubahan. Revisi harus menyebutkan masalah yang ditemukan, bukti yang digunakan, perubahan yang dilakukan, dan alasan mempertahankan atau menghapus butir.

Pengembangan dianggap belum selesai ketika hanya menghasilkan butir yang tampak baik. Instrumen perlu diuji pada peserta didik yang serupa dengan pengguna sasaran, lalu dianalisis secara konseptual dan empiris sebelum dipakai untuk keputusan penting.

- 1. memilih tugas diagnostik.
- 2. memetakan keputusan kunci.
- 3. menyiapkan probe.
- 4. menguji alat.
- 5. menetapkan urutan.

Dalam praktik kelas, Wawancara Berbasis Tugas dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.9.7 Pelaksanaan dan Pengumpulan Data

Pelaksanaan Wawancara Berbasis Tugas harus distandardisasi secukupnya agar perbedaan hasil lebih mungkin berasal dari pemahaman, bukan variasi instruksi, waktu, alat, atau bantuan. Petunjuk perlu menjelaskan tujuan umum tanpa mengajarkan jawaban. (Posner & Gertzog, 1982; Ericsson & Simon, 1993; AERA, APA, & NCME, 2014)

Kondisi administrasi mencakup waktu, urutan tugas, kesempatan bertanya, penggunaan kalkulator, akses tabel periodik, kualitas gambar, dan perangkat digital. Pada instrumen kualitatif, standardisasi berarti konsistensi fokus dan prinsip probing, bukan membaca naskah secara kaku.

Data yang dikumpulkan ialah langkah kerja, pilihan representasi, ucapan, revisi, kesalahan, dan produk tugas. Data mentah harus disimpan bersama metadata yang diperlukan, seperti kode butir, versi instrumen, waktu administrasi, dan kondisi khusus, tetapi hanya data yang relevan yang boleh dikumpulkan.

Sebelum administrasi luas, lakukan uji teknis. Pastikan tautan, urutan, penguncian respons, rekaman, penyimpanan, serta tampilan rumus dan subskrip bekerja. Gangguan teknis perlu dicatat sebagai bagian dari interpretasi.

- berikan instruksi yang sama dan mudah dipahami;
- catat kondisi yang dapat memengaruhi respons;
- hindari umpan balik sebelum seluruh respons diagnostik selesai;
- sediakan prosedur penanganan data tidak lengkap.

Dalam praktik kelas, Wawancara Berbasis Tugas dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.9.8 Penskoran, Pengodean, dan Pengolahan Data

Penskoran Wawancara Berbasis Tugas dapat diringkas sebagai analisis episode, strategi, titik keputusan, dan kesesuaian tindakan dengan penjelasan verbal. Aturan ini harus ditulis sebelum melihat hasil utama agar keputusan tidak disesuaikan secara oportunistik dengan pola data. (Posner & Gertzog, 1982; Ericsson & Simon, 1993; van Someren et al., 1994; Chi, 1997; AERA, APA, & NCME, 2014)

Skor total jarang cukup untuk diagnosis. Analisis perlu mempertahankan informasi tingkat butir, konsep, opsi, alasan, keyakinan, proses, atau representasi. Penggabungan data terlalu dini dapat menyembunyikan pola yang secara pedagogis penting.

Untuk respons terbuka atau kualitatif, buku kode harus mencantumkan definisi, contoh positif, contoh batas, noncontoh, dan aturan ketika satu respons memuat beberapa gagasan. Sebagian data perlu dinilai oleh lebih dari satu penilai.

Setiap transformasi data—misalnya dari respons ke kode, dari kode ke kategori, dan dari kategori ke persentase—harus dapat diaudit. Simpan respons asli dan keputusan pengodean agar kasus ambigu dapat ditinjau kembali.

- pisahkan skor ketepatan dari skor keyakinan atau kualitas alasan;
- laporkan kasus tidak lengkap dan aturan penanganannya;
- uji konsistensi penilai untuk data berkode;
- hindari mengubah data ordinal menjadi ukuran interval tanpa alasan.

Dalam praktik kelas, Wawancara Berbasis Tugas dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.9.9 Interpretasi dan Penetapan Kategori

Interpretasi hasil Wawancara Berbasis Tugas harus bergerak dari respons menuju pola, bukan langsung dari satu jawaban menuju label. Status konseptual menjadi lebih kuat apabila gagasan yang sama muncul pada beberapa butir, representasi, atau konteks. (Posner & Gertzog, 1982; Ericsson & Simon, 1993; van Someren et al., 1994; Chi, 1997; Kane, 2013)

Kategori dapat mencakup pemahaman ilmiah, pemahaman parsial, kurang pengetahuan, respons tidak konsisten, false positive, false negative, dan miskonsepsi. Tidak semua instrumen mampu membedakan seluruh kategori tersebut; skema harus mengikuti informasi yang benar-benar tersedia.

Ambang, seperti batas keyakinan atau persentase pemilihan, perlu memiliki rasional dan diuji melalui analisis sensitivitas. Perubahan kecil pada ambang dapat mengubah prevalensi kategori secara berarti.

Interpretasi yang bertanggung jawab menyatakan bahwa kategori berlaku pada konsep, bentuk tugas, waktu, dan kondisi tertentu. Hasil merupakan hipotesis kerja untuk pembelajaran, bukan diagnosis psikologis permanen.

- gunakan lebih dari satu butir untuk konsepsi penting;
- periksa konsistensi antara jawaban dan alasan;
- tandai kasus ambigu untuk konfirmasi;
- hubungkan kategori dengan rekomendasi yang berbeda.

Dalam praktik kelas, Wawancara Berbasis Tugas dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.9.10 Validitas dan Kualitas Inferensi

Validitas Wawancara Berbasis Tugas berkaitan dengan kualitas interpretasi dan penggunaan hasil, bukan sifat tetap yang dimiliki instrumen. Fokus khususnya adalah kesesuaian tugas dengan konsep sasaran dan bukti bahwa tindakan yang diamati benar-benar menuntut penggunaan konsep tersebut. (AERA, APA, & NCME, 2014; Messick, 1995; Kane, 2013; Haynes et al., 1995)

Bukti berdasarkan isi menilai keterwakilan domain dan akurasi konsep. Bukti proses respons memeriksa bagaimana peserta didik memahami tugas. Bukti struktur internal menilai kesesuaian pola respons. Bukti hubungan dengan variabel lain dan konsekuensi penggunaan melengkapi argumen.

Ancaman validitas dapat berasal dari konten yang kurang terwakili, tuntutan bahasa berlebihan, petunjuk opsi, rubrik yang tidak stabil, keyakinan yang tidak terkalibrasi, atau algoritma yang belajar dari label keliru.

Dokumen validasi perlu menghubungkan setiap klaim dengan bukti dan alternatif penjelasan. Menampilkan satu indeks tidak cukup; validasi adalah program argumentatif yang diperbarui setiap kali instrumen dipakai pada populasi atau tujuan baru.

- bukti isi dan kesesuaian kurikulum;
- bukti proses respons melalui wawancara kognitif;
- bukti struktur dan fungsi butir;
- bukti kegunaan serta konsekuensi keputusan.

Dalam praktik kelas, Wawancara Berbasis Tugas dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.9.11 Reliabilitas, Keajegan, dan Analisis Empiris

Keajegan Wawancara Berbasis Tugas harus dinilai sesuai jenis data. Pada respons objektif dapat digunakan konsistensi internal, stabilitas waktu, atau model pengukuran. Pada data kualitatif, perhatian diarahkan pada konsistensi pengodean, audit trail, dan ketepatan penerapan definisi kode. (Kuder & Richardson, 1937; Cronbach, 1951; Ding et al., 2006; Wren & Barbera, 2014; AERA, APA, & NCME, 2014)

Reliabilitas tinggi tidak otomatis menunjukkan diagnosis yang benar. Butir yang berulang-ulang mengukur hafalan dapat menghasilkan koefisien tinggi tetapi tidak mewakili pemahaman konseptual. Sebaliknya, konstruk yang multidimensi tidak selalu layak dipaksa menjadi satu skala.

Analisis empiris perlu memeriksa kesukaran, daya beda, fungsi opsi, pola data hilang, dan keterhubungan komponen. Untuk instrumen bertingkat, reliabilitas dapat dihitung pada setiap tier serta kombinasi yang mempunyai makna teoritis.

Ukuran sampel, distribusi kemampuan, dan homogenitas konten memengaruhi statistik. Karena itu, laporan harus menyatakan kondisi estimasi dan tidak menyalin batas angka sebagai aturan universal.

- pilih statistik yang sesuai dengan jenis skor;
- laporkan interval atau ketidakpastian bila tersedia;
- periksa butir secara substantif dan statistik;
- uji kembali setelah revisi atau perubahan platform.

Dalam praktik kelas, Wawancara Berbasis Tugas dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.9.12 Contoh Penerapan dalam Pembelajaran Kimia

Penerapan Wawancara Berbasis Tugas dalam kimia dapat dimulai dari konsep mengurutkan gambar partikel saat perubahan fase. Pengembang perlu menuliskan konsep ilmiah target, miskonsepsi yang mungkin, bukti respons yang diharapkan, dan keputusan pembelajaran setelah hasil diperoleh. (Posner & Gertzog, 1982; Ericsson & Simon, 1993; van Someren et al., 1994; Chi, 1997; Nakhleh, 1992)

Contoh kedua adalah menafsirkan kurva titrasi. Stimulus sebaiknya menghubungkan fenomena dengan model partikel atau persamaan, lalu meminta peserta didik menunjukkan hubungan sebab-akibat. Jawaban yang hanya menyebut istilah belum cukup menjadi bukti pemahaman.

Pada merancang percobaan laju reaksi, diagnosis dapat memeriksa apakah peserta didik membedakan perubahan laju, keadaan setimbang, energi, dan komposisi. Kesalahan sering muncul ketika satu aturan dipindahkan ke konteks yang tidak sesuai.

Kasus menyusun struktur Lewis dan menjelaskan muatan formal menunjukkan pentingnya representasi. Respons verbal, gambar, simbol, dan hitungan dapat tidak konsisten. Hasil instrumen sebaiknya digunakan untuk memilih demonstrasi, diskusi, contoh-noncontoh, atau tugas remediasi yang secara langsung menantang hubungan salah.

- tentukan proposisi ilmiah dan konsepsi alternatif;
- gunakan minimal dua bentuk representasi;
- minta alasan atau bukti ketika format memungkinkan;
- rancang tindak lanjut berdasarkan pola, bukan hanya skor.

Dalam praktik kelas, Wawancara Berbasis Tugas dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.9.13 Kelebihan, Keterbatasan, dan Potensi Bias

Kekuatan utama Wawancara Berbasis Tugas meliputi mengurangi ketergantungan pada pertanyaan definisi, mengungkap pengetahuan dalam tindakan, cocok untuk konsep representasional, dan memungkinkan triangulasi ucapan dan produk. Kekuatan tersebut menjelaskan situasi ketika instrumen layak dipilih dibanding teknik lain. (Posner & Gertzog, 1982; Ericsson & Simon, 1993; van Someren et al., 1994; Gurel et al., 2015; AERA, APA, & NCME, 2014)

Keterbatasannya mencakup kinerja dipengaruhi tuntutan tugas, alat dapat memberi petunjuk, perbandingan antarpartisipan sulit bila probing berbeda, dan analisis multimodal kompleks. Keterbatasan harus diperlakukan sebagai batas inferensi, bukan sekadar paragraf formal yang tidak memengaruhi analisis.

Potensi bias dapat berasal dari bahasa, pengalaman konteks, akses teknologi, kecemasan, gaya menggunakan skala, keterampilan menggambar, atau kemampuan menulis. Instrumen perlu diuji pada kelompok yang beragam dan ditinjau untuk fungsi berbeda antarkelompok bila data mencukupi.

Keputusan memilih instrumen sebaiknya mempertimbangkan kedalaman data, ukuran kelompok, waktu, keterampilan analis, tujuan diagnosis, dan konsekuensi kesalahan klasifikasi. Tidak ada satu metode yang unggul untuk semua tujuan.

- Kekuatan: mengurangi ketergantungan pada pertanyaan definisi.
- Kekuatan: mengungkap pengetahuan dalam tindakan.
- Batas: kinerja dipengaruhi tuntutan tugas.
- Batas: alat dapat memberi petunjuk.

Dalam praktik kelas, Wawancara Berbasis Tugas dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

Aspek	Temuan	Respons Metodologis
Kekuatan	mengurangi ketergantungan pada pertanyaan definisi	Pilih ketika sesuai tujuan
Keterbatasan	kinerja dipengaruhi tuntutan tugas	Tambahkan bukti atau kontrol
Bias potensial	Bahasa, konteks, perangkat, gaya respons	Uji proses respons

6.9.14 Integrasi Digital, Etika, dan Tata Kelola Data

Integrasi digital Wawancara Berbasis Tugas dapat berupa simulasi interaktif, screen recording, stylus, dan log klik yang diselaraskan dengan rekaman suara. Teknologi seharusnya memperluas kualitas bukti, bukan hanya memindahkan lembar kertas ke layar. (Baker & Inventado, 2014; Romero & Ventura, 2020; AERA, APA, & NCME, 2014; Posner & Gertzog, 1982)

Platform perlu menyimpan versi instrumen, urutan respons, timestamp yang relevan, aturan skor, dan log perubahan. Namun, hanya data yang dibutuhkan untuk tujuan diagnosis yang boleh dikumpulkan; pengumpulan berlebihan meningkatkan risiko privasi tanpa manfaat ilmiah.

Aspek etika utama adalah bahwa tugas harus aman, dapat diakses, dan tidak membuat peserta didik dinilai dari keterampilan motorik yang tidak relevan. Peserta didik dan guru perlu mengetahui data apa yang dikumpulkan, untuk tujuan apa, siapa yang dapat mengakses, berapa lama disimpan, dan bagaimana kesalahan data diperbaiki.

Algoritma otomatis harus dapat diaudit. Kategori yang dihasilkan sistem perlu ditautkan ke respons mentah dan aturan keputusan. Kasus berisiko atau ambigu harus ditinjau manusia sebelum tindak lanjut penting dilakukan.

- gunakan identitas pseudonim dan akses berbasis peran;
- uji tampilan rumus, subskrip, grafik, dan gambar;
- sediakan opsi aksesibilitas dan alternatif perangkat;
- dokumentasikan algoritma serta perubahan versi.

Dalam praktik kelas, Wawancara Berbasis Tugas dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.9.15 Sintesis dan Pedoman Implementasi

Wawancara Berbasis Tugas bermanfaat apabila dipilih berdasarkan tujuan dan digunakan sesuai batasnya. Instrumen ini menghasilkan langkah kerja, pilihan representasi, ucapan, revisi, kesalahan, dan produk tugas, lalu data diolah melalui analisis episode, strategi, titik keputusan, dan kesesuaian tindakan dengan penjelasan verbal. (Posner & Gertzog, 1982; Ericsson & Simon, 1993; van Someren et al., 1994; AERA, APA, & NCME, 2014; Kane, 2013)

Rangkaian kerja yang disarankan dimulai dari analisis konsep, identifikasi konsepsi alternatif, desain tugas, telaah ahli, uji proses respons, uji lapangan, analisis, penetapan kategori, dan penyusunan tindak lanjut. Setiap tahap harus meninggalkan dokumentasi yang dapat diperiksa.

Hasil diagnosis tidak berhenti pada tabel prevalensi. Guru perlu memilih konsepsi prioritas berdasarkan frekuensi, keyakinan, posisi sebagai konsep prasyarat, dampak terhadap topik lanjutan, dan kemungkinan diperbaiki melalui pembelajaran.

Prinsip akhirnya ialah kehati-hatian inferensial: sebutkan apa yang benar-benar ditunjukkan data, bedakan bukti kuat dari indikasi, gunakan triangulasi ketika diperlukan, dan hindari label permanen pada peserta didik.

- selaraskan tujuan, konstruk, tugas, skor, kategori, dan tindakan;
- utamakan konsepsi yang berdampak pada pembelajaran lanjutan;
- gabungkan bukti bila keputusan memiliki konsekuensi besar;
- evaluasi kembali efektivitas tindak lanjut melalui diagnosis ulang.

Dalam praktik kelas, Wawancara Berbasis Tugas dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.10 Peta Konsep

6.10.1 Landasan Konseptual dan Kedudukan Instrumen

Peta Konsep menempati posisi tertentu dalam rangkaian bukti diagnosis, dari pengungkapan respons awal sampai penentuan tindak lanjut pembelajaran. Dalam bab ini, instrumen tersebut didefinisikan sebagai diagram yang menghubungkan konsep melalui kata penghubung menjadi proposisi sehingga organisasi dan hubungan dalam pengetahuan peserta didik dapat diamati. Definisi operasional ini menegaskan bentuk bukti yang dikumpulkan sekaligus membatasi jenis kesimpulan yang layak dibuat. (Novak & Gowin, 1984; Ruiz-Primo & Shavelson, 1996; McClure et al., 1999)

Klaim utamanya adalah bahwa instrumen ini mendiagnosis kelengkapan, hierarki, integrasi, hubungan silang, dan proposisi salah dalam jaringan konsep. Klaim tersebut harus dipisahkan dari anggapan bahwa setiap jawaban salah merupakan miskonsepsi. Kesalahan dapat berasal dari kurang pengetahuan, ketidakcermatan, kesulitan bahasa, salah hitung, atau ketidakfamiliaran dengan representasi.

Data pokok yang dihasilkan berupa struktur graf, jumlah dan kualitas proposisi, tingkat hierarki, serta hubungan benar dan salah. Data itu baru memperoleh makna diagnostik setelah dikaitkan dengan peta konsep sasaran, kemungkinan konsepsi alternatif, dan aturan interpretasi yang telah ditetapkan sebelum analisis.

Pembedaan ini penting karena satu pola respons dapat memiliki lebih dari satu penjelasan yang masuk akal. Posisi terbaik Peta Konsep adalah sebagai bagian dari sistem asesmen berlapis yang dapat dipadukan dengan bukti lain ketika konsekuensi keputusan semakin besar.

- menentukan konsep dan proposisi yang hendak didiagnosis;
- menjelaskan perilaku respons yang dianggap sebagai bukti;
- membatasi klaim sesuai kekuatan instrumen;
- menautkan hasil dengan tindakan pembelajaran yang realistis.

Dalam praktik kelas, Peta Konsep dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.10.2 Perkembangan Historis dan Rasional Penggunaan

Perkembangan Peta Konsep tidak dapat dipisahkan dari perubahan pandangan tentang belajar sains. Ketika peserta didik tidak lagi dipandang sebagai penerima informasi yang kosong, instrumen mulai dirancang untuk mengungkap gagasan awal, pola penalaran, dan keyakinan yang dibawa ke kelas. (Novak & Gowin, 1984; Ruiz-Primo & Shavelson, 1996; McClure et al., 1999; Yin et al., 2005)

Pada tahap awal, penilaian banyak mengandalkan benar-salah atau skor total. Perkembangan penelitian konsepsi peserta didik kemudian mendorong penambahan alasan, keyakinan, tugas kinerja, rekaman proses, dan bukti sosial. Peta Konsep merupakan salah satu jawaban metodologis terhadap kebutuhan tersebut.

Dalam pendidikan kimia, kebutuhan diagnosis semakin kuat karena peserta didik harus mengoordinasikan fenomena makroskopik, model submikroskopik, simbol, grafik, dan perhitungan. Instrumen yang hanya menguji hafalan definisi sering gagal memperlihatkan cara representasi itu dihubungkan.

Sejarah penggunaannya juga menunjukkan bahwa format yang lebih kompleks tidak otomatis lebih baik. Setiap tambahan tier, rubrik, atau kanal data harus mempunyai fungsi yang jelas dan didukung bukti bahwa informasi tambahan benar-benar memperbaiki keputusan.

- pergeseran dari pengukuran skor menuju diagnosis pola pemahaman;
- penggunaan respons terbuka dan wawancara untuk menemukan konsepsi alternatif;
- pengembangan teknik kuantitatif agar diagnosis dapat dilakukan pada kelompok besar;
- integrasi data digital dan metakognitif pada perkembangan mutakhir.

Dalam praktik kelas, Peta Konsep dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.10.3 Konstruk, Klaim, dan Bukti Diagnosis

Konstruk yang dinilai melalui Peta Konsep bukan objek yang terlihat langsung. Yang diamati adalah struktur graf, jumlah dan kualitas proposisi, tingkat hierarki, serta hubungan benar dan salah; dari sana dibuat inferensi tentang pemahaman konsep, miskonsepsi, pengetahuan parsial, atau ketidakpastian. (Yin et al., 2005; AERA, APA, & NCME, 2014; Messick, 1995; Kane, 2013)

Klaim diagnosis perlu dirumuskan sebagai argumen. Pertama, tugas harus memerlukan konsep sasaran. Kedua, respons harus diberi skor atau kode secara konsisten. Ketiga, pola skor harus mempunyai hubungan yang masuk akal dengan struktur pengetahuan. Keempat, penggunaan hasil harus menghasilkan keputusan yang proporsional.

Sumber ancaman utama adalah ketidakjelasan konstruk. Misalnya, peserta didik dapat gagal karena istilah yang ambigu, beban membaca, representasi yang tidak dikenal, atau prosedur matematika. Kegagalan semacam itu tidak boleh langsung diperlakukan sebagai miskonsepsi kimia.

Oleh sebab itu, setiap pengembang perlu menulis pernyataan klaim-bukti: instrumen mengukur apa, perilaku apa yang menjadi indikator, alternatif penjelasan apa yang mungkin, dan bukti tambahan apa yang diperlukan untuk memperkuat kesimpulan.

- konstruk utama: pemahaman konseptual melalui representasi struktur pengetahuan;
- bukti observabel: respons, alasan, tindakan, atau representasi;
- inferensi: profil konsepsi pada konten dan konteks tertentu;
- penggunaan: perencanaan pembelajaran, remediasi, atau penelitian.

Dalam praktik kelas, Peta Konsep dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.10.4 Struktur, Komponen, dan Alur Respons

Struktur Peta Konsep dibangun dari komponen yang saling bergantung. Komponen inti meliputi daftar konsep atau konsep bebas, garis hubungan, kata penghubung, proposisi, hierarki, dan hubungan silang. Hilangnya satu komponen dapat mengubah makna data yang diperoleh. (Novak & Gowin, 1984; Ruiz-Primo & Shavelson, 1996; McClure et al., 1999; Yin et al., 2005)

Urutan penyajian harus ditetapkan secara sadar. Komponen yang muncul lebih awal dapat memberi petunjuk pada komponen berikutnya, sedangkan permintaan alasan atau keyakinan dapat mengubah cara peserta didik memeriksa jawabannya. Penguncian respons dan uji urutan diperlukan pada format digital.

Setiap komponen membutuhkan spesifikasi. Stimulus harus memuat informasi cukup, opsi atau prompt harus sejajar dengan indikator, skala harus memiliki jangkar, dan rubrik harus menjelaskan perbedaan kategori yang berdekatan.

Struktur yang baik memungkinkan peneliti menelusuri kembali setiap kategori hasil ke respons mentah. Transparansi ini penting agar guru tidak hanya menerima label otomatis tanpa memahami alasan di baliknya.

- daftar konsep atau konsep bebas;
- garis hubungan;
- kata penghubung;
- proposisi;
- hierarki;
- hubungan silang;

Dalam praktik kelas, Peta Konsep dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

No.	Komponen	Fungsi Diagnostik
1	daftar konsep atau konsep bebas	Bukti yang harus ditelusuri
2	garis hubungan	Bukti yang harus ditelusuri
3	kata penghubung	Bukti yang harus ditelusuri
4	proposisi	Bukti yang harus ditelusuri

6.10.5 Desain Stimulus, Tugas, atau Pertanyaan

Desain stimulus pada Peta Konsep harus memunculkan penggunaan konsep, bukan sekadar pengenalan istilah. Situasi dapat berbentuk fenomena, gambar partikel, grafik, data eksperimen, persamaan, atau masalah kontekstual, tetapi seluruh informasi yang tidak relevan perlu dikendalikan. (Haladyna et al., 2002; Novak & Gowin, 1984; Ruiz-Primo & Shavelson, 1996; McClure et al., 1999)

Dalam kimia, satu stimulus sering menggabungkan beberapa tingkat representasi. Pengembang harus memutuskan apakah tujuan butir adalah menerjemahkan representasi, menjelaskan hubungan sebab-akibat, membedakan contoh dan noncontoh, atau menerapkan prinsip pada konteks baru.

Bahasa perlu ringkas tanpa menghilangkan syarat ilmiah. Kata seperti kuat, stabil, jenuh, spontan, atau setimbang mempunyai arti teknis dan sehari-hari yang berbeda. Ketidakjelasan istilah dapat menghasilkan kesalahan yang berasal dari bahasa, bukan struktur konsep.

Desain yang baik juga menghindari petunjuk format, pilihan terlalu panjang, negasi bertumpuk, gambar dekoratif, dan data yang tidak diperlukan. Setiap elemen harus mempunyai alasan diagnostik.

- gunakan konteks yang cukup dikenal tetapi tidak menghafalkan contoh;
- pastikan hanya satu interpretasi ilmiah yang menjadi target;
- uji semua simbol, satuan, grafik, dan gambar;
- sediakan ruang bagi respons yang tidak diprediksi bila tujuan masih eksploratif.

Dalam praktik kelas, Peta Konsep dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.10.6 Prosedur Pengembangan Secara Sistematis

Pengembangan Peta Konsep harus mengikuti alur yang terdokumentasi. Langkah yang lazim ialah menentukan format tugas, memilih konsep, menjelaskan aturan, mengembangkan rubrik, kemudian menyediakan latihan, mengumpulkan peta, menilai proposisi, menganalisis perubahan struktur. Urutan dapat berulang karena temuan uji coba sering memaksa revisi konstruk atau redaksi. (Novak & Gowin, 1984; Ruiz-Primo & Shavelson, 1996; McClure et al., 1999; Adams & Wieman, 2011; AERA, APA, & NCME, 2014)

Tahap konseptual dimulai dengan analisis kurikulum, literatur miskonsepsi, dan peta konsep. Tahap empiris awal menggunakan wawancara, respons terbuka, atau observasi untuk memastikan bahwa kategori yang dirancang benar-benar muncul pada populasi sasaran.

Setiap versi instrumen perlu diberi nomor, tanggal, dan catatan perubahan. Revisi harus menyebutkan masalah yang ditemukan, bukti yang digunakan, perubahan yang dilakukan, dan alasan mempertahankan atau menghapus butir.

Pengembangan dianggap belum selesai ketika hanya menghasilkan butir yang tampak baik. Instrumen perlu diuji pada peserta didik yang serupa dengan pengguna sasaran, lalu dianalisis secara konseptual dan empiris sebelum dipakai untuk keputusan penting.

- 1. menentukan format tugas.
- 2. memilih konsep.
- 3. menjelaskan aturan.
- 4. mengembangkan rubrik.
- 5. menyediakan latihan.

Dalam praktik kelas, Peta Konsep dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.10.7 Pelaksanaan dan Pengumpulan Data

Pelaksanaan Peta Konsep harus distandardisasi secukupnya agar perbedaan hasil lebih mungkin berasal dari pemahaman, bukan variasi instruksi, waktu, alat, atau bantuan. Petunjuk perlu menjelaskan tujuan umum tanpa mengajarkan jawaban. (Novak & Gowin, 1984; Ruiz-Primo & Shavelson, 1996; AERA, APA, & NCME, 2014)

Kondisi administrasi mencakup waktu, urutan tugas, kesempatan bertanya, penggunaan kalkulator, akses tabel periodik, kualitas gambar, dan perangkat digital. Pada instrumen kualitatif, standardisasi berarti konsistensi fokus dan prinsip probing, bukan membaca naskah secara kaku.

Data yang dikumpulkan ialah struktur graf, jumlah dan kualitas proposisi, tingkat hierarki, serta hubungan benar dan salah. Data mentah harus disimpan bersama metadata yang diperlukan, seperti kode butir, versi instrumen, waktu administrasi, dan kondisi khusus, tetapi hanya data yang relevan yang boleh dikumpulkan.

Sebelum administrasi luas, lakukan uji teknis. Pastikan tautan, urutan, penguncian respons, rekaman, penyimpanan, serta tampilan rumus dan subskrip bekerja. Gangguan teknis perlu dicatat sebagai bagian dari interpretasi.

- berikan instruksi yang sama dan mudah dipahami;
- catat kondisi yang dapat memengaruhi respons;
- hindari umpan balik sebelum seluruh respons diagnostik selesai;
- sediakan prosedur penanganan data tidak lengkap.

Dalam praktik kelas, Peta Konsep dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.10.8 Penskoran, Pengodean, dan Pengolahan Data

Penskoran Peta Konsep dapat diringkas sebagai penskoran proposisional, struktural, holistik, atau perbandingan dengan peta kriteria disertai analisis miskonsepsi spesifik. Aturan ini harus ditulis sebelum melihat hasil utama agar keputusan tidak disesuaikan secara oportunistik dengan pola data. (Novak & Gowin, 1984; Ruiz-Primo & Shavelson, 1996; McClure et al., 1999; Chi, 1997; AERA, APA, & NCME, 2014)

Skor total jarang cukup untuk diagnosis. Analisis perlu mempertahankan informasi tingkat butir, konsep, opsi, alasan, keyakinan, proses, atau representasi. Penggabungan data terlalu dini dapat menyembunyikan pola yang secara pedagogis penting.

Untuk respons terbuka atau kualitatif, buku kode harus mencantumkan definisi, contoh positif, contoh batas, noncontoh, dan aturan ketika satu respons memuat beberapa gagasan. Sebagian data perlu dinilai oleh lebih dari satu penilai.

Setiap transformasi data—misalnya dari respons ke kode, dari kode ke kategori, dan dari kategori ke persentase—harus dapat diaudit. Simpan respons asli dan keputusan pengodean agar kasus ambigu dapat ditinjau kembali.

- pisahkan skor ketepatan dari skor keyakinan atau kualitas alasan;
- laporkan kasus tidak lengkap dan aturan penanganannya;
- uji konsistensi penilai untuk data berkode;
- hindari mengubah data ordinal menjadi ukuran interval tanpa alasan.

Dalam praktik kelas, Peta Konsep dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.10.9 Interpretasi dan Penetapan Kategori

Interpretasi hasil Peta Konsep harus bergerak dari respons menuju pola, bukan langsung dari satu jawaban menuju label. Status konseptual menjadi lebih kuat apabila gagasan yang sama muncul pada beberapa butir, representasi, atau konteks. (Novak & Gowin, 1984; Ruiz-Primo & Shavelson, 1996; McClure et al., 1999; Yin et al., 2005; Kane, 2013)

Kategori dapat mencakup pemahaman ilmiah, pemahaman parsial, kurang pengetahuan, respons tidak konsisten, false positive, false negative, dan miskonsepsi. Tidak semua instrumen mampu membedakan seluruh kategori tersebut; skema harus mengikuti informasi yang benar-benar tersedia.

Ambang, seperti batas keyakinan atau persentase pemilihan, perlu memiliki rasional dan diuji melalui analisis sensitivitas. Perubahan kecil pada ambang dapat mengubah prevalensi kategori secara berarti.

Interpretasi yang bertanggung jawab menyatakan bahwa kategori berlaku pada konsep, bentuk tugas, waktu, dan kondisi tertentu. Hasil merupakan hipotesis kerja untuk pembelajaran, bukan diagnosis psikologis permanen.

- gunakan lebih dari satu butir untuk konsepsi penting;
- periksa konsistensi antara jawaban dan alasan;
- tandai kasus ambigu untuk konfirmasi;
- hubungkan kategori dengan rekomendasi yang berbeda.

Dalam praktik kelas, Peta Konsep dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.10.10 Validitas dan Kualitas Inferensi

Validitas Peta Konsep berkaitan dengan kualitas interpretasi dan penggunaan hasil, bukan sifat tetap yang dimiliki instrumen. Fokus khususnya adalah kesesuaian format peta dengan klaim tentang struktur pengetahuan dan konsistensi penskoran proposisi. (AERA, APA, & NCME, 2014; Messick, 1995; Kane, 2013; Haynes et al., 1995)

Bukti berdasarkan isi menilai keterwakilan domain dan akurasi konsep. Bukti proses respons memeriksa bagaimana peserta didik memahami tugas. Bukti struktur internal menilai kesesuaian pola respons. Bukti hubungan dengan variabel lain dan konsekuensi penggunaan melengkapi argumen.

Ancaman validitas dapat berasal dari konten yang kurang terwakili, tuntutan bahasa berlebihan, petunjuk opsi, rubrik yang tidak stabil, keyakinan yang tidak terkalibrasi, atau algoritma yang belajar dari label keliru.

Dokumen validasi perlu menghubungkan setiap klaim dengan bukti dan alternatif penjelasan. Menampilkan satu indeks tidak cukup; validasi adalah program argumentatif yang diperbarui setiap kali instrumen dipakai pada populasi atau tujuan baru.

- bukti isi dan kesesuaian kurikulum;
- bukti proses respons melalui wawancara kognitif;
- bukti struktur dan fungsi butir;
- bukti kegunaan serta konsekuensi keputusan.

Dalam praktik kelas, Peta Konsep dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.10.11 Reliabilitas, Keajegan, dan Analisis Empiris

Keajegan Peta Konsep harus dinilai sesuai jenis data. Pada respons objektif dapat digunakan konsistensi internal, stabilitas waktu, atau model pengukuran. Pada data kualitatif, perhatian diarahkan pada konsistensi pengodean, audit trail, dan ketepatan penerapan definisi kode. (Kuder & Richardson, 1937; Cronbach, 1951; Ding et al., 2006; Wren & Barbera, 2014; AERA, APA, & NCME, 2014)

Reliabilitas tinggi tidak otomatis menunjukkan diagnosis yang benar. Butir yang berulang-ulang mengukur hafalan dapat menghasilkan koefisien tinggi tetapi tidak mewakili pemahaman konseptual. Sebaliknya, konstruk yang multidimensi tidak selalu layak dipaksa menjadi satu skala.

Analisis empiris perlu memeriksa kesukaran, daya beda, fungsi opsi, pola data hilang, dan keterhubungan komponen. Untuk instrumen bertingkat, reliabilitas dapat dihitung pada setiap tier serta kombinasi yang mempunyai makna teoritis.

Ukuran sampel, distribusi kemampuan, dan homogenitas konten memengaruhi statistik. Karena itu, laporan harus menyatakan kondisi estimasi dan tidak menyalin batas angka sebagai aturan universal.

- pilih statistik yang sesuai dengan jenis skor;
- laporkan interval atau ketidakpastian bila tersedia;
- periksa butir secara substantif dan statistik;
- uji kembali setelah revisi atau perubahan platform.

Dalam praktik kelas, Peta Konsep dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.10.12 Contoh Penerapan dalam Pembelajaran Kimia

Penerapan Peta Konsep dalam kimia dapat dimulai dari konsep jaringan asam-basa. Pengembang perlu menuliskan konsep ilmiah target, miskonsepsi yang mungkin, bukti respons yang diharapkan, dan keputusan pembelajaran setelah hasil diperoleh. (Novak & Gowin, 1984; Ruiz-Primo & Shavelson, 1996; McClure et al., 1999; Yin et al., 2005; Nakhleh, 1992)

Contoh kedua adalah hubungan struktur, ikatan, dan sifat. Stimulus sebaiknya menghubungkan fenomena dengan model partikel atau persamaan, lalu meminta peserta didik menunjukkan hubungan sebab-akibat. Jawaban yang hanya menyebut istilah belum cukup menjadi bukti pemahaman.

Pada peta termokimia, diagnosis dapat memeriksa apakah peserta didik membedakan perubahan laju, keadaan setimbang, energi, dan komposisi. Kesalahan sering muncul ketika satu aturan dipindahkan ke konteks yang tidak sesuai.

Kasus keterkaitan kesetimbangan, laju, dan energi menunjukkan pentingnya representasi. Respons verbal, gambar, simbol, dan hitungan dapat tidak konsisten. Hasil instrumen sebaiknya digunakan untuk memilih demonstrasi, diskusi, contoh-noncontoh, atau tugas remediasi yang secara langsung menantang hubungan salah.

- tentukan proposisi ilmiah dan konsepsi alternatif;
- gunakan minimal dua bentuk representasi;
- minta alasan atau bukti ketika format memungkinkan;
- rancang tindak lanjut berdasarkan pola, bukan hanya skor.

Dalam praktik kelas, Peta Konsep dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.10.13 Kelebihan, Keterbatasan, dan Potensi Bias

Kekuatan utama Peta Konsep meliputi menampilkan jaringan pengetahuan, mengungkap hubungan antarkonsep, baik untuk diagnosis prasyarat, dan dapat digunakan sebelum dan sesudah pembelajaran. Kekuatan tersebut menjelaskan situasi ketika instrumen layak dipilih dibanding teknik lain. (Novak & Gowin, 1984; Ruiz-Primo & Shavelson, 1996; McClure et al., 1999; Gurel et al., 2015; AERA, APA, & NCME, 2014)

Keterbatasannya mencakup hasil dipengaruhi keterampilan membuat peta, peta kriteria bukan satu-satunya struktur benar, skor total dapat menyembunyikan proposisi salah, dan tugas terlalu terarah membatasi ekspresi. Keterbatasan harus diperlakukan sebagai batas inferensi, bukan sekadar paragraf formal yang tidak memengaruhi analisis.

Potensi bias dapat berasal dari bahasa, pengalaman konteks, akses teknologi, kecemasan, gaya menggunakan skala, keterampilan menggambar, atau kemampuan menulis. Instrumen perlu diuji pada kelompok yang beragam dan ditinjau untuk fungsi berbeda antarkelompok bila data mencukupi.

Keputusan memilih instrumen sebaiknya mempertimbangkan kedalaman data, ukuran kelompok, waktu, keterampilan analisis, tujuan diagnosis, dan konsekuensi kesalahan klasifikasi. Tidak ada satu metode yang unggul untuk semua tujuan.

- Kekuatan: menampilkan jaringan pengetahuan.
- Kekuatan: mengungkap hubungan antarkonsep.
- Batas: hasil dipengaruhi keterampilan membuat peta.
- Batas: peta kriteria bukan satu-satunya struktur benar.

Dalam praktik kelas, Peta Konsep dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

Aspek	Temuan	Respons Metodologis
Kekuatan	menampilkan jaringan pengetahuan	Pilih ketika sesuai tujuan
Keterbatasan	hasil dipengaruhi keterampilan membuat peta	Tambahkan bukti atau kontrol
Bias potensial	Bahasa, konteks, perangkat, gaya respons	Uji proses respons

6.10.14 Integrasi Digital, Etika, dan Tata Kelola Data

Integrasi digital Peta Konsep dapat berupa perangkat pemetaan yang menyimpan urutan pembuatan, revisi hubungan, dan versi peta. Teknologi seharusnya memperluas kualitas bukti, bukan hanya memindahkan lembar kertas ke layar. (Baker & Inventado, 2014; Romero & Ventura, 2020; AERA, APA, & NCME, 2014; Novak & Gowin, 1984)

Platform perlu menyimpan versi instrumen, urutan respons, timestamp yang relevan, aturan skor, dan log perubahan. Namun, hanya data yang dibutuhkan untuk tujuan diagnosis yang boleh dikumpulkan; pengumpulan berlebihan meningkatkan risiko privasi tanpa manfaat ilmiah.

Aspek etika utama adalah bahwa peserta didik perlu kesempatan latihan agar diagnosis konsep tidak tercampur dengan ketidakfamiliaran terhadap teknik peta konsep. Peserta didik dan guru perlu mengetahui data apa yang dikumpulkan, untuk tujuan apa, siapa yang dapat mengakses, berapa lama disimpan, dan bagaimana kesalahan data diperbaiki.

Algoritma otomatis harus dapat diaudit. Kategori yang dihasilkan sistem perlu ditautkan ke respons mentah dan aturan keputusan. Kasus berisiko atau ambigu harus ditinjau manusia sebelum tindak lanjut penting dilakukan.

- gunakan identitas pseudonim dan akses berbasis peran;
- uji tampilan rumus, subskrip, grafik, dan gambar;
- sediakan opsi aksesibilitas dan alternatif perangkat;
- dokumentasikan algoritma serta perubahan versi.

Dalam praktik kelas, Peta Konsep dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.10.15 Sintesis dan Pedoman Implementasi

Peta Konsep bermanfaat apabila dipilih berdasarkan tujuan dan digunakan sesuai batasnya. Instrumen ini menghasilkan struktur graf, jumlah dan kualitas proposisi, tingkat hierarki, serta hubungan benar dan salah, lalu data diolah melalui penskoran proposisional, struktural, holistik, atau perbandingan dengan peta kriteria disertai analisis miskonsepsi spesifik. (Novak & Gowin, 1984; Ruiz-Primo & Shavelson, 1996; McClure et al., 1999; AERA, APA, & NCME, 2014; Kane, 2013)

Rangkaian kerja yang disarankan dimulai dari analisis konsep, identifikasi konsepsi alternatif, desain tugas, telaah ahli, uji proses respons, uji lapangan, analisis, penetapan kategori, dan penyusunan tindak lanjut. Setiap tahap harus meninggalkan dokumentasi yang dapat diperiksa.

Hasil diagnosis tidak berhenti pada tabel prevalensi. Guru perlu memilih konsepsi prioritas berdasarkan frekuensi, keyakinan, posisi sebagai konsep prasyarat, dampak terhadap topik lanjutan, dan kemungkinan diperbaiki melalui pembelajaran.

Prinsip akhirnya ialah kehati-hatian inferensial: sebutkan apa yang benar-benar ditunjukkan data, bedakan bukti kuat dari indikasi, gunakan triangulasi ketika diperlukan, dan hindari label permanen pada peserta didik.

- selaraskan tujuan, konstruk, tugas, skor, kategori, dan tindakan;
- utamakan konsepsi yang berdampak pada pembelajaran lanjutan;
- gabungkan bukti bila keputusan memiliki konsekuensi besar;
- evaluasi kembali efektivitas tindak lanjut melalui diagnosis ulang.

Dalam praktik kelas, Peta Konsep dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.11 Open-Ended Diagnostic Test

6.11.1 Landasan Konseptual dan Kedudukan Instrumen

Open-Ended Diagnostic Test menempati posisi tertentu dalam rangkaian bukti diagnosis, dari pengungkapan respons awal sampai penentuan tindak lanjut pembelajaran. Dalam bab ini, instrumen tersebut didefinisikan sebagai serangkaian pertanyaan terbuka yang meminta peserta didik menjelaskan, memprediksi, menggambar, membandingkan, atau membenarkan tanpa dibatasi opsi jawaban. Definisi operasional ini menegaskan bentuk bukti yang dikumpulkan sekaligus membatasi jenis kesimpulan yang layak dibuat. (Chi, 1997; Treagust, 1988; Taber, 2002)

Klaim utamanya adalah bahwa instrumen ini menangkap konsepsi yang belum diperkirakan peneliti dan menunjukkan bahasa serta struktur penalaran peserta didik. Klaim tersebut harus dipisahkan dari anggapan bahwa setiap jawaban salah merupakan miskonsepsi. Kesalahan dapat berasal dari kurang pengetahuan, ketidakcermatan, kesulitan bahasa, salah hitung, atau ketidakfamiliaran dengan representasi.

Data pokok yang dihasilkan berupa teks, simbol, perhitungan, gambar, dan argumentasi yang dihasilkan peserta didik. Data itu baru memperoleh makna diagnostik setelah dikaitkan dengan peta konsep sasaran, kemungkinan konsepsi alternatif, dan aturan interpretasi yang telah ditetapkan sebelum analisis.

Karena miskonsepsi bersifat kontekstual dan kadang tersembunyi, bukti tunggal perlu ditafsirkan secara terbatas. Posisi terbaik Open-Ended Diagnostic Test adalah sebagai bagian dari sistem asesmen berlapis yang dapat dipadukan dengan bukti lain ketika konsekuensi keputusan semakin besar.

- menentukan konsep dan proposisi yang hendak didiagnosis;
- menjelaskan perilaku respons yang dianggap sebagai bukti;
- membatasi klaim sesuai kekuatan instrumen;
- menautkan hasil dengan tindakan pembelajaran yang realistis.

Dalam praktik kelas, Open-Ended Diagnostic Test dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.11.2 Perkembangan Historis dan Rasional Penggunaan

Perkembangan Open-Ended Diagnostic Test tidak dapat dipisahkan dari perubahan pandangan tentang belajar sains. Ketika peserta didik tidak lagi dipandang sebagai penerima informasi yang kosong, instrumen mulai dirancang untuk mengungkap gagasan awal, pola penalaran, dan keyakinan yang dibawa ke kelas. (Chi, 1997; Treagust, 1988; Taber, 2002; AERA, APA, & NCME, 2014)

Pada tahap awal, penilaian banyak mengandalkan benar-salah atau skor total. Perkembangan penelitian konsepsi peserta didik kemudian mendorong penambahan alasan, keyakinan, tugas kinerja, rekaman proses, dan bukti sosial. Open-Ended Diagnostic Test merupakan salah satu jawaban metodologis terhadap kebutuhan tersebut.

Dalam pendidikan kimia, kebutuhan diagnosis semakin kuat karena peserta didik harus mengoordinasikan fenomena makroskopik, model submikroskopik, simbol, grafik, dan perhitungan. Instrumen yang hanya menguji hafalan definisi sering gagal memperlihatkan cara representasi itu dihubungkan.

Sejarah penggunaannya juga menunjukkan bahwa format yang lebih kompleks tidak otomatis lebih baik. Setiap tambahan tier, rubrik, atau kanal data harus mempunyai fungsi yang jelas dan didukung bukti bahwa informasi tambahan benar-benar memperbaiki keputusan.

- pergeseran dari pengukuran skor menuju diagnosis pola pemahaman;
- penggunaan respons terbuka dan wawancara untuk menemukan konsepsi alternatif;
- pengembangan teknik kuantitatif agar diagnosis dapat dilakukan pada kelompok besar;
- integrasi data digital dan metakognitif pada perkembangan mutakhir.

Dalam praktik kelas, Open-Ended Diagnostic Test dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.11.3 Konstruk, Klaim, dan Bukti Diagnosis

Konstruk yang dinilai melalui Open-Ended Diagnostic Test bukan objek yang terlihat langsung. Yang diamati adalah teks, simbol, perhitungan, gambar, dan argumentasi yang dihasilkan peserta didik; dari sana dibuat inferensi tentang pemahaman konsep, miskonsepsi, pengetahuan parsial, atau ketidakpastian. (AERA, APA, & NCME, 2014; Flake & Fried, 2020; Messick, 1995; Kane, 2013)

Klaim diagnosis perlu dirumuskan sebagai argumen. Pertama, tugas harus memerlukan konsep sasaran. Kedua, respons harus diberi skor atau kode secara konsisten. Ketiga, pola skor harus mempunyai hubungan yang masuk akal dengan struktur pengetahuan. Keempat, penggunaan hasil harus menghasilkan keputusan yang proporsional.

Sumber ancaman utama adalah ketidakjelasan konstruk. Misalnya, peserta didik dapat gagal karena istilah yang ambigu, beban membaca, representasi yang tidak dikenal, atau prosedur matematika. Kegagalan semacam itu tidak boleh langsung diperlakukan sebagai miskonsepsi kimia.

Oleh sebab itu, setiap pengembang perlu menulis pernyataan klaim-bukti: instrumen mengukur apa, perilaku apa yang menjadi indikator, alternatif penjelasan apa yang mungkin, dan bukti tambahan apa yang diperlukan untuk memperkuat kesimpulan.

- konstruk utama: pemahaman konseptual melalui tes uraian diagnostik;
- bukti observabel: respons, alasan, tindakan, atau representasi;
- inferensi: profil konsepsi pada konten dan konteks tertentu;
- penggunaan: perencanaan pembelajaran, remediasi, atau penelitian.

Dalam praktik kelas, Open-Ended Diagnostic Test dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.11.4 Struktur, Komponen, dan Alur Respons

Struktur Open-Ended Diagnostic Test dibangun dari komponen yang saling bergantung. Komponen inti meliputi stimulus, prompt terbuka, ruang respons, rubrik analitik, kode miskonsepsi, dan contoh jangkar. Hilangnya satu komponen dapat mengubah makna data yang diperoleh. (Chi, 1997; Treagust, 1988; Taber, 2002; AERA, APA, & NCME, 2014)

Urutan penyajian harus ditetapkan secara sadar. Komponen yang muncul lebih awal dapat memberi petunjuk pada komponen berikutnya, sedangkan permintaan alasan atau keyakinan dapat mengubah cara peserta didik memeriksa jawabannya. Penguncian respons dan uji urutan diperlukan pada format digital.

Setiap komponen membutuhkan spesifikasi. Stimulus harus memuat informasi cukup, opsi atau prompt harus sejajar dengan indikator, skala harus memiliki jangkar, dan rubrik harus menjelaskan perbedaan kategori yang berdekatan.

Struktur yang baik memungkinkan peneliti menelusuri kembali setiap kategori hasil ke respons mentah. Transparansi ini penting agar guru tidak hanya menerima label otomatis tanpa memahami alasan di baliknya.

- stimulus;
- prompt terbuka;
- ruang respons;
- rubrik analitik;
- kode miskonsepsi;
- contoh jangkar;

Dalam praktik kelas, Open-Ended Diagnostic Test dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

No.	Komponen	Fungsi Diagnostik
1	stimulus	Bukti yang harus ditelusuri
2	prompt terbuka	Bukti yang harus ditelusuri
3	ruang respons	Bukti yang harus ditelusuri
4	rubrik analitik	Bukti yang harus ditelusuri

6.11.5 Desain Stimulus, Tugas, atau Pertanyaan

Desain stimulus pada Open-Ended Diagnostic Test harus memunculkan penggunaan konsep, bukan sekadar pengenalan istilah. Situasi dapat berbentuk fenomena, gambar partikel, grafik, data eksperimen, persamaan, atau masalah kontekstual, tetapi seluruh informasi yang tidak relevan perlu dikendalikan. (Haladyna et al., 2002; Chi, 1997; Treagust, 1988; Taber, 2002)

Dalam kimia, satu stimulus sering menggabungkan beberapa tingkat representasi. Pengembang harus memutuskan apakah tujuan butir adalah menerjemahkan representasi, menjelaskan hubungan sebab-akibat, membedakan contoh dan noncontoh, atau menerapkan prinsip pada konteks baru.

Bahasa perlu ringkas tanpa menghilangkan syarat ilmiah. Kata seperti kuat, stabil, jenuh, spontan, atau setimbang mempunyai arti teknis dan sehari-hari yang berbeda. Ketidakjelasan istilah dapat menghasilkan kesalahan yang berasal dari bahasa, bukan struktur konsep.

Desain yang baik juga menghindari petunjuk format, pilihan terlalu panjang, negasi bertumpuk, gambar dekoratif, dan data yang tidak diperlukan. Setiap elemen harus mempunyai alasan diagnostik.

- gunakan konteks yang cukup dikenal tetapi tidak menghafalkan contoh;
- pastikan hanya satu interpretasi ilmiah yang menjadi target;
- uji semua simbol, satuan, grafik, dan gambar;
- sediakan ruang bagi respons yang tidak diprediksi bila tujuan masih eksploratif.

Dalam praktik kelas, Open-Ended Diagnostic Test dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.11.6 Prosedur Pengembangan Secara Sistematis

Pengembangan Open-Ended Diagnostic Test harus mengikuti alur yang terdokumentasi. Langkah yang lazim ialah menetapkan konstruk, menulis prompt nonmenuntun, menyiapkan rubrik awal, mengumpulkan sampel respons, kemudian merevisi kode, melatih penilai, menghitung kesepakatan, menyempurnakan rubrik. Urutan dapat berulang karena temuan uji coba sering memaksa revisi konstruk atau redaksi. (Chi, 1997; Treagust, 1988; Taber, 2002; Adams & Wieman, 2011; AERA, APA, & NCME, 2014)

Tahap konseptual dimulai dengan analisis kurikulum, literatur miskonsepsi, dan peta konsep. Tahap empiris awal menggunakan wawancara, respons terbuka, atau observasi untuk memastikan bahwa kategori yang dirancang benar-benar muncul pada populasi sasaran.

Setiap versi instrumen perlu diberi nomor, tanggal, dan catatan perubahan. Revisi harus menyebutkan masalah yang ditemukan, bukti yang digunakan, perubahan yang dilakukan, dan alasan mempertahankan atau menghapus butir.

Pengembangan dianggap belum selesai ketika hanya menghasilkan butir yang tampak baik. Instrumen perlu diuji pada peserta didik yang serupa dengan pengguna sasaran, lalu dianalisis secara konseptual dan empiris sebelum dipakai untuk keputusan penting.

- 1. menetapkan konstruk.
- 2. menulis prompt nonmenuntun.
- 3. menyiapkan rubrik awal.
- 4. mengumpulkan sampel respons.
- 5. merevisi kode.

Dalam praktik kelas, Open-Ended Diagnostic Test dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.11.7 Pelaksanaan dan Pengumpulan Data

Pelaksanaan Open-Ended Diagnostic Test harus distandardisasi secukupnya agar perbedaan hasil lebih mungkin berasal dari pemahaman, bukan variasi instruksi, waktu, alat, atau bantuan. Petunjuk perlu menjelaskan tujuan umum tanpa mengajarkan jawaban. (Chi, 1997; Treagust, 1988; AERA, APA, & NCME, 2014)

Kondisi administrasi mencakup waktu, urutan tugas, kesempatan bertanya, penggunaan kalkulator, akses tabel periodik, kualitas gambar, dan perangkat digital. Pada instrumen kualitatif, standardisasi berarti konsistensi fokus dan prinsip probing, bukan membaca naskah secara kaku.

Data yang dikumpulkan ialah teks, simbol, perhitungan, gambar, dan argumentasi yang dihasilkan peserta didik. Data mentah harus disimpan bersama metadata yang diperlukan, seperti kode butir, versi instrumen, waktu administrasi, dan kondisi khusus, tetapi hanya data yang relevan yang boleh dikumpulkan.

Sebelum administrasi luas, lakukan uji teknis. Pastikan tautan, urutan, penguncian respons, rekaman, penyimpanan, serta tampilan rumus dan subskrip bekerja. Gangguan teknis perlu dicatat sebagai bagian dari interpretasi.

- berikan instruksi yang sama dan mudah dipahami;
- catat kondisi yang dapat memengaruhi respons;
- hindari umpan balik sebelum seluruh respons diagnostik selesai;
- sediakan prosedur penanganan data tidak lengkap.

Dalam praktik kelas, Open-Ended Diagnostic Test dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.11.8 Penskoran, Pengodean, dan Pengolahan Data

Penskoran Open-Ended Diagnostic Test dapat diringkas sebagai rubrik bertingkat untuk ketepatan konsep, kelengkapan, koherensi, penggunaan representasi, dan jenis kesalahan. Aturan ini harus ditulis sebelum melihat hasil utama agar keputusan tidak disesuaikan secara oportunistik dengan pola data. (Chi, 1997; Treagust, 1988; Taber, 2002; AERA, APA, & NCME, 2014)

Skor total jarang cukup untuk diagnosis. Analisis perlu mempertahankan informasi tingkat butir, konsep, opsi, alasan, keyakinan, proses, atau representasi. Penggabungan data terlalu dini dapat menyembunyikan pola yang secara pedagogis penting.

Untuk respons terbuka atau kualitatif, buku kode harus mencantumkan definisi, contoh positif, contoh batas, noncontoh, dan aturan ketika satu respons memuat beberapa gagasan. Sebagian data perlu dinilai oleh lebih dari satu penilai.

Setiap transformasi data—misalnya dari respons ke kode, dari kode ke kategori, dan dari kategori ke persentase—harus dapat diaudit. Simpan respons asli dan keputusan pengodean agar kasus ambigu dapat ditinjau kembali.

- pisahkan skor ketepatan dari skor keyakinan atau kualitas alasan;
- laporkan kasus tidak lengkap dan aturan penanganannya;
- uji konsistensi penilai untuk data berkode;
- hindari mengubah data ordinal menjadi ukuran interval tanpa alasan.

Dalam praktik kelas, Open-Ended Diagnostic Test dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.11.9 Interpretasi dan Penetapan Kategori

Interpretasi hasil Open-Ended Diagnostic Test harus bergerak dari respons menuju pola, bukan langsung dari satu jawaban menuju label. Status konseptual menjadi lebih kuat apabila gagasan yang sama muncul pada beberapa butir, representasi, atau konteks. (Chi, 1997; Treagust, 1988; Taber, 2002; AERA, APA, & NCME, 2014; Kane, 2013)

Kategori dapat mencakup pemahaman ilmiah, pemahaman parsial, kurang pengetahuan, respons tidak konsisten, false positive, false negative, dan miskonsepsi. Tidak semua instrumen mampu membedakan seluruh kategori tersebut; skema harus mengikuti informasi yang benar-benar tersedia.

Ambang, seperti batas keyakinan atau persentase pemilihan, perlu memiliki rasional dan diuji melalui analisis sensitivitas. Perubahan kecil pada ambang dapat mengubah prevalensi kategori secara berarti.

Interpretasi yang bertanggung jawab menyatakan bahwa kategori berlaku pada konsep, bentuk tugas, waktu, dan kondisi tertentu. Hasil merupakan hipotesis kerja untuk pembelajaran, bukan diagnosis psikologis permanen.

- gunakan lebih dari satu butir untuk konsepsi penting;
- periksa konsistensi antara jawaban dan alasan;
- tandai kasus ambigu untuk konfirmasi;
- hubungkan kategori dengan rekomendasi yang berbeda.

Dalam praktik kelas, Open-Ended Diagnostic Test dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.11.10 Validitas dan Kualitas Inferensi

Validitas Open-Ended Diagnostic Test berkaitan dengan kualitas interpretasi dan penggunaan hasil, bukan sifat tetap yang dimiliki instrumen. Fokus khususnya adalah kecukupan prompt dalam memunculkan penalaran sasaran dan kecocokan kategori rubrik dengan variasi respons nyata. (AERA, APA, & NCME, 2014; Messick, 1995; Kane, 2013; Haynes et al., 1995)

Bukti berdasarkan isi menilai keterwakilan domain dan akurasi konsep. Bukti proses respons memeriksa bagaimana peserta didik memahami tugas. Bukti struktur internal menilai kesesuaian pola respons. Bukti hubungan dengan variabel lain dan konsekuensi penggunaan melengkapi argumen.

Ancaman validitas dapat berasal dari konten yang kurang terwakili, tuntutan bahasa berlebihan, petunjuk opsi, rubrik yang tidak stabil, keyakinan yang tidak terkalibrasi, atau algoritma yang belajar dari label keliru.

Dokumen validasi perlu menghubungkan setiap klaim dengan bukti dan alternatif penjelasan. Menampilkan satu indeks tidak cukup; validasi adalah program argumentatif yang diperbarui setiap kali instrumen dipakai pada populasi atau tujuan baru.

- bukti isi dan kesesuaian kurikulum;
- bukti proses respons melalui wawancara kognitif;
- bukti struktur dan fungsi butir;
- bukti kegunaan serta konsekuensi keputusan.

Dalam praktik kelas, Open-Ended Diagnostic Test dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.11.11 Reliabilitas, Keajegan, dan Analisis Empiris

Keajegan Open-Ended Diagnostic Test harus dinilai sesuai jenis data. Pada respons objektif dapat digunakan konsistensi internal, stabilitas waktu, atau model pengukuran. Pada data kualitatif, perhatian diarahkan pada konsistensi pengodean, audit trail, dan ketepatan penerapan definisi kode. (Kuder & Richardson, 1937; Cronbach, 1951; Ding et al., 2006; Wren & Barbera, 2014; AERA, APA, & NCME, 2014)

Reliabilitas tinggi tidak otomatis menunjukkan diagnosis yang benar. Butir yang berulang-ulang mengukur hafalan dapat menghasilkan koefisien tinggi tetapi tidak mewakili pemahaman konseptual. Sebaliknya, konstruk yang multidimensi tidak selalu layak dipaksa menjadi satu skala.

Analisis empiris perlu memeriksa kesukaran, daya beda, fungsi opsi, pola data hilang, dan keterhubungan komponen. Untuk instrumen bertingkat, reliabilitas dapat dihitung pada setiap tier serta kombinasi yang mempunyai makna teoritis.

Ukuran sampel, distribusi kemampuan, dan homogenitas konten memengaruhi statistik. Karena itu, laporan harus menyatakan kondisi estimasi dan tidak menyalin batas angka sebagai aturan universal.

- pilih statistik yang sesuai dengan jenis skor;
- laporkan interval atau ketidakpastian bila tersedia;
- periksa butir secara substantif dan statistik;
- uji kembali setelah revisi atau perubahan platform.

Dalam praktik kelas, Open-Ended Diagnostic Test dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.11.12 Contoh Penerapan dalam Pembelajaran Kimia

Penerapan Open-Ended Diagnostic Test dalam kimia dapat dimulai dari konsep menjelaskan mengapa es mengapung. Pengembang perlu menuliskan konsep ilmiah target, miskonsepsi yang mungkin, bukti respons yang diharapkan, dan keputusan pembelajaran setelah hasil diperoleh. (Chi, 1997; Treagust, 1988; Taber, 2002; AERA, APA, & NCME, 2014; Nakhleh, 1992)

Contoh kedua adalah menggambarkan pelarutan NaCl. Stimulus sebaiknya menghubungkan fenomena dengan model partikel atau persamaan, lalu meminta peserta didik menunjukkan hubungan sebab-akibat. Jawaban yang hanya menyebut istilah belum cukup menjadi bukti pemahaman.

Pada memprediksi pengaruh suhu terhadap kesetimbangan, diagnosis dapat memeriksa apakah peserta didik membedakan perubahan laju, keadaan setimbang, energi, dan komposisi. Kesalahan sering muncul ketika satu aturan dipindahkan ke konteks yang tidak sesuai.

Kasus menerangkan energi pada pemutusan dan pembentukan ikatan menunjukkan pentingnya representasi. Respons verbal, gambar, simbol, dan hitungan dapat tidak konsisten. Hasil instrumen sebaiknya digunakan untuk memilih demonstrasi, diskusi, contoh-noncontoh, atau tugas remediasi yang secara langsung menantang hubungan salah.

- tentukan proposisi ilmiah dan konsepsi alternatif;
- gunakan minimal dua bentuk representasi;
- minta alasan atau bukti ketika format memungkinkan;
- rancang tindak lanjut berdasarkan pola, bukan hanya skor.

Dalam praktik kelas, Open-Ended Diagnostic Test dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.11.13 Kelebihan, Keterbatasan, dan Potensi Bias

Kekuatan utama Open-Ended Diagnostic Test meliputi tidak membatasi respons pada opsi peneliti, berguna pada tahap eksplorasi, mengungkap pengetahuan parsial, dan menjadi sumber distraktor autentik. Kekuatan tersebut menjelaskan situasi ketika instrumen layak dipilih dibanding teknik lain. (Chi, 1997; Treagust, 1988; Taber, 2002; Gurel et al., 2015; AERA, APA, & NCME, 2014)

Keterbatasannya mencakup penskoran lambat, respons kosong ambigu, kemampuan menulis memengaruhi hasil, dan reliabilitas penilai perlu dijaga. Keterbatasan harus diperlakukan sebagai batas inferensi, bukan sekadar paragraf formal yang tidak memengaruhi analisis.

Potensi bias dapat berasal dari bahasa, pengalaman konteks, akses teknologi, kecemasan, gaya menggunakan skala, keterampilan menggambar, atau kemampuan menulis. Instrumen perlu diuji pada kelompok yang beragam dan ditinjau untuk fungsi berbeda antarkelompok bila data mencukupi.

Keputusan memilih instrumen sebaiknya mempertimbangkan kedalaman data, ukuran kelompok, waktu, keterampilan analisis, tujuan diagnosis, dan konsekuensi kesalahan klasifikasi. Tidak ada satu metode yang unggul untuk semua tujuan.

- Kekuatan: tidak membatasi respons pada opsi peneliti.
- Kekuatan: berguna pada tahap eksplorasi.
- Batas: penskoran lambat.
- Batas: respons kosong ambigu.

Dalam praktik kelas, Open-Ended Diagnostic Test dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

Aspek	Temuan	Respons Metodologis
Kekuatan	tidak membatasi respons pada opsi peneliti	Pilih ketika sesuai tujuan
Keterbatasan	penskoran lambat	Tambahkan bukti atau kontrol
Bias potensial	Bahasa, konteks, perangkat, gaya respons	Uji proses respons

6.11.14 Integrasi Digital, Etika, dan Tata Kelola Data

Integrasi digital Open-Ended Diagnostic Test dapat berupa jawaban teks, unggahan gambar, editor persamaan, dan pengodean semiotomatis yang tetap diaudit manusia. Teknologi seharusnya memperluas kualitas bukti, bukan hanya memindahkan lembar kertas ke layar. (Baker & Inventado, 2014; Romero & Ventura, 2020; AERA, APA, & NCME, 2014; Chi, 1997)

Platform perlu menyimpan versi instrumen, urutan respons, timestamp yang relevan, aturan skor, dan log perubahan. Namun, hanya data yang dibutuhkan untuk tujuan diagnosis yang boleh dikumpulkan; pengumpulan berlebihan meningkatkan risiko privasi tanpa manfaat ilmiah.

Aspek etika utama adalah bahwa kesalahan bahasa, ejaan, dan akses teknologi tidak boleh dikacaukan dengan miskonsepsi kimia. Peserta didik dan guru perlu mengetahui data apa yang dikumpulkan, untuk tujuan apa, siapa yang dapat mengakses, berapa lama disimpan, dan bagaimana kesalahan data diperbaiki.

Algoritma otomatis harus dapat diaudit. Kategori yang dihasilkan sistem perlu ditautkan ke respons mentah dan aturan keputusan. Kasus berisiko atau ambigu harus ditinjau manusia sebelum tindak lanjut penting dilakukan.

- gunakan identitas pseudonim dan akses berbasis peran;
- uji tampilan rumus, subskrip, grafik, dan gambar;
- sediakan opsi aksesibilitas dan alternatif perangkat;
- dokumentasikan algoritma serta perubahan versi.

Dalam praktik kelas, Open-Ended Diagnostic Test dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.11.15 Sintesis dan Pedoman Implementasi

Open-Ended Diagnostic Test bermanfaat apabila dipilih berdasarkan tujuan dan digunakan sesuai batasnya. Instrumen ini menghasilkan teks, simbol, perhitungan, gambar, dan argumentasi yang dihasilkan peserta didik, lalu data diolah melalui rubrik bertingkat untuk ketepatan konsep, kelengkapan, koherensi, penggunaan representasi, dan jenis kesalahan. (Chi, 1997; Treagust, 1988; Taber, 2002; AERA, APA, & NCME, 2014; Kane, 2013)

Rangkaian kerja yang disarankan dimulai dari analisis konsep, identifikasi konsepsi alternatif, desain tugas, telaah ahli, uji proses respons, uji lapangan, analisis, penetapan kategori, dan penyusunan tindak lanjut. Setiap tahap harus meninggalkan dokumentasi yang dapat diperiksa.

Hasil diagnosis tidak berhenti pada tabel prevalensi. Guru perlu memilih konsepsi prioritas berdasarkan frekuensi, keyakinan, posisi sebagai konsep prasyarat, dampak terhadap topik lanjutan, dan kemungkinan diperbaiki melalui pembelajaran.

Prinsip akhirnya ialah kehati-hatian inferensial: sebutkan apa yang benar-benar ditunjukkan data, bedakan bukti kuat dari indikasi, gunakan triangulasi ketika diperlukan, dan hindari label permanen pada peserta didik.

- selaraskan tujuan, konstruk, tugas, skor, kategori, dan tindakan;
- utamakan konsepsi yang berdampak pada pembelajaran lanjutan;
- gabungkan bukti bila keputusan memiliki konsekuensi besar;
- evaluasi kembali efektivitas tindak lanjut melalui diagnosis ulang.

Dalam praktik kelas, Open-Ended Diagnostic Test dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.12 Predict–Observe–Explain

6.12.1 Landasan Konseptual dan Kedudukan Instrumen

Predict–Observe–Explain menempati posisi tertentu dalam rangkaian bukti diagnosis, dari pengungkapan respons awal sampai penentuan tindak lanjut pembelajaran. Dalam bab ini, instrumen tersebut didefinisikan sebagai rangkaian kegiatan yang meminta peserta didik memprediksi hasil, mengamati fenomena atau data, kemudian menjelaskan kesesuaian atau perbedaan antara prediksi dan observasi. Definisi operasional ini menegaskan bentuk bukti yang dikumpulkan sekaligus membatasi jenis kesimpulan yang layak dibuat. (White & Gunstone, 1992; Kearney, 2004; Chi, 2009)

Klaim utamanya adalah bahwa instrumen ini mengungkap konsepsi awal dan cara peserta didik merekonsiliasi bukti dengan model mentalnya. Klaim tersebut harus dipisahkan dari anggapan bahwa setiap jawaban salah merupakan miskonsepsi. Kesalahan dapat berasal dari kurang pengetahuan, ketidakcermatan, kesulitan bahasa, salah hitung, atau ketidakfamiliaran dengan representasi.

Data pokok yang dihasilkan berupa prediksi, alasan, catatan observasi, perubahan penjelasan, dan respons terhadap ketidaksesuaian. Data itu baru memperoleh makna diagnostik setelah dikaitkan dengan peta konsep sasaran, kemungkinan konsepsi alternatif, dan aturan interpretasi yang telah ditetapkan sebelum analisis.

Dengan demikian, mutu diagnosis ditentukan oleh keselarasan antara tujuan, desain, pelaksanaan, analisis, dan keputusan yang diambil. Posisi terbaik Predict–Observe–Explain adalah sebagai bagian dari sistem asesmen berlapis yang dapat dipadukan dengan bukti lain ketika konsekuensi keputusan semakin besar.

- menentukan konsep dan proposisi yang hendak didiagnosis;
- menjelaskan perilaku respons yang dianggap sebagai bukti;
- membatasi klaim sesuai kekuatan instrumen;
- menautkan hasil dengan tindakan pembelajaran yang realistis.

Dalam praktik kelas, Predict–Observe–Explain dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.12.2 Perkembangan Historis dan Rasional Penggunaan

Perkembangan Predict–Observe–Explain tidak dapat dipisahkan dari perubahan pandangan tentang belajar sains. Ketika peserta didik tidak lagi dipandang sebagai penerima informasi yang kosong, instrumen mulai dirancang untuk mengungkap gagasan awal, pola penalaran, dan keyakinan yang dibawa ke kelas. (White & Gunstone, 1992; Kearney, 2004; Chi, 2009; Taber, 2002)

Pada tahap awal, penilaian banyak mengandalkan benar-salah atau skor total. Perkembangan penelitian konsepsi peserta didik kemudian mendorong penambahan alasan, keyakinan, tugas kinerja, rekaman proses, dan bukti sosial. Predict–Observe–Explain merupakan salah satu jawaban metodologis terhadap kebutuhan tersebut.

Dalam pendidikan kimia, kebutuhan diagnosis semakin kuat karena peserta didik harus mengoordinasikan fenomena makroskopik, model submikroskopik, simbol, grafik, dan perhitungan. Instrumen yang hanya menguji hafalan definisi sering gagal memperlihatkan cara representasi itu dihubungkan.

Sejarah penggunaannya juga menunjukkan bahwa format yang lebih kompleks tidak otomatis lebih baik. Setiap tambahan tier, rubrik, atau kanal data harus mempunyai fungsi yang jelas dan didukung bukti bahwa informasi tambahan benar-benar memperbaiki keputusan.

- pergeseran dari pengukuran skor menuju diagnosis pola pemahaman;
- penggunaan respons terbuka dan wawancara untuk menemukan konsepsi alternatif;
- pengembangan teknik kuantitatif agar diagnosis dapat dilakukan pada kelompok besar;
- integrasi data digital dan metakognitif pada perkembangan mutakhir.

Dalam praktik kelas, Predict–Observe–Explain dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.12.3 Konstruk, Klaim, dan Bukti Diagnosis

Konstruk yang dinilai melalui Predict–Observe–Explain bukan objek yang terlihat langsung. Yang diamati adalah prediksi, alasan, catatan observasi, perubahan penjelasan, dan respons terhadap ketidaksesuaian; dari sana dibuat inferensi tentang pemahaman konsep, miskonsepsi, pengetahuan parsial, atau ketidakpastian. (Chi, 2009; Taber, 2002; Messick, 1995; Kane, 2013)

Klaim diagnosis perlu dirumuskan sebagai argumen. Pertama, tugas harus memerlukan konsep sasaran. Kedua, respons harus diberi skor atau kode secara konsisten. Ketiga, pola skor harus mempunyai hubungan yang masuk akal dengan struktur pengetahuan. Keempat, penggunaan hasil harus menghasilkan keputusan yang proporsional.

Sumber ancaman utama adalah ketidakjelasan konstruk. Misalnya, peserta didik dapat gagal karena istilah yang ambigu, beban membaca, representasi yang tidak dikenal, atau prosedur matematika. Kegagalan semacam itu tidak boleh langsung diperlakukan sebagai miskonsepsi kimia.

Oleh sebab itu, setiap pengembang perlu menulis pernyataan klaim-bukti: instrumen mengukur apa, perilaku apa yang menjadi indikator, alternatif penjelasan apa yang mungkin, dan bukti tambahan apa yang diperlukan untuk memperkuat kesimpulan.

- konstruk utama: pemahaman konseptual melalui tugas diagnostik berbasis fenomena;
- bukti observabel: respons, alasan, tindakan, atau representasi;
- inferensi: profil konsepsi pada konten dan konteks tertentu;
- penggunaan: perencanaan pembelajaran, remediasi, atau penelitian.

Dalam praktik kelas, Predict–Observe–Explain dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.12.4 Struktur, Komponen, dan Alur Respons

Struktur Predict–Observe–Explain dibangun dari komponen yang saling bergantung. Komponen inti meliputi situasi fenomena, pertanyaan prediksi, alasan prediksi, observasi terarah, penjelasan pascaobservasi, dan prompt refleksi. Hilangnya satu komponen dapat mengubah makna data yang diperoleh. (White & Gunstone, 1992; Kearney, 2004; Chi, 2009; Taber, 2002)

Urutan penyajian harus ditetapkan secara sadar. Komponen yang muncul lebih awal dapat memberi petunjuk pada komponen berikutnya, sedangkan permintaan alasan atau keyakinan dapat mengubah cara peserta didik memeriksa jawabannya. Penguncian respons dan uji urutan diperlukan pada format digital.

Setiap komponen membutuhkan spesifikasi. Stimulus harus memuat informasi cukup, opsi atau prompt harus sejajar dengan indikator, skala harus memiliki jangkar, dan rubrik harus menjelaskan perbedaan kategori yang berdekatan.

Struktur yang baik memungkinkan peneliti menelusuri kembali setiap kategori hasil ke respons mentah. Transparansi ini penting agar guru tidak hanya menerima label otomatis tanpa memahami alasan di baliknya.

- situasi fenomena;
- pertanyaan prediksi;
- alasan prediksi;
- observasi terarah;
- penjelasan pascaobservasi;
- prompt refleksi;

Dalam praktik kelas, Predict–Observe–Explain dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

No.	Komponen	Fungsi Diagnostik
1	situasi fenomena	Bukti yang harus ditelusuri
2	pertanyaan prediksi	Bukti yang harus ditelusuri
3	alasan prediksi	Bukti yang harus ditelusuri
4	observasi terarah	Bukti yang harus ditelusuri

6.12.5 Desain Stimulus, Tugas, atau Pertanyaan

Desain stimulus pada Predict–Observe–Explain harus memunculkan penggunaan konsep, bukan sekadar pengenalan istilah. Situasi dapat berbentuk fenomena, gambar partikel, grafik, data eksperimen, persamaan, atau masalah kontekstual, tetapi seluruh informasi yang tidak relevan perlu dikendalikan. (Haladyna et al., 2002; White & Gunstone, 1992; Kearney, 2004; Chi, 2009)

Dalam kimia, satu stimulus sering menggabungkan beberapa tingkat representasi. Pengembang harus memutuskan apakah tujuan butir adalah menerjemahkan representasi, menjelaskan hubungan sebab-akibat, membedakan contoh dan noncontoh, atau menerapkan prinsip pada konteks baru.

Bahasa perlu ringkas tanpa menghilangkan syarat ilmiah. Kata seperti kuat, stabil, jenuh, spontan, atau setimbang mempunyai arti teknis dan sehari-hari yang berbeda. Ketidakjelasan istilah dapat menghasilkan kesalahan yang berasal dari bahasa, bukan struktur konsep.

Desain yang baik juga menghindari petunjuk format, pilihan terlalu panjang, negasi bertumpuk, gambar dekoratif, dan data yang tidak diperlukan. Setiap elemen harus mempunyai alasan diagnostik.

- gunakan konteks yang cukup dikenal tetapi tidak menghafalkan contoh;
- pastikan hanya satu interpretasi ilmiah yang menjadi target;
- uji semua simbol, satuan, grafik, dan gambar;
- sediakan ruang bagi respons yang tidak diprediksi bila tujuan masih eksploratif.

Dalam praktik kelas, Predict–Observe–Explain dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.12.6 Prosedur Pengembangan Secara Sistematis

Pengembangan Predict–Observe–Explain harus mengikuti alur yang terdokumentasi. Langkah yang lazim ialah memilih fenomena kritis, mengidentifikasi prediksi alternatif, menjamin observasi jelas, menulis prompt alasan, kemudian menyiapkan data cadangan, mengembangkan rubrik, uji coba, menganalisis pola perubahan. Urutan dapat berulang karena temuan uji coba sering memaksa revisi konstruk atau redaksi. (White & Gunstone, 1992; Kearney, 2004; Chi, 2009; Adams & Wieman, 2011; AERA, APA, & NCME, 2014)

Tahap konseptual dimulai dengan analisis kurikulum, literatur miskonsepsi, dan peta konsep. Tahap empiris awal menggunakan wawancara, respons terbuka, atau observasi untuk memastikan bahwa kategori yang dirancang benar-benar muncul pada populasi sasaran.

Setiap versi instrumen perlu diberi nomor, tanggal, dan catatan perubahan. Revisi harus menyebutkan masalah yang ditemukan, bukti yang digunakan, perubahan yang dilakukan, dan alasan mempertahankan atau menghapus butir.

Pengembangan dianggap belum selesai ketika hanya menghasilkan butir yang tampak baik. Instrumen perlu diuji pada peserta didik yang serupa dengan pengguna sasaran, lalu dianalisis secara konseptual dan empiris sebelum dipakai untuk keputusan penting.

- 1. memilih fenomena kritis.
- 2. mengidentifikasi prediksi alternatif.
- 3. menjamin observasi jelas.
- 4. menulis prompt alasan.
- 5. menyiapkan data cadangan.

Dalam praktik kelas, Predict–Observe–Explain dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.12.7 Pelaksanaan dan Pengumpulan Data

Pelaksanaan Predict–Observe–Explain harus distandardisasi secukupnya agar perbedaan hasil lebih mungkin berasal dari pemahaman, bukan variasi instruksi, waktu, alat, atau bantuan. Petunjuk perlu menjelaskan tujuan umum tanpa mengajarkan jawaban. (White & Gunstone, 1992; Kearney, 2004; AERA, APA, & NCME, 2014)

Kondisi administrasi mencakup waktu, urutan tugas, kesempatan bertanya, penggunaan kalkulator, akses tabel periodik, kualitas gambar, dan perangkat digital. Pada instrumen kualitatif, standardisasi berarti konsistensi fokus dan prinsip probing, bukan membaca naskah secara kaku.

Data yang dikumpulkan ialah prediksi, alasan, catatan observasi, perubahan penjelasan, dan respons terhadap ketidaksesuaian. Data mentah harus disimpan bersama metadata yang diperlukan, seperti kode butir, versi instrumen, waktu administrasi, dan kondisi khusus, tetapi hanya data yang relevan yang boleh dikumpulkan.

Sebelum administrasi luas, lakukan uji teknis. Pastikan tautan, urutan, penguncian respons, rekaman, penyimpanan, serta tampilan rumus dan subskrip bekerja. Gangguan teknis perlu dicatat sebagai bagian dari interpretasi.

- berikan instruksi yang sama dan mudah dipahami;
- catat kondisi yang dapat memengaruhi respons;
- hindari umpan balik sebelum seluruh respons diagnostik selesai;
- sediakan prosedur penanganan data tidak lengkap.

Dalam praktik kelas, Predict–Observe–Explain dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.12.8 Penskoran, Pengodean, dan Pengolahan Data

Penskoran Predict–Observe–Explain dapat diringkas sebagai kode hubungan prediksi-observasi-penjelasan, kualitas bukti, perubahan konsep, dan persistensi miskonsepsi. Aturan ini harus ditulis sebelum melihat hasil utama agar keputusan tidak disesuaikan secara oportunistik dengan pola data. (White & Gunstone, 1992; Kearney, 2004; Chi, 2009; Chi, 1997; AERA, APA, & NCME, 2014)

Skor total jarang cukup untuk diagnosis. Analisis perlu mempertahankan informasi tingkat butir, konsep, opsi, alasan, keyakinan, proses, atau representasi. Penggabungan data terlalu dini dapat menyembunyikan pola yang secara pedagogis penting.

Untuk respons terbuka atau kualitatif, buku kode harus mencantumkan definisi, contoh positif, contoh batas, noncontoh, dan aturan ketika satu respons memuat beberapa gagasan. Sebagian data perlu dinilai oleh lebih dari satu penilai.

Setiap transformasi data—misalnya dari respons ke kode, dari kode ke kategori, dan dari kategori ke persentase—harus dapat diaudit. Simpan respons asli dan keputusan pengodean agar kasus ambigu dapat ditinjau kembali.

- pisahkan skor ketepatan dari skor keyakinan atau kualitas alasan;
- laporkan kasus tidak lengkap dan aturan penanganannya;
- uji konsistensi penilai untuk data berkode;
- hindari mengubah data ordinal menjadi ukuran interval tanpa alasan.

Dalam praktik kelas, Predict–Observe–Explain dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.12.9 Interpretasi dan Penetapan Kategori

Interpretasi hasil Predict–Observe–Explain harus bergerak dari respons menuju pola, bukan langsung dari satu jawaban menuju label. Status konseptual menjadi lebih kuat apabila gagasan yang sama muncul pada beberapa butir, representasi, atau konteks. (White & Gunstone, 1992; Kearney, 2004; Chi, 2009; Taber, 2002; Kane, 2013)

Kategori dapat mencakup pemahaman ilmiah, pemahaman parsial, kurang pengetahuan, respons tidak konsisten, false positive, false negative, dan miskonsepsi. Tidak semua instrumen mampu membedakan seluruh kategori tersebut; skema harus mengikuti informasi yang benar-benar tersedia.

Ambang, seperti batas keyakinan atau persentase pemilihan, perlu memiliki rasional dan diuji melalui analisis sensitivitas. Perubahan kecil pada ambang dapat mengubah prevalensi kategori secara berarti.

Interpretasi yang bertanggung jawab menyatakan bahwa kategori berlaku pada konsep, bentuk tugas, waktu, dan kondisi tertentu. Hasil merupakan hipotesis kerja untuk pembelajaran, bukan diagnosis psikologis permanen.

- gunakan lebih dari satu butir untuk konsepsi penting;
- periksa konsistensi antara jawaban dan alasan;
- tandai kasus ambigu untuk konfirmasi;
- hubungkan kategori dengan rekomendasi yang berbeda.

Dalam praktik kelas, Predict–Observe–Explain dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.12.10 Validitas dan Kualitas Inferensi

Validitas Predict–Observe–Explain berkaitan dengan kualitas interpretasi dan penggunaan hasil, bukan sifat tetap yang dimiliki instrumen. Fokus khususnya adalah keterkaitan fenomena dengan konsep target dan bukti bahwa observasi mampu membedakan model penjelasan yang bersaing. (AERA, APA, & NCME, 2014; Messick, 1995; Kane, 2013; Haynes et al., 1995)

Bukti berdasarkan isi menilai keterwakilan domain dan akurasi konsep. Bukti proses respons memeriksa bagaimana peserta didik memahami tugas. Bukti struktur internal menilai kesesuaian pola respons. Bukti hubungan dengan variabel lain dan konsekuensi penggunaan melengkapi argumen.

Ancaman validitas dapat berasal dari konten yang kurang terwakili, tuntutan bahasa berlebihan, petunjuk opsi, rubrik yang tidak stabil, keyakinan yang tidak terkalibrasi, atau algoritma yang belajar dari label keliru.

Dokumen validasi perlu menghubungkan setiap klaim dengan bukti dan alternatif penjelasan. Menampilkan satu indeks tidak cukup; validasi adalah program argumentatif yang diperbarui setiap kali instrumen dipakai pada populasi atau tujuan baru.

- bukti isi dan kesesuaian kurikulum;
- bukti proses respons melalui wawancara kognitif;
- bukti struktur dan fungsi butir;
- bukti kegunaan serta konsekuensi keputusan.

Dalam praktik kelas, Predict–Observe–Explain dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.12.11 Reliabilitas, Keajegan, dan Analisis Empiris

Keajegan Predict–Observe–Explain harus dinilai sesuai jenis data. Pada respons objektif dapat digunakan konsistensi internal, stabilitas waktu, atau model pengukuran. Pada data kualitatif, perhatian diarahkan pada konsistensi pengodean, audit trail, dan ketepatan penerapan definisi kode. (Kuder & Richardson, 1937; Cronbach, 1951; Ding et al., 2006; Wren & Barbera, 2014; AERA, APA, & NCME, 2014)

Reliabilitas tinggi tidak otomatis menunjukkan diagnosis yang benar. Butir yang berulang-ulang mengukur hafalan dapat menghasilkan koefisien tinggi tetapi tidak mewakili pemahaman konseptual. Sebaliknya, konstruk yang multidimensi tidak selalu layak dipaksa menjadi satu skala.

Analisis empiris perlu memeriksa kesukaran, daya beda, fungsi opsi, pola data hilang, dan keterhubungan komponen. Untuk instrumen bertingkat, reliabilitas dapat dihitung pada setiap tier serta kombinasi yang mempunyai makna teoritis.

Ukuran sampel, distribusi kemampuan, dan homogenitas konten memengaruhi statistik. Karena itu, laporan harus menyatakan kondisi estimasi dan tidak menyalin batas angka sebagai aturan universal.

- pilih statistik yang sesuai dengan jenis skor;
- laporkan interval atau ketidakpastian bila tersedia;
- periksa butir secara substantif dan statistik;
- uji kembali setelah revisi atau perubahan platform.

Dalam praktik kelas, Predict–Observe–Explain dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.12.12 Contoh Penerapan dalam Pembelajaran Kimia

Penerapan Predict–Observe–Explain dalam kimia dapat dimulai dari konsep pengaruh katalis pada kesetimbangan. Pengembang perlu menuliskan konsep ilmiah target, miskonsepsi yang mungkin, bukti respons yang diharapkan, dan keputusan pembelajaran setelah hasil diperoleh. (White & Gunstone, 1992; Kearney, 2004; Chi, 2009; Taber, 2002; Nakhleh, 1992)

Contoh kedua adalah pelarutan dan perubahan suhu. Stimulus sebaiknya menghubungkan fenomena dengan model partikel atau persamaan, lalu meminta peserta didik menunjukkan hubungan sebab-akibat. Jawaban yang hanya menyebut istilah belum cukup menjadi bukti pemahaman.

Pada konduktivitas padatan dan larutan, diagnosis dapat memeriksa apakah peserta didik membedakan perubahan laju, keadaan setimbang, energi, dan komposisi. Kesalahan sering muncul ketika satu aturan dipindahkan ke konteks yang tidak sesuai.

Kasus perubahan warna indikator asam-basa menunjukkan pentingnya representasi. Respons verbal, gambar, simbol, dan hitungan dapat tidak konsisten. Hasil instrumen sebaiknya digunakan untuk memilih demonstrasi, diskusi, contoh-noncontoh, atau tugas remediasi yang secara langsung menantang hubungan salah.

- tentukan proposisi ilmiah dan konsepsi alternatif;
- gunakan minimal dua bentuk representasi;
- minta alasan atau bukti ketika format memungkinkan;
- rancang tindak lanjut berdasarkan pola, bukan hanya skor.

Dalam praktik kelas, Predict–Observe–Explain dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.12.13 Kelebihan, Keterbatasan, dan Potensi Bias

Kekuatan utama Predict–Observe–Explain meliputi mengaktifkan pengetahuan awal, menghubungkan fenomena dengan penjelasan, menunjukkan respons terhadap bukti, dan dapat menjadi diagnosis sekaligus pembelajaran. Kekuatan tersebut menjelaskan situasi ketika instrumen layak dipilih dibanding teknik lain. (White & Gunstone, 1992; Kearney, 2004; Chi, 2009; Gurel et al., 2015; AERA, APA, & NCME, 2014)

Keterbatasannya mencakup observasi dapat disalahartikan, hasil eksperimen yang tidak jelas mengganggu diagnosis, peserta didik dapat menyesuaikan jawaban tanpa perubahan konsep, dan kegiatan memerlukan waktu. Keterbatasan harus diperlakukan sebagai batas inferensi, bukan sekadar paragraf formal yang tidak memengaruhi analisis.

Potensi bias dapat berasal dari bahasa, pengalaman konteks, akses teknologi, kecemasan, gaya menggunakan skala, keterampilan menggambar, atau kemampuan menulis. Instrumen perlu diuji pada kelompok yang beragam dan ditinjau untuk fungsi berbeda antarkelompok bila data mencukupi.

Keputusan memilih instrumen sebaiknya mempertimbangkan kedalaman data, ukuran kelompok, waktu, keterampilan analisis, tujuan diagnosis, dan konsekuensi kesalahan klasifikasi. Tidak ada satu metode yang unggul untuk semua tujuan.

- Kekuatan: mengaktifkan pengetahuan awal.
- Kekuatan: menghubungkan fenomena dengan penjelasan.
- Batas: observasi dapat disalahartikan.
- Batas: hasil eksperimen yang tidak jelas mengganggu diagnosis.

Dalam praktik kelas, Predict–Observe–Explain dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

Aspek	Temuan	Respons Metodologis
Kekuatan	mengaktifkan pengetahuan awal	Pilih ketika sesuai tujuan
Keterbatasan	observasi dapat disalahartikan	Tambahkan bukti atau kontrol
Bias potensial	Bahasa, konteks, perangkat, gaya respons	Uji proses respons

6.12.14 Integrasi Digital, Etika, dan Tata Kelola Data

Integrasi digital Predict–Observe–Explain dapat berupa video bercabang, simulasi, sensor, dan formulir yang mengunci prediksi sebelum observasi dibuka. Teknologi seharusnya memperluas kualitas bukti, bukan hanya memindahkan lembar kertas ke layar. (Baker & Inventado, 2014; Romero & Ventura, 2020; AERA, APA, & NCME, 2014; White & Gunstone, 1992)

Platform perlu menyimpan versi instrumen, urutan respons, timestamp yang relevan, aturan skor, dan log perubahan. Namun, hanya data yang dibutuhkan untuk tujuan diagnosis yang boleh dikumpulkan; pengumpulan berlebihan meningkatkan risiko privasi tanpa manfaat ilmiah.

Aspek etika utama adalah bahwa fenomena harus aman; konflik kognitif tidak boleh berubah menjadi pengalaman memperlakukan peserta didik. Peserta didik dan guru perlu mengetahui data apa yang dikumpulkan, untuk tujuan apa, siapa yang dapat mengakses, berapa lama disimpan, dan bagaimana kesalahan data diperbaiki.

Algoritma otomatis harus dapat diaudit. Kategori yang dihasilkan sistem perlu ditautkan ke respons mentah dan aturan keputusan. Kasus berisiko atau ambigu harus ditinjau manusia sebelum tindak lanjut penting dilakukan.

- gunakan identitas pseudonim dan akses berbasis peran;
- uji tampilan rumus, subskrip, grafik, dan gambar;
- sediakan opsi aksesibilitas dan alternatif perangkat;
- dokumentasikan algoritma serta perubahan versi.

Dalam praktik kelas, Predict–Observe–Explain dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.12.15 Sintesis dan Pedoman Implementasi

Predict–Observe–Explain bermanfaat apabila dipilih berdasarkan tujuan dan digunakan sesuai batasnya. Instrumen ini menghasilkan prediksi, alasan, catatan observasi, perubahan penjelasan, dan respons terhadap ketidaksesuaian, lalu data diolah melalui kode hubungan prediksi-observasi-penjelasan, kualitas bukti, perubahan konsep, dan persistensi miskonsepsi. (White & Gunstone, 1992; Kearney, 2004; Chi, 2009; AERA, APA, & NCME, 2014; Kane, 2013)

Rangkaian kerja yang disarankan dimulai dari analisis konsep, identifikasi konsepsi alternatif, desain tugas, telaah ahli, uji proses respons, uji lapangan, analisis, penetapan kategori, dan penyusunan tindak lanjut. Setiap tahap harus meninggalkan dokumentasi yang dapat diperiksa.

Hasil diagnosis tidak berhenti pada tabel prevalensi. Guru perlu memilih konsepsi prioritas berdasarkan frekuensi, keyakinan, posisi sebagai konsep prasyarat, dampak terhadap topik lanjutan, dan kemungkinan diperbaiki melalui pembelajaran.

Prinsip akhirnya ialah kehati-hatian inferensial: sebutkan apa yang benar-benar ditunjukkan data, bedakan bukti kuat dari indikasi, gunakan triangulasi ketika diperlukan, dan hindari label permanen pada peserta didik.

- selaraskan tujuan, konstruk, tugas, skor, kategori, dan tindakan;
- utamakan konsepsi yang berdampak pada pembelajaran lanjutan;
- gabungkan bukti bila keputusan memiliki konsekuensi besar;
- evaluasi kembali efektivitas tindak lanjut melalui diagnosis ulang.

Dalam praktik kelas, Predict–Observe–Explain dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.13 Drawing Test dan Representasi Gambar

6.13.1 Landasan Konseptual dan Kedudukan Instrumen

Drawing Test dan Representasi Gambar menempati posisi tertentu dalam rangkaian bukti diagnosis, dari pengungkapan respons awal sampai penentuan tindak lanjut pembelajaran. Dalam bab ini, instrumen tersebut didefinisikan sebagai tugas yang meminta peserta didik menggambar entitas, proses, susunan partikel, atau hubungan spasial untuk mengeksternalisasi model mentalnya. Definisi operasional ini menegaskan bentuk bukti yang dikumpulkan sekaligus membatasi jenis kesimpulan yang layak dibuat. (Chandrasegaran et al., 2007; Bodner & Domin, 2000; Taber, 2002)

Klaim utamanya adalah bahwa instrumen ini mengungkap aspek konsepsi submikroskopik dan spasial yang sulit dinyatakan hanya melalui kata-kata. Klaim tersebut harus dipisahkan dari anggapan bahwa setiap jawaban salah merupakan miskonsepsi. Kesalahan dapat berasal dari kurang pengetahuan, ketidakcermatan, kesulitan bahasa, salah hitung, atau ketidakfamiliaran dengan representasi.

Data pokok yang dihasilkan berupa objek gambar, posisi, ukuran relatif, koneksi, simbol, label, dan penjelasan pendamping. Data itu baru memperoleh makna diagnostik setelah dikaitkan dengan peta konsep sasaran, kemungkinan konsepsi alternatif, dan aturan interpretasi yang telah ditetapkan sebelum analisis.

Dengan demikian, mutu diagnosis ditentukan oleh keselarasan antara tujuan, desain, pelaksanaan, analisis, dan keputusan yang diambil. Posisi terbaik Drawing Test dan Representasi Gambar adalah sebagai bagian dari sistem asesmen berlapis yang dapat dipadukan dengan bukti lain ketika konsekuensi keputusan semakin besar.

- menentukan konsep dan proposisi yang hendak didiagnosis;
- menjelaskan perilaku respons yang dianggap sebagai bukti;
- membatasi klaim sesuai kekuatan instrumen;
- menautkan hasil dengan tindakan pembelajaran yang realistis.

Dalam praktik kelas, Drawing Test dan Representasi Gambar dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.13.2 Perkembangan Historis dan Rasional Penggunaan

Perkembangan Drawing Test dan Representasi Gambar tidak dapat dipisahkan dari perubahan pandangan tentang belajar sains. Ketika peserta didik tidak lagi dipandang sebagai penerima informasi yang kosong, instrumen mulai dirancang untuk mengungkap gagasan awal, pola penalaran, dan keyakinan yang dibawa ke kelas. (Chandrasegaran et al., 2007; Bodner & Domin, 2000; Taber, 2002; AERA, APA, & NCME, 2014)

Pada tahap awal, penilaian banyak mengandalkan benar-salah atau skor total. Perkembangan penelitian konsepsi peserta didik kemudian mendorong penambahan alasan, keyakinan, tugas kinerja, rekaman proses, dan bukti sosial. Drawing Test dan Representasi Gambar merupakan salah satu jawaban metodologis terhadap kebutuhan tersebut.

Dalam pendidikan kimia, kebutuhan diagnosis semakin kuat karena peserta didik harus mengoordinasikan fenomena makroskopik, model submikroskopik, simbol, grafik, dan perhitungan. Instrumen yang hanya menguji hafalan definisi sering gagal memperlihatkan cara representasi itu dihubungkan.

Sejarah penggunaannya juga menunjukkan bahwa format yang lebih kompleks tidak otomatis lebih baik. Setiap tambahan tier, rubrik, atau kanal data harus mempunyai fungsi yang jelas dan didukung bukti bahwa informasi tambahan benar-benar memperbaiki keputusan.

- pergeseran dari pengukuran skor menuju diagnosis pola pemahaman;
- penggunaan respons terbuka dan wawancara untuk menemukan konsepsi alternatif;
- pengembangan teknik kuantitatif agar diagnosis dapat dilakukan pada kelompok besar;
- integrasi data digital dan metakognitif pada perkembangan mutakhir.

Dalam praktik kelas, Drawing Test dan Representasi Gambar dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.13.3 Konstruk, Klaim, dan Bukti Diagnosis

Konstruk yang dinilai melalui Drawing Test dan Representasi Gambar bukan objek yang terlihat langsung. Yang diamati adalah objek gambar, posisi, ukuran relatif, koneksi, simbol, label, dan penjelasan pendamping; dari sana dibuat inferensi tentang pemahaman konsep, miskonsepsi, pengetahuan parsial, atau ketidakpastian. (Taber, 2002; AERA, APA, & NCME, 2014; Messick, 1995; Kane, 2013)

Klaim diagnosis perlu dirumuskan sebagai argumen. Pertama, tugas harus memerlukan konsep sasaran. Kedua, respons harus diberi skor atau kode secara konsisten. Ketiga, pola skor harus mempunyai hubungan yang masuk akal dengan struktur pengetahuan. Keempat, penggunaan hasil harus menghasilkan keputusan yang proporsional.

Sumber ancaman utama adalah ketidakjelasan konstruk. Misalnya, peserta didik dapat gagal karena istilah yang ambigu, beban membaca, representasi yang tidak dikenal, atau prosedur matematika. Kegagalan semacam itu tidak boleh langsung diperlakukan sebagai miskonsepsi kimia.

Oleh sebab itu, setiap pengembang perlu menulis pernyataan klaim-bukti: instrumen mengukur apa, perilaku apa yang menjadi indikator, alternatif penjelasan apa yang mungkin, dan bukti tambahan apa yang diperlukan untuk memperkuat kesimpulan.

- konstruk utama: pemahaman konseptual melalui diagnosis visual dan representasional;
- bukti observabel: respons, alasan, tindakan, atau representasi;
- inferensi: profil konsepsi pada konten dan konteks tertentu;
- penggunaan: perencanaan pembelajaran, remediasi, atau penelitian.

Dalam praktik kelas, Drawing Test dan Representasi Gambar dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.13.4 Struktur, Komponen, dan Alur Respons

Struktur Drawing Test dan Representasi Gambar dibangun dari komponen yang saling bergantung. Komponen inti meliputi prompt gambar, batas konteks, label atau legenda, permintaan penjelasan, rubrik fitur kritis, dan kode representasional. Hilangnya satu komponen dapat mengubah makna data yang diperoleh. (Chandrasegaran et al., 2007; Bodner & Domin, 2000; Taber, 2002; AERA, APA, & NCME, 2014)

Urutan penyajian harus ditetapkan secara sadar. Komponen yang muncul lebih awal dapat memberi petunjuk pada komponen berikutnya, sedangkan permintaan alasan atau keyakinan dapat mengubah cara peserta didik memeriksa jawabannya. Penguncian respons dan uji urutan diperlukan pada format digital.

Setiap komponen membutuhkan spesifikasi. Stimulus harus memuat informasi cukup, opsi atau prompt harus sejajar dengan indikator, skala harus memiliki jangkar, dan rubrik harus menjelaskan perbedaan kategori yang berdekatan.

Struktur yang baik memungkinkan peneliti menelusuri kembali setiap kategori hasil ke respons mentah. Transparansi ini penting agar guru tidak hanya menerima label otomatis tanpa memahami alasan di baliknya.

- prompt gambar;
- batas konteks;
- label atau legenda;
- permintaan penjelasan;
- rubrik fitur kritis;
- kode representasional;

Dalam praktik kelas, Drawing Test dan Representasi Gambar dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

No.	Komponen	Fungsi Diagnostik
1	prompt gambar	Bukti yang harus ditelusuri
2	batas konteks	Bukti yang harus ditelusuri
3	label atau legenda	Bukti yang harus ditelusuri
4	permintaan penjelasan	Bukti yang harus ditelusuri

6.13.5 Desain Stimulus, Tugas, atau Pertanyaan

Desain stimulus pada Drawing Test dan Representasi Gambar harus memunculkan penggunaan konsep, bukan sekadar pengenalan istilah. Situasi dapat berbentuk fenomena, gambar partikel, grafik, data eksperimen, persamaan, atau masalah kontekstual, tetapi seluruh informasi yang tidak relevan perlu dikendalikan. (Haladyna et al., 2002; Chandrasegaran et al., 2007; Bodner & Domin, 2000; Taber, 2002)

Dalam kimia, satu stimulus sering menggabungkan beberapa tingkat representasi. Pengembang harus memutuskan apakah tujuan butir adalah menerjemahkan representasi, menjelaskan hubungan sebab-akibat, membedakan contoh dan noncontoh, atau menerapkan prinsip pada konteks baru.

Bahasa perlu ringkas tanpa menghilangkan syarat ilmiah. Kata seperti kuat, stabil, jenuh, spontan, atau setimbang mempunyai arti teknis dan sehari-hari yang berbeda. Ketidakjelasan istilah dapat menghasilkan kesalahan yang berasal dari bahasa, bukan struktur konsep.

Desain yang baik juga menghindari petunjuk format, pilihan terlalu panjang, negasi bertumpuk, gambar dekoratif, dan data yang tidak diperlukan. Setiap elemen harus mempunyai alasan diagnostik.

- gunakan konteks yang cukup dikenal tetapi tidak menghafalkan contoh;
- pastikan hanya satu interpretasi ilmiah yang menjadi target;
- uji semua simbol, satuan, grafik, dan gambar;
- sediakan ruang bagi respons yang tidak diprediksi bila tujuan masih eksploratif.

Dalam praktik kelas, Drawing Test dan Representasi Gambar dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.13.6 Prosedur Pengembangan Secara Sistematis

Pengembangan Drawing Test dan Representasi Gambar harus mengikuti alur yang terdokumentasi. Langkah yang lazim ialah menentukan fitur konsep, menulis prompt, menyediakan aturan label, mengumpulkan gambar awal, kemudian menyusun kode, melakukan penilaian ganda, menguji konsistensi, mengintegrasikan dengan wawancara. Urutan dapat berulang karena temuan uji coba sering memaksa revisi konstruk atau redaksi. (Chandrasegaran et al., 2007; Bodner & Domin, 2000; Taber, 2002; Adams & Wieman, 2011; AERA, APA, & NCME, 2014)

Tahap konseptual dimulai dengan analisis kurikulum, literatur miskonsepsi, dan peta konsep. Tahap empiris awal menggunakan wawancara, respons terbuka, atau observasi untuk memastikan bahwa kategori yang dirancang benar-benar muncul pada populasi sasaran.

Setiap versi instrumen perlu diberi nomor, tanggal, dan catatan perubahan. Revisi harus menyebutkan masalah yang ditemukan, bukti yang digunakan, perubahan yang dilakukan, dan alasan mempertahankan atau menghapus butir.

Pengembangan dianggap belum selesai ketika hanya menghasilkan butir yang tampak baik. Instrumen perlu diuji pada peserta didik yang serupa dengan pengguna sasaran, lalu dianalisis secara konseptual dan empiris sebelum dipakai untuk keputusan penting.

- 1. menentukan fitur konsep.
- 2. menulis prompt.
- 3. menyediakan aturan label.
- 4. mengumpulkan gambar awal.
- 5. menyusun kode.

Dalam praktik kelas, Drawing Test dan Representasi Gambar dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.13.7 Pelaksanaan dan Pengumpulan Data

Pelaksanaan Drawing Test dan Representasi Gambar harus distandardisasi secukupnya agar perbedaan hasil lebih mungkin berasal dari pemahaman, bukan variasi instruksi, waktu, alat, atau bantuan. Petunjuk perlu menjelaskan tujuan umum tanpa mengajarkan jawaban. (Chandrasegaran et al., 2007; Bodner & Domin, 2000; AERA, APA, & NCME, 2014)

Kondisi administrasi mencakup waktu, urutan tugas, kesempatan bertanya, penggunaan kalkulator, akses tabel periodik, kualitas gambar, dan perangkat digital. Pada instrumen kualitatif, standardisasi berarti konsistensi fokus dan prinsip probing, bukan membaca naskah secara kaku.

Data yang dikumpulkan ialah objek gambar, posisi, ukuran relatif, koneksi, simbol, label, dan penjelasan pendamping. Data mentah harus disimpan bersama metadata yang diperlukan, seperti kode butir, versi instrumen, waktu administrasi, dan kondisi khusus, tetapi hanya data yang relevan yang boleh dikumpulkan.

Sebelum administrasi luas, lakukan uji teknis. Pastikan tautan, urutan, penguncian respons, rekaman, penyimpanan, serta tampilan rumus dan subskrip bekerja. Gangguan teknis perlu dicatat sebagai bagian dari interpretasi.

- berikan instruksi yang sama dan mudah dipahami;
- catat kondisi yang dapat memengaruhi respons;
- hindari umpan balik sebelum seluruh respons diagnostik selesai;
- sediakan prosedur penanganan data tidak lengkap.

Dalam praktik kelas, Drawing Test dan Representasi Gambar dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.13.8 Penskoran, Pengodean, dan Pengolahan Data

Penskoran Drawing Test dan Representasi Gambar dapat diringkas sebagai analisis fitur, proposisi visual, kesalahan ontologis, konsistensi gambar-teks, dan tingkat representasi. Aturan ini harus ditulis sebelum melihat hasil utama agar keputusan tidak disesuaikan secara oportunistik dengan pola data. (Chandrasegaran et al., 2007; Bodner & Domin, 2000; Taber, 2002; Chi, 1997; AERA, APA, & NCME, 2014)

Skor total jarang cukup untuk diagnosis. Analisis perlu mempertahankan informasi tingkat butir, konsep, opsi, alasan, keyakinan, proses, atau representasi. Penggabungan data terlalu dini dapat menyembunyikan pola yang secara pedagogis penting.

Untuk respons terbuka atau kualitatif, buku kode harus mencantumkan definisi, contoh positif, contoh batas, noncontoh, dan aturan ketika satu respons memuat beberapa gagasan. Sebagian data perlu dinilai oleh lebih dari satu penilai.

Setiap transformasi data—misalnya dari respons ke kode, dari kode ke kategori, dan dari kategori ke persentase—harus dapat diaudit. Simpan respons asli dan keputusan pengodean agar kasus ambigu dapat ditinjau kembali.

- pisahkan skor ketepatan dari skor keyakinan atau kualitas alasan;
- laporkan kasus tidak lengkap dan aturan penanganannya;
- uji konsistensi penilai untuk data berkode;
- hindari mengubah data ordinal menjadi ukuran interval tanpa alasan.

Dalam praktik kelas, Drawing Test dan Representasi Gambar dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.13.9 Interpretasi dan Penetapan Kategori

Interpretasi hasil Drawing Test dan Representasi Gambar harus bergerak dari respons menuju pola, bukan langsung dari satu jawaban menuju label. Status konseptual menjadi lebih kuat apabila gagasan yang sama muncul pada beberapa butir, representasi, atau konteks. (Chandrasegaran et al., 2007; Bodner & Domin, 2000; Taber, 2002; AERA, APA, & NCME, 2014; Kane, 2013)

Kategori dapat mencakup pemahaman ilmiah, pemahaman parsial, kurang pengetahuan, respons tidak konsisten, false positive, false negative, dan miskonsepsi. Tidak semua instrumen mampu membedakan seluruh kategori tersebut; skema harus mengikuti informasi yang benar-benar tersedia.

Ambang, seperti batas keyakinan atau persentase pemilihan, perlu memiliki rasional dan diuji melalui analisis sensitivitas. Perubahan kecil pada ambang dapat mengubah prevalensi kategori secara berarti.

Interpretasi yang bertanggung jawab menyatakan bahwa kategori berlaku pada konsep, bentuk tugas, waktu, dan kondisi tertentu. Hasil merupakan hipotesis kerja untuk pembelajaran, bukan diagnosis psikologis permanen.

- gunakan lebih dari satu butir untuk konsepsi penting;
- periksa konsistensi antara jawaban dan alasan;
- tandai kasus ambigu untuk konfirmasi;
- hubungkan kategori dengan rekomendasi yang berbeda.

Dalam praktik kelas, Drawing Test dan Representasi Gambar dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.13.10 Validitas dan Kualitas Inferensi

Validitas Drawing Test dan Representasi Gambar berkaitan dengan kualitas interpretasi dan penggunaan hasil, bukan sifat tetap yang dimiliki instrumen. Fokus khususnya adalah bukti bahwa fitur visual berkaitan dengan konsep sasaran dan tidak semata-mata mencerminkan keterampilan artistik. (AERA, APA, & NCME, 2014; Messick, 1995; Kane, 2013; Haynes et al., 1995)

Bukti berdasarkan isi menilai keterwakilan domain dan akurasi konsep. Bukti proses respons memeriksa bagaimana peserta didik memahami tugas. Bukti struktur internal menilai kesesuaian pola respons. Bukti hubungan dengan variabel lain dan konsekuensi penggunaan melengkapi argumen.

Ancaman validitas dapat berasal dari konten yang kurang terwakili, tuntutan bahasa berlebihan, petunjuk opsi, rubrik yang tidak stabil, keyakinan yang tidak terkalibrasi, atau algoritma yang belajar dari label keliru.

Dokumen validasi perlu menghubungkan setiap klaim dengan bukti dan alternatif penjelasan. Menampilkan satu indeks tidak cukup; validasi adalah program argumentatif yang diperbarui setiap kali instrumen dipakai pada populasi atau tujuan baru.

- bukti isi dan kesesuaian kurikulum;
- bukti proses respons melalui wawancara kognitif;
- bukti struktur dan fungsi butir;
- bukti kegunaan serta konsekuensi keputusan.

Dalam praktik kelas, Drawing Test dan Representasi Gambar dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.13.11 Reliabilitas, Keajegan, dan Analisis Empiris

Keajegan Drawing Test dan Representasi Gambar harus dinilai sesuai jenis data. Pada respons objektif dapat digunakan konsistensi internal, stabilitas waktu, atau model pengukuran. Pada data kualitatif, perhatian diarahkan pada konsistensi pengodean, audit trail, dan ketepatan penerapan definisi kode. (Kuder & Richardson, 1937; Cronbach, 1951; Ding et al., 2006; Wren & Barbera, 2014; AERA, APA, & NCME, 2014)

Reliabilitas tinggi tidak otomatis menunjukkan diagnosis yang benar. Butir yang berulang-ulang mengukur hafalan dapat menghasilkan koefisien tinggi tetapi tidak mewakili pemahaman konseptual. Sebaliknya, konstruk yang multidimensi tidak selalu layak dipaksa menjadi satu skala.

Analisis empiris perlu memeriksa kesukaran, daya beda, fungsi opsi, pola data hilang, dan keterhubungan komponen. Untuk instrumen bertingkat, reliabilitas dapat dihitung pada setiap tier serta kombinasi yang mempunyai makna teoritis.

Ukuran sampel, distribusi kemampuan, dan homogenitas konten memengaruhi statistik. Karena itu, laporan harus menyatakan kondisi estimasi dan tidak menyalin batas angka sebagai aturan universal.

- pilih statistik yang sesuai dengan jenis skor;
- laporkan interval atau ketidakpastian bila tersedia;
- periksa butir secara substantif dan statistik;
- uji kembali setelah revisi atau perubahan platform.

Dalam praktik kelas, Drawing Test dan Representasi Gambar dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.13.12 Contoh Penerapan dalam Pembelajaran Kimia

Penerapan Drawing Test dan Representasi Gambar dalam kimia dapat dimulai dari konsep partikel gas dalam wadah. Pengembang perlu menuliskan konsep ilmiah target, miskonsepsi yang mungkin, bukti respons yang diharapkan, dan keputusan pembelajaran setelah hasil diperoleh. (Chandrasegaran et al., 2007; Bodner & Domin, 2000; Taber, 2002; AERA, APA, & NCME, 2014; Nakhleh, 1992)

Contoh kedua adalah ion pada kristal dan larutan. Stimulus sebaiknya menghubungkan fenomena dengan model partikel atau persamaan, lalu meminta peserta didik menunjukkan hubungan sebab-akibat. Jawaban yang hanya menyebut istilah belum cukup menjadi bukti pemahaman.

Pada distribusi elektron dan orbital, diagnosis dapat memeriksa apakah peserta didik membedakan perubahan laju, keadaan setimbang, energi, dan komposisi. Kesalahan sering muncul ketika satu aturan dipindahkan ke konteks yang tidak sesuai.

Kasus mekanisme tumbukan pada laju reaksi menunjukkan pentingnya representasi. Respons verbal, gambar, simbol, dan hitungan dapat tidak konsisten. Hasil instrumen sebaiknya digunakan untuk memilih demonstrasi, diskusi, contoh-noncontoh, atau tugas remediasi yang secara langsung menantang hubungan salah.

- tentukan proposisi ilmiah dan konsepsi alternatif;
- gunakan minimal dua bentuk representasi;
- minta alasan atau bukti ketika format memungkinkan;
- rancang tindak lanjut berdasarkan pola, bukan hanya skor.

Dalam praktik kelas, Drawing Test dan Representasi Gambar dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.13.13 Kelebihan, Keterbatasan, dan Potensi Bias

Kekuatan utama Drawing Test dan Representasi Gambar meliputi efektif untuk model partikel, mengungkap gambaran literal terhadap model, dapat digunakan lintas kemampuan verbal, dan mendukung diskusi representasi. Kekuatan tersebut menjelaskan situasi ketika instrumen layak dipilih dibanding teknik lain. (Chandrasegaran et al., 2007; Bodner & Domin, 2000; Taber, 2002; Gurel et al., 2015; AERA, APA, & NCME, 2014)

Keterbatasannya mencakup kemampuan menggambar dapat mengaburkan konsep, gambar tanpa penjelasan ambigu, konvensi visual berbeda, dan rubrik membutuhkan contoh jangkar. Keterbatasan harus diperlakukan sebagai batas inferensi, bukan sekadar paragraf formal yang tidak memengaruhi analisis.

Potensi bias dapat berasal dari bahasa, pengalaman konteks, akses teknologi, kecemasan, gaya menggunakan skala, keterampilan menggambar, atau kemampuan menulis. Instrumen perlu diuji pada kelompok yang beragam dan ditinjau untuk fungsi berbeda antarkelompok bila data mencukupi.

Keputusan memilih instrumen sebaiknya mempertimbangkan kedalaman data, ukuran kelompok, waktu, keterampilan analisis, tujuan diagnosis, dan konsekuensi kesalahan klasifikasi. Tidak ada satu metode yang unggul untuk semua tujuan.

- Kekuatan: efektif untuk model partikel.
- Kekuatan: mengungkap gambaran literal terhadap model.
- Batas: kemampuan menggambar dapat mengaburkan konsep.
- Batas: gambar tanpa penjelasan ambigu.

Dalam praktik kelas, Drawing Test dan Representasi Gambar dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

Aspek	Temuan	Respons Metodologis
Kekuatan	efektif untuk model partikel	Pilih ketika sesuai tujuan
Keterbatasan	kemampuan menggambar dapat mengaburkan konsep	Tambahkan bukti atau kontrol
Bias potensial	Bahasa, konteks, perangkat, gaya respons	Uji proses respons

6.13.14 Integrasi Digital, Etika, dan Tata Kelola Data

Integrasi digital Drawing Test dan Representasi Gambar dapat berupa kanvas gambar, stylus, pilihan simbol, dan rekaman urutan goresan untuk melihat proses konstruksi. Teknologi seharusnya memperluas kualitas bukti, bukan hanya memindahkan lembar kertas ke layar. (Baker & Inventado, 2014; Romero & Ventura, 2020; AERA, APA, & NCME, 2014; Chandrasegaran et al., 2007)

Platform perlu menyimpan versi instrumen, urutan respons, timestamp yang relevan, aturan skor, dan log perubahan. Namun, hanya data yang dibutuhkan untuk tujuan diagnosis yang boleh dikumpulkan; pengumpulan berlebihan meningkatkan risiko privasi tanpa manfaat ilmiah.

Aspek etika utama adalah bahwa peserta didik harus diberi tahu bahwa kualitas seni tidak dinilai; akses alat bantu perlu setara. Peserta didik dan guru perlu mengetahui data apa yang dikumpulkan, untuk tujuan apa, siapa yang dapat mengakses, berapa lama disimpan, dan bagaimana kesalahan data diperbaiki.

Algoritma otomatis harus dapat diaudit. Kategori yang dihasilkan sistem perlu ditautkan ke respons mentah dan aturan keputusan. Kasus berisiko atau ambigu harus ditinjau manusia sebelum tindak lanjut penting dilakukan.

- gunakan identitas pseudonim dan akses berbasis peran;
- uji tampilan rumus, subskrip, grafik, dan gambar;
- sediakan opsi aksesibilitas dan alternatif perangkat;
- dokumentasikan algoritma serta perubahan versi.

Dalam praktik kelas, Drawing Test dan Representasi Gambar dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.13.15 Sintesis dan Pedoman Implementasi

Drawing Test dan Representasi Gambar bermanfaat apabila dipilih berdasarkan tujuan dan digunakan sesuai batasnya. Instrumen ini menghasilkan objek gambar, posisi, ukuran relatif, koneksi, simbol, label, dan penjelasan pendamping, lalu data diolah melalui analisis fitur, proposisi visual, kesalahan ontologis, konsistensi gambar-teks, dan tingkat representasi. (Chandrasegaran et al., 2007; Bodner & Domin, 2000; Taber, 2002; AERA, APA, & NCME, 2014; Kane, 2013)

Rangkaian kerja yang disarankan dimulai dari analisis konsep, identifikasi konsepsi alternatif, desain tugas, telaah ahli, uji proses respons, uji lapangan, analisis, penetapan kategori, dan penyusunan tindak lanjut. Setiap tahap harus meninggalkan dokumentasi yang dapat diperiksa.

Hasil diagnosis tidak berhenti pada tabel prevalensi. Guru perlu memilih konsepsi prioritas berdasarkan frekuensi, keyakinan, posisi sebagai konsep prasyarat, dampak terhadap topik lanjutan, dan kemungkinan diperbaiki melalui pembelajaran.

Prinsip akhirnya ialah kehati-hatian inferensial: sebutkan apa yang benar-benar ditunjukkan data, bedakan bukti kuat dari indikasi, gunakan triangulasi ketika diperlukan, dan hindari label permanen pada peserta didik.

- selaraskan tujuan, konstruk, tugas, skor, kategori, dan tindakan;
- utamakan konsepsi yang berdampak pada pembelajaran lanjutan;
- gabungkan bukti bila keputusan memiliki konsekuensi besar;
- evaluasi kembali efektivitas tindak lanjut melalui diagnosis ulang.

Dalam praktik kelas, Drawing Test dan Representasi Gambar dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.14 Analisis Kesalahan Penyelesaian Masalah

6.14.1 Landasan Konseptual dan Kedudukan Instrumen

Analisis Kesalahan Penyelesaian Masalah menempati posisi tertentu dalam rangkaian bukti diagnosis, dari pengungkapan respons awal sampai penentuan tindak lanjut pembelajaran. Dalam bab ini, instrumen tersebut didefinisikan sebagai pemeriksaan sistematis terhadap langkah, representasi, operasi, dan keputusan peserta didik ketika menyelesaikan masalah untuk membedakan kesalahan konsep, prosedur, simbol, dan hitungan. Definisi operasional ini menegaskan bentuk bukti yang dikumpulkan sekaligus membatasi jenis kesimpulan yang layak dibuat. (Nurrenbern & Pickering, 1987; Bodner & Domin, 2000; Ericsson & Simon, 1993)

Klaim utamanya adalah bahwa instrumen ini menunjukkan titik tempat penalaran menyimpang serta hubungan antara pengetahuan konseptual dan prosedural. Klaim tersebut harus dipisahkan dari anggapan bahwa setiap jawaban salah merupakan miskonsepsi. Kesalahan dapat berasal dari kurang pengetahuan, ketidakcermatan, kesulitan bahasa, salah hitung, atau ketidakfamiliaran dengan representasi.

Data pokok yang dihasilkan berupa urutan langkah, persamaan, gambar, unit, substitusi, revisi, dan jawaban akhir. Data itu baru memperoleh makna diagnostik setelah dikaitkan dengan peta konsep sasaran, kemungkinan konsepsi alternatif, dan aturan interpretasi yang telah ditetapkan sebelum analisis.

Pembedaan ini penting karena satu pola respons dapat memiliki lebih dari satu penjelasan yang masuk akal. Posisi terbaik Analisis Kesalahan Penyelesaian Masalah adalah sebagai bagian dari sistem asesmen berlapis yang dapat dipadukan dengan bukti lain ketika konsekuensi keputusan semakin besar.

- menentukan konsep dan proposisi yang hendak didiagnosis;
- menjelaskan perilaku respons yang dianggap sebagai bukti;
- membatasi klaim sesuai kekuatan instrumen;
- menautkan hasil dengan tindakan pembelajaran yang realistis.

Dalam praktik kelas, Analisis Kesalahan Penyelesaian Masalah dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.14.2 Perkembangan Historis dan Rasional Penggunaan

Perkembangan Analisis Kesalahan Penyelesaian Masalah tidak dapat dipisahkan dari perubahan pandangan tentang belajar sains. Ketika peserta didik tidak lagi dipandang sebagai penerima informasi yang kosong, instrumen mulai dirancang untuk mengungkap gagasan awal, pola penalaran, dan keyakinan yang dibawa ke kelas. (Nurrenbern & Pickering, 1987; Bodner & Domin, 2000; Ericsson & Simon, 1993; Chi, 1997)

Pada tahap awal, penilaian banyak mengandalkan benar-salah atau skor total. Perkembangan penelitian konsepsi peserta didik kemudian mendorong penambahan alasan, keyakinan, tugas kinerja, rekaman proses, dan bukti sosial. Analisis Kesalahan Penyelesaian Masalah merupakan salah satu jawaban metodologis terhadap kebutuhan tersebut.

Dalam pendidikan kimia, kebutuhan diagnosis semakin kuat karena peserta didik harus mengoordinasikan fenomena makroskopik, model submikroskopik, simbol, grafik, dan perhitungan. Instrumen yang hanya menguji hafalan definisi sering gagal memperlihatkan cara representasi itu dihubungkan.

Sejarah penggunaannya juga menunjukkan bahwa format yang lebih kompleks tidak otomatis lebih baik. Setiap tambahan tier, rubrik, atau kanal data harus mempunyai fungsi yang jelas dan didukung bukti bahwa informasi tambahan benar-benar memperbaiki keputusan.

- pergeseran dari pengukuran skor menuju diagnosis pola pemahaman;
- penggunaan respons terbuka dan wawancara untuk menemukan konsepsi alternatif;
- pengembangan teknik kuantitatif agar diagnosis dapat dilakukan pada kelompok besar;
- integrasi data digital dan metakognitif pada perkembangan mutakhir.

Dalam praktik kelas, Analisis Kesalahan Penyelesaian Masalah dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.14.3 Konstruk, Klaim, dan Bukti Diagnosis

Konstruk yang dinilai melalui Analisis Kesalahan Penyelesaian Masalah bukan objek yang terlihat langsung. Yang diamati adalah urutan langkah, persamaan, gambar, unit, substitusi, revisi, dan jawaban akhir; dari sana dibuat inferensi tentang pemahaman konsep, miskonsepsi, pengetahuan parsial, atau ketidakpastian. (Chi, 1997; AERA, APA, & NCME, 2014; Messick, 1995; Kane, 2013)

Klaim diagnosis perlu dirumuskan sebagai argumen. Pertama, tugas harus memerlukan konsep sasaran. Kedua, respons harus diberi skor atau kode secara konsisten. Ketiga, pola skor harus mempunyai hubungan yang masuk akal dengan struktur pengetahuan. Keempat, penggunaan hasil harus menghasilkan keputusan yang proporsional.

Sumber ancaman utama adalah ketidakjelasan konstruk. Misalnya, peserta didik dapat gagal karena istilah yang ambigu, beban membaca, representasi yang tidak dikenal, atau prosedur matematika. Kegagalan semacam itu tidak boleh langsung diperlakukan sebagai miskonsepsi kimia.

Oleh sebab itu, setiap pengembang perlu menulis pernyataan klaim-bukti: instrumen mengukur apa, perilaku apa yang menjadi indikator, alternatif penjelasan apa yang mungkin, dan bukti tambahan apa yang diperlukan untuk memperkuat kesimpulan.

- konstruk utama: pemahaman konseptual melalui diagnosis melalui jejak solusi;
- bukti observabel: respons, alasan, tindakan, atau representasi;
- inferensi: profil konsepsi pada konten dan konteks tertentu;
- penggunaan: perencanaan pembelajaran, remediasi, atau penelitian.

Dalam praktik kelas, Analisis Kesalahan Penyelesaian Masalah dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.14.4 Struktur, Komponen, dan Alur Respons

Struktur Analisis Kesalahan Penyelesaian Masalah dibangun dari komponen yang saling bergantung. Komponen inti meliputi masalah diagnostik, lembar langkah, kode jenis kesalahan, pohon keputusan, rubrik strategi, dan wawancara konfirmasi. Hilangnya satu komponen dapat mengubah makna data yang diperoleh. (Nurrenbern & Pickering, 1987; Bodner & Domin, 2000; Ericsson & Simon, 1993; Chi, 1997)

Urutan penyajian harus ditetapkan secara sadar. Komponen yang muncul lebih awal dapat memberi petunjuk pada komponen berikutnya, sedangkan permintaan alasan atau keyakinan dapat mengubah cara peserta didik memeriksa jawabannya. Penguncian respons dan uji urutan diperlukan pada format digital.

Setiap komponen membutuhkan spesifikasi. Stimulus harus memuat informasi cukup, opsi atau prompt harus sejajar dengan indikator, skala harus memiliki jangkar, dan rubrik harus menjelaskan perbedaan kategori yang berdekatan.

Struktur yang baik memungkinkan peneliti menelusuri kembali setiap kategori hasil ke respons mentah. Transparansi ini penting agar guru tidak hanya menerima label otomatis tanpa memahami alasan di baliknya.

- masalah diagnostik;
- lembar langkah;
- kode jenis kesalahan;
- pohon keputusan;
- rubrik strategi;
- wawancara konfirmasi;

Dalam praktik kelas, Analisis Kesalahan Penyelesaian Masalah dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

No.	Komponen	Fungsi Diagnostik
1	masalah diagnostik	Bukti yang harus ditelusuri
2	lembar langkah	Bukti yang harus ditelusuri
3	kode jenis kesalahan	Bukti yang harus ditelusuri
4	pohon keputusan	Bukti yang harus ditelusuri

6.14.5 Desain Stimulus, Tugas, atau Pertanyaan

Desain stimulus pada Analisis Kesalahan Penyelesaian Masalah harus memunculkan penggunaan konsep, bukan sekadar pengenalan istilah. Situasi dapat berbentuk fenomena, gambar partikel, grafik, data eksperimen, persamaan, atau masalah kontekstual, tetapi seluruh informasi yang tidak relevan perlu dikendalikan. (Haladyna et al., 2002; Nurrenbern & Pickering, 1987; Bodner & Domin, 2000; Ericsson & Simon, 1993)

Dalam kimia, satu stimulus sering menggabungkan beberapa tingkat representasi. Pengembang harus memutuskan apakah tujuan butir adalah menerjemahkan representasi, menjelaskan hubungan sebab-akibat, membedakan contoh dan noncontoh, atau menerapkan prinsip pada konteks baru.

Bahasa perlu ringkas tanpa menghilangkan syarat ilmiah. Kata seperti kuat, stabil, jenuh, spontan, atau setimbang mempunyai arti teknis dan sehari-hari yang berbeda. Ketidakjelasan istilah dapat menghasilkan kesalahan yang berasal dari bahasa, bukan struktur konsep.

Desain yang baik juga menghindari petunjuk format, pilihan terlalu panjang, negasi bertumpuk, gambar dekoratif, dan data yang tidak diperlukan. Setiap elemen harus mempunyai alasan diagnostik.

- gunakan konteks yang cukup dikenal tetapi tidak menghafalkan contoh;
- pastikan hanya satu interpretasi ilmiah yang menjadi target;
- uji semua simbol, satuan, grafik, dan gambar;
- sediakan ruang bagi respons yang tidak diprediksi bila tujuan masih eksploratif.

Dalam praktik kelas, Analisis Kesalahan Penyelesaian Masalah dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.14.6 Prosedur Pengembangan Secara Sistematis

Pengembangan Analisis Kesalahan Penyelesaian Masalah harus mengikuti alur yang terdokumentasi. Langkah yang lazim ialah memilih masalah kaya langkah, mengidentifikasi jalur benar, memprediksi error, menyusun kode, kemudian mengumpulkan solusi, melakukan think-aloud pada sampel, merevisi kode, menyusun profil kesalahan. Urutan dapat berulang karena temuan uji coba sering memaksa revisi konstruk atau redaksi. (Nurrenbern & Pickering, 1987; Bodner & Domin, 2000; Ericsson & Simon, 1993; Adams & Wieman, 2011; AERA, APA, & NCME, 2014)

Tahap konseptual dimulai dengan analisis kurikulum, literatur miskonsepsi, dan peta konsep. Tahap empiris awal menggunakan wawancara, respons terbuka, atau observasi untuk memastikan bahwa kategori yang dirancang benar-benar muncul pada populasi sasaran.

Setiap versi instrumen perlu diberi nomor, tanggal, dan catatan perubahan. Revisi harus menyebutkan masalah yang ditemukan, bukti yang digunakan, perubahan yang dilakukan, dan alasan mempertahankan atau menghapus butir.

Pengembangan dianggap belum selesai ketika hanya menghasilkan butir yang tampak baik. Instrumen perlu diuji pada peserta didik yang serupa dengan pengguna sasaran, lalu dianalisis secara konseptual dan empiris sebelum dipakai untuk keputusan penting.

- 1. memilih masalah kaya langkah.
- 2. mengidentifikasi jalur benar.
- 3. memprediksi error.
- 4. menyusun kode.
- 5. mengumpulkan solusi.

Dalam praktik kelas, Analisis Kesalahan Penyelesaian Masalah dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.14.7 Pelaksanaan dan Pengumpulan Data

Pelaksanaan Analisis Kesalahan Penyelesaian Masalah harus distandardisasi secukupnya agar perbedaan hasil lebih mungkin berasal dari pemahaman, bukan variasi instruksi, waktu, alat, atau bantuan. Petunjuk perlu menjelaskan tujuan umum tanpa mengajarkan jawaban. (Nurrenbern & Pickering, 1987; Bodner & Domin, 2000; AERA, APA, & NCME, 2014)

Kondisi administrasi mencakup waktu, urutan tugas, kesempatan bertanya, penggunaan kalkulator, akses tabel periodik, kualitas gambar, dan perangkat digital. Pada instrumen kualitatif, standardisasi berarti konsistensi fokus dan prinsip probing, bukan membaca naskah secara kaku.

Data yang dikumpulkan ialah urutan langkah, persamaan, gambar, unit, substitusi, revisi, dan jawaban akhir. Data mentah harus disimpan bersama metadata yang diperlukan, seperti kode butir, versi instrumen, waktu administrasi, dan kondisi khusus, tetapi hanya data yang relevan yang boleh dikumpulkan.

Sebelum administrasi luas, lakukan uji teknis. Pastikan tautan, urutan, penguncian respons, rekaman, penyimpanan, serta tampilan rumus dan subskrip bekerja. Gangguan teknis perlu dicatat sebagai bagian dari interpretasi.

- berikan instruksi yang sama dan mudah dipahami;
- catat kondisi yang dapat memengaruhi respons;
- hindari umpan balik sebelum seluruh respons diagnostik selesai;
- sediakan prosedur penanganan data tidak lengkap.

Dalam praktik kelas, Analisis Kesalahan Penyelesaian Masalah dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.14.8 Penskoran, Pengodean, dan Pengolahan Data

Penskoran Analisis Kesalahan Penyelesaian Masalah dapat diringkas sebagai pengodean lokasi dan jenis kesalahan, ketergantungan antarlangkah, serta pemetaan error terhadap konsep prasyarat. Aturan ini harus ditulis sebelum melihat hasil utama agar keputusan tidak disesuaikan secara oportunistik dengan pola data. (Nurrenbern & Pickering, 1987; Bodner & Domin, 2000; Ericsson & Simon, 1993; Chi, 1997; AERA, APA, & NCME, 2014)

Skor total jarang cukup untuk diagnosis. Analisis perlu mempertahankan informasi tingkat butir, konsep, opsi, alasan, keyakinan, proses, atau representasi. Penggabungan data terlalu dini dapat menyembunyikan pola yang secara pedagogis penting.

Untuk respons terbuka atau kualitatif, buku kode harus mencantumkan definisi, contoh positif, contoh batas, noncontoh, dan aturan ketika satu respons memuat beberapa gagasan. Sebagian data perlu dinilai oleh lebih dari satu penilai.

Setiap transformasi data—misalnya dari respons ke kode, dari kode ke kategori, dan dari kategori ke persentase—harus dapat diaudit. Simpan respons asli dan keputusan pengodean agar kasus ambigu dapat ditinjau kembali.

- pisahkan skor ketepatan dari skor keyakinan atau kualitas alasan;
- laporkan kasus tidak lengkap dan aturan penanganannya;
- uji konsistensi penilai untuk data berkode;
- hindari mengubah data ordinal menjadi ukuran interval tanpa alasan.

Dalam praktik kelas, Analisis Kesalahan Penyelesaian Masalah dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.14.9 Interpretasi dan Penetapan Kategori

Interpretasi hasil Analisis Kesalahan Penyelesaian Masalah harus bergerak dari respons menuju pola, bukan langsung dari satu jawaban menuju label. Status konseptual menjadi lebih kuat apabila gagasan yang sama muncul pada beberapa butir, representasi, atau konteks. (Nurrenbern & Pickering, 1987; Bodner & Domin, 2000; Ericsson & Simon, 1993; Chi, 1997; Kane, 2013)

Kategori dapat mencakup pemahaman ilmiah, pemahaman parsial, kurang pengetahuan, respons tidak konsisten, false positive, false negative, dan miskonsepsi. Tidak semua instrumen mampu membedakan seluruh kategori tersebut; skema harus mengikuti informasi yang benar-benar tersedia.

Ambang, seperti batas keyakinan atau persentase pemilihan, perlu memiliki rasional dan diuji melalui analisis sensitivitas. Perubahan kecil pada ambang dapat mengubah prevalensi kategori secara berarti.

Interpretasi yang bertanggung jawab menyatakan bahwa kategori berlaku pada konsep, bentuk tugas, waktu, dan kondisi tertentu. Hasil merupakan hipotesis kerja untuk pembelajaran, bukan diagnosis psikologis permanen.

- gunakan lebih dari satu butir untuk konsepsi penting;
- periksa konsistensi antara jawaban dan alasan;
- tandai kasus ambigu untuk konfirmasi;
- hubungkan kategori dengan rekomendasi yang berbeda.

Dalam praktik kelas, Analisis Kesalahan Penyelesaian Masalah dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.14.10 Validitas dan Kualitas Inferensi

Validitas Analisis Kesalahan Penyelesaian Masalah berkaitan dengan kualitas interpretasi dan penggunaan hasil, bukan sifat tetap yang dimiliki instrumen. Fokus khususnya adalah kecocokan kode error dengan proses kognitif dan bukti triangulasi bahwa langkah tertentu memang berasal dari konsepsi yang ditafsirkan. (AERA, APA, & NCME, 2014; Messick, 1995; Kane, 2013; Haynes et al., 1995)

Bukti berdasarkan isi menilai keterwakilan domain dan akurasi konsep. Bukti proses respons memeriksa bagaimana peserta didik memahami tugas. Bukti struktur internal menilai kesesuaian pola respons. Bukti hubungan dengan variabel lain dan konsekuensi penggunaan melengkapi argumen.

Ancaman validitas dapat berasal dari konten yang kurang terwakili, tuntutan bahasa berlebihan, petunjuk opsi, rubrik yang tidak stabil, keyakinan yang tidak terkalibrasi, atau algoritma yang belajar dari label keliru.

Dokumen validasi perlu menghubungkan setiap klaim dengan bukti dan alternatif penjelasan. Menampilkan satu indeks tidak cukup; validasi adalah program argumentatif yang diperbarui setiap kali instrumen dipakai pada populasi atau tujuan baru.

- bukti isi dan kesesuaian kurikulum;
- bukti proses respons melalui wawancara kognitif;
- bukti struktur dan fungsi butir;
- bukti kegunaan serta konsekuensi keputusan.

Dalam praktik kelas, Analisis Kesalahan Penyelesaian Masalah dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.14.11 Reliabilitas, Keajegan, dan Analisis Empiris

Keajegan Analisis Kesalahan Penyelesaian Masalah harus dinilai sesuai jenis data. Pada respons objektif dapat digunakan konsistensi internal, stabilitas waktu, atau model pengukuran. Pada data kualitatif, perhatian diarahkan pada konsistensi pengodean, audit trail, dan ketepatan penerapan definisi kode. (Kuder & Richardson, 1937; Cronbach, 1951; Ding et al., 2006; Wren & Barbera, 2014; AERA, APA, & NCME, 2014)

Reliabilitas tinggi tidak otomatis menunjukkan diagnosis yang benar. Butir yang berulang-ulang mengukur hafalan dapat menghasilkan koefisien tinggi tetapi tidak mewakili pemahaman konseptual. Sebaliknya, konstruk yang multidimensi tidak selalu layak dipaksa menjadi satu skala.

Analisis empiris perlu memeriksa kesukaran, daya beda, fungsi opsi, pola data hilang, dan keterhubungan komponen. Untuk instrumen bertingkat, reliabilitas dapat dihitung pada setiap tier serta kombinasi yang mempunyai makna teoritis.

Ukuran sampel, distribusi kemampuan, dan homogenitas konten memengaruhi statistik. Karena itu, laporan harus menyatakan kondisi estimasi dan tidak menyalin batas angka sebagai aturan universal.

- pilih statistik yang sesuai dengan jenis skor;
- laporkan interval atau ketidakpastian bila tersedia;
- periksa butir secara substantif dan statistik;
- uji kembali setelah revisi atau perubahan platform.

Dalam praktik kelas, Analisis Kesalahan Penyelesaian Masalah dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.14.12 Contoh Penerapan dalam Pembelajaran Kimia

Penerapan Analisis Kesalahan Penyelesaian Masalah dalam kimia dapat dimulai dari konsep stoikiometri dan pereaksi pembatas. Pengembang perlu menuliskan konsep ilmiah target, miskonsepsi yang mungkin, bukti respons yang diharapkan, dan keputusan pembelajaran setelah hasil diperoleh. (Nurrenbern & Pickering, 1987; Bodner & Domin, 2000; Ericsson & Simon, 1993; Chi, 1997; Nakhleh, 1992)

Contoh kedua adalah perhitungan pH. Stimulus sebaiknya menghubungkan fenomena dengan model partikel atau persamaan, lalu meminta peserta didik menunjukkan hubungan sebab-akibat. Jawaban yang hanya menyebut istilah belum cukup menjadi bukti pemahaman.

Pada termokimia Hess, diagnosis dapat memeriksa apakah peserta didik membedakan perubahan laju, keadaan setimbang, energi, dan komposisi. Kesalahan sering muncul ketika satu aturan dipindahkan ke konteks yang tidak sesuai.

Kasus kesetimbangan dan kuosien reaksi menunjukkan pentingnya representasi. Respons verbal, gambar, simbol, dan hitungan dapat tidak konsisten. Hasil instrumen sebaiknya digunakan untuk memilih demonstrasi, diskusi, contoh-noncontoh, atau tugas remediasi yang secara langsung menantang hubungan salah.

- tentukan proposisi ilmiah dan konsepsi alternatif;
- gunakan minimal dua bentuk representasi;
- minta alasan atau bukti ketika format memungkinkan;
- rancang tindak lanjut berdasarkan pola, bukan hanya skor.

Dalam praktik kelas, Analisis Kesalahan Penyelesaian Masalah dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.14.13 Kelebihan, Keterbatasan, dan Potensi Bias

Kekuatan utama Analisis Kesalahan Penyelesaian Masalah meliputi menghindari fokus hanya pada jawaban akhir, membedakan kesalahan aritmetika dari miskonsepsi, menemukan aturan algoritmik keliru, dan mendukung umpan balik spesifik. Kekuatan tersebut menjelaskan situasi ketika instrumen layak dipilih dibanding teknik lain. (Nurrenbern & Pickering, 1987; Bodner & Domin, 2000; Ericsson & Simon, 1993; Gurel et al., 2015; AERA, APA, & NCME, 2014)

Keterbatasannya mencakup langkah tertulis mungkin tidak lengkap, peserta didik dapat menghapus jejak, satu kesalahan awal menghasilkan banyak kesalahan turunan, dan kode error perlu definisi operasional. Keterbatasan harus diperlakukan sebagai batas inferensi, bukan sekadar paragraf formal yang tidak memengaruhi analisis.

Potensi bias dapat berasal dari bahasa, pengalaman konteks, akses teknologi, kecemasan, gaya menggunakan skala, keterampilan menggambar, atau kemampuan menulis. Instrumen perlu diuji pada kelompok yang beragam dan ditinjau untuk fungsi berbeda antarkelompok bila data mencukupi.

Keputusan memilih instrumen sebaiknya mempertimbangkan kedalaman data, ukuran kelompok, waktu, keterampilan analisis, tujuan diagnosis, dan konsekuensi kesalahan klasifikasi. Tidak ada satu metode yang unggul untuk semua tujuan.

- Kekuatan: menghindari fokus hanya pada jawaban akhir.
- Kekuatan: membedakan kesalahan aritmetika dari miskonsepsi.
- Batas: langkah tertulis mungkin tidak lengkap.
- Batas: peserta didik dapat menghapus jejak.

Dalam praktik kelas, Analisis Kesalahan Penyelesaian Masalah dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

Aspek	Temuan	Respons Metodologis
Kekuatan	menghindari fokus hanya pada jawaban akhir	Pilih ketika sesuai tujuan
Keterbatasan	langkah tertulis mungkin tidak lengkap	Tambahkan bukti atau kontrol
Bias potensial	Bahasa, konteks, perangkat, gaya respons	Uji proses respons

6.14.14 Integrasi Digital, Etika, dan Tata Kelola Data

Integrasi digital Analisis Kesalahan Penyelesaian Masalah dapat berupa pencatatan langkah pada tablet, log revisi, analisis ekspresi, dan replay proses tanpa hanya menyimpan jawaban akhir. Teknologi seharusnya memperluas kualitas bukti, bukan hanya memindahkan lembar kertas ke layar. (Baker & Inventado, 2014; Romero & Ventura, 2020; AERA, APA, & NCME, 2014; Nurrenbern & Pickering, 1987)

Platform perlu menyimpan versi instrumen, urutan respons, timestamp yang relevan, aturan skor, dan log perubahan. Namun, hanya data yang dibutuhkan untuk tujuan diagnosis yang boleh dikumpulkan; pengumpulan berlebihan meningkatkan risiko privasi tanpa manfaat ilmiah.

Aspek etika utama adalah bahwa diagnosis harus digunakan untuk dukungan, bukan menghukum strategi alternatif yang benar atau gaya penyelesaian berbeda. Peserta didik dan guru perlu mengetahui data apa yang dikumpulkan, untuk tujuan apa, siapa yang dapat mengakses, berapa lama disimpan, dan bagaimana kesalahan data diperbaiki.

Algoritma otomatis harus dapat diaudit. Kategori yang dihasilkan sistem perlu ditautkan ke respons mentah dan aturan keputusan. Kasus berisiko atau ambigu harus ditinjau manusia sebelum tindak lanjut penting dilakukan.

- gunakan identitas pseudonim dan akses berbasis peran;
- uji tampilan rumus, subskrip, grafik, dan gambar;
- sediakan opsi aksesibilitas dan alternatif perangkat;
- dokumentasikan algoritma serta perubahan versi.

Dalam praktik kelas, Analisis Kesalahan Penyelesaian Masalah dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.14.15 Sintesis dan Pedoman Implementasi

Analisis Kesalahan Penyelesaian Masalah bermanfaat apabila dipilih berdasarkan tujuan dan digunakan sesuai batasnya. Instrumen ini menghasilkan urutan langkah, persamaan, gambar, unit, substitusi, revisi, dan jawaban akhir, lalu data diolah melalui pengodean lokasi dan jenis kesalahan, ketergantungan antarlangkah, serta pemetaan error terhadap konsep prasyarat. (Nurrenbern & Pickering, 1987; Bodner & Domin, 2000; Ericsson & Simon, 1993; AERA, APA, & NCME, 2014; Kane, 2013)

Rangkaian kerja yang disarankan dimulai dari analisis konsep, identifikasi konsepsi alternatif, desain tugas, telaah ahli, uji proses respons, uji lapangan, analisis, penetapan kategori, dan penyusunan tindak lanjut. Setiap tahap harus meninggalkan dokumentasi yang dapat diperiksa.

Hasil diagnosis tidak berhenti pada tabel prevalensi. Guru perlu memilih konsepsi prioritas berdasarkan frekuensi, keyakinan, posisi sebagai konsep prasyarat, dampak terhadap topik lanjutan, dan kemungkinan diperbaiki melalui pembelajaran.

Prinsip akhirnya ialah kehati-hatian inferensial: sebutkan apa yang benar-benar ditunjukkan data, bedakan bukti kuat dari indikasi, gunakan triangulasi ketika diperlukan, dan hindari label permanen pada peserta didik.

- selaraskan tujuan, konstruk, tugas, skor, kategori, dan tindakan;
- utamakan konsepsi yang berdampak pada pembelajaran lanjutan;
- gabungkan bukti bila keputusan memiliki konsekuensi besar;
- evaluasi kembali efektivitas tindak lanjut melalui diagnosis ulang.

Dalam praktik kelas, Analisis Kesalahan Penyelesaian Masalah dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.15 Think-Aloud Protocol

6.15.1 Landasan Konseptual dan Kedudukan Instrumen

Pemilihan Think-Aloud Protocol harus berangkat dari pertanyaan diagnosis yang jelas, bukan dari kebiasaan menggunakan instrumen yang paling mudah tersedia. Dalam bab ini, instrumen tersebut didefinisikan sebagai metode yang meminta peserta didik mengucapkan apa yang sedang dipikirkan ketika menyelesaikan tugas tanpa menjelaskan secara retrospektif lebih dari yang diperlukan. Definisi operasional ini menegaskan bentuk bukti yang dikumpulkan sekaligus membatasi jenis kesimpulan yang layak dibuat. (Ericsson & Simon, 1993; van Someren et al., 1994; Chi, 1997)

Klaim utamanya adalah bahwa instrumen ini menyediakan data proses tentang perhatian, pemilihan informasi, strategi, representasi, dan perubahan keputusan. Klaim tersebut harus dipisahkan dari anggapan bahwa setiap jawaban salah merupakan miskonsepsi. Kesalahan dapat berasal dari kurang pengetahuan, ketidakcermatan, kesulitan bahasa, salah hitung, atau ketidakfamiliaran dengan representasi.

Data pokok yang dihasilkan berupa verbalization yang berurutan, jeda, tindakan, layar, dan perubahan solusi. Data itu baru memperoleh makna diagnostik setelah dikaitkan dengan peta konsep sasaran, kemungkinan konsepsi alternatif, dan aturan interpretasi yang telah ditetapkan sebelum analisis.

Konsekuensinya, peneliti dan guru perlu membedakan data mentah, skor, kategori, dan klaim konseptual yang dibuat dari hasil tersebut. Posisi terbaik Think-Aloud Protocol adalah sebagai bagian dari sistem asesmen berlapis yang dapat dipadukan dengan bukti lain ketika konsekuensi keputusan semakin besar.

- menentukan konsep dan proposisi yang hendak didiagnosis;
- menjelaskan perilaku respons yang dianggap sebagai bukti;
- membatasi klaim sesuai kekuatan instrumen;
- menautkan hasil dengan tindakan pembelajaran yang realistis.

Dalam praktik kelas, Think-Aloud Protocol dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.15.2 Perkembangan Historis dan Rasional Penggunaan

Perkembangan Think-Aloud Protocol tidak dapat dipisahkan dari perubahan pandangan tentang belajar sains. Ketika peserta didik tidak lagi dipandang sebagai penerima informasi yang kosong, instrumen mulai dirancang untuk mengungkap gagasan awal, pola penalaran, dan keyakinan yang dibawa ke kelas. (Ericsson & Simon, 1993; van Someren et al., 1994; Chi, 1997; AERA, APA, & NCME, 2014)

Pada tahap awal, penilaian banyak mengandalkan benar-salah atau skor total. Perkembangan penelitian konsepsi peserta didik kemudian mendorong penambahan alasan, keyakinan, tugas kinerja, rekaman proses, dan bukti sosial. Think-Aloud Protocol merupakan salah satu jawaban metodologis terhadap kebutuhan tersebut.

Dalam pendidikan kimia, kebutuhan diagnosis semakin kuat karena peserta didik harus mengoordinasikan fenomena makroskopik, model submikroskopik, simbol, grafik, dan perhitungan. Instrumen yang hanya menguji hafalan definisi sering gagal memperlihatkan cara representasi itu dihubungkan.

Sejarah penggunaannya juga menunjukkan bahwa format yang lebih kompleks tidak otomatis lebih baik. Setiap tambahan tier, rubrik, atau kanal data harus mempunyai fungsi yang jelas dan didukung bukti bahwa informasi tambahan benar-benar memperbaiki keputusan.

- pergeseran dari pengukuran skor menuju diagnosis pola pemahaman;
- penggunaan respons terbuka dan wawancara untuk menemukan konsepsi alternatif;
- pengembangan teknik kuantitatif agar diagnosis dapat dilakukan pada kelompok besar;
- integrasi data digital dan metakognitif pada perkembangan mutakhir.

Dalam praktik kelas, Think-Aloud Protocol dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.15.3 Konstruk, Klaim, dan Bukti Diagnosis

Konstruk yang dinilai melalui Think-Aloud Protocol bukan objek yang terlihat langsung. Yang diamati adalah verbalization yang berurutan, jeda, tindakan, layar, dan perubahan solusi; dari sana dibuat inferensi tentang pemahaman konsep, miskonsepsi, pengetahuan parsial, atau ketidakpastian. (Chi, 1997; AERA, APA, & NCME, 2014; Messick, 1995; Kane, 2013)

Klaim diagnosis perlu dirumuskan sebagai argumen. Pertama, tugas harus memerlukan konsep sasaran. Kedua, respons harus diberi skor atau kode secara konsisten. Ketiga, pola skor harus mempunyai hubungan yang masuk akal dengan struktur pengetahuan. Keempat, penggunaan hasil harus menghasilkan keputusan yang proporsional.

Sumber ancaman utama adalah ketidakjelasan konstruk. Misalnya, peserta didik dapat gagal karena istilah yang ambigu, beban membaca, representasi yang tidak dikenal, atau prosedur matematika. Kegagalan semacam itu tidak boleh langsung diperlakukan sebagai miskonsepsi kimia.

Oleh sebab itu, setiap pengembang perlu menulis pernyataan klaim-bukti: instrumen mengukur apa, perilaku apa yang menjadi indikator, alternatif penjelasan apa yang mungkin, dan bukti tambahan apa yang diperlukan untuk memperkuat kesimpulan.

- konstruk utama: pemahaman konseptual melalui protokol verbal proses kognitif;
- bukti observabel: respons, alasan, tindakan, atau representasi;
- inferensi: profil konsepsi pada konten dan konteks tertentu;
- penggunaan: perencanaan pembelajaran, remediasi, atau penelitian.

Dalam praktik kelas, Think-Aloud Protocol dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.15.4 Struktur, Komponen, dan Alur Respons

Struktur Think-Aloud Protocol dibangun dari komponen yang saling bergantung. Komponen inti meliputi tugas, instruksi berpikir keras, prompt pengingat netral, rekaman, transkripsi bertanda waktu, dan skema pengodean proses. Hilangnya satu komponen dapat mengubah makna data yang diperoleh. (Ericsson & Simon, 1993; van Someren et al., 1994; Chi, 1997; AERA, APA, & NCME, 2014)

Urutan penyajian harus ditetapkan secara sadar. Komponen yang muncul lebih awal dapat memberi petunjuk pada komponen berikutnya, sedangkan permintaan alasan atau keyakinan dapat mengubah cara peserta didik memeriksa jawabannya. Penguncian respons dan uji urutan diperlukan pada format digital.

Setiap komponen membutuhkan spesifikasi. Stimulus harus memuat informasi cukup, opsi atau prompt harus sejajar dengan indikator, skala harus memiliki jangkar, dan rubrik harus menjelaskan perbedaan kategori yang berdekatan.

Struktur yang baik memungkinkan peneliti menelusuri kembali setiap kategori hasil ke respons mentah. Transparansi ini penting agar guru tidak hanya menerima label otomatis tanpa memahami alasan di baliknya.

- tugas;
- instruksi berpikir keras;
- prompt pengingat netral;
- rekaman;
- transkripsi bertanda waktu;
- skema pengodean proses;

Dalam praktik kelas, Think-Aloud Protocol dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

No.	Komponen	Fungsi Diagnostik
1	tugas	Bukti yang harus ditelusuri
2	instruksi berpikir keras	Bukti yang harus ditelusuri
3	prompt pengingat netral	Bukti yang harus ditelusuri
4	rekaman	Bukti yang harus ditelusuri

6.15.5 Desain Stimulus, Tugas, atau Pertanyaan

Desain stimulus pada Think-Aloud Protocol harus memunculkan penggunaan konsep, bukan sekadar pengenalan istilah. Situasi dapat berbentuk fenomena, gambar partikel, grafik, data eksperimen, persamaan, atau masalah kontekstual, tetapi seluruh informasi yang tidak relevan perlu dikendalikan. (Haladyna et al., 2002; Ericsson & Simon, 1993; van Someren et al., 1994; Chi, 1997)

Dalam kimia, satu stimulus sering menggabungkan beberapa tingkat representasi. Pengembang harus memutuskan apakah tujuan butir adalah menerjemahkan representasi, menjelaskan hubungan sebab-akibat, membedakan contoh dan noncontoh, atau menerapkan prinsip pada konteks baru.

Bahasa perlu ringkas tanpa menghilangkan syarat ilmiah. Kata seperti kuat, stabil, jenuh, spontan, atau setimbang mempunyai arti teknis dan sehari-hari yang berbeda. Ketidakjelasan istilah dapat menghasilkan kesalahan yang berasal dari bahasa, bukan struktur konsep.

Desain yang baik juga menghindari petunjuk format, pilihan terlalu panjang, negasi bertumpuk, gambar dekoratif, dan data yang tidak diperlukan. Setiap elemen harus mempunyai alasan diagnostik.

- gunakan konteks yang cukup dikenal tetapi tidak menghafalkan contoh;
- pastikan hanya satu interpretasi ilmiah yang menjadi target;
- uji semua simbol, satuan, grafik, dan gambar;
- sediakan ruang bagi respons yang tidak diprediksi bila tujuan masih eksploratif.

Dalam praktik kelas, Think-Aloud Protocol dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.15.6 Prosedur Pengembangan Secara Sistematis

Pengembangan Think-Aloud Protocol harus mengikuti alur yang terdokumentasi. Langkah yang lazim ialah menetapkan tugas, menulis instruksi, melatih peserta, menguji reaktivitas, kemudian merekam multimodal, mentranskripsi, mengodekan, menguji kesepakatan. Urutan dapat berulang karena temuan uji coba sering memaksa revisi konstruk atau redaksi. (Ericsson & Simon, 1993; van Someren et al., 1994; Chi, 1997; Adams & Wieman, 2011; AERA, APA, & NCME, 2014)

Tahap konseptual dimulai dengan analisis kurikulum, literatur miskonsepsi, dan peta konsep. Tahap empiris awal menggunakan wawancara, respons terbuka, atau observasi untuk memastikan bahwa kategori yang dirancang benar-benar muncul pada populasi sasaran.

Setiap versi instrumen perlu diberi nomor, tanggal, dan catatan perubahan. Revisi harus menyebutkan masalah yang ditemukan, bukti yang digunakan, perubahan yang dilakukan, dan alasan mempertahankan atau menghapus butir.

Pengembangan dianggap belum selesai ketika hanya menghasilkan butir yang tampak baik. Instrumen perlu diuji pada peserta didik yang serupa dengan pengguna sasaran, lalu dianalisis secara konseptual dan empiris sebelum dipakai untuk keputusan penting.

- 1. menetapkan tugas.
- 2. menulis instruksi.
- 3. melatih peserta.
- 4. menguji reaktivitas.
- 5. merekam multimodal.

Dalam praktik kelas, Think-Aloud Protocol dapat digunakan sebelum pembelajaran untuk menetapkan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.15.7 Pelaksanaan dan Pengumpulan Data

Pelaksanaan Think-Aloud Protocol harus distandardisasi secukupnya agar perbedaan hasil lebih mungkin berasal dari pemahaman, bukan variasi instruksi, waktu, alat, atau bantuan. Petunjuk perlu menjelaskan tujuan umum tanpa mengajarkan jawaban. (Ericsson & Simon, 1993; van Someren et al., 1994; AERA, APA, & NCME, 2014)

Kondisi administrasi mencakup waktu, urutan tugas, kesempatan bertanya, penggunaan kalkulator, akses tabel periodik, kualitas gambar, dan perangkat digital. Pada instrumen kualitatif, standardisasi berarti konsistensi fokus dan prinsip probing, bukan membaca naskah secara kaku.

Data yang dikumpulkan ialah verbalization yang berurutan, jeda, tindakan, layar, dan perubahan solusi. Data mentah harus disimpan bersama metadata yang diperlukan, seperti kode butir, versi instrumen, waktu administrasi, dan kondisi khusus, tetapi hanya data yang relevan yang boleh dikumpulkan.

Sebelum administrasi luas, lakukan uji teknis. Pastikan tautan, urutan, penguncian respons, rekaman, penyimpanan, serta tampilan rumus dan subskrip bekerja. Gangguan teknis perlu dicatat sebagai bagian dari interpretasi.

- berikan instruksi yang sama dan mudah dipahami;
- catat kondisi yang dapat memengaruhi respons;
- hindari umpan balik sebelum seluruh respons diagnostik selesai;
- sediakan prosedur penanganan data tidak lengkap.

Dalam praktik kelas, Think-Aloud Protocol dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.15.8 Penskoran, Pengodean, dan Pengolahan Data

Penskoran Think-Aloud Protocol dapat diringkas sebagai segmentasi episode, kode strategi, inferensi terbatas pada informasi yang diverbalkan, dan perbandingan lintaspenyelesai. Aturan ini harus ditulis sebelum melihat hasil utama agar keputusan tidak disesuaikan secara oportunistik dengan pola data. (Ericsson & Simon, 1993; van Someren et al., 1994; Chi, 1997; AERA, APA, & NCME, 2014)

Skor total jarang cukup untuk diagnosis. Analisis perlu mempertahankan informasi tingkat butir, konsep, opsi, alasan, keyakinan, proses, atau representasi. Penggabungan data terlalu dini dapat menyembunyikan pola yang secara pedagogis penting.

Untuk respons terbuka atau kualitatif, buku kode harus mencantumkan definisi, contoh positif, contoh batas, noncontoh, dan aturan ketika satu respons memuat beberapa gagasan. Sebagian data perlu dinilai oleh lebih dari satu penilai.

Setiap transformasi data—misalnya dari respons ke kode, dari kode ke kategori, dan dari kategori ke persentase—harus dapat diaudit. Simpan respons asli dan keputusan pengodean agar kasus ambigu dapat ditinjau kembali.

- pisahkan skor ketepatan dari skor keyakinan atau kualitas alasan;
- laporkan kasus tidak lengkap dan aturan penanganannya;
- uji konsistensi penilai untuk data berkode;
- hindari mengubah data ordinal menjadi ukuran interval tanpa alasan.

Dalam praktik kelas, Think-Aloud Protocol dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.15.9 Interpretasi dan Penetapan Kategori

Interpretasi hasil Think-Aloud Protocol harus bergerak dari respons menuju pola, bukan langsung dari satu jawaban menuju label. Status konseptual menjadi lebih kuat apabila gagasan yang sama muncul pada beberapa butir, representasi, atau konteks. (Ericsson & Simon, 1993; van Someren et al., 1994; Chi, 1997; AERA, APA, & NCME, 2014; Kane, 2013)

Kategori dapat mencakup pemahaman ilmiah, pemahaman parsial, kurang pengetahuan, respons tidak konsisten, false positive, false negative, dan miskonsepsi. Tidak semua instrumen mampu membedakan seluruh kategori tersebut; skema harus mengikuti informasi yang benar-benar tersedia.

Ambang, seperti batas keyakinan atau persentase pemilihan, perlu memiliki rasional dan diuji melalui analisis sensitivitas. Perubahan kecil pada ambang dapat mengubah prevalensi kategori secara berarti.

Interpretasi yang bertanggung jawab menyatakan bahwa kategori berlaku pada konsep, bentuk tugas, waktu, dan kondisi tertentu. Hasil merupakan hipotesis kerja untuk pembelajaran, bukan diagnosis psikologis permanen.

- gunakan lebih dari satu butir untuk konsepsi penting;
- periksa konsistensi antara jawaban dan alasan;
- tandai kasus ambigu untuk konfirmasi;
- hubungkan kategori dengan rekomendasi yang berbeda.

Dalam praktik kelas, Think-Aloud Protocol dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.15.10 Validitas dan Kualitas Inferensi

Validitas Think-Aloud Protocol berkaitan dengan kualitas interpretasi dan penggunaan hasil, bukan sifat tetap yang dimiliki instrumen. Fokus khususnya adalah kesesuaian antara jenis verbal report dan proses yang diklaim, serta pembatasan inferensi pada data yang tersedia. (AERA, APA, & NCME, 2014; Messick, 1995; Kane, 2013; Haynes et al., 1995)

Bukti berdasarkan isi menilai keterwakilan domain dan akurasi konsep. Bukti proses respons memeriksa bagaimana peserta didik memahami tugas. Bukti struktur internal menilai kesesuaian pola respons. Bukti hubungan dengan variabel lain dan konsekuensi penggunaan melengkapi argumen.

Ancaman validitas dapat berasal dari konten yang kurang terwakili, tuntutan bahasa berlebihan, petunjuk opsi, rubrik yang tidak stabil, keyakinan yang tidak terkalibrasi, atau algoritma yang belajar dari label keliru.

Dokumen validasi perlu menghubungkan setiap klaim dengan bukti dan alternatif penjelasan. Menampilkan satu indeks tidak cukup; validasi adalah program argumentatif yang diperbarui setiap kali instrumen dipakai pada populasi atau tujuan baru.

- bukti isi dan kesesuaian kurikulum;
- bukti proses respons melalui wawancara kognitif;
- bukti struktur dan fungsi butir;
- bukti kegunaan serta konsekuensi keputusan.

Dalam praktik kelas, Think-Aloud Protocol dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.15.11 Reliabilitas, Keajegan, dan Analisis Empiris

Keajegan Think-Aloud Protocol harus dinilai sesuai jenis data. Pada respons objektif dapat digunakan konsistensi internal, stabilitas waktu, atau model pengukuran. Pada data kualitatif, perhatian diarahkan pada konsistensi pengodean, audit trail, dan ketepatan penerapan definisi kode. (Kuder & Richardson, 1937; Cronbach, 1951; Ding et al., 2006; Wren & Barbera, 2014; AERA, APA, & NCME, 2014)

Reliabilitas tinggi tidak otomatis menunjukkan diagnosis yang benar. Butir yang berulang-ulang mengukur hafalan dapat menghasilkan koefisien tinggi tetapi tidak mewakili pemahaman konseptual. Sebaliknya, konstruk yang multidimensi tidak selalu layak dipaksa menjadi satu skala.

Analisis empiris perlu memeriksa kesukaran, daya beda, fungsi opsi, pola data hilang, dan keterhubungan komponen. Untuk instrumen bertingkat, reliabilitas dapat dihitung pada setiap tier serta kombinasi yang mempunyai makna teoritis.

Ukuran sampel, distribusi kemampuan, dan homogenitas konten memengaruhi statistik. Karena itu, laporan harus menyatakan kondisi estimasi dan tidak menyalin batas angka sebagai aturan universal.

- pilih statistik yang sesuai dengan jenis skor;
- laporkan interval atau ketidakpastian bila tersedia;
- periksa butir secara substantif dan statistik;
- uji kembali setelah revisi atau perubahan platform.

Dalam praktik kelas, Think-Aloud Protocol dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.15.12 Contoh Penerapan dalam Pembelajaran Kimia

Penerapan Think-Aloud Protocol dalam kimia dapat dimulai dari konsep menafsirkan grafik laju reaksi. Pengembang perlu menuliskan konsep ilmiah target, miskonsepsi yang mungkin, bukti respons yang diharapkan, dan keputusan pembelajaran setelah hasil diperoleh. (Ericsson & Simon, 1993; van Someren et al., 1994; Chi, 1997; AERA, APA, & NCME, 2014; Nakhleh, 1992)

Contoh kedua adalah memilih persamaan kesetimbangan. Stimulus sebaiknya menghubungkan fenomena dengan model partikel atau persamaan, lalu meminta peserta didik menunjukkan hubungan sebab-akibat. Jawaban yang hanya menyebut istilah belum cukup menjadi bukti pemahaman.

Pada menentukan struktur resonansi, diagnosis dapat memeriksa apakah peserta didik membedakan perubahan laju, keadaan setimbang, energi, dan komposisi. Kesalahan sering muncul ketika satu aturan dipindahkan ke konteks yang tidak sesuai.

Kasus menganalisis spektrum sederhana menunjukkan pentingnya representasi. Respons verbal, gambar, simbol, dan hitungan dapat tidak konsisten. Hasil instrumen sebaiknya digunakan untuk memilih demonstrasi, diskusi, contoh-noncontoh, atau tugas remediasi yang secara langsung menantang hubungan salah.

- tentukan proposisi ilmiah dan konsepsi alternatif;
- gunakan minimal dua bentuk representasi;
- minta alasan atau bukti ketika format memungkinkan;
- rancang tindak lanjut berdasarkan pola, bukan hanya skor.

Dalam praktik kelas, Think-Aloud Protocol dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.15.13 Kelebihan, Keterbatasan, dan Potensi Bias

Kekuatan utama Think-Aloud Protocol meliputi mengungkapkan proses bukan hanya produk, mendeteksi titik kebingungan, berguna untuk masalah kompleks, dan mendukung desain antarmuka dan item. Kekuatan tersebut menjelaskan situasi ketika instrumen layak dipilih dibanding teknik lain. (Ericsson & Simon, 1993; van Someren et al., 1994; Chi, 1997; Gurel et al., 2015; AERA, APA, & NCME, 2014)

Keterbatasannya mencakup tidak semua proses dapat diverbalkan, berbicara dapat mengubah tempo, protokol retrospektif rentan rekonstruksi, dan analisis sangat intensif. Keterbatasan harus diperlakukan sebagai batas inferensi, bukan sekadar paragraf formal yang tidak memengaruhi analisis.

Potensi bias dapat berasal dari bahasa, pengalaman konteks, akses teknologi, kecemasan, gaya menggunakan skala, keterampilan menggambar, atau kemampuan menulis. Instrumen perlu diuji pada kelompok yang beragam dan ditinjau untuk fungsi berbeda antarkelompok bila data mencukupi.

Keputusan memilih instrumen sebaiknya mempertimbangkan kedalaman data, ukuran kelompok, waktu, keterampilan analisis, tujuan diagnosis, dan konsekuensi kesalahan klasifikasi. Tidak ada satu metode yang unggul untuk semua tujuan.

- Kekuatan: mengungkapkan proses bukan hanya produk.
- Kekuatan: mendeteksi titik kebingungan.
- Batas: tidak semua proses dapat diverbalkan.
- Batas: berbicara dapat mengubah tempo.

Dalam praktik kelas, Think-Aloud Protocol dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

Aspek	Temuan	Respons Metodologis
Kekuatan	mengungkapkan proses bukan hanya produk	Pilih ketika sesuai tujuan
Keterbatasan	tidak semua proses dapat diverbalkan	Tambahkan bukti atau kontrol
Bias potensial	Bahasa, konteks, perangkat, gaya respons	Uji proses respons

6.15.14 Integrasi Digital, Etika, dan Tata Kelola Data

Integrasi digital Think-Aloud Protocol dapat berupa screen recording, audio sinkron, event log, dan perangkat anotasi episode. Teknologi seharusnya memperluas kualitas bukti, bukan hanya memindahkan lembar kertas ke layar. (Baker & Inventado, 2014; Romero & Ventura, 2020; AERA, APA, & NCME, 2014; Ericsson & Simon, 1993)

Platform perlu menyimpan versi instrumen, urutan respons, timestamp yang relevan, aturan skor, dan log perubahan. Namun, hanya data yang dibutuhkan untuk tujuan diagnosis yang boleh dikumpulkan; pengumpulan berlebihan meningkatkan risiko privasi tanpa manfaat ilmiah.

Aspek etika utama adalah bahwa rekaman proses dapat terasa sangat pribadi; peserta berhak menghapus bagian dan identitas harus disamarkan. Peserta didik dan guru perlu mengetahui data apa yang dikumpulkan, untuk tujuan apa, siapa yang dapat mengakses, berapa lama disimpan, dan bagaimana kesalahan data diperbaiki.

Algoritma otomatis harus dapat diaudit. Kategori yang dihasilkan sistem perlu ditautkan ke respons mentah dan aturan keputusan. Kasus berisiko atau ambigu harus ditinjau manusia sebelum tindak lanjut penting dilakukan.

- gunakan identitas pseudonim dan akses berbasis peran;
- uji tampilan rumus, subskrip, grafik, dan gambar;
- sediakan opsi aksesibilitas dan alternatif perangkat;
- dokumentasikan algoritma serta perubahan versi.

Dalam praktik kelas, Think-Aloud Protocol dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.15.15 Sintesis dan Pedoman Implementasi

Think-Aloud Protocol bermanfaat apabila dipilih berdasarkan tujuan dan digunakan sesuai batasnya. Instrumen ini menghasilkan verbalization yang berurutan, jeda, tindakan, layar, dan perubahan solusi, lalu data diolah melalui segmentasi episode, kode strategi, inferensi terbatas pada informasi yang diverbalkan, dan perbandingan lintaspenyelesai. (Ericsson & Simon, 1993; van Someren et al., 1994; Chi, 1997; AERA, APA, & NCME, 2014; Kane, 2013)

Rangkaian kerja yang disarankan dimulai dari analisis konsep, identifikasi konsepsi alternatif, desain tugas, telaah ahli, uji proses respons, uji lapangan, analisis, penetapan kategori, dan penyusunan tindak lanjut. Setiap tahap harus meninggalkan dokumentasi yang dapat diperiksa.

Hasil diagnosis tidak berhenti pada tabel prevalensi. Guru perlu memilih konsepsi prioritas berdasarkan frekuensi, keyakinan, posisi sebagai konsep prasyarat, dampak terhadap topik lanjutan, dan kemungkinan diperbaiki melalui pembelajaran.

Prinsip akhirnya ialah kehati-hatian inferensial: sebutkan apa yang benar-benar ditunjukkan data, bedakan bukti kuat dari indikasi, gunakan triangulasi ketika diperlukan, dan hindari label permanen pada peserta didik.

- selaraskan tujuan, konstruk, tugas, skor, kategori, dan tindakan;
- utamakan konsepsi yang berdampak pada pembelajaran lanjutan;
- gabungkan bukti bila keputusan memiliki konsekuensi besar;
- evaluasi kembali efektivitas tindak lanjut melalui diagnosis ulang.

Dalam praktik kelas, Think-Aloud Protocol dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.16 Analisis Percakapan Kelas

6.16.1 Landasan Konseptual dan Kedudukan Instrumen

Analisis Percakapan Kelas menempati posisi tertentu dalam rangkaian bukti diagnosis, dari pengungkapan respons awal sampai penentuan tindak lanjut pembelajaran. Dalam bab ini, instrumen tersebut didefinisikan sebagai analisis sistematis terhadap dialog guru-peserta didik dan antarpeserta didik untuk melihat bagaimana konsep, kesalahan, otoritas bukti, dan makna dinegosiasikan. Definisi operasional ini menegaskan bentuk bukti yang dikumpulkan sekaligus membatasi jenis kesimpulan yang layak dibuat. (Mortimer & Scott, 2003; Chi, 2009; Chi, 1997)

Klaim utamanya adalah bahwa instrumen ini mengungkap miskonsepsi sebagai fenomena yang dapat muncul, diperkuat, ditantang, atau direvisi melalui interaksi kelas. Klaim tersebut harus dipisahkan dari anggapan bahwa setiap jawaban salah merupakan miskonsepsi. Kesalahan dapat berasal dari kurang pengetahuan, ketidakcermatan, kesulitan bahasa, salah hitung, atau ketidakfamiliaran dengan representasi.

Data pokok yang dihasilkan berupa pertanyaan, jawaban, respons guru, perbaikan, jeda, gestur, penggunaan papan, dan rujukan pada representasi. Data itu baru memperoleh makna diagnostik setelah dikaitkan dengan peta konsep sasaran, kemungkinan konsepsi alternatif, dan aturan interpretasi yang telah ditetapkan sebelum analisis.

Konsekuensinya, peneliti dan guru perlu membedakan data mentah, skor, kategori, dan klaim konseptual yang dibuat dari hasil tersebut. Posisi terbaik Analisis Percakapan Kelas adalah sebagai bagian dari sistem asesmen berlapis yang dapat dipadukan dengan bukti lain ketika konsekuensi keputusan semakin besar.

- menentukan konsep dan proposisi yang hendak didiagnosis;
- menjelaskan perilaku respons yang dianggap sebagai bukti;
- membatasi klaim sesuai kekuatan instrumen;
- menautkan hasil dengan tindakan pembelajaran yang realistis.

Dalam praktik kelas, Analisis Percakapan Kelas dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.16.2 Perkembangan Historis dan Rasional Penggunaan

Perkembangan Analisis Percakapan Kelas tidak dapat dipisahkan dari perubahan pandangan tentang belajar sains. Ketika peserta didik tidak lagi dipandang sebagai penerima informasi yang kosong, instrumen mulai dirancang untuk mengungkap gagasan awal, pola penalaran, dan keyakinan yang dibawa ke kelas. (Mortimer & Scott, 2003; Chi, 2009; Chi, 1997; AERA, APA, & NCME, 2014)

Pada tahap awal, penilaian banyak mengandalkan benar-salah atau skor total. Perkembangan penelitian konsepsi peserta didik kemudian mendorong penambahan alasan, keyakinan, tugas kinerja, rekaman proses, dan bukti sosial. Analisis Percakapan Kelas merupakan salah satu jawaban metodologis terhadap kebutuhan tersebut.

Dalam pendidikan kimia, kebutuhan diagnosis semakin kuat karena peserta didik harus mengoordinasikan fenomena makroskopik, model submikroskopik, simbol, grafik, dan perhitungan. Instrumen yang hanya menguji hafalan definisi sering gagal memperlihatkan cara representasi itu dihubungkan.

Sejarah penggunaannya juga menunjukkan bahwa format yang lebih kompleks tidak otomatis lebih baik. Setiap tambahan tier, rubrik, atau kanal data harus mempunyai fungsi yang jelas dan didukung bukti bahwa informasi tambahan benar-benar memperbaiki keputusan.

- pergeseran dari pengukuran skor menuju diagnosis pola pemahaman;
- penggunaan respons terbuka dan wawancara untuk menemukan konsepsi alternatif;
- pengembangan teknik kuantitatif agar diagnosis dapat dilakukan pada kelompok besar;
- integrasi data digital dan metakognitif pada perkembangan mutakhir.

Dalam praktik kelas, Analisis Percakapan Kelas dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.16.3 Konstruk, Klaim, dan Bukti Diagnosis

Konstruk yang dinilai melalui Analisis Percakapan Kelas bukan objek yang terlihat langsung. Yang diamati adalah pertanyaan, jawaban, respons guru, perbaikan, jeda, gestur, penggunaan papan, dan rujukan pada representasi; dari sana dibuat inferensi tentang pemahaman konsep, miskonsepsi, pengetahuan parsial, atau ketidakpastian. (Chi, 1997; AERA, APA, & NCME, 2014; Messick, 1995; Kane, 2013)

Klaim diagnosis perlu dirumuskan sebagai argumen. Pertama, tugas harus memerlukan konsep sasaran. Kedua, respons harus diberi skor atau kode secara konsisten. Ketiga, pola skor harus mempunyai hubungan yang masuk akal dengan struktur pengetahuan. Keempat, penggunaan hasil harus menghasilkan keputusan yang proporsional.

Sumber ancaman utama adalah ketidakjelasan konstruk. Misalnya, peserta didik dapat gagal karena istilah yang ambigu, beban membaca, representasi yang tidak dikenal, atau prosedur matematika. Kegagalan semacam itu tidak boleh langsung diperlakukan sebagai miskonsepsi kimia.

Oleh sebab itu, setiap pengembang perlu menulis pernyataan klaim-bukti: instrumen mengukur apa, perilaku apa yang menjadi indikator, alternatif penjelasan apa yang mungkin, dan bukti tambahan apa yang diperlukan untuk memperkuat kesimpulan.

- konstruk utama: pemahaman konseptual melalui diagnosis sosial-interaksional;
- bukti observabel: respons, alasan, tindakan, atau representasi;
- inferensi: profil konsepsi pada konten dan konteks tertentu;
- penggunaan: perencanaan pembelajaran, remediasi, atau penelitian.

Dalam praktik kelas, Analisis Percakapan Kelas dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.16.4 Struktur, Komponen, dan Alur Respons

Struktur Analisis Percakapan Kelas dibangun dari komponen yang saling bergantung. Komponen inti meliputi rekaman kelas, unit episode, transkripsi, kode giliran bicara, kode jenis dialog, dan analisis isi konseptual. Hilangnya satu komponen dapat mengubah makna data yang diperoleh. (Mortimer & Scott, 2003; Chi, 2009; Chi, 1997; AERA, APA, & NCME, 2014)

Urutan penyajian harus ditetapkan secara sadar. Komponen yang muncul lebih awal dapat memberi petunjuk pada komponen berikutnya, sedangkan permintaan alasan atau keyakinan dapat mengubah cara peserta didik memeriksa jawabannya. Penguncian respons dan uji urutan diperlukan pada format digital.

Setiap komponen membutuhkan spesifikasi. Stimulus harus memuat informasi cukup, opsi atau prompt harus sejajar dengan indikator, skala harus memiliki jangkar, dan rubrik harus menjelaskan perbedaan kategori yang berdekatan.

Struktur yang baik memungkinkan peneliti menelusuri kembali setiap kategori hasil ke respons mentah. Transparansi ini penting agar guru tidak hanya menerima label otomatis tanpa memahami alasan di baliknya.

- rekaman kelas;
- unit episode;
- transkripsi;
- kode giliran bicara;
- kode jenis dialog;
- analisis isi konseptual;

Dalam praktik kelas, Analisis Percakapan Kelas dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

No.	Komponen	Fungsi Diagnostik
1	rekaman kelas	Bukti yang harus ditelusuri
2	unit episode	Bukti yang harus ditelusuri
3	transkripsi	Bukti yang harus ditelusuri
4	kode giliran bicara	Bukti yang harus ditelusuri

6.16.5 Desain Stimulus, Tugas, atau Pertanyaan

Desain stimulus pada Analisis Percakapan Kelas harus memunculkan penggunaan konsep, bukan sekadar pengenalan istilah. Situasi dapat berbentuk fenomena, gambar partikel, grafik, data eksperimen, persamaan, atau masalah kontekstual, tetapi seluruh informasi yang tidak relevan perlu dikendalikan. (Haladyna et al., 2002; Mortimer & Scott, 2003; Chi, 2009; Chi, 1997)

Dalam kimia, satu stimulus sering menggabungkan beberapa tingkat representasi. Pengembang harus memutuskan apakah tujuan butir adalah menerjemahkan representasi, menjelaskan hubungan sebab-akibat, membedakan contoh dan noncontoh, atau menerapkan prinsip pada konteks baru.

Bahasa perlu ringkas tanpa menghilangkan syarat ilmiah. Kata seperti kuat, stabil, jenuh, spontan, atau setimbang mempunyai arti teknis dan sehari-hari yang berbeda. Ketidakjelasan istilah dapat menghasilkan kesalahan yang berasal dari bahasa, bukan struktur konsep.

Desain yang baik juga menghindari petunjuk format, pilihan terlalu panjang, negasi bertumpuk, gambar dekoratif, dan data yang tidak diperlukan. Setiap elemen harus mempunyai alasan diagnostik.

- gunakan konteks yang cukup dikenal tetapi tidak menghafalkan contoh;
- pastikan hanya satu interpretasi ilmiah yang menjadi target;
- uji semua simbol, satuan, grafik, dan gambar;
- sediakan ruang bagi respons yang tidak diprediksi bila tujuan masih eksploratif.

Dalam praktik kelas, Analisis Percakapan Kelas dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.16.6 Prosedur Pengembangan Secara Sistematis

Pengembangan Analisis Percakapan Kelas harus mengikuti alur yang terdokumentasi. Langkah yang lazim ialah menentukan fokus, memilih episode, merekam beberapa kanal, menetapkan konvensi transkripsi, kemudian mengembangkan kode, menganalisis lintasan, memeriksa kasus negatif, menghubungkan dengan artefak. Urutan dapat berulang karena temuan uji coba sering memaksa revisi konstruk atau redaksi. (Mortimer & Scott, 2003; Chi, 2009; Chi, 1997; Adams & Wieman, 2011; AERA, APA, & NCME, 2014)

Tahap konseptual dimulai dengan analisis kurikulum, literatur miskonsepsi, dan peta konsep. Tahap empiris awal menggunakan wawancara, respons terbuka, atau observasi untuk memastikan bahwa kategori yang dirancang benar-benar muncul pada populasi sasaran.

Setiap versi instrumen perlu diberi nomor, tanggal, dan catatan perubahan. Revisi harus menyebutkan masalah yang ditemukan, bukti yang digunakan, perubahan yang dilakukan, dan alasan mempertahankan atau menghapus butir.

Pengembangan dianggap belum selesai ketika hanya menghasilkan butir yang tampak baik. Instrumen perlu diuji pada peserta didik yang serupa dengan pengguna sasaran, lalu dianalisis secara konseptual dan empiris sebelum dipakai untuk keputusan penting.

- 1. menentukan fokus.
- 2. memilih episode.
- 3. merekam beberapa kanal.
- 4. menetapkan konvensi transkripsi.
- 5. mengembangkan kode.

Dalam praktik kelas, Analisis Percakapan Kelas dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.16.7 Pelaksanaan dan Pengumpulan Data

Pelaksanaan Analisis Percakapan Kelas harus distandardisasi secukupnya agar perbedaan hasil lebih mungkin berasal dari pemahaman, bukan variasi instruksi, waktu, alat, atau bantuan. Petunjuk perlu menjelaskan tujuan umum tanpa mengajarkan jawaban. (Mortimer & Scott, 2003; Chi, 2009; AERA, APA, & NCME, 2014)

Kondisi administrasi mencakup waktu, urutan tugas, kesempatan bertanya, penggunaan kalkulator, akses tabel periodik, kualitas gambar, dan perangkat digital. Pada instrumen kualitatif, standardisasi berarti konsistensi fokus dan prinsip probing, bukan membaca naskah secara kaku.

Data yang dikumpulkan ialah pertanyaan, jawaban, respons guru, perbaikan, jeda, gestur, penggunaan papan, dan rujukan pada representasi. Data mentah harus disimpan bersama metadata yang diperlukan, seperti kode butir, versi instrumen, waktu administrasi, dan kondisi khusus, tetapi hanya data yang relevan yang boleh dikumpulkan.

Sebelum administrasi luas, lakukan uji teknis. Pastikan tautan, urutan, penguncian respons, rekaman, penyimpanan, serta tampilan rumus dan subskrip bekerja. Gangguan teknis perlu dicatat sebagai bagian dari interpretasi.

- berikan instruksi yang sama dan mudah dipahami;
- catat kondisi yang dapat memengaruhi respons;
- hindari umpan balik sebelum seluruh respons diagnostik selesai;
- sediakan prosedur penanganan data tidak lengkap.

Dalam praktik kelas, Analisis Percakapan Kelas dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.16.8 Penskoran, Pengodean, dan Pengolahan Data

Penskoran Analisis Percakapan Kelas dapat diringkas sebagai pengodean pola interaksi, fungsi ujaran, posisi konseptual, serta lintasan gagasan sepanjang episode. Aturan ini harus ditulis sebelum melihat hasil utama agar keputusan tidak disesuaikan secara oportunistik dengan pola data. (Mortimer & Scott, 2003; Chi, 2009; Chi, 1997; AERA, APA, & NCME, 2014)

Skor total jarang cukup untuk diagnosis. Analisis perlu mempertahankan informasi tingkat butir, konsep, opsi, alasan, keyakinan, proses, atau representasi. Penggabungan data terlalu dini dapat menyembunyikan pola yang secara pedagogis penting.

Untuk respons terbuka atau kualitatif, buku kode harus mencantumkan definisi, contoh positif, contoh batas, noncontoh, dan aturan ketika satu respons memuat beberapa gagasan. Sebagian data perlu dinilai oleh lebih dari satu penilai.

Setiap transformasi data—misalnya dari respons ke kode, dari kode ke kategori, dan dari kategori ke persentase—harus dapat diaudit. Simpan respons asli dan keputusan pengodean agar kasus ambigu dapat ditinjau kembali.

- pisahkan skor ketepatan dari skor keyakinan atau kualitas alasan;
- laporkan kasus tidak lengkap dan aturan penanganannya;
- uji konsistensi penilai untuk data berkode;
- hindari mengubah data ordinal menjadi ukuran interval tanpa alasan.

Dalam praktik kelas, Analisis Percakapan Kelas dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.16.9 Interpretasi dan Penetapan Kategori

Interpretasi hasil Analisis Percakapan Kelas harus bergerak dari respons menuju pola, bukan langsung dari satu jawaban menuju label. Status konseptual menjadi lebih kuat apabila gagasan yang sama muncul pada beberapa butir, representasi, atau konteks. (Mortimer & Scott, 2003; Chi, 2009; Chi, 1997; AERA, APA, & NCME, 2014; Kane, 2013)

Kategori dapat mencakup pemahaman ilmiah, pemahaman parsial, kurang pengetahuan, respons tidak konsisten, false positive, false negative, dan miskonsepsi. Tidak semua instrumen mampu membedakan seluruh kategori tersebut; skema harus mengikuti informasi yang benar-benar tersedia.

Ambang, seperti batas keyakinan atau persentase pemilihan, perlu memiliki rasional dan diuji melalui analisis sensitivitas. Perubahan kecil pada ambang dapat mengubah prevalensi kategori secara berarti.

Interpretasi yang bertanggung jawab menyatakan bahwa kategori berlaku pada konsep, bentuk tugas, waktu, dan kondisi tertentu. Hasil merupakan hipotesis kerja untuk pembelajaran, bukan diagnosis psikologis permanen.

- gunakan lebih dari satu butir untuk konsepsi penting;
- periksa konsistensi antara jawaban dan alasan;
- tandai kasus ambigu untuk konfirmasi;
- hubungkan kategori dengan rekomendasi yang berbeda.

Dalam praktik kelas, Analisis Percakapan Kelas dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.16.10 Validitas dan Kualitas Inferensi

Validitas Analisis Percakapan Kelas berkaitan dengan kualitas interpretasi dan penggunaan hasil, bukan sifat tetap yang dimiliki instrumen. Fokus khususnya adalah ketepatan transkripsi, kecukupan konteks, transparansi unit analisis, dan triangulasi dengan pekerjaan peserta didik. (AERA, APA, & NCME, 2014; Messick, 1995; Kane, 2013; Haynes et al., 1995)

Bukti berdasarkan isi menilai keterwakilan domain dan akurasi konsep. Bukti proses respons memeriksa bagaimana peserta didik memahami tugas. Bukti struktur internal menilai kesesuaian pola respons. Bukti hubungan dengan variabel lain dan konsekuensi penggunaan melengkapi argumen.

Ancaman validitas dapat berasal dari konten yang kurang terwakili, tuntutan bahasa berlebihan, petunjuk opsi, rubrik yang tidak stabil, keyakinan yang tidak terkalibrasi, atau algoritma yang belajar dari label keliru.

Dokumen validasi perlu menghubungkan setiap klaim dengan bukti dan alternatif penjelasan. Menampilkan satu indeks tidak cukup; validasi adalah program argumentatif yang diperbarui setiap kali instrumen dipakai pada populasi atau tujuan baru.

- bukti isi dan kesesuaian kurikulum;
- bukti proses respons melalui wawancara kognitif;
- bukti struktur dan fungsi butir;
- bukti kegunaan serta konsekuensi keputusan.

Dalam praktik kelas, Analisis Percakapan Kelas dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.16.11 Reliabilitas, Keajegan, dan Analisis Empiris

Keajegan Analisis Percakapan Kelas harus dinilai sesuai jenis data. Pada respons objektif dapat digunakan konsistensi internal, stabilitas waktu, atau model pengukuran. Pada data kualitatif, perhatian diarahkan pada konsistensi pengodean, audit trail, dan ketepatan penerapan definisi kode. (Kuder & Richardson, 1937; Cronbach, 1951; Ding et al., 2006; Wren & Barbera, 2014; AERA, APA, & NCME, 2014)

Reliabilitas tinggi tidak otomatis menunjukkan diagnosis yang benar. Butir yang berulang-ulang mengukur hafalan dapat menghasilkan koefisien tinggi tetapi tidak mewakili pemahaman konseptual. Sebaliknya, konstruk yang multidimensi tidak selalu layak dipaksa menjadi satu skala.

Analisis empiris perlu memeriksa kesukaran, daya beda, fungsi opsi, pola data hilang, dan keterhubungan komponen. Untuk instrumen bertingkat, reliabilitas dapat dihitung pada setiap tier serta kombinasi yang mempunyai makna teoritis.

Ukuran sampel, distribusi kemampuan, dan homogenitas konten memengaruhi statistik. Karena itu, laporan harus menyatakan kondisi estimasi dan tidak menyalin batas angka sebagai aturan universal.

- pilih statistik yang sesuai dengan jenis skor;
- laporkan interval atau ketidakpastian bila tersedia;
- periksa butir secara substantif dan statistik;
- uji kembali setelah revisi atau perubahan platform.

Dalam praktik kelas, Analisis Percakapan Kelas dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.16.12 Contoh Penerapan dalam Pembelajaran Kimia

Penerapan Analisis Percakapan Kelas dalam kimia dapat dimulai dari konsep diskusi tentang ikatan dan kestabilan. Pengembang perlu menuliskan konsep ilmiah target, miskonsepsi yang mungkin, bukti respons yang diharapkan, dan keputusan pembelajaran setelah hasil diperoleh. (Mortimer & Scott, 2003; Chi, 2009; Chi, 1997; AERA, APA, & NCME, 2014; Nakhleh, 1992)

Contoh kedua adalah argumentasi perubahan fase. Stimulus sebaiknya menghubungkan fenomena dengan model partikel atau persamaan, lalu meminta peserta didik menunjukkan hubungan sebab-akibat. Jawaban yang hanya menyebut istilah belum cukup menjadi bukti pemahaman.

Pada percakapan kesetimbangan dinamis, diagnosis dapat memeriksa apakah peserta didik membedakan perubahan laju, keadaan setimbang, energi, dan komposisi. Kesalahan sering muncul ketika satu aturan dipindahkan ke konteks yang tidak sesuai.

Kasus negosiasi makna istilah kuat dan pekat menunjukkan pentingnya representasi. Respons verbal, gambar, simbol, dan hitungan dapat tidak konsisten. Hasil instrumen sebaiknya digunakan untuk memilih demonstrasi, diskusi, contoh-noncontoh, atau tugas remediasi yang secara langsung menantang hubungan salah.

- tentukan proposisi ilmiah dan konsepsi alternatif;
- gunakan minimal dua bentuk representasi;
- minta alasan atau bukti ketika format memungkinkan;
- rancang tindak lanjut berdasarkan pola, bukan hanya skor.

Dalam praktik kelas, Analisis Percakapan Kelas dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.16.13 Kelebihan, Keterbatasan, dan Potensi Bias

Kekuatan utama Analisis Percakapan Kelas meliputi menangkap konteks autentik, menunjukkan peran bahasa guru, mengidentifikasi miskonsepsi kolektif, dan menghubungkan diagnosis dengan pedagogi. Kekuatan tersebut menjelaskan situasi ketika instrumen layak dipilih dibanding teknik lain. (Mortimer & Scott, 2003; Chi, 2009; Chi, 1997; Gurel et al., 2015; AERA, APA, & NCME, 2014)

Keterbatasannya mencakup transkripsi rumit, sulit memisahkan konsepsi individu dari tekanan sosial, kamera memengaruhi perilaku, dan interpretasi memerlukan konteks. Keterbatasan harus diperlakukan sebagai batas inferensi, bukan sekadar paragraf formal yang tidak memengaruhi analisis.

Potensi bias dapat berasal dari bahasa, pengalaman konteks, akses teknologi, kecemasan, gaya menggunakan skala, keterampilan menggambar, atau kemampuan menulis. Instrumen perlu diuji pada kelompok yang beragam dan ditinjau untuk fungsi berbeda antarkelompok bila data mencukupi.

Keputusan memilih instrumen sebaiknya mempertimbangkan kedalaman data, ukuran kelompok, waktu, keterampilan analisis, tujuan diagnosis, dan konsekuensi kesalahan klasifikasi. Tidak ada satu metode yang unggul untuk semua tujuan.

- Kekuatan: menangkap konteks autentik.
- Kekuatan: menunjukkan peran bahasa guru.
- Batas: transkripsi rumit.
- Batas: sulit memisahkan konsepsi individu dari tekanan sosial.

Dalam praktik kelas, Analisis Percakapan Kelas dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

Aspek	Temuan	Respons Metodologis
Kekuatan	menangkap konteks autentik	Pilih ketika sesuai tujuan
Keterbatasan	transkripsi rumit	Tambahkan bukti atau kontrol
Bias potensial	Bahasa, konteks, perangkat, gaya respons	Uji proses respons

6.16.14 Integrasi Digital, Etika, dan Tata Kelola Data

Integrasi digital Analisis Percakapan Kelas dapat berupa rekaman multi-kamera, transkripsi otomatis yang dikoreksi, serta penautan ujaran dengan slide dan papan tulis. Teknologi seharusnya memperluas kualitas bukti, bukan hanya memindahkan lembar kertas ke layar. (Baker & Inventado, 2014; Romero & Ventura, 2020; AERA, APA, & NCME, 2014; Mortimer & Scott, 2003)

Platform perlu menyimpan versi instrumen, urutan respons, timestamp yang relevan, aturan skor, dan log perubahan. Namun, hanya data yang dibutuhkan untuk tujuan diagnosis yang boleh dikumpulkan; pengumpulan berlebihan meningkatkan risiko privasi tanpa manfaat ilmiah.

Aspek etika utama adalah bahwa persetujuan seluruh pihak, pengaburan identitas, pengelolaan suara dan wajah, serta pembatasan akses rekaman. Peserta didik dan guru perlu mengetahui data apa yang dikumpulkan, untuk tujuan apa, siapa yang dapat mengakses, berapa lama disimpan, dan bagaimana kesalahan data diperbaiki.

Algoritma otomatis harus dapat diaudit. Kategori yang dihasilkan sistem perlu ditautkan ke respons mentah dan aturan keputusan. Kasus berisiko atau ambigu harus ditinjau manusia sebelum tindak lanjut penting dilakukan.

- gunakan identitas pseudonim dan akses berbasis peran;
- uji tampilan rumus, subskrip, grafik, dan gambar;
- sediakan opsi aksesibilitas dan alternatif perangkat;
- dokumentasikan algoritma serta perubahan versi.

Dalam praktik kelas, Analisis Percakapan Kelas dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.16.15 Sintesis dan Pedoman Implementasi

Analisis Percakapan Kelas bermanfaat apabila dipilih berdasarkan tujuan dan digunakan sesuai batasnya. Instrumen ini menghasilkan pertanyaan, jawaban, respons guru, perbaikan, jeda, gestur, penggunaan papan, dan rujukan pada representasi, lalu data diolah melalui pengodean pola interaksi, fungsi ujaran, posisi konseptual, serta lintasan gagasan sepanjang episode. (Mortimer & Scott, 2003; Chi, 2009; Chi, 1997; AERA, APA, & NCME, 2014; Kane, 2013)

Rangkaian kerja yang disarankan dimulai dari analisis konsep, identifikasi konsepsi alternatif, desain tugas, telaah ahli, uji proses respons, uji lapangan, analisis, penetapan kategori, dan penyusunan tindak lanjut. Setiap tahap harus meninggalkan dokumentasi yang dapat diperiksa.

Hasil diagnosis tidak berhenti pada tabel prevalensi. Guru perlu memilih konsepsi prioritas berdasarkan frekuensi, keyakinan, posisi sebagai konsep prasyarat, dampak terhadap topik lanjutan, dan kemungkinan diperbaiki melalui pembelajaran.

Prinsip akhirnya ialah kehati-hatian inferensial: sebutkan apa yang benar-benar ditunjukkan data, bedakan bukti kuat dari indikasi, gunakan triangulasi ketika diperlukan, dan hindari label permanen pada peserta didik.

- selaraskan tujuan, konstruk, tugas, skor, kategori, dan tindakan;
- utamakan konsepsi yang berdampak pada pembelajaran lanjutan;
- gabungkan bukti bila keputusan memiliki konsekuensi besar;
- evaluasi kembali efektivitas tindak lanjut melalui diagnosis ulang.

Dalam praktik kelas, Analisis Percakapan Kelas dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.17 Penggunaan Log Aktivitas Digital

6.17.1 Landasan Konseptual dan Kedudukan Instrumen

Dalam diagnosis miskonsepsi, Penggunaan Log Aktivitas Digital perlu dipahami bukan sekadar sebagai bentuk soal atau teknik pengumpulan data, melainkan sebagai perangkat inferensi. Dalam bab ini, instrumen tersebut didefinisikan sebagai pemanfaatan catatan interaksi peserta didik dengan platform, simulasi, kuis, laboratorium virtual, atau sistem pembelajaran untuk merekonstruksi pola tindakan. Definisi operasional ini menegaskan bentuk bukti yang dikumpulkan sekaligus membatasi jenis kesimpulan yang layak dibuat. (Baker & Inventado, 2014; Siemens & Baker, 2012; Romero & Ventura, 2020)

Klaim utamanya adalah bahwa instrumen ini menyediakan bukti temporal dan berskala besar tentang urutan klik, revisi, waktu, penggunaan bantuan, dan jalur representasi yang berkaitan dengan konsepsi. Klaim tersebut harus dipisahkan dari anggapan bahwa setiap jawaban salah merupakan miskonsepsi. Kesalahan dapat berasal dari kurang pengetahuan, ketidakcermatan, kesulitan bahasa, salah hitung, atau ketidakfamiliaran dengan representasi.

Data pokok yang dihasilkan berupa klik, respons, durasi, urutan layar, revisi, penggunaan petunjuk, hasil simulasi, dan navigasi. Data itu baru memperoleh makna diagnostik setelah dikaitkan dengan peta konsep sasaran, kemungkinan konsepsi alternatif, dan aturan interpretasi yang telah ditetapkan sebelum analisis.

Karena miskonsepsi bersifat kontekstual dan kadang tersembunyi, bukti tunggal perlu ditafsirkan secara terbatas. Posisi terbaik Penggunaan Log Aktivitas Digital adalah sebagai bagian dari sistem asesmen berlapis yang dapat dipadukan dengan bukti lain ketika konsekuensi keputusan semakin besar.

- menentukan konsep dan proposisi yang hendak didiagnosis;
- menjelaskan perilaku respons yang dianggap sebagai bukti;
- membatasi klaim sesuai kekuatan instrumen;
- menautkan hasil dengan tindakan pembelajaran yang realistis.

Dalam praktik kelas, Penggunaan Log Aktivitas Digital dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.17.2 Perkembangan Historis dan Rasional Penggunaan

Perkembangan Penggunaan Log Aktivitas Digital tidak dapat dipisahkan dari perubahan pandangan tentang belajar sains. Ketika peserta didik tidak lagi dipandang sebagai penerima informasi yang kosong, instrumen mulai dirancang untuk mengungkap gagasan awal, pola penalaran, dan keyakinan yang dibawa ke kelas. (Baker & Inventado, 2014; Siemens & Baker, 2012; Romero & Ventura, 2020; AERA, APA, & NCME, 2014)

Pada tahap awal, penilaian banyak mengandalkan benar-salah atau skor total. Perkembangan penelitian konsepsi peserta didik kemudian mendorong penambahan alasan, keyakinan, tugas kinerja, rekaman proses, dan bukti sosial. Penggunaan Log Aktivitas Digital merupakan salah satu jawaban metodologis terhadap kebutuhan tersebut.

Dalam pendidikan kimia, kebutuhan diagnosis semakin kuat karena peserta didik harus mengoordinasikan fenomena makroskopik, model submikroskopik, simbol, grafik, dan perhitungan. Instrumen yang hanya menguji hafalan definisi sering gagal memperlihatkan cara representasi itu dihubungkan.

Sejarah penggunaannya juga menunjukkan bahwa format yang lebih kompleks tidak otomatis lebih baik. Setiap tambahan tier, rubrik, atau kanal data harus mempunyai fungsi yang jelas dan didukung bukti bahwa informasi tambahan benar-benar memperbaiki keputusan.

- pergeseran dari pengukuran skor menuju diagnosis pola pemahaman;
- penggunaan respons terbuka dan wawancara untuk menemukan konsepsi alternatif;
- pengembangan teknik kuantitatif agar diagnosis dapat dilakukan pada kelompok besar;
- integrasi data digital dan metakognitif pada perkembangan mutakhir.

Dalam praktik kelas, Penggunaan Log Aktivitas Digital dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.17.3 Konstruk, Klaim, dan Bukti Diagnosis

Konstruk yang dinilai melalui Penggunaan Log Aktivitas Digital bukan objek yang terlihat langsung. Yang diamati adalah klik, respons, durasi, urutan layar, revisi, penggunaan petunjuk, hasil simulasi, dan navigasi; dari sana dibuat inferensi tentang pemahaman konsep, miskonsepsi, pengetahuan parsial, atau ketidakpastian. (AERA, APA, & NCME, 2014; Flake & Fried, 2020; Messick, 1995; Kane, 2013)

Klaim diagnosis perlu dirumuskan sebagai argumen. Pertama, tugas harus memerlukan konsep sasaran. Kedua, respons harus diberi skor atau kode secara konsisten. Ketiga, pola skor harus mempunyai hubungan yang masuk akal dengan struktur pengetahuan. Keempat, penggunaan hasil harus menghasilkan keputusan yang proporsional.

Sumber ancaman utama adalah ketidakjelasan konstruk. Misalnya, peserta didik dapat gagal karena istilah yang ambigu, beban membaca, representasi yang tidak dikenal, atau prosedur matematika. Kegagalan semacam itu tidak boleh langsung diperlakukan sebagai miskonsepsi kimia.

Oleh sebab itu, setiap pengembang perlu menulis pernyataan klaim-bukti: instrumen mengukur apa, perilaku apa yang menjadi indikator, alternatif penjelasan apa yang mungkin, dan bukti tambahan apa yang diperlukan untuk memperkuat kesimpulan.

- konstruk utama: pemahaman konseptual melalui diagnosis berbasis jejak data;
- bukti observabel: respons, alasan, tindakan, atau representasi;
- inferensi: profil konsepsi pada konten dan konteks tertentu;
- penggunaan: perencanaan pembelajaran, remediasi, atau penelitian.

Dalam praktik kelas, Penggunaan Log Aktivitas Digital dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.17.4 Struktur, Komponen, dan Alur Respons

Struktur Penggunaan Log Aktivitas Digital dibangun dari komponen yang saling bergantung. Komponen inti meliputi event log, identitas pseudonim, timestamp, konteks aktivitas, fitur turunan, model analitik, dan prosedur validasi. Hilangnya satu komponen dapat mengubah makna data yang diperoleh. (Baker & Inventado, 2014; Siemens & Baker, 2012; Romero & Ventura, 2020; AERA, APA, & NCME, 2014)

Urutan penyajian harus ditetapkan secara sadar. Komponen yang muncul lebih awal dapat memberi petunjuk pada komponen berikutnya, sedangkan permintaan alasan atau keyakinan dapat mengubah cara peserta didik memeriksa jawabannya. Penguncian respons dan uji urutan diperlukan pada format digital.

Setiap komponen membutuhkan spesifikasi. Stimulus harus memuat informasi cukup, opsi atau prompt harus sejajar dengan indikator, skala harus memiliki jangkar, dan rubrik harus menjelaskan perbedaan kategori yang berdekatan.

Struktur yang baik memungkinkan peneliti menelusuri kembali setiap kategori hasil ke respons mentah. Transparansi ini penting agar guru tidak hanya menerima label otomatis tanpa memahami alasan di baliknya.

- event log;
- identitas pseudonim;
- timestamp;
- konteks aktivitas;
- fitur turunan;
- model analitik;
- prosedur validasi;

Dalam praktik kelas, Penggunaan Log Aktivitas Digital dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

No.	Komponen	Fungsi Diagnostik
1	event log	Bukti yang harus ditelusuri
2	identitas pseudonim	Bukti yang harus ditelusuri
3	timestamp	Bukti yang harus ditelusuri
4	konteks aktivitas	Bukti yang harus ditelusuri

6.17.5 Desain Stimulus, Tugas, atau Pertanyaan

Desain stimulus pada Penggunaan Log Aktivitas Digital harus memunculkan penggunaan konsep, bukan sekadar pengenalan istilah. Situasi dapat berbentuk fenomena, gambar partikel, grafik, data eksperimen, persamaan, atau masalah kontekstual, tetapi seluruh informasi yang tidak relevan perlu dikendalikan. (Haladyna et al., 2002; Baker & Inventado, 2014; Siemens & Baker, 2012; Romero & Ventura, 2020)

Dalam kimia, satu stimulus sering menggabungkan beberapa tingkat representasi. Pengembang harus memutuskan apakah tujuan butir adalah menerjemahkan representasi, menjelaskan hubungan sebab-akibat, membedakan contoh dan noncontoh, atau menerapkan prinsip pada konteks baru.

Bahasa perlu ringkas tanpa menghilangkan syarat ilmiah. Kata seperti kuat, stabil, jenuh, spontan, atau setimbang mempunyai arti teknis dan sehari-hari yang berbeda. Ketidakjelasan istilah dapat menghasilkan kesalahan yang berasal dari bahasa, bukan struktur konsep.

Desain yang baik juga menghindari petunjuk format, pilihan terlalu panjang, negasi bertumpuk, gambar dekoratif, dan data yang tidak diperlukan. Setiap elemen harus mempunyai alasan diagnostik.

- gunakan konteks yang cukup dikenal tetapi tidak menghafalkan contoh;
- pastikan hanya satu interpretasi ilmiah yang menjadi target;
- uji semua simbol, satuan, grafik, dan gambar;
- sediakan ruang bagi respons yang tidak diprediksi bila tujuan masih eksploratif.

Dalam praktik kelas, Penggunaan Log Aktivitas Digital dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.17.6 Prosedur Pengembangan Secara Sistematis

Pengembangan Penggunaan Log Aktivitas Digital harus mengikuti alur yang terdokumentasi. Langkah yang lazim ialah mendefinisikan pertanyaan diagnosis, merancang skema log, memastikan timestamp, membersihkan data, kemudian membangun fitur, menghubungkan dengan bukti lain, memvalidasi model, menerapkan dashboard. Urutan dapat berulang karena temuan uji coba sering memaksa revisi konstruk atau redaksi. (Baker & Inventado, 2014; Siemens & Baker, 2012; Romero & Ventura, 2020; Adams & Wieman, 2011; AERA, APA, & NCME, 2014)

Tahap konseptual dimulai dengan analisis kurikulum, literatur miskonsepsi, dan peta konsep. Tahap empiris awal menggunakan wawancara, respons terbuka, atau observasi untuk memastikan bahwa kategori yang dirancang benar-benar muncul pada populasi sasaran.

Setiap versi instrumen perlu diberi nomor, tanggal, dan catatan perubahan. Revisi harus menyebutkan masalah yang ditemukan, bukti yang digunakan, perubahan yang dilakukan, dan alasan mempertahankan atau menghapus butir.

Pengembangan dianggap belum selesai ketika hanya menghasilkan butir yang tampak baik. Instrumen perlu diuji pada peserta didik yang serupa dengan pengguna sasaran, lalu dianalisis secara konseptual dan empiris sebelum dipakai untuk keputusan penting.

- 1. mendefinisikan pertanyaan diagnosis.
- 2. merancang skema log.
- 3. memastikan timestamp.
- 4. membersihkan data.
- 5. membangun fitur.

Dalam praktik kelas, Penggunaan Log Aktivitas Digital dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.17.7 Pelaksanaan dan Pengumpulan Data

Pelaksanaan Penggunaan Log Aktivitas Digital harus distandardisasi secukupnya agar perbedaan hasil lebih mungkin berasal dari pemahaman, bukan variasi instruksi, waktu, alat, atau bantuan. Petunjuk perlu menjelaskan tujuan umum tanpa mengajarkan jawaban. (Baker & Inventado, 2014; Siemens & Baker, 2012; AERA, APA, & NCME, 2014)

Kondisi administrasi mencakup waktu, urutan tugas, kesempatan bertanya, penggunaan kalkulator, akses tabel periodik, kualitas gambar, dan perangkat digital. Pada instrumen kualitatif, standardisasi berarti konsistensi fokus dan prinsip probing, bukan membaca naskah secara kaku.

Data yang dikumpulkan ialah klik, respons, durasi, urutan layar, revisi, penggunaan petunjuk, hasil simulasi, dan navigasi. Data mentah harus disimpan bersama metadata yang diperlukan, seperti kode butir, versi instrumen, waktu administrasi, dan kondisi khusus, tetapi hanya data yang relevan yang boleh dikumpulkan.

Sebelum administrasi luas, lakukan uji teknis. Pastikan tautan, urutan, penguncian respons, rekaman, penyimpanan, serta tampilan rumus dan subskrip bekerja. Gangguan teknis perlu dicatat sebagai bagian dari interpretasi.

- berikan instruksi yang sama dan mudah dipahami;
- catat kondisi yang dapat memengaruhi respons;
- hindari umpan balik sebelum seluruh respons diagnostik selesai;
- sediakan prosedur penanganan data tidak lengkap.

Dalam praktik kelas, Penggunaan Log Aktivitas Digital dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.17.8 Penskoran, Pengodean, dan Pengolahan Data

Penskoran Penggunaan Log Aktivitas Digital dapat diringkas sebagai ekstraksi fitur, analisis urutan, aturan pola, model klasifikasi, dan pengaitan dengan hasil tes atau wawancara. Aturan ini harus ditulis sebelum melihat hasil utama agar keputusan tidak disesuaikan secara oportunistik dengan pola data. (Baker & Inventado, 2014; Siemens & Baker, 2012; Romero & Ventura, 2020; Chi, 1997; AERA, APA, & NCME, 2014)

Skor total jarang cukup untuk diagnosis. Analisis perlu mempertahankan informasi tingkat butir, konsep, opsi, alasan, keyakinan, proses, atau representasi. Penggabungan data terlalu dini dapat menyembunyikan pola yang secara pedagogis penting.

Untuk respons terbuka atau kualitatif, buku kode harus mencantumkan definisi, contoh positif, contoh batas, noncontoh, dan aturan ketika satu respons memuat beberapa gagasan. Sebagian data perlu dinilai oleh lebih dari satu penilai.

Setiap transformasi data—misalnya dari respons ke kode, dari kode ke kategori, dan dari kategori ke persentase—harus dapat diaudit. Simpan respons asli dan keputusan pengodean agar kasus ambigu dapat ditinjau kembali.

- pisahkan skor ketepatan dari skor keyakinan atau kualitas alasan;
- laporkan kasus tidak lengkap dan aturan penanganannya;
- uji konsistensi penilai untuk data berkode;
- hindari mengubah data ordinal menjadi ukuran interval tanpa alasan.

Dalam praktik kelas, Penggunaan Log Aktivitas Digital dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.17.9 Interpretasi dan Penetapan Kategori

Interpretasi hasil Penggunaan Log Aktivitas Digital harus bergerak dari respons menuju pola, bukan langsung dari satu jawaban menuju label. Status konseptual menjadi lebih kuat apabila gagasan yang sama muncul pada beberapa butir, representasi, atau konteks. (Baker & Inventado, 2014; Siemens & Baker, 2012; Romero & Ventura, 2020; AERA, APA, & NCME, 2014; Kane, 2013)

Kategori dapat mencakup pemahaman ilmiah, pemahaman parsial, kurang pengetahuan, respons tidak konsisten, false positive, false negative, dan miskonsepsi. Tidak semua instrumen mampu membedakan seluruh kategori tersebut; skema harus mengikuti informasi yang benar-benar tersedia.

Ambang, seperti batas keyakinan atau persentase pemilihan, perlu memiliki rasional dan diuji melalui analisis sensitivitas. Perubahan kecil pada ambang dapat mengubah prevalensi kategori secara berarti.

Interpretasi yang bertanggung jawab menyatakan bahwa kategori berlaku pada konsep, bentuk tugas, waktu, dan kondisi tertentu. Hasil merupakan hipotesis kerja untuk pembelajaran, bukan diagnosis psikologis permanen.

- gunakan lebih dari satu butir untuk konsepsi penting;
- periksa konsistensi antara jawaban dan alasan;
- tandai kasus ambigu untuk konfirmasi;
- hubungkan kategori dengan rekomendasi yang berbeda.

Dalam praktik kelas, Penggunaan Log Aktivitas Digital dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.17.10 Validitas dan Kualitas Inferensi

Validitas Penggunaan Log Aktivitas Digital berkaitan dengan kualitas interpretasi dan penggunaan hasil, bukan sifat tetap yang dimiliki instrumen. Fokus khususnya adalah hubungan fitur log dengan proses konseptual yang diklaim dan evaluasi model pada data baru, bukan hanya akurasi pelatihan. (AERA, APA, & NCME, 2014; Messick, 1995; Kane, 2013; Haynes et al., 1995)

Bukti berdasarkan isi menilai keterwakilan domain dan akurasi konsep. Bukti proses respons memeriksa bagaimana peserta didik memahami tugas. Bukti struktur internal menilai kesesuaian pola respons. Bukti hubungan dengan variabel lain dan konsekuensi penggunaan melengkapi argumen.

Ancaman validitas dapat berasal dari konten yang kurang terwakili, tuntutan bahasa berlebihan, petunjuk opsi, rubrik yang tidak stabil, keyakinan yang tidak terkalibrasi, atau algoritma yang belajar dari label keliru.

Dokumen validasi perlu menghubungkan setiap klaim dengan bukti dan alternatif penjelasan. Menampilkan satu indeks tidak cukup; validasi adalah program argumentatif yang diperbarui setiap kali instrumen dipakai pada populasi atau tujuan baru.

- bukti isi dan kesesuaian kurikulum;
- bukti proses respons melalui wawancara kognitif;
- bukti struktur dan fungsi butir;
- bukti kegunaan serta konsekuensi keputusan.

Dalam praktik kelas, Penggunaan Log Aktivitas Digital dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.17.11 Reliabilitas, Keajegan, dan Analisis Empiris

Keajegan Penggunaan Log Aktivitas Digital harus dinilai sesuai jenis data. Pada respons objektif dapat digunakan konsistensi internal, stabilitas waktu, atau model pengukuran. Pada data kualitatif, perhatian diarahkan pada konsistensi pengodean, audit trail, dan ketepatan penerapan definisi kode. (Kuder & Richardson, 1937; Cronbach, 1951; Ding et al., 2006; Wren & Barbera, 2014; AERA, APA, & NCME, 2014)

Reliabilitas tinggi tidak otomatis menunjukkan diagnosis yang benar. Butir yang berulang-ulang mengukur hafalan dapat menghasilkan koefisien tinggi tetapi tidak mewakili pemahaman konseptual. Sebaliknya, konstruk yang multidimensi tidak selalu layak dipaksa menjadi satu skala.

Analisis empiris perlu memeriksa kesukaran, daya beda, fungsi opsi, pola data hilang, dan keterhubungan komponen. Untuk instrumen bertingkat, reliabilitas dapat dihitung pada setiap tier serta kombinasi yang mempunyai makna teoritis.

Ukuran sampel, distribusi kemampuan, dan homogenitas konten memengaruhi statistik. Karena itu, laporan harus menyatakan kondisi estimasi dan tidak menyalin batas angka sebagai aturan universal.

- pilih statistik yang sesuai dengan jenis skor;
- laporkan interval atau ketidakpastian bila tersedia;
- periksa butir secara substantif dan statistik;
- uji kembali setelah revisi atau perubahan platform.

Dalam praktik kelas, Penggunaan Log Aktivitas Digital dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.17.12 Contoh Penerapan dalam Pembelajaran Kimia

Penerapan Penggunaan Log Aktivitas Digital dalam kimia dapat dimulai dari konsep urutan manipulasi konsentrasi pada simulasi kesetimbangan. Pengembang perlu menuliskan konsep ilmiah target, miskonsepsi yang mungkin, bukti respons yang diharapkan, dan keputusan pembelajaran setelah hasil diperoleh. (Baker & Inventado, 2014; Siemens & Baker, 2012; Romero & Ventura, 2020; AERA, APA, & NCME, 2014; Nakhleh, 1992)

Contoh kedua adalah pola penggunaan representasi partikel. Stimulus sebaiknya menghubungkan fenomena dengan model partikel atau persamaan, lalu meminta peserta didik menunjukkan hubungan sebab-akibat. Jawaban yang hanya menyebut istilah belum cukup menjadi bukti pemahaman.

Pada revisi jawaban kuis asam-basa, diagnosis dapat memeriksa apakah peserta didik membedakan perubahan laju, keadaan setimbang, energi, dan komposisi. Kesalahan sering muncul ketika satu aturan dipindahkan ke konteks yang tidak sesuai.

Kasus jalur eksperimen virtual kinetika menunjukkan pentingnya representasi. Respons verbal, gambar, simbol, dan hitungan dapat tidak konsisten. Hasil instrumen sebaiknya digunakan untuk memilih demonstrasi, diskusi, contoh-noncontoh, atau tugas remediasi yang secara langsung menantang hubungan salah.

- tentukan proposisi ilmiah dan konsepsi alternatif;
- gunakan minimal dua bentuk representasi;
- minta alasan atau bukti ketika format memungkinkan;
- rancang tindak lanjut berdasarkan pola, bukan hanya skor.

Dalam praktik kelas, Penggunaan Log Aktivitas Digital dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.17.13 Kelebihan, Keterbatasan, dan Potensi Bias

Kekuatan utama Penggunaan Log Aktivitas Digital meliputi data berskala besar, menangkap proses secara tidak mengganggu, mendukung umpan balik cepat, dan memungkinkan analisis perubahan waktu. Kekuatan tersebut menjelaskan situasi ketika instrumen layak dipilih dibanding teknik lain. (Baker & Inventado, 2014; Siemens & Baker, 2012; Romero & Ventura, 2020; Gurel et al., 2015; AERA, APA, & NCME, 2014)

Keterbatasannya mencakup klik tidak memiliki makna tanpa konteks, durasi panjang dapat berarti banyak hal, model rentan bias, dan data hilang dan perangkat berbeda mengganggu. Keterbatasan harus diperlakukan sebagai batas inferensi, bukan sekadar paragraf formal yang tidak memengaruhi analisis.

Potensi bias dapat berasal dari bahasa, pengalaman konteks, akses teknologi, kecemasan, gaya menggunakan skala, keterampilan menggambar, atau kemampuan menulis. Instrumen perlu diuji pada kelompok yang beragam dan ditinjau untuk fungsi berbeda antarkelompok bila data mencukupi.

Keputusan memilih instrumen sebaiknya mempertimbangkan kedalaman data, ukuran kelompok, waktu, keterampilan analisis, tujuan diagnosis, dan konsekuensi kesalahan klasifikasi. Tidak ada satu metode yang unggul untuk semua tujuan.

- Kekuatan: data berskala besar.
- Kekuatan: menangkap proses secara tidak mengganggu.
- Batas: klik tidak memiliki makna tanpa konteks.
- Batas: durasi panjang dapat berarti banyak hal.

Dalam praktik kelas, Penggunaan Log Aktivitas Digital dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

Aspek	Temuan	Respons Metodologis
Kekuatan	data berskala besar	Pilih ketika sesuai tujuan
Keterbatasan	klik tidak memiliki makna tanpa konteks	Tambahkan bukti atau kontrol
Bias potensial	Bahasa, konteks, perangkat, gaya respons	Uji proses respons

6.17.14 Integrasi Digital, Etika, dan Tata Kelola Data

Integrasi digital Penggunaan Log Aktivitas Digital dapat berupa merupakan medium utama; desain harus mengutamakan interoperabilitas, dokumentasi skema, dan kontrol akses. Teknologi seharusnya memperluas kualitas bukti, bukan hanya memindahkan lembar kertas ke layar. (Baker & Inventado, 2014; Romero & Ventura, 2020; AERA, APA, & NCME, 2014)

Platform perlu menyimpan versi instrumen, urutan respons, timestamp yang relevan, aturan skor, dan log perubahan. Namun, hanya data yang dibutuhkan untuk tujuan diagnosis yang boleh dikumpulkan; pengumpulan berlebihan meningkatkan risiko privasi tanpa manfaat ilmiah.

Aspek etika utama adalah bahwa minimalisasi data, persetujuan, keamanan, transparansi algoritma, serta larangan membuat keputusan penting hanya dari prediksi otomatis. Peserta didik dan guru perlu mengetahui data apa yang dikumpulkan, untuk tujuan apa, siapa yang dapat mengakses, berapa lama disimpan, dan bagaimana kesalahan data diperbaiki.

Algoritma otomatis harus dapat diaudit. Kategori yang dihasilkan sistem perlu ditautkan ke respons mentah dan aturan keputusan. Kasus berisiko atau ambigu harus ditinjau manusia sebelum tindak lanjut penting dilakukan.

- gunakan identitas pseudonim dan akses berbasis peran;
- uji tampilan rumus, subskrip, grafik, dan gambar;
- sediakan opsi aksesibilitas dan alternatif perangkat;
- dokumentasikan algoritma serta perubahan versi.

Dalam praktik kelas, Penggunaan Log Aktivitas Digital dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.17.15 Sintesis dan Pedoman Implementasi

Penggunaan Log Aktivitas Digital bermanfaat apabila dipilih berdasarkan tujuan dan digunakan sesuai batasnya. Instrumen ini menghasilkan klik, respons, durasi, urutan layar, revisi, penggunaan petunjuk, hasil simulasi, dan navigasi, lalu data diolah melalui ekstraksi fitur, analisis urutan, aturan pola, model klasifikasi, dan pengaitan dengan hasil tes atau wawancara. (Baker & Inventado, 2014; Siemens & Baker, 2012; Romero & Ventura, 2020; AERA, APA, & NCME, 2014; Kane, 2013)

Rangkaian kerja yang disarankan dimulai dari analisis konsep, identifikasi konsepsi alternatif, desain tugas, telaah ahli, uji proses respons, uji lapangan, analisis, penetapan kategori, dan penyusunan tindak lanjut. Setiap tahap harus meninggalkan dokumentasi yang dapat diperiksa.

Hasil diagnosis tidak berhenti pada tabel prevalensi. Guru perlu memilih konsepsi prioritas berdasarkan frekuensi, keyakinan, posisi sebagai konsep prasyarat, dampak terhadap topik lanjutan, dan kemungkinan diperbaiki melalui pembelajaran.

Prinsip akhirnya ialah kehati-hatian inferensial: sebutkan apa yang benar-benar ditunjukkan data, bedakan bukti kuat dari indikasi, gunakan triangulasi ketika diperlukan, dan hindari label permanen pada peserta didik.

- selaraskan tujuan, konstruk, tugas, skor, kategori, dan tindakan;
- utamakan konsepsi yang berdampak pada pembelajaran lanjutan;
- gabungkan bukti bila keputusan memiliki konsekuensi besar;
- evaluasi kembali efektivitas tindak lanjut melalui diagnosis ulang.

Dalam praktik kelas, Penggunaan Log Aktivitas Digital dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.18 Pengembangan Distraktor Berbasis Miskonsepsi

6.18.1 Landasan Konseptual dan Kedudukan Instrumen

Dalam diagnosis miskonsepsi, Pengembangan Distraktor Berbasis Miskonsepsi perlu dipahami bukan sekadar sebagai bentuk soal atau teknik pengumpulan data, melainkan sebagai perangkat inferensi. Dalam bab ini, instrumen tersebut didefinisikan sebagai proses membangun opsi salah yang masuk akal karena merepresentasikan konsepsi alternatif nyata, bukan sekadar jawaban acak atau kesalahan hitung sepele. Definisi operasional ini menegaskan bentuk bukti yang dikumpulkan sekaligus membatasi jenis kesimpulan yang layak dibuat. (Sadler, 1998; Haladyna et al., 2002; Tarrant et al., 2009)

Klaim utamanya adalah bahwa instrumen ini mengubah pilihan opsi menjadi bukti tentang pola pemikiran dengan syarat distraktor memiliki dasar empiris dan fungsi psikometrik. Klaim tersebut harus dipisahkan dari anggapan bahwa setiap jawaban salah merupakan miskonsepsi. Kesalahan dapat berasal dari kurang pengetahuan, ketidakcermatan, kesulitan bahasa, salah hitung, atau ketidakfamiliaran dengan representasi.

Data pokok yang dihasilkan berupa sumber kualitatif distraktor, frekuensi pilihan, daya tarik menurut kelompok kemampuan, dan alasan pemilihan. Data itu baru memperoleh makna diagnostik setelah dikaitkan dengan peta konsep sasaran, kemungkinan konsepsi alternatif, dan aturan interpretasi yang telah ditetapkan sebelum analisis.

Karena miskonsepsi bersifat kontekstual dan kadang tersembunyi, bukti tunggal perlu ditafsirkan secara terbatas. Posisi terbaik Pengembangan Distraktor Berbasis Miskonsepsi adalah sebagai bagian dari sistem asesmen berlapis yang dapat dipadukan dengan bukti lain ketika konsekuensi keputusan semakin besar.

- menentukan konsep dan proposisi yang hendak didiagnosis;
- menjelaskan perilaku respons yang dianggap sebagai bukti;
- membatasi klaim sesuai kekuatan instrumen;
- menautkan hasil dengan tindakan pembelajaran yang realistis.

Dalam praktik kelas, Pengembangan Distraktor Berbasis Miskonsepsi dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.18.2 Perkembangan Historis dan Rasional Penggunaan

Perkembangan Pengembangan Distraktor Berbasis Miskonsepsi tidak dapat dipisahkan dari perubahan pandangan tentang belajar sains. Ketika peserta didik tidak lagi dipandang sebagai penerima informasi yang kosong, instrumen mulai dirancang untuk mengungkap gagasan awal, pola penalaran, dan keyakinan yang dibawa ke kelas. (Sadler, 1998; Haladyna et al., 2002; Tarrant et al., 2009; Rodriguez, 2005)

Pada tahap awal, penilaian banyak mengandalkan benar-salah atau skor total. Perkembangan penelitian konsepsi peserta didik kemudian mendorong penambahan alasan, keyakinan, tugas kinerja, rekaman proses, dan bukti sosial. Pengembangan Distraktor Berbasis Miskonsepsi merupakan salah satu jawaban metodologis terhadap kebutuhan tersebut.

Dalam pendidikan kimia, kebutuhan diagnosis semakin kuat karena peserta didik harus mengoordinasikan fenomena makroskopik, model submikroskopik, simbol, grafik, dan perhitungan. Instrumen yang hanya menguji hafalan definisi sering gagal memperlihatkan cara representasi itu dihubungkan.

Sejarah penggunaannya juga menunjukkan bahwa format yang lebih kompleks tidak otomatis lebih baik. Setiap tambahan tier, rubrik, atau kanal data harus mempunyai fungsi yang jelas dan didukung bukti bahwa informasi tambahan benar-benar memperbaiki keputusan.

- pergeseran dari pengukuran skor menuju diagnosis pola pemahaman;
- penggunaan respons terbuka dan wawancara untuk menemukan konsepsi alternatif;
- pengembangan teknik kuantitatif agar diagnosis dapat dilakukan pada kelompok besar;
- integrasi data digital dan metakognitif pada perkembangan mutakhir.

Dalam praktik kelas, Pengembangan Distraktor Berbasis Miskonsepsi dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.18.3 Konstruk, Klaim, dan Bukti Diagnosis

Konstruk yang dinilai melalui Pengembangan Distraktor Berbasis Miskonsepsi bukan objek yang terlihat langsung. Yang diamati adalah sumber kualitatif distraktor, frekuensi pilihan, daya tarik menurut kelompok kemampuan, dan alasan pemilihan; dari sana dibuat inferensi tentang pemahaman konsep, miskonsepsi, pengetahuan parsial, atau ketidakpastian. (Treagust, 1988; Gurel et al., 2015; Messick, 1995; Kane, 2013)

Klaim diagnosis perlu dirumuskan sebagai argumen. Pertama, tugas harus memerlukan konsep sasaran. Kedua, respons harus diberi skor atau kode secara konsisten. Ketiga, pola skor harus mempunyai hubungan yang masuk akal dengan struktur pengetahuan. Keempat, penggunaan hasil harus menghasilkan keputusan yang proporsional.

Sumber ancaman utama adalah ketidakjelasan konstruk. Misalnya, peserta didik dapat gagal karena istilah yang ambigu, beban membaca, representasi yang tidak dikenal, atau prosedur matematika. Kegagalan semacam itu tidak boleh langsung diperlakukan sebagai miskonsepsi kimia.

Oleh sebab itu, setiap pengembang perlu menulis pernyataan klaim-bukti: instrumen mengukur apa, perilaku apa yang menjadi indikator, alternatif penjelasan apa yang mungkin, dan bukti tambahan apa yang diperlukan untuk memperkuat kesimpulan.

- konstruk utama: pemahaman konseptual melalui rekayasa opsi diagnostik;
- bukti observabel: respons, alasan, tindakan, atau representasi;
- inferensi: profil konsepsi pada konten dan konteks tertentu;
- penggunaan: perencanaan pembelajaran, remediasi, atau penelitian.

Dalam praktik kelas, Pengembangan Distraktor Berbasis Miskonsepsi dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.18.4 Struktur, Komponen, dan Alur Respons

Struktur Pengembangan Distraktor Berbasis Miskonsepsi dibangun dari komponen yang saling bergantung. Komponen inti meliputi sumber miskonsepsi, rumusan proposisi salah, opsi homogen, kunci ilmiah, kode konsepsi, dan statistik pemilihan opsi. Hilangnya satu komponen dapat mengubah makna data yang diperoleh. (Sadler, 1998; Haladyna et al., 2002; Tarrant et al., 2009; Rodriguez, 2005)

Urutan penyajian harus ditetapkan secara sadar. Komponen yang muncul lebih awal dapat memberi petunjuk pada komponen berikutnya, sedangkan permintaan alasan atau keyakinan dapat mengubah cara peserta didik memeriksa jawabannya. Penguncian respons dan uji urutan diperlukan pada format digital.

Setiap komponen membutuhkan spesifikasi. Stimulus harus memuat informasi cukup, opsi atau prompt harus sejajar dengan indikator, skala harus memiliki jangkar, dan rubrik harus menjelaskan perbedaan kategori yang berdekatan.

Struktur yang baik memungkinkan peneliti menelusuri kembali setiap kategori hasil ke respons mentah. Transparansi ini penting agar guru tidak hanya menerima label otomatis tanpa memahami alasan di baliknya.

- sumber miskonsepsi;
- rumusan proposisi salah;
- opsi homogen;
- kunci ilmiah;
- kode konsepsi;
- statistik pemilihan opsi;

Dalam praktik kelas, Pengembangan Distraktor Berbasis Miskonsepsi dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

No.	Komponen	Fungsi Diagnostik
1	sumber miskonsepsi	Bukti yang harus ditelusuri
2	rumusan proposisi salah	Bukti yang harus ditelusuri
3	opsi homogen	Bukti yang harus ditelusuri
4	kunci ilmiah	Bukti yang harus ditelusuri

6.18.5 Desain Stimulus, Tugas, atau Pertanyaan

Desain stimulus pada Pengembangan Distraktor Berbasis Miskonsepsi harus memunculkan penggunaan konsep, bukan sekadar pengenalan istilah. Situasi dapat berbentuk fenomena, gambar partikel, grafik, data eksperimen, persamaan, atau masalah kontekstual, tetapi seluruh informasi yang tidak relevan perlu dikendalikan. (Haladyna et al., 2002; Sadler, 1998; Tarrant et al., 2009)

Dalam kimia, satu stimulus sering menggabungkan beberapa tingkat representasi. Pengembang harus memutuskan apakah tujuan butir adalah menerjemahkan representasi, menjelaskan hubungan sebab-akibat, membedakan contoh dan noncontoh, atau menerapkan prinsip pada konteks baru.

Bahasa perlu ringkas tanpa menghilangkan syarat ilmiah. Kata seperti kuat, stabil, jenuh, spontan, atau setimbang mempunyai arti teknis dan sehari-hari yang berbeda. Ketidakjelasan istilah dapat menghasilkan kesalahan yang berasal dari bahasa, bukan struktur konsep.

Desain yang baik juga menghindari petunjuk format, pilihan terlalu panjang, negasi bertumpuk, gambar dekoratif, dan data yang tidak diperlukan. Setiap elemen harus mempunyai alasan diagnostik.

- gunakan konteks yang cukup dikenal tetapi tidak menghafalkan contoh;
- pastikan hanya satu interpretasi ilmiah yang menjadi target;
- uji semua simbol, satuan, grafik, dan gambar;
- sediakan ruang bagi respons yang tidak diprediksi bila tujuan masih eksploratif.

Dalam praktik kelas, Pengembangan Distraktor Berbasis Miskonsepsi dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.18.6 Prosedur Pengembangan Secara Sistematis

Pengembangan Pengembangan Distraktor Berbasis Miskonsepsi harus mengikuti alur yang terdokumentasi. Langkah yang lazim ialah menelaah literatur, melakukan wawancara, mengumpulkan respons terbuka, merumuskan proposisi, kemudian menulis opsi paralel, telaah ahli, uji kognitif, analisis fungsi. Urutan dapat berulang karena temuan uji coba sering memaksa revisi konstruk atau redaksi. (Sadler, 1998; Haladyna et al., 2002; Tarrant et al., 2009; Adams & Wieman, 2011; AERA, APA, & NCME, 2014)

Tahap konseptual dimulai dengan analisis kurikulum, literatur miskonsepsi, dan peta konsep. Tahap empiris awal menggunakan wawancara, respons terbuka, atau observasi untuk memastikan bahwa kategori yang dirancang benar-benar muncul pada populasi sasaran.

Setiap versi instrumen perlu diberi nomor, tanggal, dan catatan perubahan. Revisi harus menyebutkan masalah yang ditemukan, bukti yang digunakan, perubahan yang dilakukan, dan alasan mempertahankan atau menghapus butir.

Pengembangan dianggap belum selesai ketika hanya menghasilkan butir yang tampak baik. Instrumen perlu diuji pada peserta didik yang serupa dengan pengguna sasaran, lalu dianalisis secara konseptual dan empiris sebelum dipakai untuk keputusan penting.

- 1. menelaah literatur.
- 2. melakukan wawancara.
- 3. mengumpulkan respons terbuka.
- 4. merumuskan proposisi.
- 5. menulis opsi paralel.

Dalam praktik kelas, Pengembangan Distraktor Berbasis Miskonsepsi dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.18.7 Pelaksanaan dan Pengumpulan Data

Pelaksanaan Pengembangan Distraktor Berbasis Miskonsepsi harus distandardisasi secukupnya agar perbedaan hasil lebih mungkin berasal dari pemahaman, bukan variasi instruksi, waktu, alat, atau bantuan. Petunjuk perlu menjelaskan tujuan umum tanpa mengajarkan jawaban. (Sadler, 1998; Haladyna et al., 2002; AERA, APA, & NCME, 2014)

Kondisi administrasi mencakup waktu, urutan tugas, kesempatan bertanya, penggunaan kalkulator, akses tabel periodik, kualitas gambar, dan perangkat digital. Pada instrumen kualitatif, standardisasi berarti konsistensi fokus dan prinsip probing, bukan membaca naskah secara kaku.

Data yang dikumpulkan ialah sumber kualitatif distraktor, frekuensi pilihan, daya tarik menurut kelompok kemampuan, dan alasan pemilihan. Data mentah harus disimpan bersama metadata yang diperlukan, seperti kode butir, versi instrumen, waktu administrasi, dan kondisi khusus, tetapi hanya data yang relevan yang boleh dikumpulkan.

Sebelum administrasi luas, lakukan uji teknis. Pastikan tautan, urutan, penguncian respons, rekaman, penyimpanan, serta tampilan rumus dan subskrip bekerja. Gangguan teknis perlu dicatat sebagai bagian dari interpretasi.

- berikan instruksi yang sama dan mudah dipahami;
- catat kondisi yang dapat memengaruhi respons;
- hindari umpan balik sebelum seluruh respons diagnostik selesai;
- sediakan prosedur penanganan data tidak lengkap.

Dalam praktik kelas, Pengembangan Distraktor Berbasis Miskonsepsi dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.18.8 Penskoran, Pengodean, dan Pengolahan Data

Penskoran Pengembangan Distraktor Berbasis Miskonsepsi dapat diringkas sebagai analisis fungsi setiap distraktor, indeks pemilihan, hubungan dengan alasan, dan konsistensi lintasbutir. Aturan ini harus ditulis sebelum melihat hasil utama agar keputusan tidak disesuaikan secara oportunistik dengan pola data. (Sadler, 1998; Haladyna et al., 2002; Tarrant et al., 2009; Chi, 1997; AERA, APA, & NCME, 2014)

Skor total jarang cukup untuk diagnosis. Analisis perlu mempertahankan informasi tingkat butir, konsep, opsi, alasan, keyakinan, proses, atau representasi. Penggabungan data terlalu dini dapat menyembunyikan pola yang secara pedagogis penting.

Untuk respons terbuka atau kualitatif, buku kode harus mencantumkan definisi, contoh positif, contoh batas, noncontoh, dan aturan ketika satu respons memuat beberapa gagasan. Sebagian data perlu dinilai oleh lebih dari satu penilai.

Setiap transformasi data—misalnya dari respons ke kode, dari kode ke kategori, dan dari kategori ke persentase—harus dapat diaudit. Simpan respons asli dan keputusan pengodean agar kasus ambigu dapat ditinjau kembali.

- pisahkan skor ketepatan dari skor keyakinan atau kualitas alasan;
- laporkan kasus tidak lengkap dan aturan penanganannya;
- uji konsistensi penilai untuk data berkode;
- hindari mengubah data ordinal menjadi ukuran interval tanpa alasan.

Dalam praktik kelas, Pengembangan Distraktor Berbasis Miskonsepsi dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.18.9 Interpretasi dan Penetapan Kategori

Interpretasi hasil Pengembangan Distraktor Berbasis Miskonsepsi harus bergerak dari respons menuju pola, bukan langsung dari satu jawaban menuju label. Status konseptual menjadi lebih kuat apabila gagasan yang sama muncul pada beberapa butir, representasi, atau konteks. (Sadler, 1998; Haladyna et al., 2002; Tarrant et al., 2009; Rodriguez, 2005; Kane, 2013)

Kategori dapat mencakup pemahaman ilmiah, pemahaman parsial, kurang pengetahuan, respons tidak konsisten, false positive, false negative, dan miskonsepsi. Tidak semua instrumen mampu membedakan seluruh kategori tersebut; skema harus mengikuti informasi yang benar-benar tersedia.

Ambang, seperti batas keyakinan atau persentase pemilihan, perlu memiliki rasional dan diuji melalui analisis sensitivitas. Perubahan kecil pada ambang dapat mengubah prevalensi kategori secara berarti.

Interpretasi yang bertanggung jawab menyatakan bahwa kategori berlaku pada konsep, bentuk tugas, waktu, dan kondisi tertentu. Hasil merupakan hipotesis kerja untuk pembelajaran, bukan diagnosis psikologis permanen.

- gunakan lebih dari satu butir untuk konsepsi penting;
- periksa konsistensi antara jawaban dan alasan;
- tandai kasus ambigu untuk konfirmasi;
- hubungkan kategori dengan rekomendasi yang berbeda.

Dalam praktik kelas, Pengembangan Distraktor Berbasis Miskonsepsi dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.18.10 Validitas dan Kualitas Inferensi

Validitas Pengembangan Distraktor Berbasis Miskonsepsi berkaitan dengan kualitas interpretasi dan penggunaan hasil, bukan sifat tetap yang dimiliki instrumen. Fokus khususnya adalah bukti bahwa opsi benar-benar ditafsirkan sesuai konsepsi yang dimaksud, bukan karena bahasa, panjang, atau petunjuk teknis. (AERA, APA, & NCME, 2014; Messick, 1995; Kane, 2013; Haynes et al., 1995)

Bukti berdasarkan isi menilai keterwakilan domain dan akurasi konsep. Bukti proses respons memeriksa bagaimana peserta didik memahami tugas. Bukti struktur internal menilai kesesuaian pola respons. Bukti hubungan dengan variabel lain dan konsekuensi penggunaan melengkapi argumen.

Ancaman validitas dapat berasal dari konten yang kurang terwakili, tuntutan bahasa berlebihan, petunjuk opsi, rubrik yang tidak stabil, keyakinan yang tidak terkalibrasi, atau algoritma yang belajar dari label keliru.

Dokumen validasi perlu menghubungkan setiap klaim dengan bukti dan alternatif penjelasan. Menampilkan satu indeks tidak cukup; validasi adalah program argumentatif yang diperbarui setiap kali instrumen dipakai pada populasi atau tujuan baru.

- bukti isi dan kesesuaian kurikulum;
- bukti proses respons melalui wawancara kognitif;
- bukti struktur dan fungsi butir;
- bukti kegunaan serta konsekuensi keputusan.

Dalam praktik kelas, Pengembangan Distraktor Berbasis Miskonsepsi dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.18.11 Reliabilitas, Keajegan, dan Analisis Empiris

Keajegan Pengembangan Distraktor Berbasis Miskonsepsi harus dinilai sesuai jenis data. Pada respons objektif dapat digunakan konsistensi internal, stabilitas waktu, atau model pengukuran. Pada data kualitatif, perhatian diarahkan pada konsistensi pengodean, audit trail, dan ketepatan penerapan definisi kode. (Kuder & Richardson, 1937; Cronbach, 1951; Ding et al., 2006; Wren & Barbera, 2014; AERA, APA, & NCME, 2014)

Reliabilitas tinggi tidak otomatis menunjukkan diagnosis yang benar. Butir yang berulang-ulang mengukur hafalan dapat menghasilkan koefisien tinggi tetapi tidak mewakili pemahaman konseptual. Sebaliknya, konstruk yang multidimensi tidak selalu layak dipaksa menjadi satu skala.

Analisis empiris perlu memeriksa kesukaran, daya beda, fungsi opsi, pola data hilang, dan keterhubungan komponen. Untuk instrumen bertingkat, reliabilitas dapat dihitung pada setiap tier serta kombinasi yang mempunyai makna teoritis.

Ukuran sampel, distribusi kemampuan, dan homogenitas konten memengaruhi statistik. Karena itu, laporan harus menyatakan kondisi estimasi dan tidak menyalin batas angka sebagai aturan universal.

- pilih statistik yang sesuai dengan jenis skor;
- laporkan interval atau ketidakpastian bila tersedia;
- periksa butir secara substantif dan statistik;
- uji kembali setelah revisi atau perubahan platform.

Dalam praktik kelas, Pengembangan Distraktor Berbasis Miskonsepsi dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.18.12 Contoh Penerapan dalam Pembelajaran Kimia

Penerapan Pengembangan Distraktor Berbasis Miskonsepsi dalam kimia dapat dimulai dari konsep katalis menaikkan jumlah produk setimbang. Pengembang perlu menuliskan konsep ilmiah target, miskonsepsi yang mungkin, bukti respons yang diharapkan, dan keputusan pembelajaran setelah hasil diperoleh. (Sadler, 1998; Haladyna et al., 2002; Tarrant et al., 2009; Rodriguez, 2005; Nakhleh, 1992)

Contoh kedua adalah ikatan putus melepaskan energi. Stimulus sebaiknya menghubungkan fenomena dengan model partikel atau persamaan, lalu meminta peserta didik menunjukkan hubungan sebab-akibat. Jawaban yang hanya menyebut istilah belum cukup menjadi bukti pemahaman.

Pada asam kuat selalu pekat, diagnosis dapat memeriksa apakah peserta didik membedakan perubahan laju, keadaan setimbang, energi, dan komposisi. Kesalahan sering muncul ketika satu aturan dipindahkan ke konteks yang tidak sesuai.

Kasus partikel membesar saat dipanaskan menunjukkan pentingnya representasi. Respons verbal, gambar, simbol, dan hitungan dapat tidak konsisten. Hasil instrumen sebaiknya digunakan untuk memilih demonstrasi, diskusi, contoh-noncontoh, atau tugas remediasi yang secara langsung menantang hubungan salah.

- tentukan proposisi ilmiah dan konsepsi alternatif;
- gunakan minimal dua bentuk representasi;
- minta alasan atau bukti ketika format memungkinkan;
- rancang tindak lanjut berdasarkan pola, bukan hanya skor.

Dalam praktik kelas, Pengembangan Distraktor Berbasis Miskonsepsi dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.18.13 Kelebihan, Keterbatasan, dan Potensi Bias

Kekuatan utama Pengembangan Distraktor Berbasis Miskonsepsi meliputi meningkatkan makna respons salah, mendukung diagnosis kelompok besar, menghubungkan studi kualitatif dengan tes, dan memperbaiki kualitas item. Kekuatan tersebut menjelaskan situasi ketika instrumen layak dipilih dibanding teknik lain. (Sadler, 1998; Haladyna et al., 2002; Tarrant et al., 2009; Gurel et al., 2015; AERA, APA, & NCME, 2014)

Keterbatasannya mencakup distraktor dapat memperkenalkan gagasan baru, opsi terlalu jelas tidak berfungsi, satu opsi dapat mewakili beberapa alasan, dan popularitas opsi tidak otomatis membuktikan miskonsepsi. Keterbatasan harus diperlakukan sebagai batas inferensi, bukan sekadar paragraf formal yang tidak memengaruhi analisis.

Potensi bias dapat berasal dari bahasa, pengalaman konteks, akses teknologi, kecemasan, gaya menggunakan skala, keterampilan menggambar, atau kemampuan menulis. Instrumen perlu diuji pada kelompok yang beragam dan ditinjau untuk fungsi berbeda antarkelompok bila data mencukupi.

Keputusan memilih instrumen sebaiknya mempertimbangkan kedalaman data, ukuran kelompok, waktu, keterampilan analisis, tujuan diagnosis, dan konsekuensi kesalahan klasifikasi. Tidak ada satu metode yang unggul untuk semua tujuan.

- Kekuatan: meningkatkan makna respons salah.
- Kekuatan: mendukung diagnosis kelompok besar.
- Batas: distraktor dapat memperkenalkan gagasan baru.
- Batas: opsi terlalu jelas tidak berfungsi.

Dalam praktik kelas, Pengembangan Distraktor Berbasis Miskonsepsi dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

Aspek	Temuan	Respons Metodologis
Kekuatan	meningkatkan makna respons salah	Pilih ketika sesuai tujuan
Keterbatasan	distraktor dapat memperkenalkan gagasan baru	Tambahkan bukti atau kontrol
Bias potensial	Bahasa, konteks, perangkat, gaya respons	Uji proses respons

6.18.14 Integrasi Digital, Etika, dan Tata Kelola Data

Integrasi digital Pengembangan Distraktor Berbasis Miskonsepsi dapat berupa bank distraktor dengan metadata konsepsi, statistik lintasadministrasi, dan pengacakan posisi opsi tanpa mengubah makna. Teknologi seharusnya memperluas kualitas bukti, bukan hanya memindahkan lembar kertas ke layar. (Baker & Inventado, 2014; Romero & Ventura, 2020; AERA, APA, & NCME, 2014; Sadler, 1998)

Platform perlu menyimpan versi instrumen, urutan respons, timestamp yang relevan, aturan skor, dan log perubahan. Namun, hanya data yang dibutuhkan untuk tujuan diagnosis yang boleh dikumpulkan; pengumpulan berlebihan meningkatkan risiko privasi tanpa manfaat ilmiah.

Aspek etika utama adalah bahwa distraktor tidak boleh menghina, menstereotipkan, atau sengaja menyesatkan melalui trik yang tidak relevan dengan konsep. Peserta didik dan guru perlu mengetahui data apa yang dikumpulkan, untuk tujuan apa, siapa yang dapat mengakses, berapa lama disimpan, dan bagaimana kesalahan data diperbaiki.

Algoritma otomatis harus dapat diaudit. Kategori yang dihasilkan sistem perlu ditautkan ke respons mentah dan aturan keputusan. Kasus berisiko atau ambigu harus ditinjau manusia sebelum tindak lanjut penting dilakukan.

- gunakan identitas pseudonim dan akses berbasis peran;
- uji tampilan rumus, subskrip, grafik, dan gambar;
- sediakan opsi aksesibilitas dan alternatif perangkat;
- dokumentasikan algoritma serta perubahan versi.

Dalam praktik kelas, Pengembangan Distraktor Berbasis Miskonsepsi dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.18.15 Sintesis dan Pedoman Implementasi

Pengembangan Distraktor Berbasis Miskonsepsi bermanfaat apabila dipilih berdasarkan tujuan dan digunakan sesuai batasnya. Instrumen ini menghasilkan sumber kualitatif distraktor, frekuensi pilihan, daya tarik menurut kelompok kemampuan, dan alasan pemilihan, lalu data diolah melalui analisis fungsi setiap distraktor, indeks pemilihan, hubungan dengan alasan, dan konsistensi lintasbutir. (Sadler, 1998; Haladyna et al., 2002; Tarrant et al., 2009; AERA, APA, & NCME, 2014; Kane, 2013)

Rangkaian kerja yang disarankan dimulai dari analisis konsep, identifikasi konsepsi alternatif, desain tugas, telaah ahli, uji proses respons, uji lapangan, analisis, penetapan kategori, dan penyusunan tindak lanjut. Setiap tahap harus meninggalkan dokumentasi yang dapat diperiksa.

Hasil diagnosis tidak berhenti pada tabel prevalensi. Guru perlu memilih konsepsi prioritas berdasarkan frekuensi, keyakinan, posisi sebagai konsep prasyarat, dampak terhadap topik lanjutan, dan kemungkinan diperbaiki melalui pembelajaran.

Prinsip akhirnya ialah kehati-hatian inferensial: sebutkan apa yang benar-benar ditunjukkan data, bedakan bukti kuat dari indikasi, gunakan triangulasi ketika diperlukan, dan hindari label permanen pada peserta didik.

- selaraskan tujuan, konstruk, tugas, skor, kategori, dan tindakan;
- utamakan konsepsi yang berdampak pada pembelajaran lanjutan;
- gabungkan bukti bila keputusan memiliki konsekuensi besar;
- evaluasi kembali efektivitas tindak lanjut melalui diagnosis ulang.

Dalam praktik kelas, Pengembangan Distraktor Berbasis Miskonsepsi dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.19 Validasi Isi oleh Ahli

6.19.1 Landasan Konseptual dan Kedudukan Instrumen

Dalam diagnosis miskonsepsi, Validasi Isi oleh Ahli perlu dipahami bukan sekadar sebagai bentuk soal atau teknik pengumpulan data, melainkan sebagai perangkat inferensi. Dalam bab ini, instrumen tersebut didefinisikan sebagai proses sistematis meminta ahli menilai relevansi, representativitas, kejelasan, akurasi ilmiah, dan kesesuaian pedagogis setiap komponen instrumen. Definisi operasional ini menegaskan bentuk bukti yang dikumpulkan sekaligus membatasi jenis kesimpulan yang layak dibuat. (Lawshe, 1975; Lynn, 1986; Aiken, 1985)

Klaim utamanya adalah bahwa instrumen ini memberikan bukti bahwa instrumen mencakup domain yang dimaksud dan tidak memuat distorsi konsep, tetapi bukan satu-satunya bukti validitas. Klaim tersebut harus dipisahkan dari anggapan bahwa setiap jawaban salah merupakan miskonsepsi. Kesalahan dapat berasal dari kurang pengetahuan, ketidakcermatan, kesulitan bahasa, salah hitung, atau ketidakfamiliaran dengan representasi.

Data pokok yang dihasilkan berupa rating ahli, alasan, saran redaksi, tanda konflik konsep, dan keputusan revisi. Data itu baru memperoleh makna diagnostik setelah dikaitkan dengan peta konsep sasaran, kemungkinan konsepsi alternatif, dan aturan interpretasi yang telah ditetapkan sebelum analisis.

Pembedaan ini penting karena satu pola respons dapat memiliki lebih dari satu penjelasan yang masuk akal. Posisi terbaik Validasi Isi oleh Ahli adalah sebagai bagian dari sistem asesmen berlapis yang dapat dipadukan dengan bukti lain ketika konsekuensi keputusan semakin besar.

- menentukan konsep dan proposisi yang hendak didiagnosis;
- menjelaskan perilaku respons yang dianggap sebagai bukti;
- membatasi klaim sesuai kekuatan instrumen;
- menautkan hasil dengan tindakan pembelajaran yang realistis.

Dalam praktik kelas, Validasi Isi oleh Ahli dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.19.2 Perkembangan Historis dan Rasional Penggunaan

Perkembangan Validasi Isi oleh Ahli tidak dapat dipisahkan dari perubahan pandangan tentang belajar sains. Ketika peserta didik tidak lagi dipandang sebagai penerima informasi yang kosong, instrumen mulai dirancang untuk mengungkap gagasan awal, pola penalaran, dan keyakinan yang dibawa ke kelas. (Lawshe, 1975; Lynn, 1986; Aiken, 1985; Polit & Beck, 2006)

Pada tahap awal, penilaian banyak mengandalkan benar-salah atau skor total. Perkembangan penelitian konsepsi peserta didik kemudian mendorong penambahan alasan, keyakinan, tugas kinerja, rekaman proses, dan bukti sosial. Validasi Isi oleh Ahli merupakan salah satu jawaban metodologis terhadap kebutuhan tersebut.

Dalam pendidikan kimia, kebutuhan diagnosis semakin kuat karena peserta didik harus mengoordinasikan fenomena makroskopik, model submikroskopik, simbol, grafik, dan perhitungan. Instrumen yang hanya menguji hafalan definisi sering gagal memperlihatkan cara representasi itu dihubungkan.

Sejarah penggunaannya juga menunjukkan bahwa format yang lebih kompleks tidak otomatis lebih baik. Setiap tambahan tier, rubrik, atau kanal data harus mempunyai fungsi yang jelas dan didukung bukti bahwa informasi tambahan benar-benar memperbaiki keputusan.

- pergeseran dari pengukuran skor menuju diagnosis pola pemahaman;
- penggunaan respons terbuka dan wawancara untuk menemukan konsepsi alternatif;
- pengembangan teknik kuantitatif agar diagnosis dapat dilakukan pada kelompok besar;
- integrasi data digital dan metakognitif pada perkembangan mutakhir.

Dalam praktik kelas, Validasi Isi oleh Ahli dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.19.3 Konstruk, Klaim, dan Bukti Diagnosis

Konstruk yang dinilai melalui Validasi Isi oleh Ahli bukan objek yang terlihat langsung. Yang diamati adalah rating ahli, alasan, saran redaksi, tanda konflik konsep, dan keputusan revisi; dari sana dibuat inferensi tentang pemahaman konsep, miskonsepsi, pengetahuan parsial, atau ketidakpastian. (AERA, APA, & NCME, 2014; Messick, 1995; Kane, 2013)

Klaim diagnosis perlu dirumuskan sebagai argumen. Pertama, tugas harus memerlukan konsep sasaran. Kedua, respons harus diberi skor atau kode secara konsisten. Ketiga, pola skor harus mempunyai hubungan yang masuk akal dengan struktur pengetahuan. Keempat, penggunaan hasil harus menghasilkan keputusan yang proporsional.

Sumber ancaman utama adalah ketidakjelasan konstruk. Misalnya, peserta didik dapat gagal karena istilah yang ambigu, beban membaca, representasi yang tidak dikenal, atau prosedur matematika. Kegagalan semacam itu tidak boleh langsung diperlakukan sebagai miskonsepsi kimia.

Oleh sebab itu, setiap pengembang perlu menulis pernyataan klaim-bukti: instrumen mengukur apa, perilaku apa yang menjadi indikator, alternatif penjelasan apa yang mungkin, dan bukti tambahan apa yang diperlukan untuk memperkuat kesimpulan.

- konstruk utama: pemahaman konseptual melalui pengumpulan bukti validitas isi;
- bukti observabel: respons, alasan, tindakan, atau representasi;
- inferensi: profil konsepsi pada konten dan konteks tertentu;
- penggunaan: perencanaan pembelajaran, remediasi, atau penelitian.

Dalam praktik kelas, Validasi Isi oleh Ahli dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.19.4 Struktur, Komponen, dan Alur Respons

Struktur Validasi Isi oleh Ahli dibangun dari komponen yang saling bergantung. Komponen inti meliputi definisi domain, kriteria ahli, lembar penilaian, skala relevansi, komentar kualitatif, indeks kesepakatan, dan catatan revisi. Hilangnya satu komponen dapat mengubah makna data yang diperoleh. (Lawshe, 1975; Lynn, 1986; Aiken, 1985; Polit & Beck, 2006)

Urutan penyajian harus ditetapkan secara sadar. Komponen yang muncul lebih awal dapat memberi petunjuk pada komponen berikutnya, sedangkan permintaan alasan atau keyakinan dapat mengubah cara peserta didik memeriksa jawabannya. Penguncian respons dan uji urutan diperlukan pada format digital.

Setiap komponen membutuhkan spesifikasi. Stimulus harus memuat informasi cukup, opsi atau prompt harus sejajar dengan indikator, skala harus memiliki jangkar, dan rubrik harus menjelaskan perbedaan kategori yang berdekatan.

Struktur yang baik memungkinkan peneliti menelusuri kembali setiap kategori hasil ke respons mentah. Transparansi ini penting agar guru tidak hanya menerima label otomatis tanpa memahami alasan di baliknya.

- definisi domain;
- kriteria ahli;
- lembar penilaian;
- skala relevansi;
- komentar kualitatif;
- indeks kesepakatan;
- catatan revisi;

Dalam praktik kelas, Validasi Isi oleh Ahli dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

No.	Komponen	Fungsi Diagnostik
1	definisi domain	Bukti yang harus ditelusuri
2	kriteria ahli	Bukti yang harus ditelusuri
3	lembar penilaian	Bukti yang harus ditelusuri
4	skala relevansi	Bukti yang harus ditelusuri

6.19.5 Desain Stimulus, Tugas, atau Pertanyaan

Desain stimulus pada Validasi Isi oleh Ahli harus memunculkan penggunaan konsep, bukan sekadar pengenalan istilah. Situasi dapat berbentuk fenomena, gambar partikel, grafik, data eksperimen, persamaan, atau masalah kontekstual, tetapi seluruh informasi yang tidak relevan perlu dikendalikan. (Haladyna et al., 2002; Lawshe, 1975; Lynn, 1986; Aiken, 1985)

Dalam kimia, satu stimulus sering menggabungkan beberapa tingkat representasi. Pengembang harus memutuskan apakah tujuan butir adalah menerjemahkan representasi, menjelaskan hubungan sebab-akibat, membedakan contoh dan noncontoh, atau menerapkan prinsip pada konteks baru.

Bahasa perlu ringkas tanpa menghilangkan syarat ilmiah. Kata seperti kuat, stabil, jenuh, spontan, atau setimbang mempunyai arti teknis dan sehari-hari yang berbeda. Ketidakjelasan istilah dapat menghasilkan kesalahan yang berasal dari bahasa, bukan struktur konsep.

Desain yang baik juga menghindari petunjuk format, pilihan terlalu panjang, negasi bertumpuk, gambar dekoratif, dan data yang tidak diperlukan. Setiap elemen harus mempunyai alasan diagnostik.

- gunakan konteks yang cukup dikenal tetapi tidak menghafalkan contoh;
- pastikan hanya satu interpretasi ilmiah yang menjadi target;
- uji semua simbol, satuan, grafik, dan gambar;
- sediakan ruang bagi respons yang tidak diprediksi bila tujuan masih eksploratif.

Dalam praktik kelas, Validasi Isi oleh Ahli dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.19.6 Prosedur Pengembangan Secara Sistematis

Pengembangan Validasi Isi oleh Ahli harus mengikuti alur yang terdokumentasi. Langkah yang lazim ialah mendefinisikan domain, memilih panel, menyiapkan kisi-kisi, menulis pedoman rating, kemudian mengumpulkan penilaian independen, menghitung indeks, membahas konflik, mendokumentasikan revisi. Urutan dapat berulang karena temuan uji coba sering memaksa revisi konstruk atau redaksi. (Lawshe, 1975; Lynn, 1986; Aiken, 1985; Adams & Wieman, 2011; AERA, APA, & NCME, 2014)

Tahap konseptual dimulai dengan analisis kurikulum, literatur miskonsepsi, dan peta konsep. Tahap empiris awal menggunakan wawancara, respons terbuka, atau observasi untuk memastikan bahwa kategori yang dirancang benar-benar muncul pada populasi sasaran.

Setiap versi instrumen perlu diberi nomor, tanggal, dan catatan perubahan. Revisi harus menyebutkan masalah yang ditemukan, bukti yang digunakan, perubahan yang dilakukan, dan alasan mempertahankan atau menghapus butir.

Pengembangan dianggap belum selesai ketika hanya menghasilkan butir yang tampak baik. Instrumen perlu diuji pada peserta didik yang serupa dengan pengguna sasaran, lalu dianalisis secara konseptual dan empiris sebelum dipakai untuk keputusan penting.

- 1. mendefinisikan domain.
- 2. memilih panel.
- 3. menyiapkan kisi-kisi.
- 4. menulis pedoman rating.
- 5. mengumpulkan penilaian independen.

Dalam praktik kelas, Validasi Isi oleh Ahli dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.19.7 Pelaksanaan dan Pengumpulan Data

Pelaksanaan Validasi Isi oleh Ahli harus distandardisasi secukupnya agar perbedaan hasil lebih mungkin berasal dari pemahaman, bukan variasi instruksi, waktu, alat, atau bantuan. Petunjuk perlu menjelaskan tujuan umum tanpa mengajarkan jawaban. (Lawshe, 1975; Lynn, 1986; AERA, APA, & NCME, 2014)

Kondisi administrasi mencakup waktu, urutan tugas, kesempatan bertanya, penggunaan kalkulator, akses tabel periodik, kualitas gambar, dan perangkat digital. Pada instrumen kualitatif, standardisasi berarti konsistensi fokus dan prinsip probing, bukan membaca naskah secara kaku.

Data yang dikumpulkan ialah rating ahli, alasan, saran redaksi, tanda konflik konsep, dan keputusan revisi. Data mentah harus disimpan bersama metadata yang diperlukan, seperti kode butir, versi instrumen, waktu administrasi, dan kondisi khusus, tetapi hanya data yang relevan yang boleh dikumpulkan.

Sebelum administrasi luas, lakukan uji teknis. Pastikan tautan, urutan, penguncian respons, rekaman, penyimpanan, serta tampilan rumus dan subskrip bekerja. Gangguan teknis perlu dicatat sebagai bagian dari interpretasi.

- berikan instruksi yang sama dan mudah dipahami;
- catat kondisi yang dapat memengaruhi respons;
- hindari umpan balik sebelum seluruh respons diagnostik selesai;
- sediakan prosedur penanganan data tidak lengkap.

Dalam praktik kelas, Validasi Isi oleh Ahli dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.19.8 Penskoran, Pengodean, dan Pengolahan Data

Penskoran Validasi Isi oleh Ahli dapat diringkas sebagai CVI, CVR, Aiken's V, persentase kesepakatan, dan sintesis komentar dengan aturan keputusan yang dinyatakan. Aturan ini harus ditulis sebelum melihat hasil utama agar keputusan tidak disesuaikan secara oportunistik dengan pola data. (Lawshe, 1975; Lynn, 1986; Aiken, 1985; Chi, 1997; AERA, APA, & NCME, 2014)

Skor total jarang cukup untuk diagnosis. Analisis perlu mempertahankan informasi tingkat butir, konsep, opsi, alasan, keyakinan, proses, atau representasi. Penggabungan data terlalu dini dapat menyembunyikan pola yang secara pedagogis penting.

Untuk respons terbuka atau kualitatif, buku kode harus mencantumkan definisi, contoh positif, contoh batas, noncontoh, dan aturan ketika satu respons memuat beberapa gagasan. Sebagian data perlu dinilai oleh lebih dari satu penilai.

Setiap transformasi data—misalnya dari respons ke kode, dari kode ke kategori, dan dari kategori ke persentase—harus dapat diaudit. Simpan respons asli dan keputusan pengodean agar kasus ambigu dapat ditinjau kembali.

- pisahkan skor ketepatan dari skor keyakinan atau kualitas alasan;
- laporkan kasus tidak lengkap dan aturan penanganannya;
- uji konsistensi penilai untuk data berkode;
- hindari mengubah data ordinal menjadi ukuran interval tanpa alasan.

Dalam praktik kelas, Validasi Isi oleh Ahli dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.19.9 Interpretasi dan Penetapan Kategori

Interpretasi hasil Validasi Isi oleh Ahli harus bergerak dari respons menuju pola, bukan langsung dari satu jawaban menuju label. Status konseptual menjadi lebih kuat apabila gagasan yang sama muncul pada beberapa butir, representasi, atau konteks. (Lawshe, 1975; Lynn, 1986; Aiken, 1985; Polit & Beck, 2006; Kane, 2013)

Kategori dapat mencakup pemahaman ilmiah, pemahaman parsial, kurang pengetahuan, respons tidak konsisten, false positive, false negative, dan miskonsepsi. Tidak semua instrumen mampu membedakan seluruh kategori tersebut; skema harus mengikuti informasi yang benar-benar tersedia.

Ambang, seperti batas keyakinan atau persentase pemilihan, perlu memiliki rasional dan diuji melalui analisis sensitivitas. Perubahan kecil pada ambang dapat mengubah prevalensi kategori secara berarti.

Interpretasi yang bertanggung jawab menyatakan bahwa kategori berlaku pada konsep, bentuk tugas, waktu, dan kondisi tertentu. Hasil merupakan hipotesis kerja untuk pembelajaran, bukan diagnosis psikologis permanen.

- gunakan lebih dari satu butir untuk konsepsi penting;
- periksa konsistensi antara jawaban dan alasan;
- tandai kasus ambigu untuk konfirmasi;
- hubungkan kategori dengan rekomendasi yang berbeda.

Dalam praktik kelas, Validasi Isi oleh Ahli dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.19.10 Validitas dan Kualitas Inferensi

Validitas Validasi Isi oleh Ahli berkaitan dengan kualitas interpretasi dan penggunaan hasil, bukan sifat tetap yang dimiliki instrumen. Fokus khususnya adalah merupakan bukti berdasarkan isi yang harus diintegrasikan dengan proses respons, struktur internal, hubungan dengan variabel lain, dan konsekuensi penggunaan. (AERA, APA, & NCME, 2014; Messick, 1995; Kane, 2013; Haynes et al., 1995)

Bukti berdasarkan isi menilai keterwakilan domain dan akurasi konsep. Bukti proses respons memeriksa bagaimana peserta didik memahami tugas. Bukti struktur internal menilai kesesuaian pola respons. Bukti hubungan dengan variabel lain dan konsekuensi penggunaan melengkapi argumen.

Ancaman validitas dapat berasal dari konten yang kurang terwakili, tuntutan bahasa berlebihan, petunjuk opsi, rubrik yang tidak stabil, keyakinan yang tidak terkalibrasi, atau algoritma yang belajar dari label keliru.

Dokumen validasi perlu menghubungkan setiap klaim dengan bukti dan alternatif penjelasan. Menampilkan satu indeks tidak cukup; validasi adalah program argumentatif yang diperbarui setiap kali instrumen dipakai pada populasi atau tujuan baru.

- bukti isi dan kesesuaian kurikulum;
- bukti proses respons melalui wawancara kognitif;
- bukti struktur dan fungsi butir;
- bukti kegunaan serta konsekuensi keputusan.

Dalam praktik kelas, Validasi Isi oleh Ahli dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.19.11 Reliabilitas, Keajegan, dan Analisis Empiris

Keajegan Validasi Isi oleh Ahli harus dinilai sesuai jenis data. Pada respons objektif dapat digunakan konsistensi internal, stabilitas waktu, atau model pengukuran. Pada data kualitatif, perhatian diarahkan pada konsistensi pengodean, audit trail, dan ketepatan penerapan definisi kode. (Kuder & Richardson, 1937; Cronbach, 1951; Ding et al., 2006; Wren & Barbera, 2014; AERA, APA, & NCME, 2014)

Reliabilitas tinggi tidak otomatis menunjukkan diagnosis yang benar. Butir yang berulang-ulang mengukur hafalan dapat menghasilkan koefisien tinggi tetapi tidak mewakili pemahaman konseptual. Sebaliknya, konstruk yang multidimensi tidak selalu layak dipaksa menjadi satu skala.

Analisis empiris perlu memeriksa kesukaran, daya beda, fungsi opsi, pola data hilang, dan keterhubungan komponen. Untuk instrumen bertingkat, reliabilitas dapat dihitung pada setiap tier serta kombinasi yang mempunyai makna teoritis.

Ukuran sampel, distribusi kemampuan, dan homogenitas konten memengaruhi statistik. Karena itu, laporan harus menyatakan kondisi estimasi dan tidak menyalin batas angka sebagai aturan universal.

- pilih statistik yang sesuai dengan jenis skor;
- laporkan interval atau ketidakpastian bila tersedia;
- periksa butir secara substantif dan statistik;
- uji kembali setelah revisi atau perubahan platform.

Dalam praktik kelas, Validasi Isi oleh Ahli dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.19.12 Contoh Penerapan dalam Pembelajaran Kimia

Penerapan Validasi Isi oleh Ahli dalam kimia dapat dimulai dari konsep akurasi konsep orbital. Pengembang perlu menuliskan konsep ilmiah target, miskonsepsi yang mungkin, bukti respons yang diharapkan, dan keputusan pembelajaran setelah hasil diperoleh. (Lawshe, 1975; Lynn, 1986; Aiken, 1985; Polit & Beck, 2006; Nakhleh, 1992)

Contoh kedua adalah ketepatan alasan pada kesetimbangan. Stimulus sebaiknya menghubungkan fenomena dengan model partikel atau persamaan, lalu meminta peserta didik menunjukkan hubungan sebab-akibat. Jawaban yang hanya menyebut istilah belum cukup menjadi bukti pemahaman.

Pada kejelasan diagram partikel, diagnosis dapat memeriksa apakah peserta didik membedakan perubahan laju, keadaan setimbang, energi, dan komposisi. Kesalahan sering muncul ketika satu aturan dipindahkan ke konteks yang tidak sesuai.

Kasus kesesuaian bahasa untuk jenjang SMA menunjukkan pentingnya representasi. Respons verbal, gambar, simbol, dan hitungan dapat tidak konsisten. Hasil instrumen sebaiknya digunakan untuk memilih demonstrasi, diskusi, contoh-noncontoh, atau tugas remediasi yang secara langsung menantang hubungan salah.

- tentukan proposisi ilmiah dan konsepsi alternatif;
- gunakan minimal dua bentuk representasi;
- minta alasan atau bukti ketika format memungkinkan;
- rancang tindak lanjut berdasarkan pola, bukan hanya skor.

Dalam praktik kelas, Validasi Isi oleh Ahli dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.19.13 Kelebihan, Keterbatasan, dan Potensi Bias

Kekuatan utama Validasi Isi oleh Ahli meliputi mencegah kesalahan konten, menilai keterwakilan domain, mendokumentasikan revisi, dan menggabungkan perspektif materi dan pengukuran. Kekuatan tersebut menjelaskan situasi ketika instrumen layak dipilih dibanding teknik lain. (Lawshe, 1975; Lynn, 1986; Aiken, 1985; Gurel et al., 2015; AERA, APA, & NCME, 2014)

Keterbatasannya mencakup jumlah ahli sering kecil, kesepakatan tidak menjamin konstruk tepat, keahlian harus relevan, dan indeks dapat disalahgunakan sebagai ritual angka. Keterbatasan harus diperlakukan sebagai batas inferensi, bukan sekadar paragraf formal yang tidak memengaruhi analisis.

Potensi bias dapat berasal dari bahasa, pengalaman konteks, akses teknologi, kecemasan, gaya menggunakan skala, keterampilan menggambar, atau kemampuan menulis. Instrumen perlu diuji pada kelompok yang beragam dan ditinjau untuk fungsi berbeda antarkelompok bila data mencukupi.

Keputusan memilih instrumen sebaiknya mempertimbangkan kedalaman data, ukuran kelompok, waktu, keterampilan analisis, tujuan diagnosis, dan konsekuensi kesalahan klasifikasi. Tidak ada satu metode yang unggul untuk semua tujuan.

- Kekuatan: mencegah kesalahan konten.
- Kekuatan: menilai keterwakilan domain.
- Batas: jumlah ahli sering kecil.
- Batas: kesepakatan tidak menjamin konstruk tepat.

Dalam praktik kelas, Validasi Isi oleh Ahli dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

Aspek	Temuan	Respons Metodologis
Kekuatan	mencegah kesalahan konten	Pilih ketika sesuai tujuan
Keterbatasan	jumlah ahli sering kecil	Tambahkan bukti atau kontrol
Bias potensial	Bahasa, konteks, perangkat, gaya respons	Uji proses respons

6.19.14 Integrasi Digital, Etika, dan Tata Kelola Data

Integrasi digital Validasi Isi oleh Ahli dapat berupa formulir penilaian ahli, riwayat versi, komentar per butir, dan ekspor indeks otomatis yang tetap diperiksa. Teknologi seharusnya memperluas kualitas bukti, bukan hanya memindahkan lembar kertas ke layar. (Baker & Inventado, 2014; Romero & Ventura, 2020; AERA, APA, & NCME, 2014; Lawshe, 1975)

Platform perlu menyimpan versi instrumen, urutan respons, timestamp yang relevan, aturan skor, dan log perubahan. Namun, hanya data yang dibutuhkan untuk tujuan diagnosis yang boleh dikumpulkan; pengumpulan berlebihan meningkatkan risiko privasi tanpa manfaat ilmiah.

Aspek etika utama adalah bahwa kontribusi ahli perlu diakui, konflik kepentingan dinyatakan, dan peneliti tidak boleh menyembunyikan kritik substansial. Peserta didik dan guru perlu mengetahui data apa yang dikumpulkan, untuk tujuan apa, siapa yang dapat mengakses, berapa lama disimpan, dan bagaimana kesalahan data diperbaiki.

Algoritma otomatis harus dapat diaudit. Kategori yang dihasilkan sistem perlu ditautkan ke respons mentah dan aturan keputusan. Kasus berisiko atau ambigu harus ditinjau manusia sebelum tindak lanjut penting dilakukan.

- gunakan identitas pseudonim dan akses berbasis peran;
- uji tampilan rumus, subskrip, grafik, dan gambar;
- sediakan opsi aksesibilitas dan alternatif perangkat;
- dokumentasikan algoritma serta perubahan versi.

Dalam praktik kelas, Validasi Isi oleh Ahli dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.19.15 Sintesis dan Pedoman Implementasi

Validasi Isi oleh Ahli bermanfaat apabila dipilih berdasarkan tujuan dan digunakan sesuai batasnya. Instrumen ini menghasilkan rating ahli, alasan, saran redaksi, tanda konflik konsep, dan keputusan revisi, lalu data diolah melalui CVI, CVR, Aiken's V, persentase kesepakatan, dan sintesis komentar dengan aturan keputusan yang dinyatakan. (Lawshe, 1975; Lynn, 1986; Aiken, 1985; AERA, APA, & NCME, 2014; Kane, 2013)

Rangkaian kerja yang disarankan dimulai dari analisis konsep, identifikasi konsepsi alternatif, desain tugas, telaah ahli, uji proses respons, uji lapangan, analisis, penetapan kategori, dan penyusunan tindak lanjut. Setiap tahap harus meninggalkan dokumentasi yang dapat diperiksa.

Hasil diagnosis tidak berhenti pada tabel prevalensi. Guru perlu memilih konsepsi prioritas berdasarkan frekuensi, keyakinan, posisi sebagai konsep prasyarat, dampak terhadap topik lanjutan, dan kemungkinan diperbaiki melalui pembelajaran.

Prinsip akhirnya ialah kehati-hatian inferensial: sebutkan apa yang benar-benar ditunjukkan data, bedakan bukti kuat dari indikasi, gunakan triangulasi ketika diperlukan, dan hindari label permanen pada peserta didik.

- selaraskan tujuan, konstruk, tugas, skor, kategori, dan tindakan;
- utamakan konsepsi yang berdampak pada pembelajaran lanjutan;
- gabungkan bukti bila keputusan memiliki konsekuensi besar;
- evaluasi kembali efektivitas tindak lanjut melalui diagnosis ulang.

Dalam praktik kelas, Validasi Isi oleh Ahli dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.20 Uji Coba dan Analisis Butir

6.20.1 Landasan Konseptual dan Kedudukan Instrumen

Dalam diagnosis miskonsepsi, Uji Coba dan Analisis Butir perlu dipahami bukan sekadar sebagai bentuk soal atau teknik pengumpulan data, melainkan sebagai perangkat inferensi. Dalam bab ini, instrumen tersebut didefinisikan sebagai tahap administrasi awal dan lapangan untuk memeriksa keterbacaan, waktu, fungsi opsi, kesukaran, daya beda, reliabilitas, struktur, dan kesesuaian model pengukuran. Definisi operasional ini menegaskan bentuk bukti yang dikumpulkan sekaligus membatasi jenis kesimpulan yang layak dibuat. (Kuder & Richardson, 1937; Cronbach, 1951; Ding et al., 2006)

Klaim utamanya adalah bahwa instrumen ini menentukan apakah butir berfungsi sesuai rancangan pada populasi sasaran serta menyediakan dasar revisi sebelum penggunaan diagnostik. Klaim tersebut harus dipisahkan dari anggapan bahwa setiap jawaban salah merupakan miskonsepsi. Kesalahan dapat berasal dari kurang pengetahuan, ketidakcermatan, kesulitan bahasa, salah hitung, atau ketidakfamiliaran dengan representasi.

Data pokok yang dihasilkan berupa respons, waktu, komentar peserta, pola opsi, skor tier, keyakinan, dan data demografis yang relevan. Data itu baru memperoleh makna diagnostik setelah dikaitkan dengan peta konsep sasaran, kemungkinan konsepsi alternatif, dan aturan interpretasi yang telah ditetapkan sebelum analisis.

Konsekuensinya, peneliti dan guru perlu membedakan data mentah, skor, kategori, dan klaim konseptual yang dibuat dari hasil tersebut. Posisi terbaik Uji Coba dan Analisis Butir adalah sebagai bagian dari sistem asesmen berlapis yang dapat dipadukan dengan bukti lain ketika konsekuensi keputusan semakin besar.

- menentukan konsep dan proposisi yang hendak didiagnosis;
- menjelaskan perilaku respons yang dianggap sebagai bukti;
- membatasi klaim sesuai kekuatan instrumen;
- menautkan hasil dengan tindakan pembelajaran yang realistis.

Dalam praktik kelas, Uji Coba dan Analisis Butir dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.20.2 Perkembangan Historis dan Rasional Penggunaan

Perkembangan Uji Coba dan Analisis Butir tidak dapat dipisahkan dari perubahan pandangan tentang belajar sains. Ketika peserta didik tidak lagi dipandang sebagai penerima informasi yang kosong, instrumen mulai dirancang untuk mengungkap gagasan awal, pola penalaran, dan keyakinan yang dibawa ke kelas. (Kuder & Richardson, 1937; Cronbach, 1951; Ding et al., 2006; Adams & Wieman, 2011)

Pada tahap awal, penilaian banyak mengandalkan benar-salah atau skor total. Perkembangan penelitian konsepsi peserta didik kemudian mendorong penambahan alasan, keyakinan, tugas kinerja, rekaman proses, dan bukti sosial. Uji Coba dan Analisis Butir merupakan salah satu jawaban metodologis terhadap kebutuhan tersebut.

Dalam pendidikan kimia, kebutuhan diagnosis semakin kuat karena peserta didik harus mengoordinasikan fenomena makroskopik, model submikroskopik, simbol, grafik, dan perhitungan. Instrumen yang hanya menguji hafalan definisi sering gagal memperlihatkan cara representasi itu dihubungkan.

Sejarah penggunaannya juga menunjukkan bahwa format yang lebih kompleks tidak otomatis lebih baik. Setiap tambahan tier, rubrik, atau kanal data harus mempunyai fungsi yang jelas dan didukung bukti bahwa informasi tambahan benar-benar memperbaiki keputusan.

- pergeseran dari pengukuran skor menuju diagnosis pola pemahaman;
- penggunaan respons terbuka dan wawancara untuk menemukan konsepsi alternatif;
- pengembangan teknik kuantitatif agar diagnosis dapat dilakukan pada kelompok besar;
- integrasi data digital dan metakognitif pada perkembangan mutakhir.

Dalam praktik kelas, Uji Coba dan Analisis Butir dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.20.3 Konstruk, Klaim, dan Bukti Diagnosis

Konstruk yang dinilai melalui Uji Coba dan Analisis Butir bukan objek yang terlihat langsung. Yang diamati adalah respons, waktu, komentar peserta, pola opsi, skor tier, keyakinan, dan data demografis yang relevan; dari sana dibuat inferensi tentang pemahaman konsep, miskonsepsi, pengetahuan parsial, atau ketidakpastian. (Wren & Barbera, 2014; AERA, APA, & NCME, 2014; Messick, 1995; Kane, 2013)

Klaim diagnosis perlu dirumuskan sebagai argumen. Pertama, tugas harus memerlukan konsep sasaran. Kedua, respons harus diberi skor atau kode secara konsisten. Ketiga, pola skor harus mempunyai hubungan yang masuk akal dengan struktur pengetahuan. Keempat, penggunaan hasil harus menghasilkan keputusan yang proporsional.

Sumber ancaman utama adalah ketidakjelasan konstruk. Misalnya, peserta didik dapat gagal karena istilah yang ambigu, beban membaca, representasi yang tidak dikenal, atau prosedur matematika. Kegagalan semacam itu tidak boleh langsung diperlakukan sebagai miskonsepsi kimia.

Oleh sebab itu, setiap pengembang perlu menulis pernyataan klaim-bukti: instrumen mengukur apa, perilaku apa yang menjadi indikator, alternatif penjelasan apa yang mungkin, dan bukti tambahan apa yang diperlukan untuk memperkuat kesimpulan.

- konstruk utama: pemahaman konseptual melalui evaluasi empiris instrumen;
- bukti observabel: respons, alasan, tindakan, atau representasi;
- inferensi: profil konsepsi pada konten dan konteks tertentu;
- penggunaan: perencanaan pembelajaran, remediasi, atau penelitian.

Dalam praktik kelas, Uji Coba dan Analisis Butir dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.20.4 Struktur, Komponen, dan Alur Respons

Struktur Uji Coba dan Analisis Butir dibangun dari komponen yang saling bergantung. Komponen inti meliputi uji kognitif, pilot kecil, uji lapangan, data respons, statistik butir, analisis reliabilitas, analisis model, dan keputusan revisi. Hilangnya satu komponen dapat mengubah makna data yang diperoleh. (Kuder & Richardson, 1937; Cronbach, 1951; Ding et al., 2006; Adams & Wieman, 2011)

Urutan penyajian harus ditetapkan secara sadar. Komponen yang muncul lebih awal dapat memberi petunjuk pada komponen berikutnya, sedangkan permintaan alasan atau keyakinan dapat mengubah cara peserta didik memeriksa jawabannya. Penguncian respons dan uji urutan diperlukan pada format digital.

Setiap komponen membutuhkan spesifikasi. Stimulus harus memuat informasi cukup, opsi atau prompt harus sejajar dengan indikator, skala harus memiliki jangkar, dan rubrik harus menjelaskan perbedaan kategori yang berdekatan.

Struktur yang baik memungkinkan peneliti menelusuri kembali setiap kategori hasil ke respons mentah. Transparansi ini penting agar guru tidak hanya menerima label otomatis tanpa memahami alasan di baliknya.

- uji kognitif;
- pilot kecil;
- uji lapangan;
- data respons;
- statistik butir;
- analisis reliabilitas;
- analisis model;
- keputusan revisi;

Dalam praktik kelas, Uji Coba dan Analisis Butir dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

No.	Komponen	Fungsi Diagnostik
1	uji kognitif	Bukti yang harus ditelusuri
2	pilot kecil	Bukti yang harus ditelusuri
3	uji lapangan	Bukti yang harus ditelusuri
4	data respons	Bukti yang harus ditelusuri

6.20.5 Desain Stimulus, Tugas, atau Pertanyaan

Desain stimulus pada Uji Coba dan Analisis Butir harus memunculkan penggunaan konsep, bukan sekadar pengenalan istilah. Situasi dapat berbentuk fenomena, gambar partikel, grafik, data eksperimen, persamaan, atau masalah kontekstual, tetapi seluruh informasi yang tidak relevan perlu dikendalikan. (Haladyna et al., 2002; Kuder & Richardson, 1937; Cronbach, 1951; Ding et al., 2006)

Dalam kimia, satu stimulus sering menggabungkan beberapa tingkat representasi. Pengembang harus memutuskan apakah tujuan butir adalah menerjemahkan representasi, menjelaskan hubungan sebab-akibat, membedakan contoh dan noncontoh, atau menerapkan prinsip pada konteks baru.

Bahasa perlu ringkas tanpa menghilangkan syarat ilmiah. Kata seperti kuat, stabil, jenuh, spontan, atau setimbang mempunyai arti teknis dan sehari-hari yang berbeda. Ketidakjelasan istilah dapat menghasilkan kesalahan yang berasal dari bahasa, bukan struktur konsep.

Desain yang baik juga menghindari petunjuk format, pilihan terlalu panjang, negasi bertumpuk, gambar dekoratif, dan data yang tidak diperlukan. Setiap elemen harus mempunyai alasan diagnostik.

- gunakan konteks yang cukup dikenal tetapi tidak menghafalkan contoh;
- pastikan hanya satu interpretasi ilmiah yang menjadi target;
- uji semua simbol, satuan, grafik, dan gambar;
- sediakan ruang bagi respons yang tidak diprediksi bila tujuan masih eksploratif.

Dalam praktik kelas, Uji Coba dan Analisis Butir dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.20.6 Prosedur Pengembangan Secara Sistematis

Pengembangan Uji Coba dan Analisis Butir harus mengikuti alur yang terdokumentasi. Langkah yang lazim ialah uji keterbacaan, wawancara kognitif, pilot, pembersihan data, kemudian analisis klasik, analisis tier, model pengukuran, revisi dan uji ulang. Urutan dapat berulang karena temuan uji coba sering memaksa revisi konstruk atau redaksi. (Kuder & Richardson, 1937; Cronbach, 1951; Ding et al., 2006; Adams & Wieman, 2011; AERA, APA, & NCME, 2014)

Tahap konseptual dimulai dengan analisis kurikulum, literatur miskonsepsi, dan peta konsep. Tahap empiris awal menggunakan wawancara, respons terbuka, atau observasi untuk memastikan bahwa kategori yang dirancang benar-benar muncul pada populasi sasaran.

Setiap versi instrumen perlu diberi nomor, tanggal, dan catatan perubahan. Revisi harus menyebutkan masalah yang ditemukan, bukti yang digunakan, perubahan yang dilakukan, dan alasan mempertahankan atau menghapus butir.

Pengembangan dianggap belum selesai ketika hanya menghasilkan butir yang tampak baik. Instrumen perlu diuji pada peserta didik yang serupa dengan pengguna sasaran, lalu dianalisis secara konseptual dan empiris sebelum dipakai untuk keputusan penting.

- 1. uji keterbacaan.
- 2. wawancara kognitif.
- 3. pilot.
- 4. pembersihan data.
- 5. analisis klasik.

Dalam praktik kelas, Uji Coba dan Analisis Butir dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.20.7 Pelaksanaan dan Pengumpulan Data

Pelaksanaan Uji Coba dan Analisis Butir harus distandardisasi secukupnya agar perbedaan hasil lebih mungkin berasal dari pemahaman, bukan variasi instruksi, waktu, alat, atau bantuan. Petunjuk perlu menjelaskan tujuan umum tanpa mengajarkan jawaban. (Kuder & Richardson, 1937; Cronbach, 1951; AERA, APA, & NCME, 2014)

Kondisi administrasi mencakup waktu, urutan tugas, kesempatan bertanya, penggunaan kalkulator, akses tabel periodik, kualitas gambar, dan perangkat digital. Pada instrumen kualitatif, standardisasi berarti konsistensi fokus dan prinsip probing, bukan membaca naskah secara kaku.

Data yang dikumpulkan ialah respons, waktu, komentar peserta, pola opsi, skor tier, keyakinan, dan data demografis yang relevan. Data mentah harus disimpan bersama metadata yang diperlukan, seperti kode butir, versi instrumen, waktu administrasi, dan kondisi khusus, tetapi hanya data yang relevan yang boleh dikumpulkan.

Sebelum administrasi luas, lakukan uji teknis. Pastikan tautan, urutan, penguncian respons, rekaman, penyimpanan, serta tampilan rumus dan subskrip bekerja. Gangguan teknis perlu dicatat sebagai bagian dari interpretasi.

- berikan instruksi yang sama dan mudah dipahami;
- catat kondisi yang dapat memengaruhi respons;
- hindari umpan balik sebelum seluruh respons diagnostik selesai;
- sediakan prosedur penanganan data tidak lengkap.

Dalam praktik kelas, Uji Coba dan Analisis Butir dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.20.8 Penskoran, Pengodean, dan Pengolahan Data

Penskoran Uji Coba dan Analisis Butir dapat diringkas sebagai analisis kesukaran, daya beda, korelasi butir-total, fungsi distraktor, konsistensi internal, Rasch atau IRT, dan DIF bila layak. Aturan ini harus ditulis sebelum melihat hasil utama agar keputusan tidak disesuaikan secara oportunistik dengan pola data. (Kuder & Richardson, 1937; Cronbach, 1951; Ding et al., 2006; Chi, 1997; AERA, APA, & NCME, 2014)

Skor total jarang cukup untuk diagnosis. Analisis perlu mempertahankan informasi tingkat butir, konsep, opsi, alasan, keyakinan, proses, atau representasi. Penggabungan data terlalu dini dapat menyembunyikan pola yang secara pedagogis penting.

Untuk respons terbuka atau kualitatif, buku kode harus mencantumkan definisi, contoh positif, contoh batas, noncontoh, dan aturan ketika satu respons memuat beberapa gagasan. Sebagian data perlu dinilai oleh lebih dari satu penilai.

Setiap transformasi data—misalnya dari respons ke kode, dari kode ke kategori, dan dari kategori ke persentase—harus dapat diaudit. Simpan respons asli dan keputusan pengodean agar kasus ambigu dapat ditinjau kembali.

- pisahkan skor ketepatan dari skor keyakinan atau kualitas alasan;
- laporkan kasus tidak lengkap dan aturan penanganannya;
- uji konsistensi penilai untuk data berkode;
- hindari mengubah data ordinal menjadi ukuran interval tanpa alasan.

Dalam praktik kelas, Uji Coba dan Analisis Butir dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.20.9 Interpretasi dan Penetapan Kategori

Interpretasi hasil Uji Coba dan Analisis Butir harus bergerak dari respons menuju pola, bukan langsung dari satu jawaban menuju label. Status konseptual menjadi lebih kuat apabila gagasan yang sama muncul pada beberapa butir, representasi, atau konteks. (Kuder & Richardson, 1937; Cronbach, 1951; Ding et al., 2006; Adams & Wieman, 2011; Kane, 2013)

Kategori dapat mencakup pemahaman ilmiah, pemahaman parsial, kurang pengetahuan, respons tidak konsisten, false positive, false negative, dan miskonsepsi. Tidak semua instrumen mampu membedakan seluruh kategori tersebut; skema harus mengikuti informasi yang benar-benar tersedia.

Ambang, seperti batas keyakinan atau persentase pemilihan, perlu memiliki rasional dan diuji melalui analisis sensitivitas. Perubahan kecil pada ambang dapat mengubah prevalensi kategori secara berarti.

Interpretasi yang bertanggung jawab menyatakan bahwa kategori berlaku pada konsep, bentuk tugas, waktu, dan kondisi tertentu. Hasil merupakan hipotesis kerja untuk pembelajaran, bukan diagnosis psikologis permanen.

- gunakan lebih dari satu butir untuk konsepsi penting;
- periksa konsistensi antara jawaban dan alasan;
- tandai kasus ambigu untuk konfirmasi;
- hubungkan kategori dengan rekomendasi yang berbeda.

Dalam praktik kelas, Uji Coba dan Analisis Butir dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.20.10 Validitas dan Kualitas Inferensi

Validitas Uji Coba dan Analisis Butir berkaitan dengan kualitas interpretasi dan penggunaan hasil, bukan sifat tetap yang dimiliki instrumen. Fokus khususnya adalah interpretasi statistik harus kembali pada konstruk, konten, dan proses respons; angka tidak boleh menggantikan argumentasi. (AERA, APA, & NCME, 2014; Messick, 1995; Kane, 2013; Haynes et al., 1995)

Bukti berdasarkan isi menilai keterwakilan domain dan akurasi konsep. Bukti proses respons memeriksa bagaimana peserta didik memahami tugas. Bukti struktur internal menilai kesesuaian pola respons. Bukti hubungan dengan variabel lain dan konsekuensi penggunaan melengkapi argumen.

Ancaman validitas dapat berasal dari konten yang kurang terwakili, tuntutan bahasa berlebihan, petunjuk opsi, rubrik yang tidak stabil, keyakinan yang tidak terkalibrasi, atau algoritma yang belajar dari label keliru.

Dokumen validasi perlu menghubungkan setiap klaim dengan bukti dan alternatif penjelasan. Menampilkan satu indeks tidak cukup; validasi adalah program argumentatif yang diperbarui setiap kali instrumen dipakai pada populasi atau tujuan baru.

- bukti isi dan kesesuaian kurikulum;
- bukti proses respons melalui wawancara kognitif;
- bukti struktur dan fungsi butir;
- bukti kegunaan serta konsekuensi keputusan.

Dalam praktik kelas, Uji Coba dan Analisis Butir dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.20.11 Reliabilitas, Keajegan, dan Analisis Empiris

Keajegan Uji Coba dan Analisis Butir harus dinilai sesuai jenis data. Pada respons objektif dapat digunakan konsistensi internal, stabilitas waktu, atau model pengukuran. Pada data kualitatif, perhatian diarahkan pada konsistensi pengodean, audit trail, dan ketepatan penerapan definisi kode. (Kuder & Richardson, 1937; Cronbach, 1951; Ding et al., 2006; Wren & Barbera, 2014; AERA, APA, & NCME, 2014)

Reliabilitas tinggi tidak otomatis menunjukkan diagnosis yang benar. Butir yang berulang-ulang mengukur hafalan dapat menghasilkan koefisien tinggi tetapi tidak mewakili pemahaman konseptual. Sebaliknya, konstruk yang multidimensi tidak selalu layak dipaksa menjadi satu skala.

Analisis empiris perlu memeriksa kesukaran, daya beda, fungsi opsi, pola data hilang, dan keterhubungan komponen. Untuk instrumen bertingkat, reliabilitas dapat dihitung pada setiap tier serta kombinasi yang mempunyai makna teoritis.

Ukuran sampel, distribusi kemampuan, dan homogenitas konten memengaruhi statistik. Karena itu, laporan harus menyatakan kondisi estimasi dan tidak menyalin batas angka sebagai aturan universal.

- pilih statistik yang sesuai dengan jenis skor;
- laporkan interval atau ketidakpastian bila tersedia;
- periksa butir secara substantif dan statistik;
- uji kembali setelah revisi atau perubahan platform.

Dalam praktik kelas, Uji Coba dan Analisis Butir dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.20.12 Contoh Penerapan dalam Pembelajaran Kimia

Penerapan Uji Coba dan Analisis Butir dalam kimia dapat dimulai dari konsep analisis kombinasi tier kinetika. Pengembang perlu menuliskan konsep ilmiah target, miskonsepsi yang mungkin, bukti respons yang diharapkan, dan keputusan pembelajaran setelah hasil diperoleh. (Kuder & Richardson, 1937; Cronbach, 1951; Ding et al., 2006; Adams & Wieman, 2011; Nakhleh, 1992)

Contoh kedua adalah fit butir termokimia. Stimulus sebaiknya menghubungkan fenomena dengan model partikel atau persamaan, lalu meminta peserta didik menunjukkan hubungan sebab-akibat. Jawaban yang hanya menyebut istilah belum cukup menjadi bukti pemahaman.

Pada DIF berdasarkan jenjang, diagnosis dapat memeriksa apakah peserta didik membedakan perubahan laju, keadaan setimbang, energi, dan komposisi. Kesalahan sering muncul ketika satu aturan dipindahkan ke konteks yang tidak sesuai.

Kasus fungsi distraktor asam-basa menunjukkan pentingnya representasi. Respons verbal, gambar, simbol, dan hitungan dapat tidak konsisten. Hasil instrumen sebaiknya digunakan untuk memilih demonstrasi, diskusi, contoh-noncontoh, atau tugas remediasi yang secara langsung menantang hubungan salah.

- tentukan proposisi ilmiah dan konsepsi alternatif;
- gunakan minimal dua bentuk representasi;
- minta alasan atau bukti ketika format memungkinkan;
- rancang tindak lanjut berdasarkan pola, bukan hanya skor.

Dalam praktik kelas, Uji Coba dan Analisis Butir dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.20.13 Kelebihan, Keterbatasan, dan Potensi Bias

Kekuatan utama Uji Coba dan Analisis Butir meliputi mendeteksi butir bermasalah, menyediakan bukti struktur internal, meningkatkan keadilan, dan membuat revisi berbasis data. Kekuatan tersebut menjelaskan situasi ketika instrumen layak dipilih dibanding teknik lain. (Kuder & Richardson, 1937; Cronbach, 1951; Ding et al., 2006; Gurel et al., 2015; AERA, APA, & NCME, 2014)

Keterbatasannya mencakup statistik bergantung sampel, reliabilitas tinggi tidak menjamin validitas, kategori multi-tier menuntut analisis khusus, dan sampel kecil membatasi model kompleks. Keterbatasan harus diperlakukan sebagai batas inferensi, bukan sekadar paragraf formal yang tidak memengaruhi analisis.

Potensi bias dapat berasal dari bahasa, pengalaman konteks, akses teknologi, kecemasan, gaya menggunakan skala, keterampilan menggambar, atau kemampuan menulis. Instrumen perlu diuji pada kelompok yang beragam dan ditinjau untuk fungsi berbeda antarkelompok bila data mencukupi.

Keputusan memilih instrumen sebaiknya mempertimbangkan kedalaman data, ukuran kelompok, waktu, keterampilan analisis, tujuan diagnosis, dan konsekuensi kesalahan klasifikasi. Tidak ada satu metode yang unggul untuk semua tujuan.

- Kekuatan: mendeteksi butir bermasalah.
- Kekuatan: menyediakan bukti struktur internal.
- Batas: statistik bergantung sampel.
- Batas: reliabilitas tinggi tidak menjamin validitas.

Dalam praktik kelas, Uji Coba dan Analisis Butir dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

Aspek	Temuan	Respons Metodologis
Kekuatan	mendeteksi butir bermasalah	Pilih ketika sesuai tujuan
Keterbatasan	statistik bergantung sampel	Tambahkan bukti atau kontrol
Bias potensial	Bahasa, konteks, perangkat, gaya respons	Uji proses respons

6.20.14 Integrasi Digital, Etika, dan Tata Kelola Data

Integrasi digital Uji Coba dan Analisis Butir dapat berupa pengumpulan respons otomatis, pemeriksaan data, skrip analisis reproducible, dan dashboard revisi butir. Teknologi seharusnya memperluas kualitas bukti, bukan hanya memindahkan lembar kertas ke layar. (Baker & Inventado, 2014; Romero & Ventura, 2020; AERA, APA, & NCME, 2014; Kuder & Richardson, 1937)

Platform perlu menyimpan versi instrumen, urutan respons, timestamp yang relevan, aturan skor, dan log perubahan. Namun, hanya data yang dibutuhkan untuk tujuan diagnosis yang boleh dikumpulkan; pengumpulan berlebihan meningkatkan risiko privasi tanpa manfaat ilmiah.

Aspek etika utama adalah bahwa sampel pilot harus diinformasikan, data diamankan, dan butir tidak adil harus direvisi atau dihapus. Peserta didik dan guru perlu mengetahui data apa yang dikumpulkan, untuk tujuan apa, siapa yang dapat mengakses, berapa lama disimpan, dan bagaimana kesalahan data diperbaiki.

Algoritma otomatis harus dapat diaudit. Kategori yang dihasilkan sistem perlu ditautkan ke respons mentah dan aturan keputusan. Kasus berisiko atau ambigu harus ditinjau manusia sebelum tindak lanjut penting dilakukan.

- gunakan identitas pseudonim dan akses berbasis peran;
- uji tampilan rumus, subskrip, grafik, dan gambar;
- sediakan opsi aksesibilitas dan alternatif perangkat;
- dokumentasikan algoritma serta perubahan versi.

Dalam praktik kelas, Uji Coba dan Analisis Butir dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.20.15 Sintesis dan Pedoman Implementasi

Uji Coba dan Analisis Butir bermanfaat apabila dipilih berdasarkan tujuan dan digunakan sesuai batasnya. Instrumen ini menghasilkan respons, waktu, komentar peserta, pola opsi, skor tier, keyakinan, dan data demografis yang relevan, lalu data diolah melalui analisis kesukaran, daya beda, korelasi butir-total, fungsi distraktor, konsistensi internal, Rasch atau IRT, dan DIF bila layak. (Kuder & Richardson, 1937; Cronbach, 1951; Ding et al., 2006; AERA, APA, & NCME, 2014; Kane, 2013)

Rangkaian kerja yang disarankan dimulai dari analisis konsep, identifikasi konsepsi alternatif, desain tugas, telaah ahli, uji proses respons, uji lapangan, analisis, penetapan kategori, dan penyusunan tindak lanjut. Setiap tahap harus meninggalkan dokumentasi yang dapat diperiksa.

Hasil diagnosis tidak berhenti pada tabel prevalensi. Guru perlu memilih konsepsi prioritas berdasarkan frekuensi, keyakinan, posisi sebagai konsep prasyarat, dampak terhadap topik lanjutan, dan kemungkinan diperbaiki melalui pembelajaran.

Prinsip akhirnya ialah kehati-hatian inferensial: sebutkan apa yang benar-benar ditunjukkan data, bedakan bukti kuat dari indikasi, gunakan triangulasi ketika diperlukan, dan hindari label permanen pada peserta didik.

- selaraskan tujuan, konstruk, tugas, skor, kategori, dan tindakan;
- utamakan konsepsi yang berdampak pada pembelajaran lanjutan;
- gabungkan bukti bila keputusan memiliki konsekuensi besar;
- evaluasi kembali efektivitas tindak lanjut melalui diagnosis ulang.

Dalam praktik kelas, Uji Coba dan Analisis Butir dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.21 Penetapan Kategori Pemahaman Konsep

6.21.1 Landasan Konseptual dan Kedudukan Instrumen

Penetapan Kategori Pemahaman Konsep menempati posisi tertentu dalam rangkaian bukti diagnosis, dari pengungkapan respons awal sampai penentuan tindak lanjut pembelajaran. Dalam bab ini, instrumen tersebut didefinisikan sebagai proses menerjemahkan pola jawaban, alasan, keyakinan, dan bukti lain menjadi kategori seperti paham, paham parsial, tidak tahu, menebak, false positive, false negative, dan miskonsepsi. Definisi operasional ini menegaskan bentuk bukti yang dikumpulkan sekaligus membatasi jenis kesimpulan yang layak dibuat. (Caleon & Subramaniam, 2010b; Kaltakci-Gurel et al., 2017; Yan & Subramaniam, 2018)

Klaim utamanya adalah bahwa instrumen ini menyederhanakan data untuk keputusan pembelajaran tanpa kehilangan ketidakpastian dan variasi konteks. Klaim tersebut harus dipisahkan dari anggapan bahwa setiap jawaban salah merupakan miskonsepsi. Kesalahan dapat berasal dari kurang pengetahuan, ketidakcermatan, kesulitan bahasa, salah hitung, atau ketidakfamiliaran dengan representasi.

Data pokok yang dihasilkan berupa kombinasi tier, skor rubrik, keyakinan, konsistensi lintasbutir, dan bukti kualitatif. Data itu baru memperoleh makna diagnostik setelah dikaitkan dengan peta konsep sasaran, kemungkinan konsepsi alternatif, dan aturan interpretasi yang telah ditetapkan sebelum analisis.

Konsekuensinya, peneliti dan guru perlu membedakan data mentah, skor, kategori, dan klaim konseptual yang dibuat dari hasil tersebut. Posisi terbaik Penetapan Kategori Pemahaman Konsep adalah sebagai bagian dari sistem asesmen berlapis yang dapat dipadukan dengan bukti lain ketika konsekuensi keputusan semakin besar.

- menentukan konsep dan proposisi yang hendak didiagnosis;
- menjelaskan perilaku respons yang dianggap sebagai bukti;
- membatasi klaim sesuai kekuatan instrumen;
- menautkan hasil dengan tindakan pembelajaran yang realistis.

Dalam praktik kelas, Penetapan Kategori Pemahaman Konsep dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.21.2 Perkembangan Historis dan Rasional Penggunaan

Perkembangan Penetapan Kategori Pemahaman Konsep tidak dapat dipisahkan dari perubahan pandangan tentang belajar sains. Ketika peserta didik tidak lagi dipandang sebagai penerima informasi yang kosong, instrumen mulai dirancang untuk mengungkap gagasan awal, pola penalaran, dan keyakinan yang dibawa ke kelas. (Caleon & Subramaniam, 2010b; Kaltakci-Gurel et al., 2017; Yan & Subramaniam, 2018; Hasan et al., 1999)

Pada tahap awal, penilaian banyak mengandalkan benar-salah atau skor total. Perkembangan penelitian konsepsi peserta didik kemudian mendorong penambahan alasan, keyakinan, tugas kinerja, rekaman proses, dan bukti sosial. Penetapan Kategori Pemahaman Konsep merupakan salah satu jawaban metodologis terhadap kebutuhan tersebut.

Dalam pendidikan kimia, kebutuhan diagnosis semakin kuat karena peserta didik harus mengoordinasikan fenomena makroskopik, model submikroskopik, simbol, grafik, dan perhitungan. Instrumen yang hanya menguji hafalan definisi sering gagal memperlihatkan cara representasi itu dihubungkan.

Sejarah penggunaannya juga menunjukkan bahwa format yang lebih kompleks tidak otomatis lebih baik. Setiap tambahan tier, rubrik, atau kanal data harus mempunyai fungsi yang jelas dan didukung bukti bahwa informasi tambahan benar-benar memperbaiki keputusan.

- pergeseran dari pengukuran skor menuju diagnosis pola pemahaman;
- penggunaan respons terbuka dan wawancara untuk menemukan konsepsi alternatif;
- pengembangan teknik kuantitatif agar diagnosis dapat dilakukan pada kelompok besar;
- integrasi data digital dan metakognitif pada perkembangan mutakhir.

Dalam praktik kelas, Penetapan Kategori Pemahaman Konsep dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.21.3 Konstruk, Klaim, dan Bukti Diagnosis

Konstruk yang dinilai melalui Penetapan Kategori Pemahaman Konsep bukan objek yang terlihat langsung. Yang diamati adalah kombinasi tier, skor rubrik, keyakinan, konsistensi lintasbutir, dan bukti kualitatif; dari sana dibuat inferensi tentang pemahaman konsep, miskonsepsi, pengetahuan parsial, atau ketidakpastian. (AERA, APA, & NCME, 2014; Kane, 2013; Messick, 1995)

Klaim diagnosis perlu dirumuskan sebagai argumen. Pertama, tugas harus memerlukan konsep sasaran. Kedua, respons harus diberi skor atau kode secara konsisten. Ketiga, pola skor harus mempunyai hubungan yang masuk akal dengan struktur pengetahuan. Keempat, penggunaan hasil harus menghasilkan keputusan yang proporsional.

Sumber ancaman utama adalah ketidakjelasan konstruk. Misalnya, peserta didik dapat gagal karena istilah yang ambigu, beban membaca, representasi yang tidak dikenal, atau prosedur matematika. Kegagalan semacam itu tidak boleh langsung diperlakukan sebagai miskonsepsi kimia.

Oleh sebab itu, setiap pengembang perlu menulis pernyataan klaim-bukti: instrumen mengukur apa, perilaku apa yang menjadi indikator, alternatif penjelasan apa yang mungkin, dan bukti tambahan apa yang diperlukan untuk memperkuat kesimpulan.

- konstruk utama: pemahaman konseptual melalui klasifikasi hasil diagnosis;
- bukti observabel: respons, alasan, tindakan, atau representasi;
- inferensi: profil konsepsi pada konten dan konteks tertentu;
- penggunaan: perencanaan pembelajaran, remediasi, atau penelitian.

Dalam praktik kelas, Penetapan Kategori Pemahaman Konsep dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.21.4 Struktur, Komponen, dan Alur Respons

Struktur Penetapan Kategori Pemahaman Konsep dibangun dari komponen yang saling bergantung. Komponen inti meliputi definisi operasional, matriks kombinasi respons, ambang keyakinan, aturan kasus ambigu, analisis konsistensi, dan prosedur konfirmasi. Hilangnya satu komponen dapat mengubah makna data yang diperoleh. (Caleon & Subramaniam, 2010b; Kaltakci-Gurel et al., 2017; Yan & Subramaniam, 2018; Hasan et al., 1999)

Urutan penyajian harus ditetapkan secara sadar. Komponen yang muncul lebih awal dapat memberi petunjuk pada komponen berikutnya, sedangkan permintaan alasan atau keyakinan dapat mengubah cara peserta didik memeriksa jawabannya. Penguncian respons dan uji urutan diperlukan pada format digital.

Setiap komponen membutuhkan spesifikasi. Stimulus harus memuat informasi cukup, opsi atau prompt harus sejajar dengan indikator, skala harus memiliki jangkar, dan rubrik harus menjelaskan perbedaan kategori yang berdekatan.

Struktur yang baik memungkinkan peneliti menelusuri kembali setiap kategori hasil ke respons mentah. Transparansi ini penting agar guru tidak hanya menerima label otomatis tanpa memahami alasan di baliknya.

- definisi operasional;
- matriks kombinasi respons;
- ambang keyakinan;
- aturan kasus ambigu;
- analisis konsistensi;
- prosedur konfirmasi;

Dalam praktik kelas, Penetapan Kategori Pemahaman Konsep dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

No.	Komponen	Fungsi Diagnostik
1	definisi operasional	Bukti yang harus ditelusuri
2	matriks kombinasi respons	Bukti yang harus ditelusuri
3	ambang keyakinan	Bukti yang harus ditelusuri
4	aturan kasus ambigu	Bukti yang harus ditelusuri

6.21.5 Desain Stimulus, Tugas, atau Pertanyaan

Desain stimulus pada Penetapan Kategori Pemahaman Konsep harus memunculkan penggunaan konsep, bukan sekadar pengenalan istilah. Situasi dapat berbentuk fenomena, gambar partikel, grafik, data eksperimen, persamaan, atau masalah kontekstual, tetapi seluruh informasi yang tidak relevan perlu dikendalikan. (Haladyna et al., 2002; Caleon & Subramaniam, 2010b; Kaltakci-Gurel et al., 2017; Yan & Subramaniam, 2018)

Dalam kimia, satu stimulus sering menggabungkan beberapa tingkat representasi. Pengembang harus memutuskan apakah tujuan butir adalah menerjemahkan representasi, menjelaskan hubungan sebab-akibat, membedakan contoh dan noncontoh, atau menerapkan prinsip pada konteks baru.

Bahasa perlu ringkas tanpa menghilangkan syarat ilmiah. Kata seperti kuat, stabil, jenuh, spontan, atau setimbang mempunyai arti teknis dan sehari-hari yang berbeda. Ketidakjelasan istilah dapat menghasilkan kesalahan yang berasal dari bahasa, bukan struktur konsep.

Desain yang baik juga menghindari petunjuk format, pilihan terlalu panjang, negasi bertumpuk, gambar dekoratif, dan data yang tidak diperlukan. Setiap elemen harus mempunyai alasan diagnostik.

- gunakan konteks yang cukup dikenal tetapi tidak menghafalkan contoh;
- pastikan hanya satu interpretasi ilmiah yang menjadi target;
- uji semua simbol, satuan, grafik, dan gambar;
- sediakan ruang bagi respons yang tidak diprediksi bila tujuan masih eksploratif.

Dalam praktik kelas, Penetapan Kategori Pemahaman Konsep dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.21.6 Prosedur Pengembangan Secara Sistematis

Pengembangan Penetapan Kategori Pemahaman Konsep harus mengikuti alur yang terdokumentasi. Langkah yang lazim ialah menetapkan tujuan kategori, memilih dimensi, menulis aturan, menguji kasus contoh, kemudian melakukan analisis sensitivitas, membandingkan dengan wawancara, merevisi, mendokumentasikan. Urutan dapat berulang karena temuan uji coba sering memaksa revisi konstruk atau redaksi. (Caleon & Subramaniam, 2010b; Kaltakci-Gurel et al., 2017; Yan & Subramaniam, 2018; Adams & Wieman, 2011; AERA, APA, & NCME, 2014)

Tahap konseptual dimulai dengan analisis kurikulum, literatur miskonsepsi, dan peta konsep. Tahap empiris awal menggunakan wawancara, respons terbuka, atau observasi untuk memastikan bahwa kategori yang dirancang benar-benar muncul pada populasi sasaran.

Setiap versi instrumen perlu diberi nomor, tanggal, dan catatan perubahan. Revisi harus menyebutkan masalah yang ditemukan, bukti yang digunakan, perubahan yang dilakukan, dan alasan mempertahankan atau menghapus butir.

Pengembangan dianggap belum selesai ketika hanya menghasilkan butir yang tampak baik. Instrumen perlu diuji pada peserta didik yang serupa dengan pengguna sasaran, lalu dianalisis secara konseptual dan empiris sebelum dipakai untuk keputusan penting.

- 1. menetapkan tujuan kategori.
- 2. memilih dimensi.
- 3. menulis aturan.
- 4. menguji kasus contoh.
- 5. melakukan analisis sensitivitas.

Dalam praktik kelas, Penetapan Kategori Pemahaman Konsep dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.21.7 Pelaksanaan dan Pengumpulan Data

Pelaksanaan Penetapan Kategori Pemahaman Konsep harus distandardisasi secukupnya agar perbedaan hasil lebih mungkin berasal dari pemahaman, bukan variasi instruksi, waktu, alat, atau bantuan. Petunjuk perlu menjelaskan tujuan umum tanpa mengajarkan jawaban. (Caleon & Subramaniam, 2010b; Kaltakci-Gurel et al., 2017; AERA, APA, & NCME, 2014)

Kondisi administrasi mencakup waktu, urutan tugas, kesempatan bertanya, penggunaan kalkulator, akses tabel periodik, kualitas gambar, dan perangkat digital. Pada instrumen kualitatif, standardisasi berarti konsistensi fokus dan prinsip probing, bukan membaca naskah secara kaku.

Data yang dikumpulkan ialah kombinasi tier, skor rubrik, keyakinan, konsistensi lintasbutir, dan bukti kualitatif. Data mentah harus disimpan bersama metadata yang diperlukan, seperti kode butir, versi instrumen, waktu administrasi, dan kondisi khusus, tetapi hanya data yang relevan yang boleh dikumpulkan.

Sebelum administrasi luas, lakukan uji teknis. Pastikan tautan, urutan, penguncian respons, rekaman, penyimpanan, serta tampilan rumus dan subskrip bekerja. Gangguan teknis perlu dicatat sebagai bagian dari interpretasi.

- berikan instruksi yang sama dan mudah dipahami;
- catat kondisi yang dapat memengaruhi respons;
- hindari umpan balik sebelum seluruh respons diagnostik selesai;
- sediakan prosedur penanganan data tidak lengkap.

Dalam praktik kelas, Penetapan Kategori Pemahaman Konsep dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.21.8 Penskoran, Pengodean, dan Pengolahan Data

Penskoran Penetapan Kategori Pemahaman Konsep dapat diringkas sebagai aturan klasifikasi transparan yang diuji melalui analisis sensitivitas dan, bila mungkin, dibandingkan dengan wawancara. Aturan ini harus ditulis sebelum melihat hasil utama agar keputusan tidak disesuaikan secara oportunistik dengan pola data. (Caleon & Subramaniam, 2010b; Kaltakci-Gurel et al., 2017; Yan & Subramaniam, 2018; Chi, 1997; AERA, APA, & NCME, 2014)

Skor total jarang cukup untuk diagnosis. Analisis perlu mempertahankan informasi tingkat butir, konsep, opsi, alasan, keyakinan, proses, atau representasi. Penggabungan data terlalu dini dapat menyembunyikan pola yang secara pedagogis penting.

Untuk respons terbuka atau kualitatif, buku kode harus mencantumkan definisi, contoh positif, contoh batas, noncontoh, dan aturan ketika satu respons memuat beberapa gagasan. Sebagian data perlu dinilai oleh lebih dari satu penilai.

Setiap transformasi data—misalnya dari respons ke kode, dari kode ke kategori, dan dari kategori ke persentase—harus dapat diaudit. Simpan respons asli dan keputusan pengodean agar kasus ambigu dapat ditinjau kembali.

- pisahkan skor ketepatan dari skor keyakinan atau kualitas alasan;
- laporkan kasus tidak lengkap dan aturan penanganannya;
- uji konsistensi penilai untuk data berkode;
- hindari mengubah data ordinal menjadi ukuran interval tanpa alasan.

Dalam praktik kelas, Penetapan Kategori Pemahaman Konsep dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.21.9 Interpretasi dan Penetapan Kategori

Interpretasi hasil Penetapan Kategori Pemahaman Konsep harus bergerak dari respons menuju pola, bukan langsung dari satu jawaban menuju label. Status konseptual menjadi lebih kuat apabila gagasan yang sama muncul pada beberapa butir, representasi, atau konteks. (Caleon & Subramaniam, 2010b; Kaltakci-Gurel et al., 2017; Yan & Subramaniam, 2018; Hasan et al., 1999; Kane, 2013)

Kategori dapat mencakup pemahaman ilmiah, pemahaman parsial, kurang pengetahuan, respons tidak konsisten, false positive, false negative, dan miskonsepsi. Tidak semua instrumen mampu membedakan seluruh kategori tersebut; skema harus mengikuti informasi yang benar-benar tersedia.

Ambang, seperti batas keyakinan atau persentase pemilihan, perlu memiliki rasional dan diuji melalui analisis sensitivitas. Perubahan kecil pada ambang dapat mengubah prevalensi kategori secara berarti.

Interpretasi yang bertanggung jawab menyatakan bahwa kategori berlaku pada konsep, bentuk tugas, waktu, dan kondisi tertentu. Hasil merupakan hipotesis kerja untuk pembelajaran, bukan diagnosis psikologis permanen.

- gunakan lebih dari satu butir untuk konsepsi penting;
- periksa konsistensi antara jawaban dan alasan;
- tandai kasus ambigu untuk konfirmasi;
- hubungkan kategori dengan rekomendasi yang berbeda.

Dalam praktik kelas, Penetapan Kategori Pemahaman Konsep dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.21.10 Validitas dan Kualitas Inferensi

Validitas Penetapan Kategori Pemahaman Konsep berkaitan dengan kualitas interpretasi dan penggunaan hasil, bukan sifat tetap yang dimiliki instrumen. Fokus khususnya adalah kesesuaian kategori dengan bukti, stabilitas pada pengulangan, dan kegunaan untuk keputusan yang benar-benar akan dibuat. (AERA, APA, & NCME, 2014; Messick, 1995; Kane, 2013; Haynes et al., 1995)

Bukti berdasarkan isi menilai keterwakilan domain dan akurasi konsep. Bukti proses respons memeriksa bagaimana peserta didik memahami tugas. Bukti struktur internal menilai kesesuaian pola respons. Bukti hubungan dengan variabel lain dan konsekuensi penggunaan melengkapi argumen.

Ancaman validitas dapat berasal dari konten yang kurang terwakili, tuntutan bahasa berlebihan, petunjuk opsi, rubrik yang tidak stabil, keyakinan yang tidak terkalibrasi, atau algoritma yang belajar dari label keliru.

Dokumen validasi perlu menghubungkan setiap klaim dengan bukti dan alternatif penjelasan. Menampilkan satu indeks tidak cukup; validasi adalah program argumentatif yang diperbarui setiap kali instrumen dipakai pada populasi atau tujuan baru.

- bukti isi dan kesesuaian kurikulum;
- bukti proses respons melalui wawancara kognitif;
- bukti struktur dan fungsi butir;
- bukti kegunaan serta konsekuensi keputusan.

Dalam praktik kelas, Penetapan Kategori Pemahaman Konsep dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.21.11 Reliabilitas, Keajegan, dan Analisis Empiris

Keajegan Penetapan Kategori Pemahaman Konsep harus dinilai sesuai jenis data. Pada respons objektif dapat digunakan konsistensi internal, stabilitas waktu, atau model pengukuran. Pada data kualitatif, perhatian diarahkan pada konsistensi pengodean, audit trail, dan ketepatan penerapan definisi kode. (Kuder & Richardson, 1937; Cronbach, 1951; Ding et al., 2006; Wren & Barbera, 2014; AERA, APA, & NCME, 2014)

Reliabilitas tinggi tidak otomatis menunjukkan diagnosis yang benar. Butir yang berulang-ulang mengukur hafalan dapat menghasilkan koefisien tinggi tetapi tidak mewakili pemahaman konseptual. Sebaliknya, konstruk yang multidimensi tidak selalu layak dipaksa menjadi satu skala.

Analisis empiris perlu memeriksa kesukaran, daya beda, fungsi opsi, pola data hilang, dan keterhubungan komponen. Untuk instrumen bertingkat, reliabilitas dapat dihitung pada setiap tier serta kombinasi yang mempunyai makna teoritis.

Ukuran sampel, distribusi kemampuan, dan homogenitas konten memengaruhi statistik. Karena itu, laporan harus menyatakan kondisi estimasi dan tidak menyalin batas angka sebagai aturan universal.

- pilih statistik yang sesuai dengan jenis skor;
- laporkan interval atau ketidakpastian bila tersedia;
- periksa butir secara substantif dan statistik;
- uji kembali setelah revisi atau perubahan platform.

Dalam praktik kelas, Penetapan Kategori Pemahaman Konsep dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.21.12 Contoh Penerapan dalam Pembelajaran Kimia

Penerapan Penetapan Kategori Pemahaman Konsep dalam kimia dapat dimulai dari konsep paham ilmiah tetapi ragu pada konsep mol. Pengembang perlu menuliskan konsep ilmiah target, miskonsepsi yang mungkin, bukti respons yang diharapkan, dan keputusan pembelajaran setelah hasil diperoleh. (Caleon & Subramaniam, 2010b; Kaltakci-Gurel et al., 2017; Yan & Subramaniam, 2018; Hasan et al., 1999; Nakhleh, 1992)

Contoh kedua adalah miskonsepsi yakin tentang katalis. Stimulus sebaiknya menghubungkan fenomena dengan model partikel atau persamaan, lalu meminta peserta didik menunjukkan hubungan sebab-akibat. Jawaban yang hanya menyebut istilah belum cukup menjadi bukti pemahaman.

Pada false positive pada kesetimbangan, diagnosis dapat memeriksa apakah peserta didik membedakan perubahan laju, keadaan setimbang, energi, dan komposisi. Kesalahan sering muncul ketika satu aturan dipindahkan ke konteks yang tidak sesuai.

Kasus pengetahuan parsial tentang ikatan menunjukkan pentingnya representasi. Respons verbal, gambar, simbol, dan hitungan dapat tidak konsisten. Hasil instrumen sebaiknya digunakan untuk memilih demonstrasi, diskusi, contoh-noncontoh, atau tugas remediasi yang secara langsung menantang hubungan salah.

- tentukan proposisi ilmiah dan konsepsi alternatif;
- gunakan minimal dua bentuk representasi;
- minta alasan atau bukti ketika format memungkinkan;
- rancang tindak lanjut berdasarkan pola, bukan hanya skor.

Dalam praktik kelas, Penetapan Kategori Pemahaman Konsep dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.21.13 Kelebihan, Keterbatasan, dan Potensi Bias

Kekuatan utama Penetapan Kategori Pemahaman Konsep meliputi membuat hasil dapat ditindaklanjuti, membedakan jenis kebutuhan belajar, mendukung pemetaan kelas, dan memfokuskan remediasi. Kekuatan tersebut menjelaskan situasi ketika instrumen layak dipilih dibanding teknik lain. (Caleon & Subramaniam, 2010b; Kaltakci-Gurel et al., 2017; Yan & Subramaniam, 2018; Gurel et al., 2015; AERA, APA, & NCME, 2014)

Keterbatasannya mencakup kategori dapat memberi kesan kepastian palsu, ambang mengubah prevalensi, respons kontekstual tidak selalu stabil, dan kategori berbeda antarpenelitian menghambat perbandingan. Keterbatasan harus diperlakukan sebagai batas inferensi, bukan sekadar paragraf formal yang tidak memengaruhi analisis.

Potensi bias dapat berasal dari bahasa, pengalaman konteks, akses teknologi, kecemasan, gaya menggunakan skala, keterampilan menggambar, atau kemampuan menulis. Instrumen perlu diuji pada kelompok yang beragam dan ditinjau untuk fungsi berbeda antarkelompok bila data mencukupi.

Keputusan memilih instrumen sebaiknya mempertimbangkan kedalaman data, ukuran kelompok, waktu, keterampilan analisis, tujuan diagnosis, dan konsekuensi kesalahan klasifikasi. Tidak ada satu metode yang unggul untuk semua tujuan.

- Kekuatan: membuat hasil dapat ditindaklanjuti.
- Kekuatan: membedakan jenis kebutuhan belajar.
- Batas: kategori dapat memberi kesan kepastian palsu.
- Batas: ambang mengubah prevalensi.

Dalam praktik kelas, Penetapan Kategori Pemahaman Konsep dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

Aspek	Temuan	Respons Metodologis
Kekuatan	membuat hasil dapat ditindaklanjuti	Pilih ketika sesuai tujuan
Keterbatasan	kategori dapat memberi kesan kepastian palsu	Tambahkan bukti atau kontrol
Bias potensial	Bahasa, konteks, perangkat, gaya respons	Uji proses respons

6.21.14 Integrasi Digital, Etika, dan Tata Kelola Data

Integrasi digital Penetapan Kategori Pemahaman Konsep dapat berupa mesin aturan yang dapat diaudit, tampilan probabilistik, dan akses ke respons mentah di balik setiap label. Teknologi seharusnya memperluas kualitas bukti, bukan hanya memindahkan lembar kertas ke layar. (Baker & Inventado, 2014; Romero & Ventura, 2020; AERA, APA, & NCME, 2014; Caleon & Subramaniam, 2010b)

Platform perlu menyimpan versi instrumen, urutan respons, timestamp yang relevan, aturan skor, dan log perubahan. Namun, hanya data yang dibutuhkan untuk tujuan diagnosis yang boleh dikumpulkan; pengumpulan berlebihan meningkatkan risiko privasi tanpa manfaat ilmiah.

Aspek etika utama adalah bahwa label harus bersifat sementara, spesifik konsep, dan tidak digunakan untuk menghakimi kecerdasan atau potensi peserta didik. Peserta didik dan guru perlu mengetahui data apa yang dikumpulkan, untuk tujuan apa, siapa yang dapat mengakses, berapa lama disimpan, dan bagaimana kesalahan data diperbaiki.

Algoritma otomatis harus dapat diaudit. Kategori yang dihasilkan sistem perlu ditautkan ke respons mentah dan aturan keputusan. Kasus berisiko atau ambigu harus ditinjau manusia sebelum tindak lanjut penting dilakukan.

- gunakan identitas pseudonim dan akses berbasis peran;
- uji tampilan rumus, subskrip, grafik, dan gambar;
- sediakan opsi aksesibilitas dan alternatif perangkat;
- dokumentasikan algoritma serta perubahan versi.

Dalam praktik kelas, Penetapan Kategori Pemahaman Konsep dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.21.15 Sintesis dan Pedoman Implementasi

Penetapan Kategori Pemahaman Konsep bermanfaat apabila dipilih berdasarkan tujuan dan digunakan sesuai batasnya. Instrumen ini menghasilkan kombinasi tier, skor rubrik, keyakinan, konsistensi lintasbutir, dan bukti kualitatif, lalu data diolah melalui aturan klasifikasi transparan yang diuji melalui analisis sensitivitas dan, bila mungkin, dibandingkan dengan wawancara. (Caleon & Subramaniam, 2010b; Kaltakci-Gurel et al., 2017; Yan & Subramaniam, 2018; AERA, APA, & NCME, 2014; Kane, 2013)

Rangkaian kerja yang disarankan dimulai dari analisis konsep, identifikasi konsepsi alternatif, desain tugas, telaah ahli, uji proses respons, uji lapangan, analisis, penetapan kategori, dan penyusunan tindak lanjut. Setiap tahap harus meninggalkan dokumentasi yang dapat diperiksa.

Hasil diagnosis tidak berhenti pada tabel prevalensi. Guru perlu memilih konsepsi prioritas berdasarkan frekuensi, keyakinan, posisi sebagai konsep prasyarat, dampak terhadap topik lanjutan, dan kemungkinan diperbaiki melalui pembelajaran.

Prinsip akhirnya ialah kehati-hatian inferensial: sebutkan apa yang benar-benar ditunjukkan data, bedakan bukti kuat dari indikasi, gunakan triangulasi ketika diperlukan, dan hindari label permanen pada peserta didik.

- selaraskan tujuan, konstruk, tugas, skor, kategori, dan tindakan;
- utamakan konsepsi yang berdampak pada pembelajaran lanjutan;
- gabungkan bukti bila keputusan memiliki konsekuensi besar;
- evaluasi kembali efektivitas tindak lanjut melalui diagnosis ulang.

Dalam praktik kelas, Penetapan Kategori Pemahaman Konsep dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.22 Pelaporan Hasil Diagnosis

6.22.1 Landasan Konseptual dan Kedudukan Instrumen

Pemilihan Pelaporan Hasil Diagnosis harus berangkat dari pertanyaan diagnosis yang jelas, bukan dari kebiasaan menggunakan instrumen yang paling mudah tersedia. Dalam bab ini, instrumen tersebut didefinisikan sebagai penyajian temuan diagnosis secara akurat, transparan, dapat ditindaklanjuti, dan sesuai tingkat pembaca mulai dari peserta didik, guru, sekolah, peneliti, hingga pembuat kebijakan. Definisi operasional ini menegaskan bentuk bukti yang dikumpulkan sekaligus membatasi jenis kesimpulan yang layak dibuat. (AERA, APA, & NCME, 2014; Appelbaum et al., 2018; Messick, 1995)

Klaim utamanya adalah bahwa instrumen ini mengubah data instrumen menjadi informasi pembelajaran tanpa melebih-lebihkan kepastian atau menstigmatisasi peserta didik. Klaim tersebut harus dipisahkan dari anggapan bahwa setiap jawaban salah merupakan miskonsepsi. Kesalahan dapat berasal dari kurang pengetahuan, ketidakcermatan, kesulitan bahasa, salah hitung, atau ketidakfamiliaran dengan representasi.

Data pokok yang dihasilkan berupa prevalensi respons, kombinasi kategori, pola alasan, keyakinan, kutipan, visualisasi, dan catatan kualitas instrumen. Data itu baru memperoleh makna diagnostik setelah dikaitkan dengan peta konsep sasaran, kemungkinan konsepsi alternatif, dan aturan interpretasi yang telah ditetapkan sebelum analisis.

Pembedaan ini penting karena satu pola respons dapat memiliki lebih dari satu penjelasan yang masuk akal. Posisi terbaik Pelaporan Hasil Diagnosis adalah sebagai bagian dari sistem asesmen berlapis yang dapat dipadukan dengan bukti lain ketika konsekuensi keputusan semakin besar.

- menentukan konsep dan proposisi yang hendak didiagnosis;
- menjelaskan perilaku respons yang dianggap sebagai bukti;
- membatasi klaim sesuai kekuatan instrumen;
- menautkan hasil dengan tindakan pembelajaran yang realistis.

Dalam praktik kelas, Pelaporan Hasil Diagnosis dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.22.2 Perkembangan Historis dan Rasional Penggunaan

Perkembangan Pelaporan Hasil Diagnosis tidak dapat dipisahkan dari perubahan pandangan tentang belajar sains. Ketika peserta didik tidak lagi dipandang sebagai penerima informasi yang kosong, instrumen mulai dirancang untuk mengungkap gagasan awal, pola penalaran, dan keyakinan yang dibawa ke kelas. (AERA, APA, & NCME, 2014; Appelbaum et al., 2018; Messick, 1995; Kane, 2013)

Pada tahap awal, penilaian banyak mengandalkan benar-salah atau skor total. Perkembangan penelitian konsepsi peserta didik kemudian mendorong penambahan alasan, keyakinan, tugas kinerja, rekaman proses, dan bukti sosial. Pelaporan Hasil Diagnosis merupakan salah satu jawaban metodologis terhadap kebutuhan tersebut.

Dalam pendidikan kimia, kebutuhan diagnosis semakin kuat karena peserta didik harus mengoordinasikan fenomena makroskopik, model submikroskopik, simbol, grafik, dan perhitungan. Instrumen yang hanya menguji hafalan definisi sering gagal memperlihatkan cara representasi itu dihubungkan.

Sejarah penggunaannya juga menunjukkan bahwa format yang lebih kompleks tidak otomatis lebih baik. Setiap tambahan tier, rubrik, atau kanal data harus mempunyai fungsi yang jelas dan didukung bukti bahwa informasi tambahan benar-benar memperbaiki keputusan.

- pergeseran dari pengukuran skor menuju diagnosis pola pemahaman;
- penggunaan respons terbuka dan wawancara untuk menemukan konsepsi alternatif;
- pengembangan teknik kuantitatif agar diagnosis dapat dilakukan pada kelompok besar;
- integrasi data digital dan metakognitif pada perkembangan mutakhir.

Dalam praktik kelas, Pelaporan Hasil Diagnosis dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.22.3 Konstruk, Klaim, dan Bukti Diagnosis

Konstruk yang dinilai melalui Pelaporan Hasil Diagnosis bukan objek yang terlihat langsung. Yang diamati adalah prevalensi respons, kombinasi kategori, pola alasan, keyakinan, kutipan, visualisasi, dan catatan kualitas instrumen; dari sana dibuat inferensi tentang pemahaman konsep, miskonsepsi, pengetahuan parsial, atau ketidakpastian. (Flake & Fried, 2020; Gurel et al., 2015; Messick, 1995; Kane, 2013)

Klaim diagnosis perlu dirumuskan sebagai argumen. Pertama, tugas harus memerlukan konsep sasaran. Kedua, respons harus diberi skor atau kode secara konsisten. Ketiga, pola skor harus mempunyai hubungan yang masuk akal dengan struktur pengetahuan. Keempat, penggunaan hasil harus menghasilkan keputusan yang proporsional.

Sumber ancaman utama adalah ketidakjelasan konstruk. Misalnya, peserta didik dapat gagal karena istilah yang ambigu, beban membaca, representasi yang tidak dikenal, atau prosedur matematika. Kegagalan semacam itu tidak boleh langsung diperlakukan sebagai miskonsepsi kimia.

Oleh sebab itu, setiap pengembang perlu menulis pernyataan klaim-bukti: instrumen mengukur apa, perilaku apa yang menjadi indikator, alternatif penjelasan apa yang mungkin, dan bukti tambahan apa yang diperlukan untuk memperkuat kesimpulan.

- konstruk utama: pemahaman konseptual melalui komunikasi bukti diagnostik;
- bukti observabel: respons, alasan, tindakan, atau representasi;
- inferensi: profil konsepsi pada konten dan konteks tertentu;
- penggunaan: perencanaan pembelajaran, remediasi, atau penelitian.

Dalam praktik kelas, Pelaporan Hasil Diagnosis dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.22.4 Struktur, Komponen, dan Alur Respons

Struktur Pelaporan Hasil Diagnosis dibangun dari komponen yang saling bergantung. Komponen inti meliputi tujuan laporan, deskripsi instrumen, kualitas bukti, hasil tingkat butir dan konsep, ketidakpastian, rekomendasi, dan batas interpretasi. Hilangnya satu komponen dapat mengubah makna data yang diperoleh. (AERA, APA, & NCME, 2014; Appelbaum et al., 2018; Messick, 1995; Kane, 2013)

Urutan penyajian harus ditetapkan secara sadar. Komponen yang muncul lebih awal dapat memberi petunjuk pada komponen berikutnya, sedangkan permintaan alasan atau keyakinan dapat mengubah cara peserta didik memeriksa jawabannya. Penguncian respons dan uji urutan diperlukan pada format digital.

Setiap komponen membutuhkan spesifikasi. Stimulus harus memuat informasi cukup, opsi atau prompt harus sejajar dengan indikator, skala harus memiliki jangkar, dan rubrik harus menjelaskan perbedaan kategori yang berdekatan.

Struktur yang baik memungkinkan peneliti menelusuri kembali setiap kategori hasil ke respons mentah. Transparansi ini penting agar guru tidak hanya menerima label otomatis tanpa memahami alasan di baliknya.

- tujuan laporan;
- deskripsi instrumen;
- kualitas bukti;
- hasil tingkat butir dan konsep;
- ketidakpastian;
- rekomendasi;
- batas interpretasi;

Dalam praktik kelas, Pelaporan Hasil Diagnosis dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

No.	Komponen	Fungsi Diagnostik
1	tujuan laporan	Bukti yang harus ditelusuri
2	deskripsi instrumen	Bukti yang harus ditelusuri
3	kualitas bukti	Bukti yang harus ditelusuri
4	hasil tingkat butir dan konsep	Bukti yang harus ditelusuri

6.22.5 Desain Stimulus, Tugas, atau Pertanyaan

Desain stimulus pada Pelaporan Hasil Diagnosis harus memunculkan penggunaan konsep, bukan sekadar pengenalan istilah. Situasi dapat berbentuk fenomena, gambar partikel, grafik, data eksperimen, persamaan, atau masalah kontekstual, tetapi seluruh informasi yang tidak relevan perlu dikendalikan. (Haladyna et al., 2002; AERA, APA, & NCME, 2014; Appelbaum et al., 2018; Messick, 1995)

Dalam kimia, satu stimulus sering menggabungkan beberapa tingkat representasi. Pengembang harus memutuskan apakah tujuan butir adalah menerjemahkan representasi, menjelaskan hubungan sebab-akibat, membedakan contoh dan noncontoh, atau menerapkan prinsip pada konteks baru.

Bahasa perlu ringkas tanpa menghilangkan syarat ilmiah. Kata seperti kuat, stabil, jenuh, spontan, atau setimbang mempunyai arti teknis dan sehari-hari yang berbeda. Ketidakjelasan istilah dapat menghasilkan kesalahan yang berasal dari bahasa, bukan struktur konsep.

Desain yang baik juga menghindari petunjuk format, pilihan terlalu panjang, negasi bertumpuk, gambar dekoratif, dan data yang tidak diperlukan. Setiap elemen harus mempunyai alasan diagnostik.

- gunakan konteks yang cukup dikenal tetapi tidak menghafalkan contoh;
- pastikan hanya satu interpretasi ilmiah yang menjadi target;
- uji semua simbol, satuan, grafik, dan gambar;
- sediakan ruang bagi respons yang tidak diprediksi bila tujuan masih eksploratif.

Dalam praktik kelas, Pelaporan Hasil Diagnosis dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.22.6 Prosedur Pengembangan Secara Sistematis

Pengembangan Pelaporan Hasil Diagnosis harus mengikuti alur yang terdokumentasi. Langkah yang lazim ialah menetapkan audiens, memilih tingkat agregasi, menyajikan metode, melaporkan kualitas instrumen, kemudian menampilkan hasil, menyatakan ketidakpastian, merumuskan tindak lanjut, mengarsipkan data dan skrip. Urutan dapat berulang karena temuan uji coba sering memaksa revisi konstruk atau redaksi. (AERA, APA, & NCME, 2014; Appelbaum et al., 2018; Messick, 1995; Adams & Wieman, 2011)

Tahap konseptual dimulai dengan analisis kurikulum, literatur miskonsepsi, dan peta konsep. Tahap empiris awal menggunakan wawancara, respons terbuka, atau observasi untuk memastikan bahwa kategori yang dirancang benar-benar muncul pada populasi sasaran.

Setiap versi instrumen perlu diberi nomor, tanggal, dan catatan perubahan. Revisi harus menyebutkan masalah yang ditemukan, bukti yang digunakan, perubahan yang dilakukan, dan alasan mempertahankan atau menghapus butir.

Pengembangan dianggap belum selesai ketika hanya menghasilkan butir yang tampak baik. Instrumen perlu diuji pada peserta didik yang serupa dengan pengguna sasaran, lalu dianalisis secara konseptual dan empiris sebelum dipakai untuk keputusan penting.

- 1. menetapkan audiens.
- 2. memilih tingkat agregasi.
- 3. menyajikan metode.
- 4. melaporkan kualitas instrumen.
- 5. menampilkan hasil.

Dalam praktik kelas, Pelaporan Hasil Diagnosis dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.22.7 Pelaksanaan dan Pengumpulan Data

Pelaksanaan Pelaporan Hasil Diagnosis harus distandardisasi secukupnya agar perbedaan hasil lebih mungkin berasal dari pemahaman, bukan variasi instruksi, waktu, alat, atau bantuan. Petunjuk perlu menjelaskan tujuan umum tanpa mengajarkan jawaban. (AERA, APA, & NCME, 2014; Appelbaum et al., 2018)

Kondisi administrasi mencakup waktu, urutan tugas, kesempatan bertanya, penggunaan kalkulator, akses tabel periodik, kualitas gambar, dan perangkat digital. Pada instrumen kualitatif, standardisasi berarti konsistensi fokus dan prinsip probing, bukan membaca naskah secara kaku.

Data yang dikumpulkan ialah prevalensi respons, kombinasi kategori, pola alasan, keyakinan, kutipan, visualisasi, dan catatan kualitas instrumen. Data mentah harus disimpan bersama metadata yang diperlukan, seperti kode butir, versi instrumen, waktu administrasi, dan kondisi khusus, tetapi hanya data yang relevan yang boleh dikumpulkan.

Sebelum administrasi luas, lakukan uji teknis. Pastikan tautan, urutan, penguncian respons, rekaman, penyimpanan, serta tampilan rumus dan subskrip bekerja. Gangguan teknis perlu dicatat sebagai bagian dari interpretasi.

- berikan instruksi yang sama dan mudah dipahami;
- catat kondisi yang dapat memengaruhi respons;
- hindari umpan balik sebelum seluruh respons diagnostik selesai;
- sediakan prosedur penanganan data tidak lengkap.

Dalam praktik kelas, Pelaporan Hasil Diagnosis dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.22.8 Penskoran, Pengodean, dan Pengolahan Data

Penskoran Pelaporan Hasil Diagnosis dapat diringkas sebagai bukan hanya skor total; laporan menampilkan distribusi, pola konsepsi, indikator kualitas, serta aturan kategori. Aturan ini harus ditulis sebelum melihat hasil utama agar keputusan tidak disesuaikan secara oportunistik dengan pola data. (AERA, APA, & NCME, 2014; Appelbaum et al., 2018; Messick, 1995; Chi, 1997)

Skor total jarang cukup untuk diagnosis. Analisis perlu mempertahankan informasi tingkat butir, konsep, opsi, alasan, keyakinan, proses, atau representasi. Penggabungan data terlalu dini dapat menyembunyikan pola yang secara pedagogis penting.

Untuk respons terbuka atau kualitatif, buku kode harus mencantumkan definisi, contoh positif, contoh batas, noncontoh, dan aturan ketika satu respons memuat beberapa gagasan. Sebagian data perlu dinilai oleh lebih dari satu penilai.

Setiap transformasi data—misalnya dari respons ke kode, dari kode ke kategori, dan dari kategori ke persentase—harus dapat diaudit. Simpan respons asli dan keputusan pengodean agar kasus ambigu dapat ditinjau kembali.

- pisahkan skor ketepatan dari skor keyakinan atau kualitas alasan;
- laporkan kasus tidak lengkap dan aturan penanganannya;
- uji konsistensi penilai untuk data berkode;
- hindari mengubah data ordinal menjadi ukuran interval tanpa alasan.

Dalam praktik kelas, Pelaporan Hasil Diagnosis dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.22.9 Interpretasi dan Penetapan Kategori

Interpretasi hasil Pelaporan Hasil Diagnosis harus bergerak dari respons menuju pola, bukan langsung dari satu jawaban menuju label. Status konseptual menjadi lebih kuat apabila gagasan yang sama muncul pada beberapa butir, representasi, atau konteks. (AERA, APA, & NCME, 2014; Appelbaum et al., 2018; Messick, 1995; Kane, 2013)

Kategori dapat mencakup pemahaman ilmiah, pemahaman parsial, kurang pengetahuan, respons tidak konsisten, false positive, false negative, dan miskonsepsi. Tidak semua instrumen mampu membedakan seluruh kategori tersebut; skema harus mengikuti informasi yang benar-benar tersedia.

Ambang, seperti batas keyakinan atau persentase pemilihan, perlu memiliki rasional dan diuji melalui analisis sensitivitas. Perubahan kecil pada ambang dapat mengubah prevalensi kategori secara berarti.

Interpretasi yang bertanggung jawab menyatakan bahwa kategori berlaku pada konsep, bentuk tugas, waktu, dan kondisi tertentu. Hasil merupakan hipotesis kerja untuk pembelajaran, bukan diagnosis psikologis permanen.

- gunakan lebih dari satu butir untuk konsepsi penting;
- periksa konsistensi antara jawaban dan alasan;
- tandai kasus ambigu untuk konfirmasi;
- hubungkan kategori dengan rekomendasi yang berbeda.

Dalam praktik kelas, Pelaporan Hasil Diagnosis dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.22.10 Validitas dan Kualitas Inferensi

Validitas Pelaporan Hasil Diagnosis berkaitan dengan kualitas interpretasi dan penggunaan hasil, bukan sifat tetap yang dimiliki instrumen. Fokus khususnya adalah setiap klaim dalam laporan harus sebanding dengan bukti dan penggunaan yang dimaksud; batas generalisasi perlu dinyatakan. (AERA, APA, & NCME, 2014; Messick, 1995; Kane, 2013; Haynes et al., 1995)

Bukti berdasarkan isi menilai keterwakilan domain dan akurasi konsep. Bukti proses respons memeriksa bagaimana peserta didik memahami tugas. Bukti struktur internal menilai kesesuaian pola respons. Bukti hubungan dengan variabel lain dan konsekuensi penggunaan melengkapi argumen.

Ancaman validitas dapat berasal dari konten yang kurang terwakili, tuntutan bahasa berlebihan, petunjuk opsi, rubrik yang tidak stabil, keyakinan yang tidak terkalibrasi, atau algoritma yang belajar dari label keliru.

Dokumen validasi perlu menghubungkan setiap klaim dengan bukti dan alternatif penjelasan. Menampilkan satu indeks tidak cukup; validasi adalah program argumentatif yang diperbarui setiap kali instrumen dipakai pada populasi atau tujuan baru.

- bukti isi dan kesesuaian kurikulum;
- bukti proses respons melalui wawancara kognitif;
- bukti struktur dan fungsi butir;
- bukti kegunaan serta konsekuensi keputusan.

Dalam praktik kelas, Pelaporan Hasil Diagnosis dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.22.11 Reliabilitas, Keajegan, dan Analisis Empiris

Keajegan Pelaporan Hasil Diagnosis harus dinilai sesuai jenis data. Pada respons objektif dapat digunakan konsistensi internal, stabilitas waktu, atau model pengukuran. Pada data kualitatif, perhatian diarahkan pada konsistensi pengodean, audit trail, dan ketepatan penerapan definisi kode. (Kuder & Richardson, 1937; Cronbach, 1951; Ding et al., 2006; Wren & Barbera, 2014; AERA, APA, & NCME, 2014)

Reliabilitas tinggi tidak otomatis menunjukkan diagnosis yang benar. Butir yang berulang-ulang mengukur hafalan dapat menghasilkan koefisien tinggi tetapi tidak mewakili pemahaman konseptual. Sebaliknya, konstruk yang multidimensi tidak selalu layak dipaksa menjadi satu skala.

Analisis empiris perlu memeriksa kesukaran, daya beda, fungsi opsi, pola data hilang, dan keterhubungan komponen. Untuk instrumen bertingkat, reliabilitas dapat dihitung pada setiap tier serta kombinasi yang mempunyai makna teoritis.

Ukuran sampel, distribusi kemampuan, dan homogenitas konten memengaruhi statistik. Karena itu, laporan harus menyatakan kondisi estimasi dan tidak menyalin batas angka sebagai aturan universal.

- pilih statistik yang sesuai dengan jenis skor;
- laporkan interval atau ketidakpastian bila tersedia;
- periksa butir secara substantif dan statistik;
- uji kembali setelah revisi atau perubahan platform.

Dalam praktik kelas, Pelaporan Hasil Diagnosis dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.22.12 Contoh Penerapan dalam Pembelajaran Kimia

Penerapan Pelaporan Hasil Diagnosis dalam kimia dapat dimulai dari konsep profil miskonsepsi kesetimbangan per indikator. Pengembang perlu menuliskan konsep ilmiah target, miskonsepsi yang mungkin, bukti respons yang diharapkan, dan keputusan pembelajaran setelah hasil diperoleh. (AERA, APA, & NCME, 2014; Appelbaum et al., 2018; Messick, 1995; Kane, 2013; Nakhleh, 1992)

Contoh kedua adalah peta konsepsi kelas tentang asam-basa. Stimulus sebaiknya menghubungkan fenomena dengan model partikel atau persamaan, lalu meminta peserta didik menunjukkan hubungan sebab-akibat. Jawaban yang hanya menyebut istilah belum cukup menjadi bukti pemahaman.

Pada laporan keyakinan dan akurasi, diagnosis dapat memeriksa apakah peserta didik membedakan perubahan laju, keadaan setimbang, energi, dan komposisi. Kesalahan sering muncul ketika satu aturan dipindahkan ke konteks yang tidak sesuai.

Kasus rekomendasi remediasi representasi partikel menunjukkan pentingnya representasi. Respons verbal, gambar, simbol, dan hitungan dapat tidak konsisten. Hasil instrumen sebaiknya digunakan untuk memilih demonstrasi, diskusi, contoh-noncontoh, atau tugas remediasi yang secara langsung menantang hubungan salah.

- tentukan proposisi ilmiah dan konsepsi alternatif;
- gunakan minimal dua bentuk representasi;
- minta alasan atau bukti ketika format memungkinkan;
- rancang tindak lanjut berdasarkan pola, bukan hanya skor.

Dalam praktik kelas, Pelaporan Hasil Diagnosis dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.22.13 Kelebihan, Keterbatasan, dan Potensi Bias

Kekuatan utama Pelaporan Hasil Diagnosis meliputi mendukung keputusan pembelajaran, menjaga akuntabilitas interpretasi, memudahkan tindak lanjut, dan memungkinkan audit dan replikasi. Kekuatan tersebut menjelaskan situasi ketika instrumen layak dipilih dibanding teknik lain. (AERA, APA, & NCME, 2014; Appelbaum et al., 2018; Messick, 1995; Gurel et al., 2015)

Keterbatasannya mencakup persentase dapat disalahartikan, dashboard dapat menyembunyikan asumsi, label individu berisiko stigma, dan laporan panjang dapat tidak digunakan guru. Keterbatasan harus diperlakukan sebagai batas inferensi, bukan sekadar paragraf formal yang tidak memengaruhi analisis.

Potensi bias dapat berasal dari bahasa, pengalaman konteks, akses teknologi, kecemasan, gaya menggunakan skala, keterampilan menggambar, atau kemampuan menulis. Instrumen perlu diuji pada kelompok yang beragam dan ditinjau untuk fungsi berbeda antarkelompok bila data mencukupi.

Keputusan memilih instrumen sebaiknya mempertimbangkan kedalaman data, ukuran kelompok, waktu, keterampilan analisis, tujuan diagnosis, dan konsekuensi kesalahan klasifikasi. Tidak ada satu metode yang unggul untuk semua tujuan.

- Kekuatan: mendukung keputusan pembelajaran.
- Kekuatan: menjaga akuntabilitas interpretasi.
- Batas: persentase dapat disalahartikan.
- Batas: dashboard dapat menyembunyikan asumsi.

Dalam praktik kelas, Pelaporan Hasil Diagnosis dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

Aspek	Temuan	Respons Metodologis
Kekuatan	mendukung keputusan pembelajaran	Pilih ketika sesuai tujuan
Keterbatasan	persentase dapat disalahartikan	Tambahkan bukti atau kontrol
Bias potensial	Bahasa, konteks, perangkat, gaya respons	Uji proses respons

6.22.14 Integrasi Digital, Etika, dan Tata Kelola Data

Integrasi digital Pelaporan Hasil Diagnosis dapat berupa dashboard berlapis, ekspor PDF, akses berbasis peran, visualisasi yang dapat dijelaskan, dan jejak perubahan. Teknologi seharusnya memperluas kualitas bukti, bukan hanya memindahkan lembar kertas ke layar. (Baker & Inventado, 2014; Romero & Ventura, 2020; AERA, APA, & NCME, 2014)

Platform perlu menyimpan versi instrumen, urutan respons, timestamp yang relevan, aturan skor, dan log perubahan. Namun, hanya data yang dibutuhkan untuk tujuan diagnosis yang boleh dikumpulkan; pengumpulan berlebihan meningkatkan risiko privasi tanpa manfaat ilmiah.

Aspek etika utama adalah bahwa kerahasiaan, bahasa nonstigmatis, kontrol akses, dan larangan publikasi peringkat individu berdasarkan diagnosis miskonsepsi. Peserta didik dan guru perlu mengetahui data apa yang dikumpulkan, untuk tujuan apa, siapa yang dapat mengakses, berapa lama disimpan, dan bagaimana kesalahan data diperbaiki.

Algoritma otomatis harus dapat diaudit. Kategori yang dihasilkan sistem perlu ditautkan ke respons mentah dan aturan keputusan. Kasus berisiko atau ambigu harus ditinjau manusia sebelum tindak lanjut penting dilakukan.

- gunakan identitas pseudonim dan akses berbasis peran;
- uji tampilan rumus, subskrip, grafik, dan gambar;
- sediakan opsi aksesibilitas dan alternatif perangkat;
- dokumentasikan algoritma serta perubahan versi.

Dalam praktik kelas, Pelaporan Hasil Diagnosis dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

6.22.15 Sintesis dan Pedoman Implementasi

Pelaporan Hasil Diagnosis bermanfaat apabila dipilih berdasarkan tujuan dan digunakan sesuai batasnya. Instrumen ini menghasilkan prevalensi respons, kombinasi kategori, pola alasan, keyakinan, kutipan, visualisasi, dan catatan kualitas instrumen, lalu data diolah melalui bukan hanya skor total; laporan menampilkan distribusi, pola konsepsi, indikator kualitas, serta aturan kategori. (AERA, APA, & NCME, 2014; Appelbaum et al., 2018; Messick, 1995; Kane, 2013)

Rangkaian kerja yang disarankan dimulai dari analisis konsep, identifikasi konsepsi alternatif, desain tugas, telaah ahli, uji proses respons, uji lapangan, analisis, penetapan kategori, dan penyusunan tindak lanjut. Setiap tahap harus meninggalkan dokumentasi yang dapat diperiksa.

Hasil diagnosis tidak berhenti pada tabel prevalensi. Guru perlu memilih konsepsi prioritas berdasarkan frekuensi, keyakinan, posisi sebagai konsep prasyarat, dampak terhadap topik lanjutan, dan kemungkinan diperbaiki melalui pembelajaran.

Prinsip akhirnya ialah kehati-hatian inferensial: sebutkan apa yang benar-benar ditunjukkan data, bedakan bukti kuat dari indikasi, gunakan triangulasi ketika diperlukan, dan hindari label permanen pada peserta didik.

- selaraskan tujuan, konstruk, tugas, skor, kategori, dan tindakan;
- utamakan konsepsi yang berdampak pada pembelajaran lanjutan;
- gabungkan bukti bila keputusan memiliki konsekuensi besar;
- evaluasi kembali efektivitas tindak lanjut melalui diagnosis ulang.

Dalam praktik kelas, Pelaporan Hasil Diagnosis dapat digunakan sebelum pembelajaran untuk memetakan prasyarat, selama pembelajaran untuk memeriksa perubahan penalaran, atau setelah pembelajaran untuk menilai retensi dan transfer. Pemilihan waktu menentukan cara hasil ditafsirkan dan jenis respons pedagogis yang paling tepat.

Daftar Pustaka Bab 6

Adams, W. K., & Wieman, C. E. (2011). Development and validation of instruments to measure learning of expert-like thinking. *International Journal of Science Education*, 33(9), 1289–1312. doi: [10.1080/09500693.2010.511275](https://doi.org/10.1080/09500693.2010.511275).

Aiken, L. R. (1985). Three coefficients for analyzing the reliability and validity of ratings. *Educational and Psychological Measurement*, 45(1), 131–142. doi: [10.1177/0013164485451012](https://doi.org/10.1177/0013164485451012).

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. American Educational Research Association. Tautan aktif: <https://www.testingstandards.net/open-access-files.html>.

Appelbaum, M., Cooper, H., Kline, R. B., Mayo-Wilson, E., Nezu, A. M., & Rao, S. M. (2018). Journal article reporting standards for quantitative research in psychology. *American Psychologist*, 73(1), 3–25. doi: [10.1037/amp0000191](https://doi.org/10.1037/amp0000191).

Arslan, H. O., Cigdemoglu, C., & Moseley, C. (2012). A three-tier diagnostic test to assess pre-service teachers' misconceptions about global warming, greenhouse effect, ozone layer depletion, and acid rain. *International Journal of Science Education*, 34(11), 1667–1686. doi: [10.1080/09500693.2012.680618](https://doi.org/10.1080/09500693.2012.680618).

Baker, R. S., & Inventado, P. S. (2014). Educational data mining and learning analytics. In J. A. Larusson & B. White (Eds.), *Learning analytics: From research to practice* (pp. 61–75). Springer. doi: [10.1007/978-1-4614-3305-7_4](https://doi.org/10.1007/978-1-4614-3305-7_4).

Bodner, G. M., & Domin, D. S. (2000). Mental models: The role of representations in problem solving in chemistry. *University Chemistry Education*, 4(1), 24–30. Tautan aktif: https://chemed.chem.purdue.edu/chemed/bodnergrouppdf/35_Mental_Models.pdf.

Brandriet, A. R., & Bretz, S. L. (2014). Measuring meta-ignorance through the lens of confidence: Examining students' redox misconceptions about oxidation numbers, charge, and electron transfer. *Chemistry Education Research and Practice*, 15(4), 729–746. doi: [10.1039/C4RP00129J](https://doi.org/10.1039/C4RP00129J).

Caleon, I., & Subramaniam, R. (2010a). Development and application of a three-tier diagnostic test to assess secondary students' understanding of waves. *International Journal of Science Education*, 32(7), 939–961. doi: [10.1080/09500690902890130](https://doi.org/10.1080/09500690902890130).

Caleon, I., & Subramaniam, R. (2010b). Do students know what they know and what they don't know? Using a four-tier diagnostic test to assess the nature of students' alternative conceptions. *Research in Science Education*, 40(3), 313–337. doi: [10.1007/s11165-009-9122-4](https://doi.org/10.1007/s11165-009-9122-4).

- Chandrasegaran, A. L., Treagust, D. F., & Mocerino, M. (2007). The development of a two-tier multiple-choice diagnostic instrument for evaluating secondary school students' ability to describe and explain chemical reactions using multiple levels of representation. *Chemistry Education Research and Practice*, 8(3), 293–307. doi: [10.1039/B7RP90006F](https://doi.org/10.1039/B7RP90006F).
- Chi, M. T. H. (2009). Active-constructive-interactive: A conceptual framework for differentiating learning activities. *Topics in Cognitive Science*, 1(1), 73–105. doi: [10.1111/j.1756-8765.2008.01005.x](https://doi.org/10.1111/j.1756-8765.2008.01005.x).
- Chi, M. T. H. (1997). Quantifying qualitative analyses of verbal data: A practical guide. *The Journal of the Learning Sciences*, 6(3), 271–315. doi: [10.1207/s15327809jls0603_1](https://doi.org/10.1207/s15327809jls0603_1).
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3), 297–334. doi: [10.1007/BF02310555](https://doi.org/10.1007/BF02310555).
- Ding, L., Chabay, R., Sherwood, B., & Beichner, R. (2006). Evaluating an electricity and magnetism assessment tool: Brief electricity and magnetism assessment. *Physical Review Special Topics - Physics Education Research*, 2(1), 010105. doi: [10.1103/PhysRevSTPER.2.010105](https://doi.org/10.1103/PhysRevSTPER.2.010105).
- Ericsson, K. A., & Simon, H. A. (1993). *Protocol analysis: Verbal reports as data* (Rev. ed.). MIT Press. doi: [10.7551/mitpress/5657.001.0001](https://doi.org/10.7551/mitpress/5657.001.0001).
- Flake, J. K., & Fried, E. I. (2020). Measurement schmeasurement: Questionable measurement practices and how to avoid them. *Advances in Methods and Practices in Psychological Science*, 3(4), 456–465. doi: [10.1177/2515245920952393](https://doi.org/10.1177/2515245920952393).
- Ginsburg, H. P. (1997). *Entering the child's mind: The clinical interview in psychological research and practice*. Cambridge University Press. doi: [10.1017/CBO9780511527777](https://doi.org/10.1017/CBO9780511527777).
- Gurcay, D., & Gulbas, E. (2015). Development of three-tier heat, temperature and internal energy diagnostic test. *Research in Science & Technological Education*, 33(2), 197–217. doi: [10.1080/02635143.2015.1018154](https://doi.org/10.1080/02635143.2015.1018154).
- Gurel, D. K., Eryilmaz, A., & McDermott, L. C. (2015). A review and comparison of diagnostic instruments to identify students' misconceptions in science. *Eurasia Journal of Mathematics, Science and Technology Education*, 11(5), 989–1008. doi: [10.12973/eurasia.2015.1369a](https://doi.org/10.12973/eurasia.2015.1369a).
- Haladyna, T. M., Downing, S. M., & Rodriguez, M. C. (2002). A review of multiple-choice item-writing guidelines for classroom assessment. *Applied Measurement in Education*, 15(3), 309–334. doi: [10.1207/S15324818AME1503_5](https://doi.org/10.1207/S15324818AME1503_5).
- Hasan, S., Bagayoko, D., & Kelley, E. L. (1999). Misconceptions and the certainty of response index (CRI). *Physics Education*, 34(5), 294–299. doi: [10.1088/0031-9120/34/5/304](https://doi.org/10.1088/0031-9120/34/5/304).

Haynes, S. N., Richard, D. C. S., & Kubany, E. S. (1995). Content validity in psychological assessment: A functional approach to concepts and methods. *Psychological Assessment*, 7(3), 238–247. doi: [10.1037/1040-3590.7.3.238](https://doi.org/10.1037/1040-3590.7.3.238).

Hoe, K. Y., & Subramaniam, R. (2016). On the prevalence of alternative conceptions on acid-base chemistry among secondary students: Insights from cognitive and confidence measures. *Chemistry Education Research and Practice*, 17(2), 263–282. doi: [10.1039/C5RP00146C](https://doi.org/10.1039/C5RP00146C).

Çil, E. (2015). Effect of two-tier diagnostic tests on promoting learners' conceptual understanding of variables in conducting scientific experiments. *Applied Measurement in Education*, 28(4), 253–273. doi: [10.1080/08957347.2015.1064124](https://doi.org/10.1080/08957347.2015.1064124).

Kaltakci-Gurel, D., Eryilmaz, A., & McDermott, L. C. (2017). Development and application of a four-tier test to assess pre-service physics teachers' misconceptions about geometrical optics. *Research in Science & Technological Education*, 35(2), 238–260. doi: [10.1080/02635143.2017.1310094](https://doi.org/10.1080/02635143.2017.1310094).

Kane, M. T. (2013). Validating the interpretations and uses of test scores. *Journal of Educational Measurement*, 50(1), 1–73. doi: [10.1111/jedm.12000](https://doi.org/10.1111/jedm.12000).

Kearney, M. (2004). Classroom use of multimedia-supported predict-observe-explain tasks in a social constructivist learning environment. *Research in Science Education*, 34(4), 427–453. doi: [10.1007/s11165-004-8795-y](https://doi.org/10.1007/s11165-004-8795-y).

Kuder, G. F., & Richardson, M. W. (1937). The theory of the estimation of test reliability. *Psychometrika*, 2(3), 151–160. doi: [10.1007/BF02288391](https://doi.org/10.1007/BF02288391).

Lawshe, C. H. (1975). A quantitative approach to content validity. *Personnel Psychology*, 28(4), 563–575. doi: [10.1111/j.1744-6570.1975.tb01393.x](https://doi.org/10.1111/j.1744-6570.1975.tb01393.x).

Lindsey, B. A., Phillips, A. C., & Asato, J. M. (2015). Development and application of a novel Rasch-based methodology for evaluating multi-tiered assessment instruments: Validation and utilization of an undergraduate diagnostic test of the water cycle. *International Journal of Science Education*, 37(16), 2740–2768. doi: [10.1080/09500693.2015.1105398](https://doi.org/10.1080/09500693.2015.1105398).

Lynn, M. R. (1986). Determination and quantification of content validity. *Nursing Research*, 35(6), 382–385. doi: [10.1097/00006199-198611000-00017](https://doi.org/10.1097/00006199-198611000-00017).

Ma, H., Liu, W., & Li, G. (2025). Development and application of a five-tier diagnostic test to assess misconceptions on respiration and photosynthesis among senior high school students in Mainland China. *Research in Science Education*, 55(5), 1427–1452. doi: [10.1007/s11165-025-10232-6](https://doi.org/10.1007/s11165-025-10232-6).

Mathabathe, K. C., & Potgieter, M. (2014). Metacognitive monitoring and learning gain in foundation chemistry. *Chemistry Education Research and Practice*, 15(1), 94–104. doi: [10.1039/C3RP00119A](https://doi.org/10.1039/C3RP00119A).

- McClure, J. R., Sonak, B., & Suen, H. K. (1999). Concept map assessment of classroom learning: Reliability, validity, and logistical practicality. *Journal of Research in Science Teaching*, 36(4), 475–492. doi: [10.1002/\(SICI\)1098-2736\(199904\)36:4%3C475::AID-TEA5%3E3.0.CO;2-O](https://doi.org/10.1002/(SICI)1098-2736(199904)36:4%3C475::AID-TEA5%3E3.0.CO;2-O).
- Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist*, 50(9), 741–749. doi: [10.1037/0003-066X.50.9.741](https://doi.org/10.1037/0003-066X.50.9.741).
- Mortimer, E. F., & Scott, P. H. (2003). *Meaning making in secondary science classrooms*. Open University Press. Tautan aktif: <https://www.mheducation.co.uk/meaning-making-in-secondary-science-classrooms-9780335212071-emea-group>.
- Mulford, D. R., & Robinson, W. R. (2002). An inventory for alternate conceptions among first-semester general chemistry students. *Journal of Chemical Education*, 79(6), 739–744. doi: [10.1021/ed079p739](https://doi.org/10.1021/ed079p739).
- Nakhleh, M. B. (1992). Why some students don't learn chemistry: Chemical misconceptions. *Journal of Chemical Education*, 69(3), 191–196. doi: [10.1021/ed069p191](https://doi.org/10.1021/ed069p191).
- Novak, J. D., & Gowin, D. B. (1984). *Learning how to learn*. Cambridge University Press. doi: [10.1017/CBO9781139173469](https://doi.org/10.1017/CBO9781139173469).
- Nurrenbern, S. C., & Pickering, M. (1987). Concept learning versus problem solving: Is there a difference? *Journal of Chemical Education*, 64(6), 508–510. doi: [10.1021/ed064p508](https://doi.org/10.1021/ed064p508).
- Odom, A. L., & Barrow, L. H. (1995). Development and application of a two-tier diagnostic test measuring college biology students' understanding of diffusion and osmosis after a course of instruction. *Journal of Research in Science Teaching*, 32(1), 45–61. doi: [10.1002/tea.3660320106](https://doi.org/10.1002/tea.3660320106).
- Peşman, H., & Eryılmaz, A. (2010). Development of a three-tier test to assess misconceptions about simple electric circuits. *The Journal of Educational Research*, 103(3), 208–222. doi: [10.1080/00220670903383002](https://doi.org/10.1080/00220670903383002).
- Polit, D. F., & Beck, C. T. (2006). The content validity index: Are you sure you know what's being reported? Critique and recommendations. *Research in Nursing & Health*, 29(5), 489–497. doi: [10.1002/nur.20147](https://doi.org/10.1002/nur.20147).
- Posner, G. J., & Gertzog, W. A. (1982). The clinical interview and the measurement of conceptual change. *Science Education*, 66(2), 195–209. doi: [10.1002/sce.3730660203](https://doi.org/10.1002/sce.3730660203).
- Putra, A. S. U., Hamidah, I., & Nahadi. (2020). The development of five-tier diagnostic test to identify misconceptions and causes of students' misconceptions in waves and optics materials. *Journal of Physics: Conference Series*, 1521(2), 022020. doi: [10.1088/1742-6596/1521/2/022020](https://doi.org/10.1088/1742-6596/1521/2/022020).

Rodriguez, M. C. (2005). Three options are optimal for multiple-choice items: A meta-analysis of 80 years of research. *Educational Measurement: Issues and Practice*, 24(2), 3–13. doi: [10.1111/j.1745-3992.2005.00006.x](https://doi.org/10.1111/j.1745-3992.2005.00006.x).

Romero, C., & Ventura, S. (2020). Educational data mining and learning analytics: An updated survey. *WIREs Data Mining and Knowledge Discovery*, 10(3), e1355. doi: [10.1002/widm.1355](https://doi.org/10.1002/widm.1355).

Ruiz-Primo, M. A., & Shavelson, R. J. (1996). Problems and issues in the use of concept maps in science assessment. *Journal of Research in Science Teaching*, 33(6), 569–600. doi: [10.1002/\(SICI\)1098-2736\(199608\)33:6%3C569::AID-TEA1%3E3.0.CO;2-M](https://doi.org/10.1002/(SICI)1098-2736(199608)33:6%3C569::AID-TEA1%3E3.0.CO;2-M).

Sadler, P. M. (1998). Psychometric models of student conceptions in science: Reconciling qualitative studies and distractor-driven assessment instruments. *Journal of Research in Science Teaching*, 35(3), 265–296. doi: [10.1002/\(SICI\)1098-2736\(199803\)35:3%3C265::AID-TEA3%3E3.0.CO;2-P](https://doi.org/10.1002/(SICI)1098-2736(199803)35:3%3C265::AID-TEA3%3E3.0.CO;2-P).

Siemens, G., & Baker, R. S. J. d. (2012). Learning analytics and educational data mining: Towards communication and collaboration. In *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge* (pp. 252–254). ACM. doi: [10.1145/2330601.2330661](https://doi.org/10.1145/2330601.2330661).

Taber, K. S. (2002). *Chemical misconceptions: Prevention, diagnosis and cure. Volume II: Classroom resources*. Royal Society of Chemistry. Tautan aktif: <https://edu.rsc.org/resources/chemical-misconceptions/4012558.article>.

Tarrant, M., Ware, J., & Mohammed, A. M. (2009). An assessment of functioning and non-functioning distractors in multiple-choice questions: A descriptive analysis. *BMC Medical Education*, 9, 40. doi: [10.1186/1472-6920-9-40](https://doi.org/10.1186/1472-6920-9-40).

Taslidere, E. (2016). Development and use of a three-tier diagnostic test to assess high school students' misconceptions about the photoelectric effect. *Research in Science & Technological Education*, 34(2), 164–186. doi: [10.1080/02635143.2015.1124409](https://doi.org/10.1080/02635143.2015.1124409).

Treagust, D. F. (1988). Development and use of diagnostic tests to evaluate students' misconceptions in science. *International Journal of Science Education*, 10(2), 159–169. doi: [10.1080/0950069880100204](https://doi.org/10.1080/0950069880100204).

van Someren, M. W., Barnard, Y. F., & Sandberg, J. A. C. (1994). *The think aloud method: A practical guide to modelling cognitive processes*. Academic Press. Tautan aktif: <https://staff.fnwi.uva.nl/m.w.vansomeren/papers/ThinkAloudMethod.pdf>.

White, R. T., & Gunstone, R. F. (1992). *Probing understanding*. Falmer Press. doi: [10.4324/9780203761342](https://doi.org/10.4324/9780203761342).

Wren, D., & Barbera, J. (2014). Psychometric analysis of the thermochemistry concept inventory. *Chemistry Education Research and Practice*, 15(3), 380–390. doi: [10.1039/C3RP00170A](https://doi.org/10.1039/C3RP00170A).

Yang, D.-C., & Lin, Y.-C. (2015). Assessing 10- to 11-year-old children's performance and misconceptions in number sense using a four-tier diagnostic test. *Educational Research*, 57(4), 368–388. doi: [10.1080/00131881.2015.1085235](https://doi.org/10.1080/00131881.2015.1085235).

Yan, Y. K., & Subramaniam, R. (2018). Using a multi-tier diagnostic test to explore the nature of students' alternative conceptions on reaction kinetics. *Chemistry Education Research and Practice*, 19(1), 213–226. doi: [10.1039/C7RP00143F](https://doi.org/10.1039/C7RP00143F).

Yin, Y., Vanides, J., Ruiz-Primo, M. A., Ayala, C. C., & Shavelson, R. J. (2005). Comparison of two concept-mapping techniques: Implications for scoring, interpretation, and use. *Journal of Research in Science Teaching*, 42(2), 166–184. doi: [10.1002/tea.20049](https://doi.org/10.1002/tea.20049).

BAB 7

METODOLOGI PENELITIAN MISKONSEPSI KIMIA

Pengantar Bab

Penelitian miskonsepsi kimia bukan sekadar kegiatan menghitung jawaban salah. Penelitian ini berusaha membuat inferensi tentang cara peserta didik memaknai konsep, menghubungkan representasi, menggunakan aturan, membangun model mental, dan mempertahankan keyakinan ketika berhadapan dengan bukti. Karena objek kajiannya berupa struktur pemahaman yang tidak dapat diamati secara langsung, kualitas penelitian sangat bergantung pada ketepatan metodologi. Jawaban, alasan, gambar, persamaan, percakapan, dan tindakan eksperimen merupakan indikator yang harus ditafsirkan secara hati-hati, bukan bukti yang otomatis berbicara sendiri.

Kompleksitas kimia menambah tuntutan tersebut. Satu konsep dapat muncul pada tingkat makroskopik, submikroskopik, simbolik, matematis, dan kontekstual. Peserta didik mungkin benar pada satu tingkat tetapi keliru pada tingkat lain. Mereka juga dapat memberikan jawaban yang sama melalui jalur penalaran berbeda. Oleh sebab itu, penelitian perlu membedakan ketidaktahuan, pengetahuan parsial, kesalahan bahasa, kesalahan prosedural, tebakan, dan miskonsepsi yang relatif stabil. Perbedaan ini menentukan instrumen, teknik pengumpulan data, analisis, serta bahasa klaim.

Bab ini menyajikan metodologi penelitian miskonsepsi kimia secara menyeluruh. Pembahasan dimulai dari paradigma penelitian, kemudian bergerak menuju desain deskriptif, studi kasus, survei, eksperimen, kuasi-eksperimen, *mixed methods*, pengembangan instrumen, *design-based research*, penelitian tindakan kelas, dan *systematic literature review*. Bagian berikutnya menguraikan pemilihan subjek, teknik pengumpulan data, analisis kuantitatif dan kualitatif, triangulasi, analisis pola respons, pemetaan jaringan konsep, etika, penyajian temuan, serta keterbatasan penelitian.

Setiap desain memiliki pertanyaan yang dapat dijawab dan batas yang tidak boleh dilampaui. Studi deskriptif dapat memetakan pola, tetapi tidak dengan sendirinya membuktikan sebab. Eksperimen dapat memperkuat inferensi kausal, tetapi tetap memerlukan bukti bahwa instrumen benar-benar mengukur perubahan konsep. Wawancara menghasilkan kedalaman, tetapi interpretasinya bergantung pada kualitas pertanyaan, konteks, dan analisis. *Mixed methods* memberi peluang integrasi, tetapi tidak cukup hanya menempatkan data kuantitatif dan kualitatif dalam satu laporan.

Prinsip utama yang digunakan sepanjang bab adalah keselarasan antara pertanyaan, konstruk, desain, sampel, data, analisis, dan klaim. Validitas diperlakukan sebagai argumen yang

didukung berbagai sumber bukti. Reliabilitas dipahami sebagai ketepatan dan konsistensi skor untuk penggunaan tertentu, bukan sekadar koefisien yang harus melewati angka baku. Ketidakpastian, kasus yang bertentangan, dan keterbatasan dipandang sebagai bagian dari pengetahuan, bukan gangguan yang perlu disembunyikan.

Dalam praktiknya, metodologi penelitian miskonsepsi perlu tetap dekat dengan kebutuhan pembelajaran. Temuan yang kuat harus dapat membantu guru mengenali konsep prasyarat yang rapuh, memilih representasi, merancang pertanyaan diagnostik, dan mengevaluasi perubahan konsep. Namun, penelitian juga harus melindungi peserta didik dari stigmatisasi. Istilah miskonsepsi tidak boleh berubah menjadi label identitas; yang dinilai adalah pola pemahaman pada tugas dan kondisi tertentu.

Rujukan pada bab ini dipilih dari standar pengukuran, pedoman pelaporan, literatur metodologi penelitian, serta penelitian pendidikan kimia yang memiliki identitas bibliografis jelas. DOI dicantumkan ketika tersedia. Dengan landasan tersebut, bab ini diharapkan menjadi panduan untuk menyusun proposal, melaksanakan penelitian, menganalisis data, menulis laporan, dan menilai kualitas penelitian miskonsepsi kimia secara kritis.

7.1 Paradigma Penelitian Miskonsepsi

Subbab ini membahas kerangka filosofis yang menentukan apa yang dianggap sebagai realitas konseptual, bagaimana bukti mengenai miskonsepsi diperoleh, dan sejauh mana peneliti boleh membuat klaim tentang pikiran peserta didik. Fokus utamanya adalah menjaga keselarasan antara pertanyaan penelitian, bentuk bukti, teknik analisis, dan kekuatan klaim. Dalam konteks miskonsepsi kimia, keputusan metodologis harus mampu membedakan pola konseptual dari respons yang timbul karena keterbatasan bahasa, kesalahan hitung, atau kondisi administrasi. Pembahasan berikut dipecah menjadi lima belas bagian agar setiap keputusan dapat ditelaah secara operasional.

7.1.1 Hakikat Paradigma dalam Penelitian Miskonsepsi

Pembahasan mengenai hakikat paradigma dalam penelitian miskonsepsi diperlukan agar paradigma penelitian miskonsepsi tidak berhenti pada prosedur administratif. Fokus ini berada dalam kerangka filosofis yang menentukan apa yang dianggap sebagai realitas konseptual, bagaimana bukti mengenai miskonsepsi diperoleh, dan sejauh mana peneliti boleh membuat klaim tentang pikiran peserta didik. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. AERA, APA, dan NCME (2014) serta Nakhleh (1992) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Penerapan praktis dapat disusun dalam urutan: mendefinisikan hubungan antara respons teramati dan struktur kognitif, mengoperasionalkan peran konteks sosial dan bahasa, mengumpulkan alasan tertulis, memeriksa keterlacakan inferensi, lalu melaporkan model inferensi konseptual. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Misalnya, pada ikatan kimia, perubahan skor sesudah pembelajaran belum otomatis menandakan perubahan konsep. Peserta didik mungkin mempelajari pola jawaban tes. Bukti yang lebih meyakinkan mencakup transfer ke situasi baru, penjelasan kausal, representasi partikel, dan retensi setelah jeda waktu.

Keputusan yang harus didokumentasikan:

- Definisikan hubungan antara respons teramati dan struktur kognitif dalam istilah operasional.
- Tentukan bukti utama dari wawancara klinis dan bukti pendukung dari rekaman diskusi kelas.
- Periksa reflektivitas sebelum membuat kategori atau model.
- Catat potensi menyederhanakan konstruktivisme menjadi relativisme beserta strategi mitigasinya.
- Sajikan batas generalisasi bersama batas interpretasi dan ketidakpastian.

Temuan yang baik tidak selalu berarti hasil yang besar atau sesuai harapan. Temuan bernilai ketika keterlacakan inferensi terjaga, kecukupan bukti dijelaskan, dan ketidakpastian disajikan. Kerangka AERA, APA, dan NCME (2014) serta Nakhleh (1992) menolong peneliti membedakan kekuatan bukti dari daya tarik narasi.

Inti metodologis subbagian ini adalah keterlacakan. Pembaca harus dapat mengikuti perjalanan dari pertanyaan, data, pengodean, analisis, sampai model inferensi konseptual. Ketika jejak itu jelas, temuan dapat digunakan untuk memperbaiki pembelajaran tanpa mengubah diagnosis konseptual menjadi label bagi peserta didik.

7.1.2 Ontologi Konsep, Konsepsi, dan Miskonsepsi

Dalam metodologi penelitian miskonsepsi kimia, ontologi konsep, konsepsi, dan miskonsepsi menentukan mutu hubungan antara data dan kesimpulan. Fokus ini berada dalam kerangka filosofis yang menentukan apa yang dianggap sebagai realitas konseptual, bagaimana bukti mengenai miskonsepsi diperoleh, dan sejauh mana peneliti boleh membuat klaim tentang pikiran peserta didik. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan

objek yang dapat dilihat secara langsung. Messick (1995) serta AERA, APA, dan NCME (2014) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Prosedur minimal mencakup definisi status pengetahuan ilmiah sebagai rujukan, alasan memilih wawancara klinis, aturan mengubah respons menjadi kategori, serta mekanisme pemeriksaan reflektivitas. Semua penyimpangan dari rencana awal perlu dicatat. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Pada penelitian mengenai asam-basa, peneliti dapat menyusun beberapa tugas ekuivalen yang berubah konteks, representasi, atau tuntutan penalarannya. Jika pola respons tetap berulang dan disertai keyakinan tinggi, terdapat dasar yang lebih kuat untuk menyebut indikasi miskonsepsi. Jika pola berubah-ubah, interpretasi yang lebih hati-hati adalah pengetahuan parsial atau ketergantungan konteks.

Pertanyaan audit metodologis:

- Definisikan status pengetahuan ilmiah sebagai rujukan dalam istilah operasional.
- Tentukan bukti utama dari gambar partikel dan bukti pendukung dari jejak perubahan keyakinan.
- Periksa kecukupan bukti sebelum membuat kategori atau model.
- Catat potensi mengklaim kausalitas dari data deskriptif beserta strategi mitigasinya.
- Sajikan peta hubungan data dan klaim bersama batas interpretasi dan ketidakpastian.

Pemeriksaan mutu hendaknya dilakukan selama penelitian, bukan hanya pada akhir. Catatan tentang reflektivitas, keputusan terkait kejelasan batas klaim, dan kejadian yang berpotensi menimbulkan mengabaikan konteks tugas perlu dipelihara dalam *audit trail*. Hal ini memungkinkan koreksi dini dan pelaporan yang jujur.

Kesimpulannya, ontologi konsep, konsepsi, dan miskonsepsi perlu dirancang sebagai proses argumentatif. Data tidak berbicara sendiri; maknanya bergantung pada definisi konstruk, kualitas prosedur, dan konteks. Oleh sebab itu, klaim akhir sebaiknya menggunakan bahasa yang sesuai dengan kekuatan bukti, misalnya “menunjukkan pola”, “mengindikasikan”, atau “mendukung penjelasan”.

7.1.3 Epistemologi Bukti tentang Pemahaman Peserta Didik

Keputusan tentang epistemologi bukti tentang pemahaman peserta didik harus dibuat sebelum data utama dikumpulkan, bukan ditambahkan setelah hasil diketahui. Fokus ini berada dalam kerangka filosofis yang menentukan apa yang dianggap sebagai realitas konseptual, bagaimana bukti mengenai miskonsepsi diperoleh, dan sejauh mana peneliti boleh membuat klaim tentang pikiran peserta didik. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut

dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Treagust (1988) serta Nakhleh (1992) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Tahap kerja dimulai dengan penegasan peran konteks sosial dan bahasa. Setelah itu, peneliti menyusun aturan untuk bentuk klaim yang dapat dipertanggungjawabkan dan menentukan bagaimana hubungan antara respons teramati dan struktur kognitif akan diperiksa. Bukti sebaiknya berasal dari lebih dari satu bentuk, misalnya gambar partikel dan jejak perubahan keyakinan. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Pada topik redoks, peneliti juga perlu memperhatikan bahasa. Istilah yang sama dapat diberi makna berbeda oleh peserta didik. Klarifikasi melalui pertanyaan tindak lanjut, gambar, atau contoh tandingan membantu membedakan masalah terminologi dari kesalahan konseptual.

Bukti minimum yang perlu tersedia:

- Definisikan peran konteks sosial dan bahasa dalam istilah operasional.
- Tentukan bukti utama dari rekaman diskusi kelas dan bukti pendukung dari jawaban diagnostik.
- Periksa kejelasan batas klaim sebelum membuat kategori atau model.
- Catat potensi melabeli peserta didik secara esensial beserta strategi mitigasinya.
- Sajikan catatan refleksivitas bersama batas interpretasi dan ketidakpastian.

Kualitas penelitian meningkat ketika peneliti menunjukkan rantai bukti dari gambar partikel menuju batas generalisasi, menyatakan asumsi, dan memeriksa alternatif penjelasan. Sebaliknya, menyederhanakan konstruktivisme menjadi relativisme dapat menghasilkan klaim yang tampak tegas tetapi tidak stabil. Rekomendasi praktis harus proporsional dengan bukti.

Pada akhirnya, keputusan tentang epistemologi bukti tentang pemahaman peserta didik menentukan apakah penelitian menghasilkan daftar jawaban salah atau pemahaman yang dapat ditindaklanjuti mengenai cara peserta didik berpikir. Sasaran utamanya ialah bukti yang cukup untuk merancang diagnosis, pembelajaran, dan penelitian lanjutan secara bertanggung jawab.

7.1.4 Postpositivisme dan Inferensi Berbasis Indikator

Postpositivisme dan inferensi berbasis indikator perlu dipahami sebagai keputusan metodologis yang mempunyai konsekuensi pada desain, analisis, dan pelaporan. Fokus ini berada dalam kerangka filosofis yang menentukan apa yang dianggap sebagai realitas konseptual, bagaimana bukti mengenai miskonsepsi diperoleh, dan sejauh mana peneliti

boleh membuat klaim tentang pikiran peserta didik. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Nakhleh (1992) serta AERA, APA, dan NCME (2014) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Peneliti sebaiknya menyiapkan matriks yang menghubungkan pertanyaan penelitian dengan derajat keterlibatan peneliti, rekaman diskusi kelas, teknik analisis, dan peta hubungan data dan klaim. Matriks ini mencegah pengumpulan data yang banyak tetapi tidak relevan. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Ilustrasi termokimia juga menunjukkan bahwa kualitas instrumen memengaruhi temuan. Distraktor yang tidak berasal dari pemikiran peserta didik dapat menghasilkan pilihan acak, sedangkan pertanyaan yang terlalu menuntut matematika dapat mengukur keterampilan hitung alih-alih konsepsi. Peneliti perlu menilai kesesuaian bukti dengan konstruk yang dimaksud.

Kriteria pemeriksaan sebelum analisis:

- Definisikan derajat keterlibatan peneliti dalam istilah operasional.
- Tentukan bukti utama dari jejak perubahan keyakinan dan bukti pendukung dari alasan tertulis.
- Periksa konsistensi istilah sebelum membuat kategori atau model.
- Catat potensi menganggap respons tunggal sebagai cermin langsung pikiran beserta strategi mitigasinya.
- Sajikan pernyataan posisi filosofis bersama batas interpretasi dan ketidakpastian.

Peneliti sebaiknya memisahkan keputusan apriori dari keputusan eksploratori. Pemisahan tersebut membantu mengendalikan mengklaim kausalitas dari data deskriptif dan mencegah mencampur paradigma tanpa alasan. Selanjutnya, kejelasan batas klaim dan koherensi paradigma dan pertanyaan penelitian dilaporkan bersama ketidakpastian, sehingga pembaca dapat menilai kekuatan temuan tanpa bergantung pada kesimpulan penulis semata.

Dengan demikian, postpositivisme dan inferensi berbasis indikator berfungsi menghubungkan keputusan desain dengan interpretasi. Produk yang diharapkan bukan hanya peta hubungan data dan klaim, tetapi juga penjelasan mengenai bagaimana produk tersebut diperoleh dan dalam kondisi apa ia dapat digunakan. Penelitian miskonsepsi yang baik selalu menyisakan ruang bagi penjelasan alternatif dan verifikasi lanjutan.

7.1.5 Konstruktivisme dalam Kajian Konsepsi Alternatif

Ketelitian penelitian pada bagian konstruktivisme dalam kajian konsepsi alternatif sangat menentukan apakah pola jawaban dapat ditafsirkan sebagai bukti konseptual. Fokus ini berada dalam kerangka filosofis yang menentukan apa yang dianggap sebagai realitas konseptual, bagaimana bukti mengenai miskonsepsi diperoleh, dan sejauh mana peneliti boleh membuat klaim tentang pikiran peserta didik. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Talanquer (2011) serta Nakhleh (1992) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Dokumentasi metodologis hendaknya memuat siapa yang melakukan jejak perubahan keyakinan, kapan data diperoleh, bagaimana hubungan antara respons teramati dan struktur kognitif dikendalikan, dan atas dasar apa catatan reflektivitas disusun. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Contoh pada kajian struktur atom memperlihatkan pentingnya pemisahan antara ketidaktahuan, kesalahan prosedural, dan miskonsepsi. Seorang peserta dapat gagal karena lupa rumus, sedangkan peserta lain memilih jawaban yang sama karena menggunakan model mental yang tidak ilmiah. Perbedaan ini baru terlihat ketika jejak perubahan keyakinan dibaca bersama alasan tertulis dan bukti tentang alasan.

Langkah operasional yang disarankan:

- Definisikan bentuk klaim yang dapat dipertanggungjawabkan dalam istilah operasional.
- Tentukan bukti utama dari jawaban diagnostik dan bukti pendukung dari wawancara klinis.
- Periksa koherensi paradigma dan pertanyaan penelitian sebelum membuat kategori atau model.
- Catat potensi mencampur paradigma tanpa alasan beserta strategi mitigasinya.
- Sajikan model inferensi konseptual bersama batas interpretasi dan ketidakpastian.

Standar metodologis menuntut koherensi antara konstruk, prosedur, analisis, dan klaim. Dalam konteks ini, konsistensi istilah tidak dapat digantikan oleh indeks tunggal. Demikian pula, keterlacakan inferensi harus dinilai dengan bukti yang sesuai. Talanquer (2011) dan Nakhleh (1992) memberikan dasar untuk pelaporan yang lengkap dan dapat ditelaah.

Inti metodologis subbagian ini adalah keterlacakan. Pembaca harus dapat mengikuti perjalanan dari pertanyaan, data, pengodean, analisis, sampai catatan reflektivitas. Ketika

jejak itu jelas, temuan dapat digunakan untuk memperbaiki pembelajaran tanpa mengubah diagnosis konseptual menjadi label bagi peserta didik.

7.1.6 Interpretivisme dan Makna Kontekstual

Peneliti sering memusatkan perhatian pada teknik, padahal interpretivisme dan makna kontekstual justru mengatur batas inferensi yang dapat dibuat. Fokus ini berada dalam kerangka filosofis yang menentukan apa yang dianggap sebagai realitas konseptual, bagaimana bukti mengenai miskonsepsi diperoleh, dan sejauh mana peneliti boleh membuat klaim tentang pikiran peserta didik. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Creswell dan Creswell (2018) serta AERA, APA, dan NCME (2014) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Rancangan yang dapat diaudit mencantumkan alasan pemilihan posisi ontologis terhadap konsepsi peserta didik, prosedur penanganan status pengetahuan ilmiah sebagai rujukan, dan kriteria keputusan untuk derajat keterlibatan peneliti. Tanpa penjelasan itu, pembaca sulit menilai apakah jawaban diagnostik benar-benar mendukung klaim. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Kasus stoikiometri dapat digunakan untuk memperlihatkan hubungan antara pola individual dan pola kelas. Sebagian miskonsepsi mungkin hanya dimiliki beberapa peserta tetapi sangat kuat, sedangkan pola lain tersebar luas karena bersumber dari analogi atau buku teks yang sama. Analisis perlu menjaga kedua tingkat tersebut.

Keputusan yang harus didokumentasikan:

- Definisikan posisi ontologis terhadap konsepsi peserta didik dalam istilah operasional.
- Tentukan bukti utama dari alasan tertulis dan bukti pendukung dari gambar partikel.
- Periksa keterlacakan inferensi sebelum membuat kategori atau model.
- Catat potensi mengabaikan konteks tugas beserta strategi mitigasinya.
- Sajikan definisi operasional miskonsepsi bersama batas interpretasi dan ketidakpastian.

Pemeriksaan kualitas dapat dilakukan melalui audit terhadap koherensi paradigma dan pertanyaan penelitian, analisis kasus yang bertentangan, dan evaluasi reflektivitas. Apabila ditemukan menganggap respons tunggal sebagai cermin langsung pikiran, peneliti perlu menurunkan kekuatan klaim atau menambah bukti. Transparansi seperti ini meningkatkan kegunaan temuan bagi penelitian lanjutan dan praktik kelas.

Kesimpulannya, interpretivisme dan makna kontekstual perlu dirancang sebagai proses argumentatif. Data tidak berbicara sendiri; maknanya bergantung pada definisi konstruk, kualitas prosedur, dan konteks. Oleh sebab itu, klaim akhir sebaiknya menggunakan bahasa yang sesuai dengan kekuatan bukti, misalnya “menunjukkan pola”, “mengindikasikan”, atau “mendukung penjelasan”.

7.1.7 Pragmatisme sebagai Dasar Pemilihan Metode

Kajian yang kuat memperlakukan pragmatisme sebagai dasar pemilihan metode sebagai bagian dari rantai argumentasi, bukan sebagai unsur pelengkap. Fokus ini berada dalam kerangka filosofis yang menentukan apa yang dianggap sebagai realitas konseptual, bagaimana bukti mengenai miskonsepsi diperoleh, dan sejauh mana peneliti boleh membuat klaim tentang pikiran peserta didik. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. AERA, APA, dan NCME (2014) serta Nakhleh (1992) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Pada tahap perencanaan, hubungan antara respons teramati dan struktur kognitif dan peran konteks sosial dan bahasa harus dibedakan secara eksplisit. Pada tahap pelaksanaan, konsistensi alasan tertulis perlu diperiksa; pada tahap interpretasi, keterlacakan inferensi menjadi dasar pembatasan klaim. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Dalam kelas yang mempelajari sifat koligatif, data agregat mungkin menunjukkan mayoritas jawaban salah. Akan tetapi, wawancara dapat memperlihatkan beberapa mekanisme berbeda: bahasa sehari-hari, generalisasi aturan, salah membaca simbol, atau model partikel yang tidak tepat. Karena itu, satu persentase tidak boleh menggantikan analisis struktur alasan.

Pertanyaan audit metodologis:

- Definisikan hubungan antara respons teramati dan struktur kognitif dalam istilah operasional.
- Tentukan bukti utama dari wawancara klinis dan bukti pendukung dari rekaman diskusi kelas.
- Periksa reflektivitas sebelum membuat kategori atau model.
- Catat potensi menyederhanakan konstruktivisme menjadi relativisme beserta strategi mitigasinya.

- Sajikan batas generalisasi bersama batas interpretasi dan ketidakpastian.

Kriteria keberhasilan tidak cukup berupa banyaknya data. Yang lebih penting adalah apakah keterlacakan inferensi dapat ditelusuri, apakah kecukupan bukti diperiksa, dan apakah konsistensi istilah dijelaskan. Peneliti harus menghindari mencampur paradigma tanpa alasan; jika risiko itu tidak dapat dihilangkan, dampaknya perlu dianalisis dan dicantumkan sebagai batas penelitian.

Pada akhirnya, keputusan tentang pragmatisme sebagai dasar pemilihan metode menentukan apakah penelitian menghasilkan daftar jawaban salah atau pemahaman yang dapat ditindaklanjuti mengenai cara peserta didik berpikir. Sasaran utamanya ialah bukti yang cukup untuk merancang diagnosis, pembelajaran, dan penelitian lanjutan secara bertanggung jawab.

7.1.8 Realisme Kritis dan Mekanisme Pembentukan Miskonsepsi

Subbagian ini menempatkan realisme kritis dan mekanisme pembentukan miskonsepsi sebagai simpul penting dalam paradigma penelitian miskonsepsi. Fokus ini berada dalam kerangka filosofis yang menentukan apa yang dianggap sebagai realitas konseptual, bagaimana bukti mengenai miskonsepsi diperoleh, dan sejauh mana peneliti boleh membuat klaim tentang pikiran peserta didik. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Messick (1995) serta AERA, APA, dan NCME (2014) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Secara operasional, peneliti perlu menetapkan status pengetahuan ilmiah sebagai rujukan, menjelaskan derajat keterlibatan peneliti, dan mendokumentasikan posisi ontologis terhadap konsepsi peserta didik. Keputusan tersebut kemudian diterjemahkan ke dalam pilihan wawancara klinis, rekaman diskusi kelas, serta jawaban diagnostik. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Sebagai ilustrasi, pada topik kesetimbangan kimia, peserta didik dapat memberikan jawaban benar pada tingkat simbolik tetapi alasan submikroskopiknya keliru. Peneliti tidak cukup menghitung benar-salah; ia perlu membandingkan wawancara klinis dengan rekaman diskusi kelas, memeriksa tingkat keyakinan, dan menguji konsistensi pada konteks lain. Hasil analisis mungkin menunjukkan bahwa kesalahan hanya muncul pada satu representasi, atau justru merupakan hubungan konseptual yang stabil.

Bukti minimum yang perlu tersedia:

- Definisikan status pengetahuan ilmiah sebagai rujukan dalam istilah operasional.
- Tentukan bukti utama dari gambar partikel dan bukti pendukung dari jejak perubahan keyakinan.
- Periksa kecukupan bukti sebelum membuat kategori atau model.
- Catat potensi mengklaim kausalitas dari data deskriptif beserta strategi mitigasinya.
- Sajikan peta hubungan data dan klaim bersama batas interpretasi dan ketidakpastian.

Mutu inferensi terutama ditentukan oleh reflektivitas, kejelasan batas klaim, dan koherensi paradigma dan pertanyaan penelitian. Ancaman utama ialah mengabaikan konteks tugas serta melabeli peserta didik secara esensial. Karena itu, laporan perlu menyajikan bukti yang mendukung sekaligus bukti yang membatasi interpretasi. Prinsip ini konsisten dengan penekanan Messick (1995) dan AERA, APA, dan NCME (2014) bahwa validitas melekat pada interpretasi dan penggunaan hasil, bukan sekadar pada nama instrumen atau desain.

Dengan demikian, realisme kritis dan mekanisme pembentukan miskonsepsi berfungsi menghubungkan keputusan desain dengan interpretasi. Produk yang diharapkan bukan hanya definisi operasional miskonsepsi, tetapi juga penjelasan mengenai bagaimana produk tersebut diperoleh dan dalam kondisi apa ia dapat digunakan. Penelitian miskonsepsi yang baik selalu menyisakan ruang bagi penjelasan alternatif dan verifikasi lanjutan.

7.1.9 Paradigma Transformasional dan Keadilan Pendidikan

Pembahasan mengenai paradigma transformasional dan keadilan pendidikan diperlukan agar paradigma penelitian miskonsepsi tidak berhenti pada prosedur administratif. Fokus ini berada dalam kerangka filosofis yang menentukan apa yang dianggap sebagai realitas konseptual, bagaimana bukti mengenai miskonsepsi diperoleh, dan sejauh mana peneliti boleh membuat klaim tentang pikiran peserta didik. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Treagust (1988) serta Nakhleh (1992) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Penerapan praktis dapat disusun dalam urutan: mendefinisikan peran konteks sosial dan bahasa, mengoperasionalkan bentuk klaim yang dapat dipertanggungjawabkan, mengumpulkan gambar partikel, memeriksa kecukupan bukti, lalu melaporkan batas generalisasi. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Misalnya, pada ikatan kimia, perubahan skor sesudah pembelajaran belum otomatis menandakan perubahan konsep. Peserta didik mungkin mempelajari pola jawaban tes. Bukti yang lebih meyakinkan mencakup transfer ke situasi baru, penjelasan kausal, representasi partikel, dan retensi setelah jeda waktu.

Kriteria pemeriksaan sebelum analisis:

- Definisikan peran konteks sosial dan bahasa dalam istilah operasional.
- Tentukan bukti utama dari rekaman diskusi kelas dan bukti pendukung dari jawaban diagnostik.
- Periksa kejelasan batas klaim sebelum membuat kategori atau model.
- Catat potensi melabeli peserta didik secara esensial beserta strategi mitigasinya.
- Sajikan catatan reflektivitas bersama batas interpretasi dan ketidakpastian.

Temuan yang baik tidak selalu berarti hasil yang besar atau sesuai harapan. Temuan bernilai ketika kecukupan bukti terjaga, konsistensi istilah dijelaskan, dan ketidakpastian disajikan. Kerangka Treagust (1988) serta Nakhleh (1992) menolong peneliti membedakan kekuatan bukti dari daya tarik narasi.

Inti metodologis subbagian ini adalah keterlacakan. Pembaca harus dapat mengikuti perjalanan dari pertanyaan, data, pengodean, analisis, sampai batas generalisasi. Ketika jejak itu jelas, temuan dapat digunakan untuk memperbaiki pembelajaran tanpa mengubah diagnosis konseptual menjadi label bagi peserta didik.

7.1.10 Posisi serta Peran Peneliti

Dalam metodologi penelitian miskonsepsi kimia, posisi serta peran peneliti menentukan mutu hubungan antara data dan kesimpulan. Fokus ini berada dalam kerangka filosofis yang menentukan apa yang dianggap sebagai realitas konseptual, bagaimana bukti mengenai miskonsepsi diperoleh, dan sejauh mana peneliti boleh membuat klaim tentang pikiran peserta didik. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Nakhleh (1992) serta AERA, APA, dan NCME (2014) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Prosedur minimal mencakup definisi derajat keterlibatan peneliti, alasan memilih rekaman diskusi kelas, aturan mengubah respons menjadi kategori, serta mekanisme pemeriksaan kejelasan batas klaim. Semua penyimpangan dari rencana awal perlu dicatat. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Pada penelitian mengenai asam-basa, peneliti dapat menyusun beberapa tugas ekuivalen yang berubah konteks, representasi, atau tuntutan penalarannya. Jika pola respons tetap berulang dan disertai keyakinan tinggi, terdapat dasar yang lebih kuat untuk menyebut indikasi miskonsepsi. Jika pola berubah-ubah, interpretasi yang lebih hati-hati adalah pengetahuan parsial atau ketergantungan konteks.

Langkah operasional yang disarankan:

- Definisikan derajat keterlibatan peneliti dalam istilah operasional.
- Tentukan bukti utama dari jejak perubahan keyakinan dan bukti pendukung dari alasan tertulis.
- Periksa konsistensi istilah sebelum membuat kategori atau model.
- Catat potensi menganggap respons tunggal sebagai cermin langsung pikiran beserta strategi mitigasinya.
- Sajikan pernyataan posisi filosofis bersama batas interpretasi dan ketidakpastian.

Pemeriksaan mutu hendaknya dilakukan selama penelitian, bukan hanya pada akhir. Catatan tentang kejelasan batas klaim, keputusan terkait koherensi paradigma dan pertanyaan penelitian, dan kejadian yang berpotensi menimbulkan mengklaim kausalitas dari data deskriptif perlu dipelihara dalam *audit trail*. Hal ini memungkinkan koreksi dini dan pelaporan yang jujur.

Kesimpulannya, posisi serta peran peneliti perlu dirancang sebagai proses argumentatif. Data tidak berbicara sendiri; maknanya bergantung pada definisi konstruk, kualitas prosedur, dan konteks. Oleh sebab itu, klaim akhir sebaiknya menggunakan bahasa yang sesuai dengan kekuatan bukti, misalnya “menunjukkan pola”, “mengindikasikan”, atau “mendukung penjelasan”.

7.1.11 Unit Analisis dan Tingkat Realitas

Keputusan tentang unit analisis dan tingkat realitas harus dibuat sebelum data utama dikumpulkan, bukan ditambahkan setelah hasil diketahui. Fokus ini berada dalam kerangka filosofis yang menentukan apa yang dianggap sebagai realitas konseptual, bagaimana bukti mengenai miskonsepsi diperoleh, dan sejauh mana peneliti boleh membuat klaim tentang pikiran peserta didik. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Talanquer (2011) serta Nakhleh (1992) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Tahap kerja dimulai dengan penegasan bentuk klaim yang dapat dipertanggungjawabkan. Setelah itu, peneliti menyusun aturan untuk hubungan antara respons teramati dan struktur kognitif dan menentukan bagaimana peran konteks sosial dan bahasa akan diperiksa. Bukti

sebaiknya berasal dari lebih dari satu bentuk, misalnya jejak perubahan keyakinan dan alasan tertulis. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Pada topik redoks, peneliti juga perlu memperhatikan bahasa. Istilah yang sama dapat diberi makna berbeda oleh peserta didik. Klarifikasi melalui pertanyaan tindak lanjut, gambar, atau contoh tandingan membantu membedakan masalah terminologi dari kesalahan konseptual.

Keputusan yang harus didokumentasikan:

- Definisikan bentuk klaim yang dapat dipertanggungjawabkan dalam istilah operasional.
- Tentukan bukti utama dari jawaban diagnostik dan bukti pendukung dari wawancara klinis.
- Periksa koherensi paradigma dan pertanyaan penelitian sebelum membuat kategori atau model.
- Catat potensi mencampur paradigma tanpa alasan beserta strategi mitigasinya.
- Sajikan model inferensi konseptual bersama batas interpretasi dan ketidakpastian.

Kualitas penelitian meningkat ketika peneliti menunjukkan rantai bukti dari jejak perubahan keyakinan menuju catatan refleksivitas, menyatakan asumsi, dan memeriksa alternatif penjelasan. Sebaliknya, melabeli peserta didik secara esensial dapat menghasilkan klaim yang tampak tegas tetapi tidak stabil. Rekomendasi praktis harus proporsional dengan bukti.

Pada akhirnya, keputusan tentang unit analisis dan tingkat realitas menentukan apakah penelitian menghasilkan daftar jawaban salah atau pemahaman yang dapat ditindaklanjuti mengenai cara peserta didik berpikir. Sasaran utamanya ialah bukti yang cukup untuk merancang diagnosis, pembelajaran, dan penelitian lanjutan secara bertanggung jawab.

7.1.12 Bahasa Klaim: Indikasi, Pola, dan Penjelasan

Bahasa klaim: indikasi, pola, dan penjelasan perlu dipahami sebagai keputusan metodologis yang mempunyai konsekuensi pada desain, analisis, dan pelaporan. Fokus ini berada dalam kerangka filosofis yang menentukan apa yang dianggap sebagai realitas konseptual, bagaimana bukti mengenai miskonsepsi diperoleh, dan sejauh mana peneliti boleh membuat klaim tentang pikiran peserta didik. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Creswell dan Creswell (2018) serta AERA, APA, dan NCME (2014) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Peneliti sebaiknya menyiapkan matriks yang menghubungkan pertanyaan penelitian dengan posisi ontologis terhadap konsepsi peserta didik, jawaban diagnostik, teknik analisis, dan pernyataan posisi filosofis. Matriks ini mencegah pengumpulan data yang banyak tetapi tidak relevan. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Ilustrasi termokimia juga menunjukkan bahwa kualitas instrumen memengaruhi temuan. Distraktor yang tidak berasal dari pemikiran peserta didik dapat menghasilkan pilihan acak, sedangkan pertanyaan yang terlalu menuntut matematika dapat mengukur keterampilan hitung alih-alih konsepsi. Peneliti perlu menilai kesesuaian bukti dengan konstruk yang dimaksud.

Pertanyaan audit metodologis:

- Definisikan posisi ontologis terhadap konsepsi peserta didik dalam istilah operasional.
- Tentukan bukti utama dari alasan tertulis dan bukti pendukung dari gambar partikel.
- Periksa keterlacakan inferensi sebelum membuat kategori atau model.
- Catat potensi mengabaikan konteks tugas beserta strategi mitigasinya.
- Sajikan definisi operasional miskonsepsi bersama batas interpretasi dan ketidakpastian.

Peneliti sebaiknya memisahkan keputusan apriori dari keputusan eksploratori. Pemisahan tersebut membantu mengendalikan menganggap respons tunggal sebagai cermin langsung pikiran dan mencegah menyederhanakan konstruktivisme menjadi relativisme. Selanjutnya, koherensi paradigma dan pertanyaan penelitian dan reflektivitas dilaporkan bersama ketidakpastian, sehingga pembaca dapat menilai kekuatan temuan tanpa bergantung pada kesimpulan penulis semata.

Dengan demikian, bahasa klaim: indikasi, pola, dan penjelasan berfungsi menghubungkan keputusan desain dengan interpretasi. Produk yang diharapkan bukan hanya pernyataan posisi filosofis, tetapi juga penjelasan mengenai bagaimana produk tersebut diperoleh dan dalam kondisi apa ia dapat digunakan. Penelitian miskonsepsi yang baik selalu menyisakan ruang bagi penjelasan alternatif dan verifikasi lanjutan.

7.1.13 Keselarasan Paradigma, Pertanyaan, dan Desain

Ketelitian penelitian pada bagian keselarasan paradigma, pertanyaan, dan desain sangat menentukan apakah pola jawaban dapat ditafsirkan sebagai bukti konseptual. Fokus ini berada dalam kerangka filosofis yang menentukan apa yang dianggap sebagai realitas konseptual, bagaimana bukti mengenai miskonsepsi diperoleh, dan sejauh mana peneliti boleh membuat klaim tentang pikiran peserta didik. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah

membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. AERA, APA, dan NCME (2014) serta Nakhleh (1992) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Dokumentasi metodologis hendaknya memuat siapa yang melakukan alasan tertulis, kapan data diperoleh, bagaimana peran konteks sosial dan bahasa dikendalikan, dan atas dasar apa model inferensi konseptual disusun. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Contoh pada kajian struktur atom memperlihatkan pentingnya pemisahan antara ketidaktahuan, kesalahan prosedural, dan miskonsepsi. Seorang peserta dapat gagal karena lupa rumus, sedangkan peserta lain memilih jawaban yang sama karena menggunakan model mental yang tidak ilmiah. Perbedaan ini baru terlihat ketika alasan tertulis dibaca bersama gambar partikel dan bukti tentang alasan.

Bukti minimum yang perlu tersedia:

- Definisikan hubungan antara respons teramati dan struktur kognitif dalam istilah operasional.
- Tentukan bukti utama dari wawancara klinis dan bukti pendukung dari rekaman diskusi kelas.
- Periksa reflektivitas sebelum membuat kategori atau model.
- Catat potensi menyederhanakan konstruktivisme menjadi relativisme beserta strategi mitigasinya.
- Sajikan batas generalisasi bersama batas interpretasi dan ketidakpastian.

Standar metodologis menuntut koherensi antara konstruk, prosedur, analisis, dan klaim. Dalam konteks ini, keterlacakan inferensi tidak dapat digantikan oleh indeks tunggal. Demikian pula, kecukupan bukti harus dinilai dengan bukti yang sesuai. AERA, APA, dan NCME (2014) dan Nakhleh (1992) memberikan dasar untuk pelaporan yang lengkap dan dapat ditelaah.

Inti metodologis subbagian ini adalah keterlacakan. Pembaca harus dapat mengikuti perjalanan dari pertanyaan, data, pengodean, analisis, sampai model inferensi konseptual. Ketika jejak itu jelas, temuan dapat digunakan untuk memperbaiki pembelajaran tanpa mengubah diagnosis konseptual menjadi label bagi peserta didik.

7.1.14 Refleksivitas serta Audit Asumsi

Peneliti sering memusatkan perhatian pada teknik, padahal refleksivitas serta audit asumsi justru mengatur batas inferensi yang dapat dibuat. Fokus ini berada dalam kerangka filosofis yang menentukan apa yang dianggap sebagai realitas konseptual, bagaimana bukti mengenai miskonsepsi diperoleh, dan sejauh mana peneliti boleh membuat klaim tentang pikiran

peserta didik. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Messick (1995) serta AERA, APA, dan NCME (2014) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Rancangan yang dapat diaudit mencantumkan alasan pemilihan status pengetahuan ilmiah sebagai rujukan, prosedur penanganan derajat keterlibatan peneliti, dan kriteria keputusan untuk posisi ontologis terhadap konsepsi peserta didik. Tanpa penjelasan itu, pembaca sulit menilai apakah wawancara klinis benar-benar mendukung klaim. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Kasus stoikiometri dapat digunakan untuk memperlihatkan hubungan antara pola individual dan pola kelas. Sebagian miskonsepsi mungkin hanya dimiliki beberapa peserta tetapi sangat kuat, sedangkan pola lain tersebar luas karena bersumber dari analogi atau buku teks yang sama. Analisis perlu menjaga kedua tingkat tersebut.

Kriteria pemeriksaan sebelum analisis:

- Definisikan status pengetahuan ilmiah sebagai rujukan dalam istilah operasional.
- Tentukan bukti utama dari gambar partikel dan bukti pendukung dari jejak perubahan keyakinan.
- Periksa kecukupan bukti sebelum membuat kategori atau model.
- Catat potensi mengklaim kausalitas dari data deskriptif beserta strategi mitigasinya.
- Sajikan peta hubungan data dan klaim bersama batas interpretasi dan ketidakpastian.

Pemeriksaan kualitas dapat dilakukan melalui audit terhadap reflektivitas, analisis kasus yang bertentangan, dan evaluasi kejelasan batas klaim. Apabila ditemukan mengabaikan konteks tugas, peneliti perlu menurunkan kekuatan klaim atau menambah bukti. Transparansi seperti ini meningkatkan kegunaan temuan bagi penelitian lanjutan dan praktik kelas.

Kesimpulannya, reflektivitas serta audit asumsi perlu dirancang sebagai proses argumentatif. Data tidak berbicara sendiri; maknanya bergantung pada definisi konstruk, kualitas prosedur, dan konteks. Oleh sebab itu, klaim akhir sebaiknya menggunakan bahasa yang sesuai dengan kekuatan bukti, misalnya “menunjukkan pola”, “mengindikasikan”, atau “mendukung penjelasan”.

7.1.15 Sintesis Kerangka Paradigmatik Penelitian Miskonsepsi

Kajian yang kuat memperlakukan sintesis kerangka paradigmatik penelitian miskonsepsi sebagai bagian dari rantai argumentasi, bukan sebagai unsur pelengkap. Fokus ini berada

dalam kerangka filosofis yang menentukan apa yang dianggap sebagai realitas konseptual, bagaimana bukti mengenai miskonsepsi diperoleh, dan sejauh mana peneliti boleh membuat klaim tentang pikiran peserta didik. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Treagust (1988) serta Nakhleh (1992) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Pada tahap perencanaan, peran konteks sosial dan bahasa dan bentuk klaim yang dapat dipertanggungjawabkan harus dibedakan secara eksplisit. Pada tahap pelaksanaan, konsistensi gambar partikel perlu diperiksa; pada tahap interpretasi, kecukupan bukti menjadi dasar pembatasan klaim. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Dalam kelas yang mempelajari sifat koligatif, data agregat mungkin menunjukkan mayoritas jawaban salah. Akan tetapi, wawancara dapat memperlihatkan beberapa mekanisme berbeda: bahasa sehari-hari, generalisasi aturan, salah membaca simbol, atau model partikel yang tidak tepat. Karena itu, satu persentase tidak boleh menggantikan analisis struktur alasan.

Langkah operasional yang disarankan:

- Definisikan peran konteks sosial dan bahasa dalam istilah operasional.
- Tentukan bukti utama dari rekaman diskusi kelas dan bukti pendukung dari jawaban diagnostik.
- Periksa kejelasan batas klaim sebelum membuat kategori atau model.
- Catat potensi melabeli peserta didik secara esensial beserta strategi mitigasinya.
- Sajikan catatan reflektivitas bersama batas interpretasi dan ketidakpastian.

Kriteria keberhasilan tidak cukup berupa banyaknya data. Yang lebih penting adalah apakah kecukupan bukti dapat ditelusuri, apakah konsistensi istilah diperiksa, dan apakah keterlacakan inferensi dijelaskan. Peneliti harus menghindari menyederhanakan konstruktivisme menjadi relativisme; jika risiko itu tidak dapat dihilangkan, dampaknya perlu dianalisis dan dicantumkan sebagai batas penelitian.

Pada akhirnya, keputusan tentang sintesis kerangka paradigmatis penelitian miskonsepsi menentukan apakah penelitian menghasilkan daftar jawaban salah atau pemahaman yang dapat ditindaklanjuti mengenai cara peserta didik berpikir. Sasaran utamanya ialah bukti yang cukup untuk merancang diagnosis, pembelajaran, dan penelitian lanjutan secara bertanggung jawab.

7.2 Penelitian Deskriptif

Subbab ini membahas desain yang memetakan prevalensi, distribusi, bentuk, keyakinan, dan konteks kemunculan miskonsepsi tanpa memanipulasi kondisi pembelajaran. Fokus utamanya adalah menjaga keselarasan antara pertanyaan penelitian, bentuk bukti, teknik analisis, dan kekuatan klaim. Dalam konteks miskonsepsi kimia, keputusan metodologis harus mampu membedakan pola konseptual dari respons yang timbul karena keterbatasan bahasa, kesalahan hitung, atau kondisi administrasi. Pembahasan berikut dipecah menjadi lima belas bagian agar setiap keputusan dapat ditelaah secara operasional.

7.2.1 Tujuan dan Ruang Lingkup Studi Deskriptif

Subbagian ini menempatkan tujuan dan ruang lingkup studi deskriptif sebagai simpul penting dalam penelitian deskriptif. Fokus ini berada dalam desain yang memetakan prevalensi, distribusi, bentuk, keyakinan, dan konteks kemunculan miskonsepsi tanpa memanipulasi kondisi pembelajaran. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Hasan et al. (1999) serta von Elm et al. (2007) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Secara operasional, peneliti perlu menetapkan pemilihan pembanding, menjelaskan definisi populasi target, dan mendokumentasikan kategori konsepsi. Keputusan tersebut kemudian diterjemahkan ke dalam pilihan profil demografis akademik, tes diagnostik bertingkat, serta skala keyakinan. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Sebagai ilustrasi, pada topik konsep mol, peserta didik dapat memberikan jawaban benar pada tingkat simbolik tetapi alasan submikroskopiknya keliru. Peneliti tidak cukup menghitung benar-salah; ia perlu membandingkan profil demografis akademik dengan tes diagnostik bertingkat, memeriksa tingkat keyakinan, dan menguji konsistensi pada konteks lain. Hasil analisis mungkin menunjukkan bahwa kesalahan hanya muncul pada satu representasi, atau justru merupakan hubungan konseptual yang stabil.

Keputusan yang harus didokumentasikan:

- Definisikan pemilihan pembanding dalam istilah operasional.
- Tentukan bukti utama dari catatan konteks pembelajaran dan bukti pendukung dari respons terbuka.
- Periksa keterbandingan konteks sebelum membuat kategori atau model.

- Catat potensi menyamakan persentase salah dengan miskonsepsi beserta strategi mitigasinya.
- Sajikan tabel prevalensi bersama batas interpretasi dan ketidakpastian.

Mutu inferensi terutama ditentukan oleh transparansi denominator, representativitas, dan presisi estimasi. Ancaman utama ialah menafsirkan korelasi sebagai sebab serta mengabaikan missing data. Karena itu, laporan perlu menyajikan bukti yang mendukung sekaligus bukti yang membatasi interpretasi. Prinsip ini konsisten dengan penekanan Hasan et al. (1999) dan von Elm et al. (2007) bahwa validitas melekat pada interpretasi dan penggunaan hasil, bukan sekadar pada nama instrumen atau desain.

Dengan demikian, tujuan dan ruang lingkup studi deskriptif berfungsi menghubungkan keputusan desain dengan interpretasi. Produk yang diharapkan bukan hanya perbandingan subkelompok, tetapi juga penjelasan mengenai bagaimana produk tersebut diperoleh dan dalam kondisi apa ia dapat digunakan. Penelitian miskonsepsi yang baik selalu menyisakan ruang bagi penjelasan alternatif dan verifikasi lanjutan.

7.2.2 Desain Potong Lintang

Pembahasan mengenai desain potong lintang diperlukan agar penelitian deskriptif tidak berhenti pada prosedur administratif. Fokus ini berada dalam desain yang memetakan prevalensi, distribusi, bentuk, keyakinan, dan konteks kemunculan miskonsepsi tanpa memanipulasi kondisi pembelajaran. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Appelbaum et al. (2018) serta Hasan et al. (1999) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Penerapan praktis dapat disusun dalam urutan: mendefinisikan cara menyatakan prevalensi, mengoperasionalkan waktu pengukuran, mengumpulkan catatan konteks pembelajaran, memeriksa keterbandingan konteks, lalu melaporkan narasi pola dominan. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Misalnya, pada hukum gas, perubahan skor sesudah pembelajaran belum otomatis menandakan perubahan konsep. Peserta didik mungkin mempelajari pola jawaban tes. Bukti yang lebih meyakinkan mencakup transfer ke situasi baru, penjelasan kausal, representasi partikel, dan retensi setelah jeda waktu.

Pertanyaan audit metodologis:

- Definisikan cara menyatakan prevalensi dalam istilah operasional.

- Tentukan bukti utama dari tes diagnostik bertingkat dan bukti pendukung dari skala keyakinan.
- Periksa representativitas sebelum membuat kategori atau model.
- Catat potensi mengabaikan missing data beserta strategi mitigasinya.
- Sajikan profil kategori konsepsi bersama batas interpretasi dan ketidakpastian.

Temuan yang baik tidak selalu berarti hasil yang besar atau sesuai harapan. Temuan bernilai ketika keterbandingan konteks terjaga, ketepatan klasifikasi dijelaskan, dan ketidakpastian disajikan. Kerangka Appelbaum et al. (2018) serta Hasan et al. (1999) menolong peneliti membedakan kekuatan bukti dari daya tarik narasi.

Inti metodologis subbagian ini adalah keterlacakan. Pembaca harus dapat mengikuti perjalanan dari pertanyaan, data, pengodean, analisis, sampai narasi pola dominan. Ketika jejak itu jelas, temuan dapat digunakan untuk memperbaiki pembelajaran tanpa mengubah diagnosis konseptual menjadi label bagi peserta didik.

7.2.3 Deskripsi Prevalensi Miskonsepsi

Dalam metodologi penelitian miskonsepsi kimia, deskripsi prevalensi miskonsepsi menentukan mutu hubungan antara data dan kesimpulan. Fokus ini berada dalam desain yang memetakan prevalensi, distribusi, bentuk, keyakinan, dan konteks kemunculan miskonsepsi tanpa memanipulasi kondisi pembelajaran. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Creswell dan Creswell (2018) serta von Elm et al. (2007) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Prosedur minimal mencakup definisi definisi populasi target, alasan memilih tes diagnostik bertingkat, aturan mengubah respons menjadi kategori, serta mekanisme pemeriksaan representativitas. Semua penyimpangan dari rencana awal perlu dicatat. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

$$\hat{p} = x/n; SE(\hat{p}) = \sqrt{[\hat{p}(1 - \hat{p})/n]}$$

Pada penelitian mengenai laju reaksi, peneliti dapat menyusun beberapa tugas ekuivalen yang berubah konteks, representasi, atau tuntutan penalarannya. Jika pola respons tetap berulang dan disertai keyakinan tinggi, terdapat dasar yang lebih kuat untuk menyebut indikasi miskonsepsi. Jika pola berubah-ubah, interpretasi yang lebih hati-hati adalah pengetahuan parsial atau ketergantungan konteks.

Bukti minimum yang perlu tersedia:

- Definisikan definisi populasi target dalam istilah operasional.
- Tentukan bukti utama dari respons terbuka dan bukti pendukung dari gambar atau model.
- Periksa ketepatan klasifikasi sebelum membuat kategori atau model.
- Catat potensi membuat ranking tanpa interval ketidakpastian beserta strategi mitigasinya.
- Sajikan interval kepercayaan bersama batas interpretasi dan ketidakpastian.

Pemeriksaan mutu hendaknya dilakukan selama penelitian, bukan hanya pada akhir. Catatan tentang representativitas, keputusan terkait presisi estimasi, dan kejadian yang berpotensi menimbulkan menyamakan persentase salah dengan miskonsepsi perlu dipelihara dalam *audit trail*. Hal ini memungkinkan koreksi dini dan pelaporan yang jujur.

Kesimpulannya, deskripsi prevalensi miskonsepsi perlu dirancang sebagai proses argumentatif. Data tidak berbicara sendiri; maknanya bergantung pada definisi konstruk, kualitas prosedur, dan konteks. Oleh sebab itu, klaim akhir sebaiknya menggunakan bahasa yang sesuai dengan kekuatan bukti, misalnya “menunjukkan pola”, “mengindikasikan”, atau “mendukung penjelasan”.

7.2.4 Penyusunan Kategori Konsepsi

Keputusan tentang penyusunan kategori konsepsi harus dibuat sebelum data utama dikumpulkan, bukan ditambahkan setelah hasil diketahui. Fokus ini berada dalam desain yang memetakan prevalensi, distribusi, bentuk, keyakinan, dan konteks kemunculan miskonsepsi tanpa memanipulasi kondisi pembelajaran. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. von Elm et al. (2007) serta Hasan et al. (1999) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Tahap kerja dimulai dengan penegasan waktu pengukuran. Setelah itu, peneliti menyusun aturan untuk tingkat analisis butir dan peserta dan menentukan bagaimana cara menyatakan prevalensi akan diperiksa. Bukti sebaiknya berasal dari lebih dari satu bentuk, misalnya respons terbuka dan gambar atau model. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Pada topik kesetimbangan, peneliti juga perlu memperhatikan bahasa. Istilah yang sama dapat diberi makna berbeda oleh peserta didik. Klarifikasi melalui pertanyaan tindak lanjut, gambar,

atau contoh tandingan membantu membedakan masalah terminologi dari kesalahan konseptual.

Kriteria pemeriksaan sebelum analisis:

- Definisikan waktu pengukuran dalam istilah operasional.
- Tentukan bukti utama dari skala keyakinan dan bukti pendukung dari profil demografis akademik.
- Periksa presisi estimasi sebelum membuat kategori atau model.
- Catat potensi menggabungkan kategori yang berbeda beserta strategi mitigasinya.
- Sajikan peta butir sulit bersama batas interpretasi dan ketidakpastian.

Kualitas penelitian meningkat ketika peneliti menunjukkan rantai bukti dari respons terbuka menuju profil kategori konsepsi, menyatakan asumsi, dan memeriksa alternatif penjelasan. Sebaliknya, mengabaikan missing data dapat menghasilkan klaim yang tampak tegas tetapi tidak stabil. Rekomendasi praktis harus proporsional dengan bukti.

Pada akhirnya, keputusan tentang penyusunan kategori konsepsi menentukan apakah penelitian menghasilkan daftar jawaban salah atau pemahaman yang dapat ditindaklanjuti mengenai cara peserta didik berpikir. Sasaran utamanya ialah bukti yang cukup untuk merancang diagnosis, pembelajaran, dan penelitian lanjutan secara bertanggung jawab.

7.2.5 Analisis Profil Individu dan Kelas

Analisis profil individu dan kelas perlu dipahami sebagai keputusan metodologis yang mempunyai konsekuensi pada desain, analisis, dan pelaporan. Fokus ini berada dalam desain yang memetakan prevalensi, distribusi, bentuk, keyakinan, dan konteks kemunculan miskonsepsi tanpa memanipulasi kondisi pembelajaran. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Treagust (1988) serta von Elm et al. (2007) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Peneliti sebaiknya menyiapkan matriks yang menghubungkan pertanyaan penelitian dengan kategori konsepsi, skala keyakinan, teknik analisis, dan interval kepercayaan. Matriks ini mencegah pengumpulan data yang banyak tetapi tidak relevan. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Ilustrasi ikatan ionik juga menunjukkan bahwa kualitas instrumen memengaruhi temuan. Distraktor yang tidak berasal dari pemikiran peserta didik dapat menghasilkan pilihan acak, sedangkan pertanyaan yang terlalu menuntut matematika dapat mengukur keterampilan

hitung alih-alih konsepsi. Peneliti perlu menilai kesesuaian bukti dengan konstruk yang dimaksud.

Langkah operasional yang disarankan:

- Definisikan kategori konsepsi dalam istilah operasional.
- Tentukan bukti utama dari gambar atau model dan bukti pendukung dari catatan konteks pembelajaran.
- Periksa konsistensi pengodean sebelum membuat kategori atau model.
- Catat potensi menafsirkan korelasi sebagai sebab beserta strategi mitigasinya.
- Sajikan perbandingan subkelompok bersama batas interpretasi dan ketidakpastian.

Peneliti sebaiknya memisahkan keputusan apriori dari keputusan eksploratori. Pemisahan tersebut membantu mengendalikan membuat ranking tanpa interval ketidakpastian dan mencegah menghilangkan variasi antarbutir. Selanjutnya, presisi estimasi dan transparansi denominator dilaporkan bersama ketidakpastian, sehingga pembaca dapat menilai kekuatan temuan tanpa bergantung pada kesimpulan penulis semata.

Dengan demikian, analisis profil individu dan kelas berfungsi menghubungkan keputusan desain dengan interpretasi. Produk yang diharapkan bukan hanya interval kepercayaan, tetapi juga penjelasan mengenai bagaimana produk tersebut diperoleh dan dalam kondisi apa ia dapat digunakan. Penelitian miskonsepsi yang baik selalu menyisakan ruang bagi penjelasan alternatif dan verifikasi lanjutan.

7.2.6 Analisis Tingkat Butir

Ketelitian penelitian pada bagian analisis tingkat butir sangat menentukan apakah pola jawaban dapat ditafsirkan sebagai bukti konseptual. Fokus ini berada dalam desain yang memetakan prevalensi, distribusi, bentuk, keyakinan, dan konteks kemunculan miskonsepsi tanpa memanipulasi kondisi pembelajaran. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Kaltakci Gurel et al. (2015) serta Hasan et al. (1999) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Dokumentasi metodologis hendaknya memuat siapa yang melakukan gambar atau model, kapan data diperoleh, bagaimana cara menyatakan prevalensi dikendalikan, dan atas dasar apa peta butir sulit disusun. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Contoh pada kajian elektrolit memperlihatkan pentingnya pemisahan antara ketidaktahuan, kesalahan prosedural, dan miskonsepsi. Seorang peserta dapat gagal karena lupa rumus, sedangkan peserta lain memilih jawaban yang sama karena menggunakan model mental yang tidak ilmiah. Perbedaan ini baru terlihat ketika gambar atau model dibaca bersama catatan konteks pembelajaran dan bukti tentang alasan.

Keputusan yang harus didokumentasikan:

- Definisikan tingkat analisis butir dan peserta dalam istilah operasional.
- Tentukan bukti utama dari profil demografis akademik dan bukti pendukung dari tes diagnostik bertingkat.
- Periksa transparansi denominator sebelum membuat kategori atau model.
- Catat potensi menghilangkan variasi antarbutir beserta strategi mitigasinya.
- Sajikan narasi pola dominan bersama batas interpretasi dan ketidakpastian.

Standar metodologis menuntut koherensi antara konstruk, prosedur, analisis, dan klaim. Dalam konteks ini, konsistensi pengodean tidak dapat digantikan oleh indeks tunggal. Demikian pula, keterbandingan konteks harus dinilai dengan bukti yang sesuai. Kaltakci Gurel et al. (2015) dan Hasan et al. (1999) memberikan dasar untuk pelaporan yang lengkap dan dapat ditelaah.

Inti metodologis subbagian ini adalah keterlacakan. Pembaca harus dapat mengikuti perjalanan dari pertanyaan, data, pengodean, analisis, sampai peta butir sulit. Ketika jejak itu jelas, temuan dapat digunakan untuk memperbaiki pembelajaran tanpa mengubah diagnosis konseptual menjadi label bagi peserta didik.

7.2.7 Deskripsi Berdasarkan Tingkat Keyakinan

Peneliti sering memusatkan perhatian pada teknik, padahal deskripsi berdasarkan tingkat keyakinan justru mengatur batas inferensi yang dapat dibuat. Fokus ini berada dalam desain yang memetakan prevalensi, distribusi, bentuk, keyakinan, dan konteks kemunculan miskonsepsi tanpa memanipulasi kondisi pembelajaran. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Hasan et al. (1999) serta von Elm et al. (2007) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Rancangan yang dapat diaudit mencantumkan alasan pemilihan pemilihan pembanding, prosedur penanganan definisi populasi target, dan kriteria keputusan untuk kategori konsepsi. Tanpa penjelasan itu, pembaca sulit menilai apakah profil demografis akademik benar-benar mendukung klaim. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu

peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Kasus hidrolisis garam dapat digunakan untuk memperlihatkan hubungan antara pola individual dan pola kelas. Sebagian miskonsepsi mungkin hanya dimiliki beberapa peserta tetapi sangat kuat, sedangkan pola lain tersebar luas karena bersumber dari analogi atau buku teks yang sama. Analisis perlu menjaga kedua tingkat tersebut.

Pertanyaan audit metodologis:

- Definisikan pemilihan pembanding dalam istilah operasional.
- Tentukan bukti utama dari catatan konteks pembelajaran dan bukti pendukung dari respons terbuka.
- Periksa keterbandingan konteks sebelum membuat kategori atau model.
- Catat potensi menyamakan persentase salah dengan miskonsepsi beserta strategi mitigasinya.
- Sajikan tabel prevalensi bersama batas interpretasi dan ketidakpastian.

Pemeriksaan kualitas dapat dilakukan melalui audit terhadap transparansi denominator, analisis kasus yang bertentangan, dan evaluasi representativitas. Apabila ditemukan menafsirkan korelasi sebagai sebab, peneliti perlu menurunkan kekuatan klaim atau menambah bukti. Transparansi seperti ini meningkatkan kegunaan temuan bagi penelitian lanjutan dan praktik kelas.

Kesimpulannya, deskripsi berdasarkan tingkat keyakinan perlu dirancang sebagai proses argumentatif. Data tidak berbicara sendiri; maknanya bergantung pada definisi konstruk, kualitas prosedur, dan konteks. Oleh sebab itu, klaim akhir sebaiknya menggunakan bahasa yang sesuai dengan kekuatan bukti, misalnya “menunjukkan pola”, “mengindikasikan”, atau “mendukung penjelasan”.

7.2.8 Perbandingan Kelompok Secara Deskriptif

Kajian yang kuat memperlakukan perbandingan kelompok secara deskriptif sebagai bagian dari rantai argumentasi, bukan sebagai unsur pelengkap. Fokus ini berada dalam desain yang memetakan prevalensi, distribusi, bentuk, keyakinan, dan konteks kemunculan miskonsepsi tanpa memanipulasi kondisi pembelajaran. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Appelbaum et al. (2018) serta Hasan et al. (1999) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Pada tahap perencanaan, cara menyatakan prevalensi dan waktu pengukuran harus dibedakan secara eksplisit. Pada tahap pelaksanaan, konsistensi catatan konteks pembelajaran perlu diperiksa; pada tahap interpretasi, keterbandingan konteks menjadi dasar

pembatasan klaim. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Dalam kelas yang mempelajari energi ikatan, data agregat mungkin menunjukkan mayoritas jawaban salah. Akan tetapi, wawancara dapat memperlihatkan beberapa mekanisme berbeda: bahasa sehari-hari, generalisasi aturan, salah membaca simbol, atau model partikel yang tidak tepat. Karena itu, satu persentase tidak boleh menggantikan analisis struktur alasan.

Bukti minimum yang perlu tersedia:

- Definisikan cara menyatakan prevalensi dalam istilah operasional.
- Tentukan bukti utama dari tes diagnostik bertingkat dan bukti pendukung dari skala keyakinan.
- Periksa representativitas sebelum membuat kategori atau model.
- Catat potensi mengabaikan missing data beserta strategi mitigasinya.
- Sajikan profil kategori konsepsi bersama batas interpretasi dan ketidakpastian.

Kriteria keberhasilan tidak cukup berupa banyaknya data. Yang lebih penting adalah apakah keterbandingan konteks dapat ditelusuri, apakah ketepatan klasifikasi diperiksa, dan apakah konsistensi pengodean dijelaskan. Peneliti harus menghindari menghilangkan variasi antarbutir; jika risiko itu tidak dapat dihilangkan, dampaknya perlu dianalisis dan dicantumkan sebagai batas penelitian.

Pada akhirnya, keputusan tentang perbandingan kelompok secara deskriptif menentukan apakah penelitian menghasilkan daftar jawaban salah atau pemahaman yang dapat ditindaklanjuti mengenai cara peserta didik berpikir. Sasaran utamanya ialah bukti yang cukup untuk merancang diagnosis, pembelajaran, dan penelitian lanjutan secara bertanggung jawab.

7.2.9 Deskripsi Lintas Representasi Kimia

Subbagian ini menempatkan deskripsi lintas representasi kimia sebagai simpul penting dalam penelitian deskriptif. Fokus ini berada dalam desain yang memetakan prevalensi, distribusi, bentuk, keyakinan, dan konteks kemunculan miskonsepsi tanpa memanipulasi kondisi pembelajaran. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Creswell dan Creswell (2018) serta von Elm et al. (2007) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Secara operasional, peneliti perlu menetapkan definisi populasi target, menjelaskan kategori konsepsi, dan mendokumentasikan pemilihan pembanding. Keputusan tersebut kemudian diterjemahkan ke dalam pilihan tes diagnostik bertingkat, skala keyakinan, serta profil demografis akademik. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Sebagai ilustrasi, pada topik konsep mol, peserta didik dapat memberikan jawaban benar pada tingkat simbolik tetapi alasan submikroskopiknya keliru. Peneliti tidak cukup menghitung benar-salah; ia perlu membandingkan tes diagnostik bertingkat dengan skala keyakinan, memeriksa tingkat keyakinan, dan menguji konsistensi pada konteks lain. Hasil analisis mungkin menunjukkan bahwa kesalahan hanya muncul pada satu representasi, atau justru merupakan hubungan konseptual yang stabil.

Kriteria pemeriksaan sebelum analisis:

- Definisikan definisi populasi target dalam istilah operasional.
- Tentukan bukti utama dari respons terbuka dan bukti pendukung dari gambar atau model.
- Periksa ketepatan klasifikasi sebelum membuat kategori atau model.
- Catat potensi membuat ranking tanpa interval ketidakpastian beserta strategi mitigasinya.
- Sajikan interval kepercayaan bersama batas interpretasi dan ketidakpastian.

Mutu inferensi terutama ditentukan oleh representativitas, presisi estimasi, dan transparansi denominator. Ancaman utama ialah menyamakan persentase salah dengan miskonsepsi serta menggabungkan kategori yang berbeda. Karena itu, laporan perlu menyajikan bukti yang mendukung sekaligus bukti yang membatasi interpretasi. Prinsip ini konsisten dengan penekanan Creswell dan Creswell (2018) dan von Elm et al. (2007) bahwa validitas melekat pada interpretasi dan penggunaan hasil, bukan sekadar pada nama instrumen atau desain.

Dengan demikian, deskripsi lintas representasi kimia berfungsi menghubungkan keputusan desain dengan interpretasi. Produk yang diharapkan bukan hanya tabel prevalensi, tetapi juga penjelasan mengenai bagaimana produk tersebut diperoleh dan dalam kondisi apa ia dapat digunakan. Penelitian miskonsepsi yang baik selalu menyisakan ruang bagi penjelasan alternatif dan verifikasi lanjutan.

7.2.10 Deskripsi Miskonsepsi Berdasarkan Konteks

Pembahasan mengenai deskripsi miskonsepsi berdasarkan konteks diperlukan agar penelitian deskriptif tidak berhenti pada prosedur administratif. Fokus ini berada dalam desain yang memetakan prevalensi, distribusi, bentuk, keyakinan, dan konteks kemunculan miskonsepsi tanpa memanipulasi kondisi pembelajaran. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari

kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. von Elm et al. (2007) serta Hasan et al. (1999) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Penerapan praktis dapat disusun dalam urutan: mendefinisikan waktu pengukuran, mengoperasionalkan tingkat analisis butir dan peserta, mengumpulkan respons terbuka, memeriksa ketepatan klasifikasi, lalu melaporkan profil kategori konsepsi. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Misalnya, pada hukum gas, perubahan skor sesudah pembelajaran belum otomatis menandakan perubahan konsep. Peserta didik mungkin mempelajari pola jawaban tes. Bukti yang lebih meyakinkan mencakup transfer ke situasi baru, penjelasan kausal, representasi partikel, dan retensi setelah jeda waktu.

Langkah operasional yang disarankan:

- Definisikan waktu pengukuran dalam istilah operasional.
- Tentukan bukti utama dari skala keyakinan dan bukti pendukung dari profil demografis akademik.
- Periksa presisi estimasi sebelum membuat kategori atau model.
- Catat potensi menggabungkan kategori yang berbeda beserta strategi mitigasinya.
- Sajikan peta butir sulit bersama batas interpretasi dan ketidakpastian.

Temuan yang baik tidak selalu berarti hasil yang besar atau sesuai harapan. Temuan bernilai ketika ketepatan klasifikasi terjaga, konsistensi pengodean dijelaskan, dan ketidakpastian disajikan. Kerangka von Elm et al. (2007) serta Hasan et al. (1999) menolong peneliti membedakan kekuatan bukti dari daya tarik narasi.

Inti metodologis subbagian ini adalah keterlacakan. Pembaca harus dapat mengikuti perjalanan dari pertanyaan, data, pengodean, analisis, sampai profil kategori konsepsi. Ketika jejak itu jelas, temuan dapat digunakan untuk memperbaiki pembelajaran tanpa mengubah diagnosis konseptual menjadi label bagi peserta didik.

7.2.11 Studi Deskriptif Berulang

Dalam metodologi penelitian miskonsepsi kimia, studi deskriptif berulang menentukan mutu hubungan antara data dan kesimpulan. Fokus ini berada dalam desain yang memetakan prevalensi, distribusi, bentuk, keyakinan, dan konteks kemunculan miskonsepsi tanpa memanipulasi kondisi pembelajaran. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal.

Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Treagust (1988) serta von Elm et al. (2007) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Prosedur minimal mencakup definisi kategori konsepsi, alasan memilih skala keyakinan, aturan mengubah respons menjadi kategori, serta mekanisme pemeriksaan presisi estimasi. Semua penyimpangan dari rencana awal perlu dicatat. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Pada penelitian mengenai laju reaksi, peneliti dapat menyusun beberapa tugas ekuivalen yang berubah konteks, representasi, atau tuntutan penalarannya. Jika pola respons tetap berulang dan disertai keyakinan tinggi, terdapat dasar yang lebih kuat untuk menyebut indikasi miskonsepsi. Jika pola berubah-ubah, interpretasi yang lebih hati-hati adalah pengetahuan parsial atau ketergantungan konteks.

Keputusan yang harus didokumentasikan:

- Definisikan kategori konsepsi dalam istilah operasional.
- Tentukan bukti utama dari gambar atau model dan bukti pendukung dari catatan konteks pembelajaran.
- Periksa konsistensi pengodean sebelum membuat kategori atau model.
- Catat potensi menafsirkan korelasi sebagai sebab beserta strategi mitigasinya.
- Sajikan perbandingan subkelompok bersama batas interpretasi dan ketidakpastian.

Pemeriksaan mutu hendaknya dilakukan selama penelitian, bukan hanya pada akhir. Catatan tentang presisi estimasi, keputusan terkait transparansi denominator, dan kejadian yang berpotensi menimbulkan membuat ranking tanpa interval ketidakpastian perlu dipelihara dalam *audit trail*. Hal ini memungkinkan koreksi dini dan pelaporan yang jujur.

Kesimpulannya, studi deskriptif berulang perlu dirancang sebagai proses argumentatif. Data tidak berbicara sendiri; maknanya bergantung pada definisi konstruk, kualitas prosedur, dan konteks. Oleh sebab itu, klaim akhir sebaiknya menggunakan bahasa yang sesuai dengan kekuatan bukti, misalnya “menunjukkan pola”, “mengindikasikan”, atau “mendukung penjelasan”.

7.2.12 Penggunaan Data Sekunder dan Arsip Asesmen

Keputusan tentang penggunaan data sekunder dan arsip asesmen harus dibuat sebelum data utama dikumpulkan, bukan ditambahkan setelah hasil diketahui. Fokus ini berada dalam desain yang memetakan prevalensi, distribusi, bentuk, keyakinan, dan konteks kemunculan miskonsepsi tanpa memanipulasi kondisi pembelajaran. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut

dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Kaltakci Gurel et al. (2015) serta Hasan et al. (1999) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Tahap kerja dimulai dengan penegasan tingkat analisis butir dan peserta. Setelah itu, peneliti menyusun aturan untuk cara menyatakan prevalensi dan menentukan bagaimana waktu pengukuran akan diperiksa. Bukti sebaiknya berasal dari lebih dari satu bentuk, misalnya gambar atau model dan catatan konteks pembelajaran. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Pada topik kesetimbangan, peneliti juga perlu memperhatikan bahasa. Istilah yang sama dapat diberi makna berbeda oleh peserta didik. Klarifikasi melalui pertanyaan tindak lanjut, gambar, atau contoh tandingan membantu membedakan masalah terminologi dari kesalahan konseptual.

Pertanyaan audit metodologis:

- Definisikan tingkat analisis butir dan peserta dalam istilah operasional.
- Tentukan bukti utama dari profil demografis akademik dan bukti pendukung dari tes diagnostik bertingkat.
- Periksa transparansi denominator sebelum membuat kategori atau model.
- Catat potensi menghilangkan variasi antarbutir beserta strategi mitigasinya.
- Sajikan narasi pola dominan bersama batas interpretasi dan ketidakpastian.

Kualitas penelitian meningkat ketika peneliti menunjukkan rantai bukti dari gambar atau model menuju peta butir sulit, menyatakan asumsi, dan memeriksa alternatif penjelasan. Sebaliknya, menggabungkan kategori yang berbeda dapat menghasilkan klaim yang tampak tegas tetapi tidak stabil. Rekomendasi praktis harus proporsional dengan bukti.

Pada akhirnya, keputusan tentang penggunaan data sekunder dan arsip asesmen menentukan apakah penelitian menghasilkan daftar jawaban salah atau pemahaman yang dapat ditindaklanjuti mengenai cara peserta didik berpikir. Sasaran utamanya ialah bukti yang cukup untuk merancang diagnosis, pembelajaran, dan penelitian lanjutan secara bertanggung jawab.

7.2.13 Pengendalian Bias Deskriptif

Pengendalian bias deskriptif perlu dipahami sebagai keputusan metodologis yang mempunyai konsekuensi pada desain, analisis, dan pelaporan. Fokus ini berada dalam desain yang memetakan prevalensi, distribusi, bentuk, keyakinan, dan konteks kemunculan miskonsepsi tanpa memanipulasi kondisi pembelajaran. Peneliti perlu menyatakan sejak awal apa yang

dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Hasan et al. (1999) serta von Elm et al. (2007) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Peneliti sebaiknya menyiapkan matriks yang menghubungkan pertanyaan penelitian dengan pemilihan pembanding, profil demografis akademik, teknik analisis, dan perbandingan subkelompok. Matriks ini mencegah pengumpulan data yang banyak tetapi tidak relevan. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Ilustrasi ikatan ionik juga menunjukkan bahwa kualitas instrumen memengaruhi temuan. Distraktor yang tidak berasal dari pemikiran peserta didik dapat menghasilkan pilihan acak, sedangkan pertanyaan yang terlalu menuntut matematika dapat mengukur keterampilan hitung alih-alih konsepsi. Peneliti perlu menilai kesesuaian bukti dengan konstruk yang dimaksud.

Bukti minimum yang perlu tersedia:

- Definiskan pemilihan pembanding dalam istilah operasional.
- Tentukan bukti utama dari catatan konteks pembelajaran dan bukti pendukung dari respons terbuka.
- Periksa keterbandingan konteks sebelum membuat kategori atau model.
- Catat potensi menyamakan persentase salah dengan miskonsepsi beserta strategi mitigasinya.
- Sajikan tabel prevalensi bersama batas interpretasi dan ketidakpastian.

Peneliti sebaiknya memisahkan keputusan apriori dari keputusan eksploratori. Pemisahan tersebut membantu mengendalikan menafsirkan korelasi sebagai sebab dan mencegah mengabaikan missing data. Selanjutnya, transparansi denominator dan representativitas dilaporkan bersama ketidakpastian, sehingga pembaca dapat menilai kekuatan temuan tanpa bergantung pada kesimpulan penulis semata.

Dengan demikian, pengendalian bias deskriptif berfungsi menghubungkan keputusan desain dengan interpretasi. Produk yang diharapkan bukan hanya perbandingan subkelompok, tetapi juga penjelasan mengenai bagaimana produk tersebut diperoleh dan dalam kondisi apa ia dapat digunakan. Penelitian miskonsepsi yang baik selalu menyisakan ruang bagi penjelasan alternatif dan verifikasi lanjutan.

7.2.14 Pelaporan Temuan Deskriptif

Ketelitian penelitian pada bagian pelaporan temuan deskriptif sangat menentukan apakah pola jawaban dapat ditafsirkan sebagai bukti konseptual. Fokus ini berada dalam desain yang memetakan prevalensi, distribusi, bentuk, keyakinan, dan konteks kemunculan miskonsepsi tanpa memanipulasi kondisi pembelajaran. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Appelbaum et al. (2018) serta Hasan et al. (1999) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Dokumentasi metodologis hendaknya memuat siapa yang melakukan catatan konteks pembelajaran, kapan data diperoleh, bagaimana waktu pengukuran dikendalikan, dan atas dasar apa narasi pola dominan disusun. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Contoh pada kajian elektrolit memperlihatkan pentingnya pemisahan antara ketidaktahuan, kesalahan prosedural, dan miskonsepsi. Seorang peserta dapat gagal karena lupa rumus, sedangkan peserta lain memilih jawaban yang sama karena menggunakan model mental yang tidak ilmiah. Perbedaan ini baru terlihat ketika catatan konteks pembelajaran dibaca bersama respons terbuka dan bukti tentang alasan.

Kriteria pemeriksaan sebelum analisis:

- Definisikan cara menyatakan prevalensi dalam istilah operasional.
- Tentukan bukti utama dari tes diagnostik bertingkat dan bukti pendukung dari skala keyakinan.
- Periksa representativitas sebelum membuat kategori atau model.
- Catat potensi mengabaikan missing data beserta strategi mitigasinya.
- Sajikan profil kategori konsepsi bersama batas interpretasi dan ketidakpastian.

Standar metodologis menuntut koherensi antara konstruk, prosedur, analisis, dan klaim. Dalam konteks ini, keterbandingan konteks tidak dapat digantikan oleh indeks tunggal. Demikian pula, ketepatan klasifikasi harus dinilai dengan bukti yang sesuai. Appelbaum et al. (2018) dan Hasan et al. (1999) memberikan dasar untuk pelaporan yang lengkap dan dapat ditelaah.

Inti metodologis subbagian ini adalah keterlacakan. Pembaca harus dapat mengikuti perjalanan dari pertanyaan, data, pengodean, analisis, sampai narasi pola dominan. Ketika

jejak itu jelas, temuan dapat digunakan untuk memperbaiki pembelajaran tanpa mengubah diagnosis konseptual menjadi label bagi peserta didik.

7.2.15 Sintesis Rancangan Penelitian Deskriptif

Peneliti sering memusatkan perhatian pada teknik, padahal sintesis rancangan penelitian deskriptif justru mengatur batas inferensi yang dapat dibuat. Fokus ini berada dalam desain yang memetakan prevalensi, distribusi, bentuk, keyakinan, dan konteks kemunculan miskonsepsi tanpa memanipulasi kondisi pembelajaran. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Creswell dan Creswell (2018) serta von Elm et al. (2007) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Rancangan yang dapat diaudit mencantumkan alasan pemilihan definisi populasi target, prosedur penanganan kategori konsepsi, dan kriteria keputusan untuk pemilihan pembanding. Tanpa penjelasan itu, pembaca sulit menilai apakah tes diagnostik bertingkat benar-benar mendukung klaim. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Kasus hidrolisis garam dapat digunakan untuk memperlihatkan hubungan antara pola individual dan pola kelas. Sebagian miskonsepsi mungkin hanya dimiliki beberapa peserta tetapi sangat kuat, sedangkan pola lain tersebar luas karena bersumber dari analogi atau buku teks yang sama. Analisis perlu menjaga kedua tingkat tersebut.

Langkah operasional yang disarankan:

- Definisikan definisi populasi target dalam istilah operasional.
- Tentukan bukti utama dari respons terbuka dan bukti pendukung dari gambar atau model.
- Periksa ketepatan klasifikasi sebelum membuat kategori atau model.
- Catat potensi membuat ranking tanpa interval ketidakpastian beserta strategi mitigasinya.
- Sajikan interval kepercayaan bersama batas interpretasi dan ketidakpastian.

Pemeriksaan kualitas dapat dilakukan melalui audit terhadap representativitas, analisis kasus yang bertentangan, dan evaluasi presisi estimasi. Apabila ditemukan menyamakan persentase salah dengan miskonsepsi, peneliti perlu menurunkan kekuatan klaim atau menambah bukti. Transparansi seperti ini meningkatkan kegunaan temuan bagi penelitian lanjutan dan praktik kelas.

Kesimpulannya, sintesis rancangan penelitian deskriptif perlu dirancang sebagai proses argumentatif. Data tidak berbicara sendiri; maknanya bergantung pada definisi konstruk, kualitas prosedur, dan konteks. Oleh sebab itu, klaim akhir sebaiknya menggunakan bahasa yang sesuai dengan kekuatan bukti, misalnya “menunjukkan pola”, “mengindikasikan”, atau “mendukung penjelasan”.

7.3 Penelitian Studi Kasus

Subbab ini membahas pendekatan mendalam untuk memahami miskonsepsi sebagai fenomena yang terikat pada peserta didik, kelas, sekolah, topik, atau episode pembelajaran tertentu. Fokus utamanya adalah menjaga keselarasan antara pertanyaan penelitian, bentuk bukti, teknik analisis, dan kekuatan klaim. Dalam konteks miskonsepsi kimia, keputusan metodologis harus mampu membedakan pola konseptual dari respons yang timbul karena keterbatasan bahasa, kesalahan hitung, atau kondisi administrasi. Pembahasan berikut dipecah menjadi lima belas bagian agar setiap keputusan dapat ditelaah secara operasional.

7.3.1 Hakikat Kasus dan Batas Sistem

Kajian yang kuat memperlakukan hakikat kasus dan batas sistem sebagai bagian dari rantai argumentasi, bukan sebagai unsur pelengkap. Fokus ini berada dalam pendekatan mendalam untuk memahami miskonsepsi sebagai fenomena yang terikat pada peserta didik, kelas, sekolah, topik, atau episode pembelajaran tertentu. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Merriam dan Tisdell (2016) serta O’Brien et al. (2014) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Pada tahap perencanaan, rasional pemilihan kasus dan unit terbenam harus dibedakan secara eksplisit. Pada tahap pelaksanaan, konsistensi observasi kelas perlu diperiksa; pada tahap interpretasi, rantai bukti menjadi dasar pembatasan klaim. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Dalam kelas yang mempelajari kimia bilingual, data agregat mungkin menunjukkan mayoritas jawaban salah. Akan tetapi, wawancara dapat memperlihatkan beberapa mekanisme berbeda: bahasa sehari-hari, generalisasi aturan, salah membaca simbol, atau model partikel yang tidak tepat. Karena itu, satu persentase tidak boleh menggantikan analisis struktur alasan.

Keputusan yang harus didokumentasikan:

- Definisikan rasional pemilihan kasus dalam istilah operasional.
- Tentukan bukti utama dari artefak belajar dan bukti pendukung dari tes diagnostik.
- Periksa triangulasi sumber sebelum membuat kategori atau model.
- Catat potensi ketergantungan pada satu informan beserta strategi mitigasinya.
- Sajikan penjelasan mekanisme bersama batas interpretasi dan ketidakpastian.

Kriteria keberhasilan tidak cukup berupa banyaknya data. Yang lebih penting adalah apakah rantai bukti dapat ditelusuri, apakah deskripsi tebal diperiksa, dan apakah audit keputusan dijelaskan. Peneliti harus menghindari terlalu banyak data tanpa fokus; jika risiko itu tidak dapat dihilangkan, dampaknya perlu dianalisis dan dicantumkan sebagai batas penelitian.

Pada akhirnya, keputusan tentang hakikat kasus dan batas sistem menentukan apakah penelitian menghasilkan daftar jawaban salah atau pemahaman yang dapat ditindaklanjuti mengenai cara peserta didik berpikir. Sasaran utamanya ialah bukti yang cukup untuk merancang diagnosis, pembelajaran, dan penelitian lanjutan secara bertanggung jawab.

7.3.2 Pemilihan Kasus Tunggal atau Multikasus

Subbagian ini menempatkan pemilihan kasus tunggal atau multikasus sebagai simpul penting dalam penelitian studi kasus. Fokus ini berada dalam pendekatan mendalam untuk memahami miskonsepsi sebagai fenomena yang terikat pada peserta didik, kelas, sekolah, topik, atau episode pembelajaran tertentu. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Patton (2015) serta Merriam dan Tisdell (2016) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Secara operasional, peneliti perlu menetapkan jenis kasus tunggal atau jamak, menjelaskan proposisi awal, dan mendokumentasikan batas kasus. Keputusan tersebut kemudian diterjemahkan ke dalam pilihan artefak belajar, tes diagnostik, serta wawancara mendalam. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Sebagai ilustrasi, pada topik satu kelas pada topik kesetimbangan, peserta didik dapat memberikan jawaban benar pada tingkat simbolik tetapi alasan submikroskopiknya keliru. Peneliti tidak cukup menghitung benar-salah; ia perlu membandingkan artefak belajar dengan tes diagnostik, memeriksa tingkat keyakinan, dan menguji konsistensi pada konteks lain. Hasil

analisis mungkin menunjukkan bahwa kesalahan hanya muncul pada satu representasi, atau justru merupakan hubungan konseptual yang stabil.

Pertanyaan audit metodologis:

- Definisikan jenis kasus tunggal atau jamak dalam istilah operasional.
- Tentukan bukti utama dari rekaman video dan bukti pendukung dari dokumen kurikulum.
- Periksa deskripsi tebal sebelum membuat kategori atau model.
- Catat potensi konfirmasi dugaan awal beserta strategi mitigasinya.
- Sajikan perbandingan lintas kasus bersama batas interpretasi dan ketidakpastian.

Mutu inferensi terutama ditentukan oleh triangulasi sumber, pencarian bukti tandingan, dan konstruk yang jelas. Ancaman utama ialah generalisasi statistik yang tidak tepat serta hilangnya kronologi. Karena itu, laporan perlu menyajikan bukti yang mendukung sekaligus bukti yang membatasi interpretasi. Prinsip ini konsisten dengan penekanan Patton (2015) dan Merriam dan Tisdell (2016) bahwa validitas melekat pada interpretasi dan penggunaan hasil, bukan sekadar pada nama instrumen atau desain.

Dengan demikian, pemilihan kasus tunggal atau multikasus berfungsi menghubungkan keputusan desain dengan interpretasi. Produk yang diharapkan bukan hanya matriks bukti, tetapi juga penjelasan mengenai bagaimana produk tersebut diperoleh dan dalam kondisi apa ia dapat digunakan. Penelitian miskonsepsi yang baik selalu menyisakan ruang bagi penjelasan alternatif dan verifikasi lanjutan.

7.3.3 Kasus Tipikal, Ekstrem, dan Kritis

Pembahasan mengenai kasus tipikal, ekstrem, dan kritis diperlukan agar penelitian studi kasus tidak berhenti pada prosedur administratif. Fokus ini berada dalam pendekatan mendalam untuk memahami miskonsepsi sebagai fenomena yang terikat pada peserta didik, kelas, sekolah, topik, atau episode pembelajaran tertentu. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Lincoln dan Guba (1985) serta O'Brien et al. (2014) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Penerapan praktis dapat disusun dalam urutan: mendefinisikan unit terbenam, mengoperasionalkan logika replikasi, mengumpulkan rekaman video, memeriksa deskripsi tebal, lalu melaporkan penjelasan mekanisme. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Misalnya, pada siswa dengan keyakinan tinggi pada redoks, perubahan skor sesudah pembelajaran belum otomatis menandakan perubahan konsep. Peserta didik mungkin mempelajari pola jawaban tes. Bukti yang lebih meyakinkan mencakup transfer ke situasi baru, penjelasan kausal, representasi partikel, dan retensi setelah jeda waktu.

Bukti minimum yang perlu tersedia:

- Definisikan unit terbenam dalam istilah operasional.
- Tentukan bukti utama dari tes diagnostik dan bukti pendukung dari wawancara mendalam.
- Periksa pencarian bukti tandingan sebelum membuat kategori atau model.
- Catat potensi hilangnya kronologi beserta strategi mitigasinya.
- Sajikan generalisasi analitik bersama batas interpretasi dan ketidakpastian.

Temuan yang baik tidak selalu berarti hasil yang besar atau sesuai harapan. Temuan bernilai ketika deskripsi tebal terjaga, audit keputusan dijelaskan, dan ketidakpastian disajikan. Kerangka Lincoln dan Guba (1985) serta O'Brien et al. (2014) menolong peneliti membedakan kekuatan bukti dari daya tarik narasi.

Inti metodologis subbagian ini adalah keterlacakan. Pembaca harus dapat mengikuti perjalanan dari pertanyaan, data, pengodean, analisis, sampai penjelasan mekanisme. Ketika jejak itu jelas, temuan dapat digunakan untuk memperbaiki pembelajaran tanpa mengubah diagnosis konseptual menjadi label bagi peserta didik.

7.3.4 Perumusan Proposisi Awal

Dalam metodologi penelitian miskonsepsi kimia, perumusan proposisi awal menentukan mutu hubungan antara data dan kesimpulan. Fokus ini berada dalam pendekatan mendalam untuk memahami miskonsepsi sebagai fenomena yang terikat pada peserta didik, kelas, sekolah, topik, atau episode pembelajaran tertentu. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. O'Brien et al. (2014) serta Merriam dan Tisdell (2016) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Prosedur minimal mencakup definisi proposisi awal, alasan memilih tes diagnostik, aturan mengubah respons menjadi kategori, serta mekanisme pemeriksaan pencarian bukti tandingan. Semua penyimpangan dari rencana awal perlu dicatat. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Pada penelitian mengenai guru yang menggunakan analogi ikatan, peneliti dapat menyusun beberapa tugas ekuivalen yang berubah konteks, representasi, atau tuntutan penalarannya. Jika pola respons tetap berulang dan disertai keyakinan tinggi, terdapat dasar yang lebih kuat untuk menyebut indikasi miskonsepsi. Jika pola berubah-ubah, interpretasi yang lebih hati-hati adalah pengetahuan parsial atau ketergantungan konteks.

Kriteria pemeriksaan sebelum analisis:

- Definisikan proposisi awal dalam istilah operasional.
- Tentukan bukti utama dari dokumen kurikulum dan bukti pendukung dari observasi kelas.
- Periksa audit keputusan sebelum membuat kategori atau model.
- Catat potensi kasus tidak terbatas beserta strategi mitigasinya.
- Sajikan deskripsi kasus bersama batas interpretasi dan ketidakpastian.

Pemeriksaan mutu hendaknya dilakukan selama penelitian, bukan hanya pada akhir. Catatan tentang pencarian bukti tandingan, keputusan terkait konstruk yang jelas, dan kejadian yang berpotensi menimbulkan konfirmasi dugaan awal perlu dipelihara dalam *audit trail*. Hal ini memungkinkan koreksi dini dan pelaporan yang jujur.

Kesimpulannya, perumusan proposisi awal perlu dirancang sebagai proses argumentatif. Data tidak berbicara sendiri; maknanya bergantung pada definisi konstruk, kualitas prosedur, dan konteks. Oleh sebab itu, klaim akhir sebaiknya menggunakan bahasa yang sesuai dengan kekuatan bukti, misalnya “menunjukkan pola”, “mengindikasikan”, atau “mendukung penjelasan”.

7.3.5 Unit Analisis Terbenam

Keputusan tentang unit analisis terbenam harus dibuat sebelum data utama dikumpulkan, bukan ditambahkan setelah hasil diketahui. Fokus ini berada dalam pendekatan mendalam untuk memahami miskonsepsi sebagai fenomena yang terikat pada peserta didik, kelas, sekolah, topik, atau episode pembelajaran tertentu. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Tong et al. (2007) serta O’Brien et al. (2014) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Tahap kerja dimulai dengan penegasan logika replikasi. Setelah itu, peneliti menyusun aturan untuk rasional pemilihan kasus dan menentukan bagaimana unit terbenam akan diperiksa. Bukti sebaiknya berasal dari lebih dari satu bentuk, misalnya dokumen kurikulum dan observasi kelas. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu

peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Pada topik sekolah dengan laboratorium virtual, peneliti juga perlu memperhatikan bahasa. Istilah yang sama dapat diberi makna berbeda oleh peserta didik. Klarifikasi melalui pertanyaan tindak lanjut, gambar, atau contoh tandingan membantu membedakan masalah terminologi dari kesalahan konseptual.

Langkah operasional yang disarankan:

- Definisikan logika replikasi dalam istilah operasional.
- Tentukan bukti utama dari wawancara mendalam dan bukti pendukung dari artefak belajar.
- Periksa konstruk yang jelas sebelum membuat kategori atau model.
- Catat potensi terlalu banyak data tanpa fokus beserta strategi mitigasinya.
- Sajikan kronologi pembentukan konsepsi bersama batas interpretasi dan ketidakpastian.

Kualitas penelitian meningkat ketika peneliti menunjukkan rantai bukti dari dokumen kurikulum menuju generalisasi analitik, menyatakan asumsi, dan memeriksa alternatif penjelasan. Sebaliknya, hilangnya kronologi dapat menghasilkan klaim yang tampak tegas tetapi tidak stabil. Rekomendasi praktis harus proporsional dengan bukti.

Pada akhirnya, keputusan tentang unit analisis terbenam menentukan apakah penelitian menghasilkan daftar jawaban salah atau pemahaman yang dapat ditindaklanjuti mengenai cara peserta didik berpikir. Sasaran utamanya ialah bukti yang cukup untuk merancang diagnosis, pembelajaran, dan penelitian lanjutan secara bertanggung jawab.

7.3.6 Kronologi Perkembangan Miskonsepsi

Kronologi perkembangan miskonsepsi perlu dipahami sebagai keputusan metodologis yang mempunyai konsekuensi pada desain, analisis, dan pelaporan. Fokus ini berada dalam pendekatan mendalam untuk memahami miskonsepsi sebagai fenomena yang terikat pada peserta didik, kelas, sekolah, topik, atau episode pembelajaran tertentu. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Yin (2018) serta Merriam dan Tisdell (2016) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Peneliti sebaiknya menyiapkan matriks yang menghubungkan pertanyaan penelitian dengan batas kasus, wawancara mendalam, teknik analisis, dan deskripsi kasus. Matriks ini mencegah pengumpulan data yang banyak tetapi tidak relevan. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk

replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Ilustrasi kelompok diskusi asam-basa juga menunjukkan bahwa kualitas instrumen memengaruhi temuan. Distraktor yang tidak berasal dari pemikiran peserta didik dapat menghasilkan pilihan acak, sedangkan pertanyaan yang terlalu menuntut matematika dapat mengukur keterampilan hitung alih-alih konsepsi. Peneliti perlu menilai kesesuaian bukti dengan konstruk yang dimaksud.

Keputusan yang harus didokumentasikan:

- Definisikan batas kasus dalam istilah operasional.
- Tentukan bukti utama dari observasi kelas dan bukti pendukung dari rekaman video.
- Periksa rantai bukti sebelum membuat kategori atau model.
- Catat potensi generalisasi statistik yang tidak tepat beserta strategi mitigasinya.
- Sajikan matriks bukti bersama batas interpretasi dan ketidakpastian.

Peneliti sebaiknya memisahkan keputusan apriori dari keputusan eksploratori. Pemisahan tersebut membantu mengendalikan kasus tidak terbatas dan mencegah ketergantungan pada satu informan. Selanjutnya, konstruk yang jelas dan triangulasi sumber dilaporkan bersama ketidakpastian, sehingga pembaca dapat menilai kekuatan temuan tanpa bergantung pada kesimpulan penulis semata.

Dengan demikian, kronologi perkembangan miskonsepsi berfungsi menghubungkan keputusan desain dengan interpretasi. Produk yang diharapkan bukan hanya deskripsi kasus, tetapi juga penjelasan mengenai bagaimana produk tersebut diperoleh dan dalam kondisi apa ia dapat digunakan. Penelitian miskonsepsi yang baik selalu menyisakan ruang bagi penjelasan alternatif dan verifikasi lanjutan.

7.3.7 Wawancara Mendalam dalam Studi Kasus

Ketelitian penelitian pada bagian wawancara mendalam dalam studi kasus sangat menentukan apakah pola jawaban dapat ditafsirkan sebagai bukti konseptual. Fokus ini berada dalam pendekatan mendalam untuk memahami miskonsepsi sebagai fenomena yang terikat pada peserta didik, kelas, sekolah, topik, atau episode pembelajaran tertentu. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Merriam dan Tisdell (2016) serta O'Brien et al. (2014) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Dokumentasi metodologis hendaknya memuat siapa yang melakukan observasi kelas, kapan data diperoleh, bagaimana unit terbenam dikendalikan, dan atas dasar apa kronologi pembentukan konsepsi disusun. Setiap keputusan dicatat dalam protokol, termasuk versi

instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Contoh pada kajian episode praktikum kalorimetri memperlihatkan pentingnya pemisahan antara ketidaktahuan, kesalahan prosedural, dan miskonsepsi. Seorang peserta dapat gagal karena lupa rumus, sedangkan peserta lain memilih jawaban yang sama karena menggunakan model mental yang tidak ilmiah. Perbedaan ini baru terlihat ketika observasi kelas dibaca bersama rekaman video dan bukti tentang alasan.

Pertanyaan audit metodologis:

- Definisikan rasional pemilihan kasus dalam istilah operasional.
- Tentukan bukti utama dari artefak belajar dan bukti pendukung dari tes diagnostik.
- Periksa triangulasi sumber sebelum membuat kategori atau model.
- Catat potensi ketergantungan pada satu informan beserta strategi mitigasinya.
- Sajikan penjelasan mekanisme bersama batas interpretasi dan ketidakpastian.

Standar metodologis menuntut koherensi antara konstruk, prosedur, analisis, dan klaim. Dalam konteks ini, rantai bukti tidak dapat digantikan oleh indeks tunggal. Demikian pula, deskripsi tebal harus dinilai dengan bukti yang sesuai. Merriam dan Tisdell (2016) dan O'Brien et al. (2014) memberikan dasar untuk pelaporan yang lengkap dan dapat ditelaah.

Inti metodologis subbagian ini adalah keterlacakan. Pembaca harus dapat mengikuti perjalanan dari pertanyaan, data, pengodean, analisis, sampai kronologi pembentukan konsepsi. Ketika jejak itu jelas, temuan dapat digunakan untuk memperbaiki pembelajaran tanpa mengubah diagnosis konseptual menjadi label bagi peserta didik.

7.3.8 Observasi Episode Pembelajaran

Peneliti sering memusatkan perhatian pada teknik, padahal observasi episode pembelajaran justru mengatur batas inferensi yang dapat dibuat. Fokus ini berada dalam pendekatan mendalam untuk memahami miskonsepsi sebagai fenomena yang terikat pada peserta didik, kelas, sekolah, topik, atau episode pembelajaran tertentu. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Patton (2015) serta Merriam dan Tisdell (2016) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Rancangan yang dapat diaudit mencantumkan alasan pemilihan jenis kasus tunggal atau jamak, prosedur penanganan proposisi awal, dan kriteria keputusan untuk batas kasus. Tanpa penjelasan itu, pembaca sulit menilai apakah artefak belajar benar-benar mendukung klaim. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean,

penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Kasus transisi SMP ke SMA dapat digunakan untuk memperlihatkan hubungan antara pola individual dan pola kelas. Sebagian miskonsepsi mungkin hanya dimiliki beberapa peserta tetapi sangat kuat, sedangkan pola lain tersebar luas karena bersumber dari analogi atau buku teks yang sama. Analisis perlu menjaga kedua tingkat tersebut.

Bukti minimum yang perlu tersedia:

- Definisikan jenis kasus tunggal atau jamak dalam istilah operasional.
- Tentukan bukti utama dari rekaman video dan bukti pendukung dari dokumen kurikulum.
- Periksa deskripsi tebal sebelum membuat kategori atau model.
- Catat potensi konfirmasi dugaan awal beserta strategi mitigasinya.
- Sajikan perbandingan lintas kasus bersama batas interpretasi dan ketidakpastian.

Pemeriksaan kualitas dapat dilakukan melalui audit terhadap triangulasi sumber, analisis kasus yang bertentangan, dan evaluasi pencarian bukti tandingan. Apabila ditemukan generalisasi statistik yang tidak tepat, peneliti perlu menurunkan kekuatan klaim atau menambah bukti. Transparansi seperti ini meningkatkan kegunaan temuan bagi penelitian lanjutan dan praktik kelas.

Kesimpulannya, observasi episode pembelajaran perlu dirancang sebagai proses argumentatif. Data tidak berbicara sendiri; maknanya bergantung pada definisi konstruk, kualitas prosedur, dan konteks. Oleh sebab itu, klaim akhir sebaiknya menggunakan bahasa yang sesuai dengan kekuatan bukti, misalnya “menunjukkan pola”, “mengindikasikan”, atau “mendukung penjelasan”.

7.3.9 Analisis Artefak dan Dokumen

Kajian yang kuat memperlakukan analisis artefak dan dokumen sebagai bagian dari rantai argumentasi, bukan sebagai unsur pelengkap. Fokus ini berada dalam pendekatan mendalam untuk memahami miskonsepsi sebagai fenomena yang terikat pada peserta didik, kelas, sekolah, topik, atau episode pembelajaran tertentu. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Lincoln dan Guba (1985) serta O’Brien et al. (2014) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Pada tahap perencanaan, unit terbenam dan logika replikasi harus dibedakan secara eksplisit. Pada tahap pelaksanaan, konsistensi rekaman video perlu diperiksa; pada tahap interpretasi, deskripsi tebal menjadi dasar pembatasan klaim. Setiap keputusan dicatat dalam protokol,

termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Dalam kelas yang mempelajari kelas bilingual kimia, data agregat mungkin menunjukkan mayoritas jawaban salah. Akan tetapi, wawancara dapat memperlihatkan beberapa mekanisme berbeda: bahasa sehari-hari, generalisasi aturan, salah membaca simbol, atau model partikel yang tidak tepat. Karena itu, satu persentase tidak boleh menggantikan analisis struktur alasan.

Kriteria pemeriksaan sebelum analisis:

- Definisikan unit terbenam dalam istilah operasional.
- Tentukan bukti utama dari tes diagnostik dan bukti pendukung dari wawancara mendalam.
- Periksa pencarian bukti tandingan sebelum membuat kategori atau model.
- Catat potensi hilangnya kronologi beserta strategi mitigasinya.
- Sajikan generalisasi analitik bersama batas interpretasi dan ketidakpastian.

Kriteria keberhasilan tidak cukup berupa banyaknya data. Yang lebih penting adalah apakah deskripsi tebal dapat ditelusuri, apakah audit keputusan diperiksa, dan apakah rantai bukti dijelaskan. Peneliti harus menghindari ketergantungan pada satu informan; jika risiko itu tidak dapat dihilangkan, dampaknya perlu dianalisis dan dicantumkan sebagai batas penelitian.

Pada akhirnya, keputusan tentang analisis artefak dan dokumen menentukan apakah penelitian menghasilkan daftar jawaban salah atau pemahaman yang dapat ditindaklanjuti mengenai cara peserta didik berpikir. Sasaran utamanya ialah bukti yang cukup untuk merancang diagnosis, pembelajaran, dan penelitian lanjutan secara bertanggung jawab.

7.3.10 Penyusunan Basis Data Kasus

Subbagian ini menempatkan penyusunan basis data kasus sebagai simpul penting dalam penelitian studi kasus. Fokus ini berada dalam pendekatan mendalam untuk memahami miskonsepsi sebagai fenomena yang terikat pada peserta didik, kelas, sekolah, topik, atau episode pembelajaran tertentu. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. O'Brien et al. (2014) serta Merriam dan Tisdell (2016) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Secara operasional, peneliti perlu menetapkan proposisi awal, menjelaskan batas kasus, dan mendokumentasikan jenis kasus tunggal atau jamak. Keputusan tersebut kemudian diterjemahkan ke dalam pilihan tes diagnostik, wawancara mendalam, serta artefak belajar.

Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Sebagai ilustrasi, pada topik satu kelas pada topik kesetimbangan, peserta didik dapat memberikan jawaban benar pada tingkat simbolik tetapi alasan submikroskopiknya keliru. Peneliti tidak cukup menghitung benar-salah; ia perlu membandingkan tes diagnostik dengan wawancara mendalam, memeriksa tingkat keyakinan, dan menguji konsistensi pada konteks lain. Hasil analisis mungkin menunjukkan bahwa kesalahan hanya muncul pada satu representasi, atau justru merupakan hubungan konseptual yang stabil.

Langkah operasional yang disarankan:

- Definisikan proposisi awal dalam istilah operasional.
- Tentukan bukti utama dari dokumen kurikulum dan bukti pendukung dari observasi kelas.
- Periksa audit keputusan sebelum membuat kategori atau model.
- Catat potensi kasus tidak terbatas beserta strategi mitigasinya.
- Sajikan deskripsi kasus bersama batas interpretasi dan ketidakpastian.

Mutu inferensi terutama ditentukan oleh pencarian bukti tandingan, konstruk yang jelas, dan triangulasi sumber. Ancaman utama ialah konfirmasi dugaan awal serta terlalu banyak data tanpa fokus. Karena itu, laporan perlu menyajikan bukti yang mendukung sekaligus bukti yang membatasi interpretasi. Prinsip ini konsisten dengan penekanan O'Brien et al. (2014) dan Merriam dan Tisdell (2016) bahwa validitas melekat pada interpretasi dan penggunaan hasil, bukan sekadar pada nama instrumen atau desain.

Dengan demikian, penyusunan basis data kasus berfungsi menghubungkan keputusan desain dengan interpretasi. Produk yang diharapkan bukan hanya perbandingan lintas kasus, tetapi juga penjelasan mengenai bagaimana produk tersebut diperoleh dan dalam kondisi apa ia dapat digunakan. Penelitian miskonsepsi yang baik selalu menyisakan ruang bagi penjelasan alternatif dan verifikasi lanjutan.

7.3.11 Triangulasi dalam Studi Kasus

Pembahasan mengenai triangulasi dalam studi kasus diperlukan agar penelitian studi kasus tidak berhenti pada prosedur administratif. Fokus ini berada dalam pendekatan mendalam untuk memahami miskonsepsi sebagai fenomena yang terikat pada peserta didik, kelas, sekolah, topik, atau episode pembelajaran tertentu. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Tong et al.

(2007) serta O'Brien et al. (2014) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Penerapan praktis dapat disusun dalam urutan: mendefinisikan logika replikasi, mengoperasionalkan rasional pemilihan kasus, mengumpulkan dokumen kurikulum, memeriksa audit keputusan, lalu melaporkan generalisasi analitik. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Misalnya, pada siswa dengan keyakinan tinggi pada redoks, perubahan skor sesudah pembelajaran belum otomatis menandakan perubahan konsep. Peserta didik mungkin mempelajari pola jawaban tes. Bukti yang lebih meyakinkan mencakup transfer ke situasi baru, penjelasan kausal, representasi partikel, dan retensi setelah jeda waktu.

Keputusan yang harus didokumentasikan:

- Definisikan logika replikasi dalam istilah operasional.
- Tentukan bukti utama dari wawancara mendalam dan bukti pendukung dari artefak belajar.
- Periksa konstruk yang jelas sebelum membuat kategori atau model.
- Catat potensi terlalu banyak data tanpa fokus beserta strategi mitigasinya.
- Sajikan kronologi pembentukan konsepsi bersama batas interpretasi dan ketidakpastian.

Temuan yang baik tidak selalu berarti hasil yang besar atau sesuai harapan. Temuan bernilai ketika audit keputusan terjaga, rantai bukti dijelaskan, dan ketidakpastian disajikan. Kerangka Tong et al. (2007) serta O'Brien et al. (2014) menolong peneliti membedakan kekuatan bukti dari daya tarik narasi.

Inti metodologis subbagian ini adalah keterlacakan. Pembaca harus dapat mengikuti perjalanan dari pertanyaan, data, pengodean, analisis, sampai generalisasi analitik. Ketika jejak itu jelas, temuan dapat digunakan untuk memperbaiki pembelajaran tanpa mengubah diagnosis konseptual menjadi label bagi peserta didik.

7.3.12 Analisis Dalam Kasus

Dalam metodologi penelitian miskonsepsi kimia, analisis dalam kasus menentukan mutu hubungan antara data dan kesimpulan. Fokus ini berada dalam pendekatan mendalam untuk memahami miskonsepsi sebagai fenomena yang terikat pada peserta didik, kelas, sekolah, topik, atau episode pembelajaran tertentu. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Yin (2018) serta Merriam dan

Tisdell (2016) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Prosedur minimal mencakup definisi batas kasus, alasan memilih wawancara mendalam, aturan mengubah respons menjadi kategori, serta mekanisme pemeriksaan konstruk yang jelas. Semua penyimpangan dari rencana awal perlu dicatat. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Pada penelitian mengenai guru yang menggunakan analogi ikatan, peneliti dapat menyusun beberapa tugas ekuivalen yang berubah konteks, representasi, atau tuntutan penalarannya. Jika pola respons tetap berulang dan disertai keyakinan tinggi, terdapat dasar yang lebih kuat untuk menyebut indikasi miskonsepsi. Jika pola berubah-ubah, interpretasi yang lebih hati-hati adalah pengetahuan parsial atau ketergantungan konteks.

Pertanyaan audit metodologis:

- Definisikan batas kasus dalam istilah operasional.
- Tentukan bukti utama dari observasi kelas dan bukti pendukung dari rekaman video.
- Periksa rantai bukti sebelum membuat kategori atau model.
- Catat potensi generalisasi statistik yang tidak tepat beserta strategi mitigasinya.
- Sajikan matriks bukti bersama batas interpretasi dan ketidakpastian.

Pemeriksaan mutu hendaknya dilakukan selama penelitian, bukan hanya pada akhir. Catatan tentang konstruk yang jelas, keputusan terkait triangulasi sumber, dan kejadian yang berpotensi menimbulkan kasus tidak terbatas perlu dipelihara dalam *audit trail*. Hal ini memungkinkan koreksi dini dan pelaporan yang jujur.

Kesimpulannya, analisis dalam kasus perlu dirancang sebagai proses argumentatif. Data tidak berbicara sendiri; maknanya bergantung pada definisi konstruk, kualitas prosedur, dan konteks. Oleh sebab itu, klaim akhir sebaiknya menggunakan bahasa yang sesuai dengan kekuatan bukti, misalnya “menunjukkan pola”, “mengindikasikan”, atau “mendukung penjelasan”.

7.3.13 Analisis Lintas Kasus

Keputusan tentang analisis lintas kasus harus dibuat sebelum data utama dikumpulkan, bukan ditambahkan setelah hasil diketahui. Fokus ini berada dalam pendekatan mendalam untuk memahami miskonsepsi sebagai fenomena yang terikat pada peserta didik, kelas, sekolah, topik, atau episode pembelajaran tertentu. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Merriam dan Tisdell (2016)

serta O'Brien et al. (2014) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Tahap kerja dimulai dengan penegasan rasional pemilihan kasus. Setelah itu, peneliti menyusun aturan untuk unit terbenam dan menentukan bagaimana logika replikasi akan diperiksa. Bukti sebaiknya berasal dari lebih dari satu bentuk, misalnya observasi kelas dan rekaman video. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Pada topik sekolah dengan laboratorium virtual, peneliti juga perlu memperhatikan bahasa. Istilah yang sama dapat diberi makna berbeda oleh peserta didik. Klarifikasi melalui pertanyaan tindak lanjut, gambar, atau contoh tandingan membantu membedakan masalah terminologi dari kesalahan konseptual.

Bukti minimum yang perlu tersedia:

- Definisikan rasional pemilihan kasus dalam istilah operasional.
- Tentukan bukti utama dari artefak belajar dan bukti pendukung dari tes diagnostik.
- Periksa triangulasi sumber sebelum membuat kategori atau model.
- Catat potensi ketergantungan pada satu informan beserta strategi mitigasinya.
- Sajikan penjelasan mekanisme bersama batas interpretasi dan ketidakpastian.

Kualitas penelitian meningkat ketika peneliti menunjukkan rantai bukti dari observasi kelas menuju kronologi pembentukan konsepsi, menyatakan asumsi, dan memeriksa alternatif penjelasan. Sebaliknya, terlalu banyak data tanpa fokus dapat menghasilkan klaim yang tampak tegas tetapi tidak stabil. Rekomendasi praktis harus proporsional dengan bukti.

Pada akhirnya, keputusan tentang analisis lintas kasus menentukan apakah penelitian menghasilkan daftar jawaban salah atau pemahaman yang dapat ditindaklanjuti mengenai cara peserta didik berpikir. Sasaran utamanya ialah bukti yang cukup untuk merancang diagnosis, pembelajaran, dan penelitian lanjutan secara bertanggung jawab.

7.3.14 Generalisasi Analitik dan Transferabilitas

Generalisasi analitik dan transferabilitas perlu dipahami sebagai keputusan metodologis yang mempunyai konsekuensi pada desain, analisis, dan pelaporan. Fokus ini berada dalam pendekatan mendalam untuk memahami miskonsepsi sebagai fenomena yang terikat pada peserta didik, kelas, sekolah, topik, atau episode pembelajaran tertentu. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung.

Patton (2015) serta Merriam dan Tisdell (2016) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Peneliti sebaiknya menyiapkan matriks yang menghubungkan pertanyaan penelitian dengan jenis kasus tunggal atau jamak, artefak belajar, teknik analisis, dan matriks bukti. Matriks ini mencegah pengumpulan data yang banyak tetapi tidak relevan. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Ilustrasi kelompok diskusi asam-basa juga menunjukkan bahwa kualitas instrumen memengaruhi temuan. Distraktor yang tidak berasal dari pemikiran peserta didik dapat menghasilkan pilihan acak, sedangkan pertanyaan yang terlalu menuntut matematika dapat mengukur keterampilan hitung alih-alih konsepsi. Peneliti perlu menilai kesesuaian bukti dengan konstruk yang dimaksud.

Kriteria pemeriksaan sebelum analisis:

- Definisikan jenis kasus tunggal atau jamak dalam istilah operasional.
- Tentukan bukti utama dari rekaman video dan bukti pendukung dari dokumen kurikulum.
- Periksa deskripsi tebal sebelum membuat kategori atau model.
- Catat potensi konfirmasi dugaan awal beserta strategi mitigasinya.
- Sajikan perbandingan lintas kasus bersama batas interpretasi dan ketidakpastian.

Peneliti sebaiknya memisahkan keputusan apriori dari keputusan eksploratori. Pemisahan tersebut membantu mengendalikan generalisasi statistik yang tidak tepat dan mencegah hilangnya kronologi. Selanjutnya, triangulasi sumber dan pencarian bukti tandingan dilaporkan bersama ketidakpastian, sehingga pembaca dapat menilai kekuatan temuan tanpa bergantung pada kesimpulan penulis semata.

Dengan demikian, generalisasi analitik dan transferabilitas berfungsi menghubungkan keputusan desain dengan interpretasi. Produk yang diharapkan bukan hanya matriks bukti, tetapi juga penjelasan mengenai bagaimana produk tersebut diperoleh dan dalam kondisi apa ia dapat digunakan. Penelitian miskonsepsi yang baik selalu menyisakan ruang bagi penjelasan alternatif dan verifikasi lanjutan.

7.3.15 Sintesis Protokol Studi Kasus Miskonsepsi

Ketelitian penelitian pada bagian sintesis protokol studi kasus miskonsepsi sangat menentukan apakah pola jawaban dapat ditafsirkan sebagai bukti konseptual. Fokus ini berada dalam pendekatan mendalam untuk memahami miskonsepsi sebagai fenomena yang terikat pada peserta didik, kelas, sekolah, topik, atau episode pembelajaran tertentu. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi

merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Lincoln dan Guba (1985) serta O'Brien et al. (2014) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Dokumentasi metodologis hendaknya memuat siapa yang melakukan rekaman video, kapan data diperoleh, bagaimana logika replikasi dikendalikan, dan atas dasar apa penjelasan mekanisme disusun. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Contoh pada kajian episode praktikum kalorimetri memperlihatkan pentingnya pemisahan antara ketidaktahuan, kesalahan prosedural, dan miskonsepsi. Seorang peserta dapat gagal karena lupa rumus, sedangkan peserta lain memilih jawaban yang sama karena menggunakan model mental yang tidak ilmiah. Perbedaan ini baru terlihat ketika rekaman video dibaca bersama dokumen kurikulum dan bukti tentang alasan.

Langkah operasional yang disarankan:

- Definisikan unit terbenam dalam istilah operasional.
- Tentukan bukti utama dari tes diagnostik dan bukti pendukung dari wawancara mendalam.
- Periksa pencarian bukti tandingan sebelum membuat kategori atau model.
- Catat potensi hilangnya kronologi beserta strategi mitigasinya.
- Sajikan generalisasi analitik bersama batas interpretasi dan ketidakpastian.

Standar metodologis menuntut koherensi antara konstruk, prosedur, analisis, dan klaim. Dalam konteks ini, deskripsi tebal tidak dapat digantikan oleh indeks tunggal. Demikian pula, audit keputusan harus dinilai dengan bukti yang sesuai. Lincoln dan Guba (1985) dan O'Brien et al. (2014) memberikan dasar untuk pelaporan yang lengkap dan dapat ditelaah.

Inti metodologis subbagian ini adalah keterlacakan. Pembaca harus dapat mengikuti perjalanan dari pertanyaan, data, pengodean, analisis, sampai penjelasan mekanisme. Ketika jejak itu jelas, temuan dapat digunakan untuk memperbaiki pembelajaran tanpa mengubah diagnosis konseptual menjadi label bagi peserta didik.

7.4 Penelitian Survei

Subbab ini membahas strategi pengumpulan data terstandar dari sampel yang relatif luas untuk mengestimasi distribusi miskonsepsi, keyakinan, pengalaman belajar, dan faktor terkait. Fokus utamanya adalah menjaga keselarasan antara pertanyaan penelitian, bentuk bukti, teknik analisis, dan kekuatan klaim. Dalam konteks miskonsepsi kimia, keputusan metodologis harus mampu membedakan pola konseptual dari respons yang timbul karena keterbatasan

bahasa, kesalahan hitung, atau kondisi administrasi. Pembahasan berikut dipecah menjadi lima belas bagian agar setiap keputusan dapat ditelaah secara operasional.

7.4.1 Perbedaan Survei dan Tes Diagnostik

Peneliti sering memusatkan perhatian pada teknik, padahal perbedaan survei dan tes diagnostik justru mengatur batas inferensi yang dapat dibuat. Fokus ini berada dalam strategi pengumpulan data terstandar dari sampel yang relatif luas untuk mengestimasi distribusi miskonsepsi, keyakinan, pengalaman belajar, dan faktor terkait. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Lohr (2021) serta Fowler (2014) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Rancangan yang dapat diaudit mencantumkan alasan pemilihan strategi kontak, prosedur penanganan mode survei, dan kriteria keputusan untuk panjang instrumen. Tanpa penjelasan itu, pembaca sulit menilai apakah data latar belakang benar-benar mendukung klaim. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Kasus survei sekolah kota-desa dapat digunakan untuk memperlihatkan hubungan antara pola individual dan pola kelas. Sebagian miskonsepsi mungkin hanya dimiliki beberapa peserta tetapi sangat kuat, sedangkan pola lain tersebar luas karena bersumber dari analogi atau buku teks yang sama. Analisis perlu menjaga kedua tingkat tersebut.

Keputusan yang harus didokumentasikan:

- Definisikan strategi kontak dalam istilah operasional.
- Tentukan bukti utama dari log waktu respons dan bukti pendukung dari kuesioner kertas.
- Periksa analisis nonrespons sebelum membuat kategori atau model.
- Catat potensi sampel sukarela yang bias beserta strategi mitigasinya.
- Sajikan estimasi populasi bersama batas interpretasi dan ketidakpastian.

Pemeriksaan kualitas dapat dilakukan melalui audit terhadap kontrol duplikasi, analisis kasus yang bertentangan, dan evaluasi coverage. Apabila ditemukan ketidaksetaraan akses digital, peneliti perlu menurunkan kekuatan klaim atau menambah bukti. Transparansi seperti ini meningkatkan kegunaan temuan bagi penelitian lanjutan dan praktik kelas.

Kesimpulannya, perbedaan survei dan tes diagnostik perlu dirancang sebagai proses argumentatif. Data tidak berbicara sendiri; maknanya bergantung pada definisi konstruk, kualitas prosedur, dan konteks. Oleh sebab itu, klaim akhir sebaiknya menggunakan bahasa

yang sesuai dengan kekuatan bukti, misalnya “menunjukkan pola”, “mengindikasikan”, atau “mendukung penjelasan”.

7.4.2 Perumusan Tujuan Survei

Kajian yang kuat memperlakukan perumusan tujuan survei sebagai bagian dari rantai argumentasi, bukan sebagai unsur pelengkap. Fokus ini berada dalam strategi pengumpulan data terstandar dari sampel yang relatif luas untuk mengestimasi distribusi miskonsepsi, keyakinan, pengalaman belajar, dan faktor terkait. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. von Elm et al. (2007) serta Lohr (2021) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Pada tahap perencanaan, pengelolaan nonrespons dan kerangka sampel harus dibedakan secara eksplisit. Pada tahap pelaksanaan, konsistensi log waktu respons perlu diperiksa; pada tahap interpretasi, analisis nonrespons menjadi dasar pembatasan klaim. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Dalam kelas yang mempelajari survei setelah pembelajaran hibrida, data agregat mungkin menunjukkan mayoritas jawaban salah. Akan tetapi, wawancara dapat memperlihatkan beberapa mekanisme berbeda: bahasa sehari-hari, generalisasi aturan, salah membaca simbol, atau model partikel yang tidak tepat. Karena itu, satu persentase tidak boleh menggantikan analisis struktur alasan.

Pertanyaan audit metodologis:

- Definisikan pengelolaan nonrespons dalam istilah operasional.
- Tentukan bukti utama dari kuesioner daring dan bukti pendukung dari tes diagnostik terintegrasi.
- Periksa coverage sebelum membuat kategori atau model.
- Catat potensi survei terlalu panjang beserta strategi mitigasinya.
- Sajikan bobot survei bersama batas interpretasi dan ketidakpastian.

Kriteria keberhasilan tidak cukup berupa banyaknya data. Yang lebih penting adalah apakah analisis nonrespons dapat ditelusuri, apakah response rate diperiksa, dan apakah kesetaraan mode dijelaskan. Peneliti harus menghindari interpretasi berlebihan terhadap self-report; jika risiko itu tidak dapat dihilangkan, dampaknya perlu dianalisis dan dicantumkan sebagai batas penelitian.

Pada akhirnya, keputusan tentang perumusan tujuan survei menentukan apakah penelitian menghasilkan daftar jawaban salah atau pemahaman yang dapat ditindaklanjuti mengenai cara peserta didik berpikir. Sasaran utamanya ialah bukti yang cukup untuk merancang diagnosis, pembelajaran, dan penelitian lanjutan secara bertanggung jawab.

7.4.3 Kerangka Sampel Survei

Subbagian ini menempatkan kerangka sampel survei sebagai simpul penting dalam penelitian survei. Fokus ini berada dalam strategi pengumpulan data terstandar dari sampel yang relatif luas untuk mengestimasi distribusi miskonsepsi, keyakinan, pengalaman belajar, dan faktor terkait. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Dillman et al. (2014) serta Fowler (2014) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Secara operasional, peneliti perlu menetapkan mode survei, menjelaskan panjang instrumen, dan mendokumentasikan strategi kontak. Keputusan tersebut kemudian diterjemahkan ke dalam pilihan kuesioner daring, tes diagnostik terintegrasi, serta data latar belakang. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Sebagai ilustrasi, pada topik survei miskonsepsi stoikiometri tingkat provinsi, peserta didik dapat memberikan jawaban benar pada tingkat simbolik tetapi alasan submikroskopiknya keliru. Peneliti tidak cukup menghitung benar-salah; ia perlu membandingkan kuesioner daring dengan tes diagnostik terintegrasi, memeriksa tingkat keyakinan, dan menguji konsistensi pada konteks lain. Hasil analisis mungkin menunjukkan bahwa kesalahan hanya muncul pada satu representasi, atau justru merupakan hubungan konseptual yang stabil.

Bukti minimum yang perlu tersedia:

- Definisikan mode survei dalam istilah operasional.
- Tentukan bukti utama dari kuesioner kertas dan bukti pendukung dari skala keyakinan.
- Periksa response rate sebelum membuat kategori atau model.
- Catat potensi opsi jawaban tidak lengkap beserta strategi mitigasinya.
- Sajikan profil nonrespons bersama batas interpretasi dan ketidakpastian.

Mutu inferensi terutama ditentukan oleh coverage, validitas butir, dan kontrol duplikasi. Ancaman utama ialah sampel sukarela yang bias serta pertanyaan ganda. Karena itu, laporan perlu menyajikan bukti yang mendukung sekaligus bukti yang membatasi interpretasi. Prinsip ini konsisten dengan penekanan Dillman et al. (2014) dan Fowler (2014) bahwa validitas

melekat pada interpretasi dan penggunaan hasil, bukan sekadar pada nama instrumen atau desain.

Dengan demikian, kerangka sampel survei berfungsi menghubungkan keputusan desain dengan interpretasi. Produk yang diharapkan bukan hanya estimasi populasi, tetapi juga penjelasan mengenai bagaimana produk tersebut diperoleh dan dalam kondisi apa ia dapat digunakan. Penelitian miskonsepsi yang baik selalu menyisakan ruang bagi penjelasan alternatif dan verifikasi lanjutan.

7.4.4 Pemilihan Mode Pengumpulan

Pembahasan mengenai pemilihan mode pengumpulan diperlukan agar penelitian survei tidak berhenti pada prosedur administratif. Fokus ini berada dalam strategi pengumpulan data terstandar dari sampel yang relatif luas untuk mengestimasi distribusi miskonsepsi, keyakinan, pengalaman belajar, dan faktor terkait. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Fowler (2014) serta Lohr (2021) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Penerapan praktis dapat disusun dalam urutan: mendefinisikan kerangka sampel, mengoperasionalkan urutan butir, mengumpulkan kuesioner kertas, memeriksa response rate, lalu melaporkan bobot survei. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Misalnya, pada survei guru tentang sumber miskonsepsi, perubahan skor sesudah pembelajaran belum otomatis menandakan perubahan konsep. Peserta didik mungkin mempelajari pola jawaban tes. Bukti yang lebih meyakinkan mencakup transfer ke situasi baru, penjelasan kausal, representasi partikel, dan retensi setelah jeda waktu.

Kriteria pemeriksaan sebelum analisis:

- Definisikan kerangka sampel dalam istilah operasional.
- Tentukan bukti utama dari tes diagnostik terintegrasi dan bukti pendukung dari data latar belakang.
- Periksa validitas butir sebelum membuat kategori atau model.
- Catat potensi pertanyaan ganda beserta strategi mitigasinya.
- Sajikan tabulasi silang bersama batas interpretasi dan ketidakpastian.

Temuan yang baik tidak selalu berarti hasil yang besar atau sesuai harapan. Temuan bernilai ketika response rate terjaga, kesetaraan mode dijelaskan, dan ketidakpastian disajikan.

Kerangka Fowler (2014) serta Lohr (2021) menolong peneliti membedakan kekuatan bukti dari daya tarik narasi.

Inti metodologis subbagian ini adalah keterlacakan. Pembaca harus dapat mengikuti perjalanan dari pertanyaan, data, pengodean, analisis, sampai bobot survei. Ketika jejak itu jelas, temuan dapat digunakan untuk memperbaiki pembelajaran tanpa mengubah diagnosis konseptual menjadi label bagi peserta didik.

7.4.5 Desain Kuesioner dan Tes Terintegrasi

Dalam metodologi penelitian miskonsepsi kimia, desain kuesioner dan tes terintegrasi menentukan mutu hubungan antara data dan kesimpulan. Fokus ini berada dalam strategi pengumpulan data terstandar dari sampel yang relatif luas untuk mengestimasi distribusi miskonsepsi, keyakinan, pengalaman belajar, dan faktor terkait. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Eysenbach (2004) serta Fowler (2014) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Prosedur minimal mencakup definisi panjang instrumen, alasan memilih tes diagnostik terintegrasi, aturan mengubah respons menjadi kategori, serta mekanisme pemeriksaan validitas butir. Semua penyimpangan dari rencana awal perlu dicatat. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Pada penelitian mengenai survei daring konsep atom, peneliti dapat menyusun beberapa tugas ekuivalen yang berubah konteks, representasi, atau tuntutan penalarannya. Jika pola respons tetap berulang dan disertai keyakinan tinggi, terdapat dasar yang lebih kuat untuk menyebut indikasi miskonsepsi. Jika pola berubah-ubah, interpretasi yang lebih hati-hati adalah pengetahuan parsial atau ketergantungan konteks.

Langkah operasional yang disarankan:

- Definisikan panjang instrumen dalam istilah operasional.
- Tentukan bukti utama dari skala keyakinan dan bukti pendukung dari log waktu respons.
- Periksa kesetaraan mode sebelum membuat kategori atau model.
- Catat potensi ketidaksetaraan akses digital beserta strategi mitigasinya.
- Sajikan model asosiasi bersama batas interpretasi dan ketidakpastian.

Pemeriksaan mutu hendaknya dilakukan selama penelitian, bukan hanya pada akhir. Catatan tentang validitas butir, keputusan terkait kontrol duplikasi, dan kejadian yang berpotensi

menimbulkan opsi jawaban tidak lengkap perlu dipelihara dalam *audit trail*. Hal ini memungkinkan koreksi dini dan pelaporan yang jujur.

Kesimpulannya, desain kuesioner dan tes terintegrasi perlu dirancang sebagai proses argumentatif. Data tidak berbicara sendiri; maknanya bergantung pada definisi konstruk, kualitas prosedur, dan konteks. Oleh sebab itu, klaim akhir sebaiknya menggunakan bahasa yang sesuai dengan kekuatan bukti, misalnya “menunjukkan pola”, “mengindikasikan”, atau “mendukung penjelasan”.

7.4.6 Urutan, Tampilan, dan Beban Respons

Keputusan tentang urutan, tampilan, dan beban respons harus dibuat sebelum data utama dikumpulkan, bukan ditambahkan setelah hasil diketahui. Fokus ini berada dalam strategi pengumpulan data terstandar dari sampel yang relatif luas untuk mengestimasi distribusi miskonsepsi, keyakinan, pengalaman belajar, dan faktor terkait. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Cochran (1977) serta Lohr (2021) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Tahap kerja dimulai dengan penegasan urutan butir. Setelah itu, peneliti menyusun aturan untuk pengelolaan nonrespons dan menentukan bagaimana kerangka sampel akan diperiksa. Bukti sebaiknya berasal dari lebih dari satu bentuk, misalnya skala keyakinan dan log waktu respons. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Pada topik survei lintas jenjang asam-basa, peneliti juga perlu memperhatikan bahasa. Istilah yang sama dapat diberi makna berbeda oleh peserta didik. Klarifikasi melalui pertanyaan tindak lanjut, gambar, atau contoh tandingan membantu membedakan masalah terminologi dari kesalahan konseptual.

Keputusan yang harus didokumentasikan:

- Definisikan urutan butir dalam istilah operasional.
- Tentukan bukti utama dari data latar belakang dan bukti pendukung dari kuesioner daring.
- Periksa kontrol duplikasi sebelum membuat kategori atau model.
- Catat potensi interpretasi berlebihan terhadap self-report beserta strategi mitigasinya.
- Sajikan laporan mutu data bersama batas interpretasi dan ketidakpastian.

Kualitas penelitian meningkat ketika peneliti menunjukkan rantai bukti dari skala keyakinan menuju tabulasi silang, menyatakan asumsi, dan memeriksa alternatif penjelasan. Sebaliknya,

pertanyaan ganda dapat menghasilkan klaim yang tampak tegas tetapi tidak stabil. Rekomendasi praktis harus proporsional dengan bukti.

Pada akhirnya, keputusan tentang urutan, tampilan, dan beban respons menentukan apakah penelitian menghasilkan daftar jawaban salah atau pemahaman yang dapat ditindaklanjuti mengenai cara peserta didik berpikir. Sasaran utamanya ialah bukti yang cukup untuk merancang diagnosis, pembelajaran, dan penelitian lanjutan secara bertanggung jawab.

7.4.7 Uji Coba Kognitif Kuesioner

Uji coba kognitif kuesioner perlu dipahami sebagai keputusan metodologis yang mempunyai konsekuensi pada desain, analisis, dan pelaporan. Fokus ini berada dalam strategi pengumpulan data terstandar dari sampel yang relatif luas untuk mengestimasi distribusi miskonsepsi, keyakinan, pengalaman belajar, dan faktor terkait. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Lohr (2021) serta Fowler (2014) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Peneliti sebaiknya menyiapkan matriks yang menghubungkan pertanyaan penelitian dengan strategi kontak, data latar belakang, teknik analisis, dan model asosiasi. Matriks ini mencegah pengumpulan data yang banyak tetapi tidak relevan. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Ilustrasi survei mahasiswa kimia dasar juga menunjukkan bahwa kualitas instrumen memengaruhi temuan. Distraktor yang tidak berasal dari pemikiran peserta didik dapat menghasilkan pilihan acak, sedangkan pertanyaan yang terlalu menuntut matematika dapat mengukur keterampilan hitung alih-alih konsepsi. Peneliti perlu menilai kesesuaian bukti dengan konstruk yang dimaksud.

Pertanyaan audit metodologis:

- Definisikan strategi kontak dalam istilah operasional.
- Tentukan bukti utama dari log waktu respons dan bukti pendukung dari kuesioner kertas.
- Periksa analisis nonrespons sebelum membuat kategori atau model.
- Catat potensi sampel sukarela yang bias beserta strategi mitigasinya.
- Sajikan estimasi populasi bersama batas interpretasi dan ketidakpastian.

Peneliti sebaiknya memisahkan keputusan apriori dari keputusan eksploratori. Pemisahan tersebut membantu mengendalikan ketidaksetaraan akses digital dan mencegah survei terlalu panjang. Selanjutnya, kontrol duplikasi dan coverage dilaporkan bersama ketidakpastian,

sehingga pembaca dapat menilai kekuatan temuan tanpa bergantung pada kesimpulan penulis semata.

Dengan demikian, uji coba kognitif kuesioner berfungsi menghubungkan keputusan desain dengan interpretasi. Produk yang diharapkan bukan hanya model asosiasi, tetapi juga penjelasan mengenai bagaimana produk tersebut diperoleh dan dalam kondisi apa ia dapat digunakan. Penelitian miskonsepsi yang baik selalu menyisakan ruang bagi penjelasan alternatif dan verifikasi lanjutan.

7.4.8 Strategi Kontak dan Peningat

Ketelitian penelitian pada bagian strategi kontak dan peningat sangat menentukan apakah pola jawaban dapat ditafsirkan sebagai bukti konseptual. Fokus ini berada dalam strategi pengumpulan data terstandar dari sampel yang relatif luas untuk mengestimasi distribusi miskonsepsi, keyakinan, pengalaman belajar, dan faktor terkait. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. von Elm et al. (2007) serta Lohr (2021) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Dokumentasi metodologis hendaknya memuat siapa yang melakukan log waktu respons, kapan data diperoleh, bagaimana kerangka sampel dikendalikan, dan atas dasar apa laporan mutu data disusun. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Contoh pada kajian survei kelas unggulan dan reguler memperlihatkan pentingnya pemisahan antara ketidaktahuan, kesalahan prosedural, dan miskonsepsi. Seorang peserta dapat gagal karena lupa rumus, sedangkan peserta lain memilih jawaban yang sama karena menggunakan model mental yang tidak ilmiah. Perbedaan ini baru terlihat ketika log waktu respons dibaca bersama kuesioner kertas dan bukti tentang alasan.

Bukti minimum yang perlu tersedia:

- Definisikan pengelolaan nonrespons dalam istilah operasional.
- Tentukan bukti utama dari kuesioner daring dan bukti pendukung dari tes diagnostik terintegrasi.
- Periksa coverage sebelum membuat kategori atau model.
- Catat potensi survei terlalu panjang beserta strategi mitigasinya.
- Sajikan bobot survei bersama batas interpretasi dan ketidakpastian.

Standar metodologis menuntut koherensi antara konstruk, prosedur, analisis, dan klaim. Dalam konteks ini, analisis nonrespons tidak dapat digantikan oleh indeks tunggal. Demikian pula, response rate harus dinilai dengan bukti yang sesuai. von Elm et al. (2007) dan Lohr (2021) memberikan dasar untuk pelaporan yang lengkap dan dapat ditelaah.

Inti metodologis subbagian ini adalah keterlacakan. Pembaca harus dapat mengikuti perjalanan dari pertanyaan, data, pengodean, analisis, sampai laporan mutu data. Ketika jejak itu jelas, temuan dapat digunakan untuk memperbaiki pembelajaran tanpa mengubah diagnosis konseptual menjadi label bagi peserta didik.

7.4.9 Pengendalian Nonrespons

Peneliti sering memusatkan perhatian pada teknik, padahal pengendalian nonrespons justru mengatur batas inferensi yang dapat dibuat. Fokus ini berada dalam strategi pengumpulan data terstandar dari sampel yang relatif luas untuk mengestimasi distribusi miskonsepsi, keyakinan, pengalaman belajar, dan faktor terkait. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Dillman et al. (2014) serta Fowler (2014) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Rancangan yang dapat diaudit mencantumkan alasan pemilihan mode survei, prosedur penanganan panjang instrumen, dan kriteria keputusan untuk strategi kontak. Tanpa penjelasan itu, pembaca sulit menilai apakah kuesioner daring benar-benar mendukung klaim. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Kasus survei sekolah kota-desa dapat digunakan untuk memperlihatkan hubungan antara pola individual dan pola kelas. Sebagian miskonsepsi mungkin hanya dimiliki beberapa peserta tetapi sangat kuat, sedangkan pola lain tersebar luas karena bersumber dari analogi atau buku teks yang sama. Analisis perlu menjaga kedua tingkat tersebut.

Kriteria pemeriksaan sebelum analisis:

- Definisikan mode survei dalam istilah operasional.
- Tentukan bukti utama dari kuesioner kertas dan bukti pendukung dari skala keyakinan.
- Periksa response rate sebelum membuat kategori atau model.
- Catat potensi opsi jawaban tidak lengkap beserta strategi mitigasinya.
- Sajikan profil nonrespons bersama batas interpretasi dan ketidakpastian.

Pemeriksaan kualitas dapat dilakukan melalui audit terhadap coverage, analisis kasus yang bertentangan, dan evaluasi validitas butir. Apabila ditemukan sampel sukarela yang bias, peneliti perlu menurunkan kekuatan klaim atau menambah bukti. Transparansi seperti ini meningkatkan kegunaan temuan bagi penelitian lanjutan dan praktik kelas.

Kesimpulannya, pengendalian nonrespons perlu dirancang sebagai proses argumentatif. Data tidak berbicara sendiri; maknanya bergantung pada definisi konstruk, kualitas prosedur, dan konteks. Oleh sebab itu, klaim akhir sebaiknya menggunakan bahasa yang sesuai dengan kekuatan bukti, misalnya “menunjukkan pola”, “mengindikasikan”, atau “mendukung penjelasan”.

7.4.10 Survei Daring dan CHERRIES

Kajian yang kuat memperlakukan survei daring dan cherries sebagai bagian dari rantai argumentasi, bukan sebagai unsur pelengkap. Fokus ini berada dalam strategi pengumpulan data terstandar dari sampel yang relatif luas untuk mengestimasi distribusi miskonsepsi, keyakinan, pengalaman belajar, dan faktor terkait. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Fowler (2014) serta Lohr (2021) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Pada tahap perencanaan, kerangka sampel dan urutan butir harus dibedakan secara eksplisit. Pada tahap pelaksanaan, konsistensi kuesioner kertas perlu diperiksa; pada tahap interpretasi, response rate menjadi dasar pembatasan klaim. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Dalam kelas yang mempelajari survei setelah pembelajaran hibrida, data agregat mungkin menunjukkan mayoritas jawaban salah. Akan tetapi, wawancara dapat memperlihatkan beberapa mekanisme berbeda: bahasa sehari-hari, generalisasi aturan, salah membaca simbol, atau model partikel yang tidak tepat. Karena itu, satu persentase tidak boleh menggantikan analisis struktur alasan.

Langkah operasional yang disarankan:

- Definisikan kerangka sampel dalam istilah operasional.
- Tentukan bukti utama dari tes diagnostik terintegrasi dan bukti pendukung dari data latar belakang.
- Periksa validitas butir sebelum membuat kategori atau model.
- Catat potensi pertanyaan ganda beserta strategi mitigasinya.

- Sajikan tabulasi silang bersama batas interpretasi dan ketidakpastian.

Kriteria keberhasilan tidak cukup berupa banyaknya data. Yang lebih penting adalah apakah response rate dapat ditelusuri, apakah kesetaraan mode diperiksa, dan apakah analisis nonrespons dijelaskan. Peneliti harus menghindari survei terlalu panjang; jika risiko itu tidak dapat dihilangkan, dampaknya perlu dianalisis dan dicantumkan sebagai batas penelitian.

Pada akhirnya, keputusan tentang survei daring dan cherries menentukan apakah penelitian menghasilkan daftar jawaban salah atau pemahaman yang dapat ditindaklanjuti mengenai cara peserta didik berpikir. Sasaran utamanya ialah bukti yang cukup untuk merancang diagnosis, pembelajaran, dan penelitian lanjutan secara bertanggung jawab.

7.4.11 Pembobotan Data Survei

Subbagian ini menempatkan pembobotan data survei sebagai simpul penting dalam penelitian survei. Fokus ini berada dalam strategi pengumpulan data terstandar dari sampel yang relatif luas untuk mengestimasi distribusi miskonsepsi, keyakinan, pengalaman belajar, dan faktor terkait. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Eysenbach (2004) serta Fowler (2014) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Secara operasional, peneliti perlu menetapkan panjang instrumen, menjelaskan strategi kontak, dan mendokumentasikan mode survei. Keputusan tersebut kemudian diterjemahkan ke dalam pilihan tes diagnostik terintegrasi, data latar belakang, serta kuesioner daring. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Sebagai ilustrasi, pada topik survei miskonsepsi stoikiometri tingkat provinsi, peserta didik dapat memberikan jawaban benar pada tingkat simbolik tetapi alasan submikroskopiknya keliru. Peneliti tidak cukup menghitung benar-salah; ia perlu membandingkan tes diagnostik terintegrasi dengan data latar belakang, memeriksa tingkat keyakinan, dan menguji konsistensi pada konteks lain. Hasil analisis mungkin menunjukkan bahwa kesalahan hanya muncul pada satu representasi, atau justru merupakan hubungan konseptual yang stabil.

Keputusan yang harus didokumentasikan:

- Definisikan panjang instrumen dalam istilah operasional.
- Tentukan bukti utama dari skala keyakinan dan bukti pendukung dari log waktu respons.
- Periksa kesetaraan mode sebelum membuat kategori atau model.
- Catat potensi ketidaksetaraan akses digital beserta strategi mitigasinya.

- Sajikan model asosiasi bersama batas interpretasi dan ketidakpastian.

Mutu inferensi terutama ditentukan oleh validitas butir, kontrol duplikasi, dan coverage. Ancaman utama ialah opsi jawaban tidak lengkap serta interpretasi berlebihan terhadap self-report. Karena itu, laporan perlu menyajikan bukti yang mendukung sekaligus bukti yang membatasi interpretasi. Prinsip ini konsisten dengan penekanan Eysenbach (2004) dan Fowler (2014) bahwa validitas melekat pada interpretasi dan penggunaan hasil, bukan sekadar pada nama instrumen atau desain.

Dengan demikian, pembobotan data survei berfungsi menghubungkan keputusan desain dengan interpretasi. Produk yang diharapkan bukan hanya profil nonrespons, tetapi juga penjelasan mengenai bagaimana produk tersebut diperoleh dan dalam kondisi apa ia dapat digunakan. Penelitian miskonsepsi yang baik selalu menyisakan ruang bagi penjelasan alternatif dan verifikasi lanjutan.

7.4.12 Analisis Kesetaraan Mode

Pembahasan mengenai analisis kesetaraan mode diperlukan agar penelitian survei tidak berhenti pada prosedur administratif. Fokus ini berada dalam strategi pengumpulan data terstandar dari sampel yang relatif luas untuk mengestimasi distribusi miskonsepsi, keyakinan, pengalaman belajar, dan faktor terkait. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Cochran (1977) serta Lohr (2021) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Penerapan praktis dapat disusun dalam urutan: mendefinisikan urutan butir, mengoperasionalkan pengelolaan nonrespons, mengumpulkan skala keyakinan, memeriksa kesetaraan mode, lalu melaporkan tabulasi silang. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Misalnya, pada survei guru tentang sumber miskonsepsi, perubahan skor sesudah pembelajaran belum otomatis menandakan perubahan konsep. Peserta didik mungkin mempelajari pola jawaban tes. Bukti yang lebih meyakinkan mencakup transfer ke situasi baru, penjelasan kausal, representasi partikel, dan retensi setelah jeda waktu.

Pertanyaan audit metodologis:

- Definisikan urutan butir dalam istilah operasional.
- Tentukan bukti utama dari data latar belakang dan bukti pendukung dari kuesioner daring.
- Periksa kontrol duplikasi sebelum membuat kategori atau model.

- Catat potensi interpretasi berlebihan terhadap self-report beserta strategi mitigasinya.
- Sajikan laporan mutu data bersama batas interpretasi dan ketidakpastian.

Temuan yang baik tidak selalu berarti hasil yang besar atau sesuai harapan. Temuan bernilai ketika kesetaraan mode terjaga, analisis nonrespons dijelaskan, dan ketidakpastian disajikan. Kerangka Cochran (1977) serta Lohr (2021) menolong peneliti membedakan kekuatan bukti dari daya tarik narasi.

Inti metodologis subbagian ini adalah keterlacakan. Pembaca harus dapat mengikuti perjalanan dari pertanyaan, data, pengodean, analisis, sampai tabulasi silang. Ketika jejak itu jelas, temuan dapat digunakan untuk memperbaiki pembelajaran tanpa mengubah diagnosis konseptual menjadi label bagi peserta didik.

7.4.13 Survei Lintas Wilayah dan Budaya

Dalam metodologi penelitian miskonsepsi kimia, survei lintas wilayah dan budaya menentukan mutu hubungan antara data dan kesimpulan. Fokus ini berada dalam strategi pengumpulan data terstandar dari sampel yang relatif luas untuk mengestimasi distribusi miskonsepsi, keyakinan, pengalaman belajar, dan faktor terkait. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Lohr (2021) serta Fowler (2014) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Prosedur minimal mencakup definisi strategi kontak, alasan memilih data latar belakang, aturan mengubah respons menjadi kategori, serta mekanisme pemeriksaan kontrol duplikasi. Semua penyimpangan dari rencana awal perlu dicatat. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Pada penelitian mengenai survei daring konsep atom, peneliti dapat menyusun beberapa tugas ekuivalen yang berubah konteks, representasi, atau tuntutan penalarannya. Jika pola respons tetap berulang dan disertai keyakinan tinggi, terdapat dasar yang lebih kuat untuk menyebut indikasi miskonsepsi. Jika pola berubah-ubah, interpretasi yang lebih hati-hati adalah pengetahuan parsial atau ketergantungan konteks.

Bukti minimum yang perlu tersedia:

- Definisikan strategi kontak dalam istilah operasional.
- Tentukan bukti utama dari log waktu respons dan bukti pendukung dari kuesioner kertas.
- Periksa analisis nonrespons sebelum membuat kategori atau model.
- Catat potensi sampel sukarela yang bias beserta strategi mitigasinya.

- Sajikan estimasi populasi bersama batas interpretasi dan ketidakpastian.

Pemeriksaan mutu hendaknya dilakukan selama penelitian, bukan hanya pada akhir. Catatan tentang kontrol duplikasi, keputusan terkait coverage, dan kejadian yang berpotensi menimbulkan ketidaksetaraan akses digital perlu dipelihara dalam *audit trail*. Hal ini memungkinkan koreksi dini dan pelaporan yang jujur.

Kesimpulannya, survei lintas wilayah dan budaya perlu dirancang sebagai proses argumentatif. Data tidak berbicara sendiri; maknanya bergantung pada definisi konstruk, kualitas prosedur, dan konteks. Oleh sebab itu, klaim akhir sebaiknya menggunakan bahasa yang sesuai dengan kekuatan bukti, misalnya “menunjukkan pola”, “mengindikasikan”, atau “mendukung penjelasan”.

7.4.14 Pelaporan Survei yang Transparan

Keputusan tentang pelaporan survei yang transparan harus dibuat sebelum data utama dikumpulkan, bukan ditambahkan setelah hasil diketahui. Fokus ini berada dalam strategi pengumpulan data terstandar dari sampel yang relatif luas untuk mengestimasi distribusi miskonsepsi, keyakinan, pengalaman belajar, dan faktor terkait. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. von Elm et al. (2007) serta Lohr (2021) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Tahap kerja dimulai dengan penegasan pengelolaan nonrespons. Setelah itu, peneliti menyusun aturan untuk kerangka sampel dan menentukan bagaimana urutan butir akan diperiksa. Bukti sebaiknya berasal dari lebih dari satu bentuk, misalnya log waktu respons dan kuesioner kertas. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Pada topik survei lintas jenjang asam-basa, peneliti juga perlu memperhatikan bahasa. Istilah yang sama dapat diberi makna berbeda oleh peserta didik. Klarifikasi melalui pertanyaan tindak lanjut, gambar, atau contoh tandingan membantu membedakan masalah terminologi dari kesalahan konseptual.

Kriteria pemeriksaan sebelum analisis:

- Definisikan pengelolaan nonrespons dalam istilah operasional.
- Tentukan bukti utama dari kuesioner daring dan bukti pendukung dari tes diagnostik terintegrasi.
- Periksa coverage sebelum membuat kategori atau model.

- Catat potensi survei terlalu panjang beserta strategi mitigasinya.
- Sajikan bobot survei bersama batas interpretasi dan ketidakpastian.

Kualitas penelitian meningkat ketika peneliti menunjukkan rantai bukti dari log waktu respons menuju laporan mutu data, menyatakan asumsi, dan memeriksa alternatif penjelasan. Sebaliknya, interpretasi berlebihan terhadap self-report dapat menghasilkan klaim yang tampak tegas tetapi tidak stabil. Rekomendasi praktis harus proporsional dengan bukti.

Pada akhirnya, keputusan tentang pelaporan survei yang transparan menentukan apakah penelitian menghasilkan daftar jawaban salah atau pemahaman yang dapat ditindaklanjuti mengenai cara peserta didik berpikir. Sasaran utamanya ialah bukti yang cukup untuk merancang diagnosis, pembelajaran, dan penelitian lanjutan secara bertanggung jawab.

7.4.15 Sintesis Protokol Survei Miskonsepsi

Sintesis protokol survei miskonsepsi perlu dipahami sebagai keputusan metodologis yang mempunyai konsekuensi pada desain, analisis, dan pelaporan. Fokus ini berada dalam strategi pengumpulan data terstandar dari sampel yang relatif luas untuk mengestimasi distribusi miskonsepsi, keyakinan, pengalaman belajar, dan faktor terkait. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Dillman et al. (2014) serta Fowler (2014) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Peneliti sebaiknya menyiapkan matriks yang menghubungkan pertanyaan penelitian dengan mode survei, kuesioner daring, teknik analisis, dan estimasi populasi. Matriks ini mencegah pengumpulan data yang banyak tetapi tidak relevan. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Ilustrasi survei mahasiswa kimia dasar juga menunjukkan bahwa kualitas instrumen memengaruhi temuan. Distraktor yang tidak berasal dari pemikiran peserta didik dapat menghasilkan pilihan acak, sedangkan pertanyaan yang terlalu menuntut matematika dapat mengukur keterampilan hitung alih-alih konsepsi. Peneliti perlu menilai kesesuaian bukti dengan konstruk yang dimaksud.

Langkah operasional yang disarankan:

- Definisikan mode survei dalam istilah operasional.
- Tentukan bukti utama dari kuesioner kertas dan bukti pendukung dari skala keyakinan.
- Periksa response rate sebelum membuat kategori atau model.
- Catat potensi opsi jawaban tidak lengkap beserta strategi mitigasinya.

- Sajikan profil nonrespons bersama batas interpretasi dan ketidakpastian.

Peneliti sebaiknya memisahkan keputusan apriori dari keputusan eksploratori. Pemisahan tersebut membantu mengendalikan sampel sukarela yang bias dan mencegah pertanyaan ganda. Selanjutnya, coverage dan validitas butir dilaporkan bersama ketidakpastian, sehingga pembaca dapat menilai kekuatan temuan tanpa bergantung pada kesimpulan penulis semata.

Dengan demikian, sintesis protokol survei miskonsepsi berfungsi menghubungkan keputusan desain dengan interpretasi. Produk yang diharapkan bukan hanya estimasi populasi, tetapi juga penjelasan mengenai bagaimana produk tersebut diperoleh dan dalam kondisi apa ia dapat digunakan. Penelitian miskonsepsi yang baik selalu menyisakan ruang bagi penjelasan alternatif dan verifikasi lanjutan.

7.5 Penelitian Eksperimen dan Kuasi-Eksperimen

Subbab ini membahas desain untuk menguji apakah intervensi diagnostik atau remedial menyebabkan perubahan konsepsi, penalaran, keyakinan, transfer, dan retensi. Fokus utamanya adalah menjaga keselarasan antara pertanyaan penelitian, bentuk bukti, teknik analisis, dan kekuatan klaim. Dalam konteks miskonsepsi kimia, keputusan metodologis harus mampu membedakan pola konseptual dari respons yang timbul karena keterbatasan bahasa, kesalahan hitung, atau kondisi administrasi. Pembahasan berikut dipecah menjadi lima belas bagian agar setiap keputusan dapat ditelaah secara operasional.

7.5.1 Logika Kausal dalam Remediasi Miskonsepsi

Ketelitian penelitian pada bagian logika kausal dalam remediasi miskonsepsi sangat menentukan apakah pola jawaban dapat ditafsirkan sebagai bukti konseptual. Fokus ini berada dalam desain untuk menguji apakah intervensi diagnostik atau remedial menyebabkan perubahan konsepsi, penalaran, keyakinan, transfer, dan retensi. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Campbell dan Stanley (1963) serta Lakens (2022) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Dokumentasi metodologis hendaknya memuat siapa yang melakukan skor keyakinan, kapan data diperoleh, bagaimana kesetaraan awal dikendalikan, dan atas dasar apa ukuran efek disusun. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Contoh pada kajian laboratorium virtual redoks memperlihatkan pentingnya pemisahan antara ketidaktahuan, kesalahan prosedural, dan miskonsepsi. Seorang peserta dapat gagal karena lupa rumus, sedangkan peserta lain memilih jawaban yang sama karena menggunakan model mental yang tidak ilmiah. Perbedaan ini baru terlihat ketika skor keyakinan dibaca bersama tes retensi dan bukti tentang alasan.

Keputusan yang harus didokumentasikan:

- Definisikan kelompok pembanding dalam istilah operasional.
- Tentukan bukti utama dari tugas transfer dan bukti pendukung dari observasi fidelitas.
- Periksa validitas eksternal sebelum membuat kategori atau model.
- Catat potensi teaching to the test beserta strategi mitigasinya.
- Sajikan analisis sensitivitas bersama batas interpretasi dan ketidakpastian.

Standar metodologis menuntut koherensi antara konstruk, prosedur, analisis, dan klaim. Dalam konteks ini, validitas konstruk tidak dapat digantikan oleh indeks tunggal. Demikian pula, *power* harus dinilai dengan bukti yang sesuai. Campbell dan Stanley (1963) dan Lakens (2022) memberikan dasar untuk pelaporan yang lengkap dan dapat ditelaah.

Inti metodologis subbagian ini adalah keterlacakan. Pembaca harus dapat mengikuti perjalanan dari pertanyaan, data, pengodean, analisis, sampai ukuran efek. Ketika jejak itu jelas, temuan dapat digunakan untuk memperbaiki pembelajaran tanpa mengubah diagnosis konseptual menjadi label bagi peserta didik.

7.5.2 Eksperimen Acak Individual

Peneliti sering memusatkan perhatian pada teknik, padahal eksperimen acak individual justru mengatur batas inferensi yang dapat dibuat. Fokus ini berada dalam desain untuk menguji apakah intervensi diagnostik atau remedial menyebabkan perubahan konsepsi, penalaran, keyakinan, transfer, dan retensi. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Schulz et al. (2010) serta Campbell dan Stanley (1963) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Rancangan yang dapat diaudit mencantumkan alasan pemilihan waktu pengukuran, prosedur penanganan fidelitas implementasi, dan kriteria keputusan untuk unit alokasi. Tanpa penjelasan itu, pembaca sulit menilai apakah tugas transfer benar-benar mendukung klaim. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Kasus peta konsep pada stoikiometri dapat digunakan untuk memperlihatkan hubungan antara pola individual dan pola kelas. Sebagian miskonsepsi mungkin hanya dimiliki beberapa peserta tetapi sangat kuat, sedangkan pola lain tersebar luas karena bersumber dari analogi atau buku teks yang sama. Analisis perlu menjaga kedua tingkat tersebut.

Pertanyaan audit metodologis:

- Definisikan waktu pengukuran dalam istilah operasional.
- Tentukan bukti utama dari tes retensi dan bukti pendukung dari log aktivitas belajar.
- Periksa power sebelum membuat kategori atau model.
- Catat potensi perbedaan instruktur beserta strategi mitigasinya.
- Sajikan profil responder bersama batas interpretasi dan ketidakpastian.

Pemeriksaan kualitas dapat dilakukan melalui audit terhadap validitas eksternal, analisis kasus yang bertentangan, dan evaluasi kepatuhan intervensi. Apabila ditemukan kontaminasi kelompok, peneliti perlu menurunkan kekuatan klaim atau menambah bukti. Transparansi seperti ini meningkatkan kegunaan temuan bagi penelitian lanjutan dan praktik kelas.

Kesimpulannya, eksperimen acak individual perlu dirancang sebagai proses argumentatif. Data tidak berbicara sendiri; maknanya bergantung pada definisi konstruk, kualitas prosedur, dan konteks. Oleh sebab itu, klaim akhir sebaiknya menggunakan bahasa yang sesuai dengan kekuatan bukti, misalnya “menunjukkan pola”, “mengindikasikan”, atau “mendukung penjelasan”.

7.5.3 Eksperimen Acak Berkelompok

Kajian yang kuat memperlakukan eksperimen acak berkelompok sebagai bagian dari rantai argumentasi, bukan sebagai unsur pelengkap. Fokus ini berada dalam desain untuk menguji apakah intervensi diagnostik atau remedial menyebabkan perubahan konsepsi, penalaran, keyakinan, transfer, dan retensi. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Des Jarlais et al. (2004) serta Lakens (2022) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Pada tahap perencanaan, kesetaraan awal dan estimand efek harus dibedakan secara eksplisit. Pada tahap pelaksanaan, konsistensi tes retensi perlu diperiksa; pada tahap interpretasi, power menjadi dasar pembatasan klaim. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Dalam kelas yang mempelajari tutorial metakognitif pada mol, data agregat mungkin menunjukkan mayoritas jawaban salah. Akan tetapi, wawancara dapat memperlihatkan beberapa mekanisme berbeda: bahasa sehari-hari, generalisasi aturan, salah membaca simbol, atau model partikel yang tidak tepat. Karena itu, satu persentase tidak boleh menggantikan analisis struktur alasan.

Bukti minimum yang perlu tersedia:

- Definisikan kesetaraan awal dalam istilah operasional.
- Tentukan bukti utama dari observasi fidelitas dan bukti pendukung dari pretest-posttest diagnostik.
- Periksa kepatuhan intervensi sebelum membuat kategori atau model.
- Catat potensi mengandalkan signifikansi saja beserta strategi mitigasinya.
- Sajikan laporan fidelitas bersama batas interpretasi dan ketidakpastian.

Kriteria keberhasilan tidak cukup berupa banyaknya data. Yang lebih penting adalah apakah *power* dapat ditelusuri, apakah analisis attrition diperiksa, dan apakah validitas konstruk dijelaskan. Peneliti harus menghindari *teaching to the test*; jika risiko itu tidak dapat dihilangkan, dampaknya perlu dianalisis dan dicantumkan sebagai batas penelitian.

Pada akhirnya, keputusan tentang eksperimen acak berkelompok menentukan apakah penelitian menghasilkan daftar jawaban salah atau pemahaman yang dapat ditindaklanjuti mengenai cara peserta didik berpikir. Sasaran utamanya ialah bukti yang cukup untuk merancang diagnosis, pembelajaran, dan penelitian lanjutan secara bertanggung jawab.

7.5.4 Desain Pretest-Posttest dengan Kelompok Kontrol

Subbagian ini menempatkan desain pretest-posttest dengan kelompok kontrol sebagai simpul penting dalam penelitian eksperimen dan kuasi-eksperimen. Fokus ini berada dalam desain untuk menguji apakah intervensi diagnostik atau remedial menyebabkan perubahan konsepsi, penalaran, keyakinan, transfer, dan retensi. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Lakens (2022) serta Campbell dan Stanley (1963) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Secara operasional, peneliti perlu menetapkan fidelitas implementasi, menjelaskan unit alokasi, dan mendokumentasikan waktu pengukuran. Keputusan tersebut kemudian diterjemahkan ke dalam pilihan observasi fidelitas, pretest-posttest diagnostik, serta tugas transfer. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu

peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Sebagai ilustrasi, pada topik teks perubahan konseptual pada kalor, peserta didik dapat memberikan jawaban benar pada tingkat simbolik tetapi alasan submikroskopiknya keliru. Peneliti tidak cukup menghitung benar-salah; ia perlu membandingkan observasi fidelitas dengan *pretest-posttest* diagnostik, memeriksa tingkat keyakinan, dan menguji konsistensi pada konteks lain. Hasil analisis mungkin menunjukkan bahwa kesalahan hanya muncul pada satu representasi, atau justru merupakan hubungan konseptual yang stabil.

Kriteria pemeriksaan sebelum analisis:

- Definisikan fidelitas implementasi dalam istilah operasional.
- Tentukan bukti utama dari log aktivitas belajar dan bukti pendukung dari skor keyakinan.
- Periksa analisis attrition sebelum membuat kategori atau model.
- Catat potensi regresi menuju rerata beserta strategi mitigasinya.
- Sajikan perbedaan perubahan bersama batas interpretasi dan ketidakpastian.

Mutu inferensi terutama ditentukan oleh kepatuhan intervensi, validitas internal, dan validitas eksternal. Ancaman utama ialah perbedaan instruktur serta history dan maturation. Karena itu, laporan perlu menyajikan bukti yang mendukung sekaligus bukti yang membatasi interpretasi. Prinsip ini konsisten dengan penekanan Lakens (2022) dan Campbell dan Stanley (1963) bahwa validitas melekat pada interpretasi dan penggunaan hasil, bukan sekadar pada nama instrumen atau desain.

Dengan demikian, desain *pretest-posttest* dengan kelompok kontrol berfungsi menghubungkan keputusan desain dengan interpretasi. Produk yang diharapkan bukan hanya profil responder, tetapi juga penjelasan mengenai bagaimana produk tersebut diperoleh dan dalam kondisi apa ia dapat digunakan. Penelitian miskonsepsi yang baik selalu menyisakan ruang bagi penjelasan alternatif dan verifikasi lanjutan.

7.5.5 Kuasi-Eksperimen Kelompok Tidak Ekuivalen

Pembahasan mengenai kuasi-eksperimen kelompok tidak ekuivalen diperlukan agar penelitian eksperimen dan kuasi-eksperimen tidak berhenti pada prosedur administratif. Fokus ini berada dalam desain untuk menguji apakah intervensi diagnostik atau remedial menyebabkan perubahan konsepsi, penalaran, keyakinan, transfer, dan retensi. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Cumming (2014) serta Lakens (2022) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Penerapan praktis dapat disusun dalam urutan: mendefinisikan estimand efek, mengoperasionalkan kelompok pembanding, mengumpulkan log aktivitas belajar, memeriksa

analisis attrition, lalu melaporkan laporan fidelitas. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Misalnya, pada simulasi partikel untuk gas, perubahan skor sesudah pembelajaran belum otomatis menandakan perubahan konsep. Peserta didik mungkin mempelajari pola jawaban tes. Bukti yang lebih meyakinkan mencakup transfer ke situasi baru, penjelasan kausal, representasi partikel, dan retensi setelah jeda waktu.

Langkah operasional yang disarankan:

- Definisikan estimand efek dalam istilah operasional.
- Tentukan bukti utama dari pretest-posttest diagnostik dan bukti pendukung dari tugas transfer.
- Periksa validitas internal sebelum membuat kategori atau model.
- Catat potensi history dan maturation beserta strategi mitigasinya.
- Sajikan ukuran efek bersama batas interpretasi dan ketidakpastian.

Temuan yang baik tidak selalu berarti hasil yang besar atau sesuai harapan. Temuan bernilai ketika analisis attrition terjaga, validitas konstruk dijelaskan, dan ketidakpastian disajikan. Kerangka Cumming (2014) serta Lakens (2022) menolong peneliti membedakan kekuatan bukti dari daya tarik narasi.

Inti metodologis subbagian ini adalah keterlacakan. Pembaca harus dapat mengikuti perjalanan dari pertanyaan, data, pengodean, analisis, sampai laporan fidelitas. Ketika jejak itu jelas, temuan dapat digunakan untuk memperbaiki pembelajaran tanpa mengubah diagnosis konseptual menjadi label bagi peserta didik.

7.5.6 Desain Deret Waktu Terputus

Dalam metodologi penelitian miskonsepsi kimia, desain deret waktu terputus menentukan mutu hubungan antara data dan kesimpulan. Fokus ini berada dalam desain untuk menguji apakah intervensi diagnostik atau remedial menyebabkan perubahan konsepsi, penalaran, keyakinan, transfer, dan retensi. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Shadish et al. (2002) serta Campbell dan Stanley (1963) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Prosedur minimal mencakup definisi unit alokasi, alasan memilih pretest-posttest diagnostik, aturan mengubah respons menjadi kategori, serta mekanisme pemeriksaan validitas internal. Semua penyimpangan dari rencana awal perlu dicatat. Setiap keputusan dicatat dalam

protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Pada penelitian mengenai refutational text pada ikatan, peneliti dapat menyusun beberapa tugas ekuivalen yang berubah konteks, representasi, atau tuntutan penalarannya. Jika pola respons tetap berulang dan disertai keyakinan tinggi, terdapat dasar yang lebih kuat untuk menyebut indikasi miskonsepsi. Jika pola berubah-ubah, interpretasi yang lebih hati-hati adalah pengetahuan parsial atau ketergantungan konteks.

Keputusan yang harus didokumentasikan:

- Definisikan unit alokasi dalam istilah operasional.
- Tentukan bukti utama dari skor keyakinan dan bukti pendukung dari tes retensi.
- Periksa validitas konstruk sebelum membuat kategori atau model.
- Catat potensi kontaminasi kelompok beserta strategi mitigasinya.
- Sajikan interval kepercayaan bersama batas interpretasi dan ketidakpastian.

Pemeriksaan mutu hendaknya dilakukan selama penelitian, bukan hanya pada akhir. Catatan tentang validitas internal, keputusan terkait validitas eksternal, dan kejadian yang berpotensi menimbulkan regresi menuju rerata perlu dipelihara dalam *audit trail*. Hal ini memungkinkan koreksi dini dan pelaporan yang jujur.

Kesimpulannya, desain deret waktu terputus perlu dirancang sebagai proses argumentatif. Data tidak berbicara sendiri; maknanya bergantung pada definisi konstruk, kualitas prosedur, dan konteks. Oleh sebab itu, klaim akhir sebaiknya menggunakan bahasa yang sesuai dengan kekuatan bukti, misalnya “menunjukkan pola”, “mengindikasikan”, atau “mendukung penjelasan”.

7.5.7 Pemilihan Intervensi dan Pembandingan

Keputusan tentang pemilihan intervensi dan pembandingan harus dibuat sebelum data utama dikumpulkan, bukan ditambahkan setelah hasil diketahui. Fokus ini berada dalam desain untuk menguji apakah intervensi diagnostik atau remedial menyebabkan perubahan konsepsi, penalaran, keyakinan, transfer, dan retensi. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Campbell dan Stanley (1963) serta Lakens (2022) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Tahap kerja dimulai dengan penegasan kelompok pembandingan. Setelah itu, peneliti menyusun aturan untuk kesetaraan awal dan menentukan bagaimana estimand efek akan diperiksa. Bukti sebaiknya berasal dari lebih dari satu bentuk, misalnya skor keyakinan dan tes retensi.

Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Pada topik POE pada kesetimbangan, peneliti juga perlu memperhatikan bahasa. Istilah yang sama dapat diberi makna berbeda oleh peserta didik. Klarifikasi melalui pertanyaan tindak lanjut, gambar, atau contoh tandingan membantu membedakan masalah terminologi dari kesalahan konseptual.

Pertanyaan audit metodologis:

- Definisikan kelompok pembanding dalam istilah operasional.
- Tentukan bukti utama dari tugas transfer dan bukti pendukung dari observasi fidelitas.
- Periksa validitas eksternal sebelum membuat kategori atau model.
- Catat potensi teaching to the test beserta strategi mitigasinya.
- Sajikan analisis sensitivitas bersama batas interpretasi dan ketidakpastian.

Kualitas penelitian meningkat ketika peneliti menunjukkan rantai bukti dari skor keyakinan menuju ukuran efek, menyatakan asumsi, dan memeriksa alternatif penjelasan. Sebaliknya, history dan maturation dapat menghasilkan klaim yang tampak tegas tetapi tidak stabil. Rekomendasi praktis harus proporsional dengan bukti.

Pada akhirnya, keputusan tentang pemilihan intervensi dan pembanding menentukan apakah penelitian menghasilkan daftar jawaban salah atau pemahaman yang dapat ditindaklanjuti mengenai cara peserta didik berpikir. Sasaran utamanya ialah bukti yang cukup untuk merancang diagnosis, pembelajaran, dan penelitian lanjutan secara bertanggung jawab.

7.5.8 Pengukuran Perubahan Konsepsi

Pengukuran perubahan konsepsi perlu dipahami sebagai keputusan metodologis yang mempunyai konsekuensi pada desain, analisis, dan pelaporan. Fokus ini berada dalam desain untuk menguji apakah intervensi diagnostik atau remedial menyebabkan perubahan konsepsi, penalaran, keyakinan, transfer, dan retensi. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Schulz et al. (2010) serta Campbell dan Stanley (1963) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Peneliti sebaiknya menyiapkan matriks yang menghubungkan pertanyaan penelitian dengan waktu pengukuran, tugas transfer, teknik analisis, dan interval kepercayaan. Matriks ini mencegah pengumpulan data yang banyak tetapi tidak relevan. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda,

prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Ilustrasi analogi terarah pada asam-basa juga menunjukkan bahwa kualitas instrumen memengaruhi temuan. Distraktor yang tidak berasal dari pemikiran peserta didik dapat menghasilkan pilihan acak, sedangkan pertanyaan yang terlalu menuntut matematika dapat mengukur keterampilan hitung alih-alih konsepsi. Peneliti perlu menilai kesesuaian bukti dengan konstruk yang dimaksud.

Bukti minimum yang perlu tersedia:

- Definisikan waktu pengukuran dalam istilah operasional.
- Tentukan bukti utama dari tes retensi dan bukti pendukung dari log aktivitas belajar.
- Periksa power sebelum membuat kategori atau model.
- Catat potensi perbedaan instruktur beserta strategi mitigasinya.
- Sajikan profil responder bersama batas interpretasi dan ketidakpastian.

Peneliti sebaiknya memisahkan keputusan apriori dari keputusan eksploratori. Pemisahan tersebut membantu mengendalikan kontaminasi kelompok dan mencegah mengandalkan signifikansi saja. Selanjutnya, validitas eksternal dan kepatuhan intervensi dilaporkan bersama ketidakpastian, sehingga pembaca dapat menilai kekuatan temuan tanpa bergantung pada kesimpulan penulis semata.

Dengan demikian, pengukuran perubahan konsepsi berfungsi menghubungkan keputusan desain dengan interpretasi. Produk yang diharapkan bukan hanya interval kepercayaan, tetapi juga penjelasan mengenai bagaimana produk tersebut diperoleh dan dalam kondisi apa ia dapat digunakan. Penelitian miskonsepsi yang baik selalu menyisakan ruang bagi penjelasan alternatif dan verifikasi lanjutan.

7.5.9 Pengukuran Transfer dan Retensi

Ketelitian penelitian pada bagian pengukuran transfer dan retensi sangat menentukan apakah pola jawaban dapat ditafsirkan sebagai bukti konseptual. Fokus ini berada dalam desain untuk menguji apakah intervensi diagnostik atau remedial menyebabkan perubahan konsepsi, penalaran, keyakinan, transfer, dan retensi. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Des Jarlais et al. (2004) serta Lakens (2022) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Dokumentasi metodologis hendaknya memuat siapa yang melakukan tes retensi, kapan data diperoleh, bagaimana estimand efek dikendalikan, dan atas dasar apa analisis sensitivitas disusun. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan

pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Contoh pada kajian laboratorium virtual redoks memperlihatkan pentingnya pemisahan antara ketidaktahuan, kesalahan prosedural, dan miskonsepsi. Seorang peserta dapat gagal karena lupa rumus, sedangkan peserta lain memilih jawaban yang sama karena menggunakan model mental yang tidak ilmiah. Perbedaan ini baru terlihat ketika tes retensi dibaca bersama log aktivitas belajar dan bukti tentang alasan.

Kriteria pemeriksaan sebelum analisis:

- Definisikan kesetaraan awal dalam istilah operasional.
- Tentukan bukti utama dari observasi fidelitas dan bukti pendukung dari pretest-posttest diagnostik.
- Periksa kepatuhan intervensi sebelum membuat kategori atau model.
- Catat potensi mengandalkan signifikansi saja beserta strategi mitigasinya.
- Sajikan laporan fidelitas bersama batas interpretasi dan ketidakpastian.

Standar metodologis menuntut koherensi antara konstruk, prosedur, analisis, dan klaim. Dalam konteks ini, *power* tidak dapat digantikan oleh indeks tunggal. Demikian pula, analisis attrition harus dinilai dengan bukti yang sesuai. Des Jarlais et al. (2004) dan Lakens (2022) memberikan dasar untuk pelaporan yang lengkap dan dapat ditelaah.

Inti metodologis subbagian ini adalah keterlacakan. Pembaca harus dapat mengikuti perjalanan dari pertanyaan, data, pengodean, analisis, sampai analisis sensitivitas. Ketika jejak itu jelas, temuan dapat digunakan untuk memperbaiki pembelajaran tanpa mengubah diagnosis konseptual menjadi label bagi peserta didik.

7.5.10 Fidelitas Implementasi

Peneliti sering memusatkan perhatian pada teknik, padahal fidelitas implementasi justru mengatur batas inferensi yang dapat dibuat. Fokus ini berada dalam desain untuk menguji apakah intervensi diagnostik atau remedial menyebabkan perubahan konsepsi, penalaran, keyakinan, transfer, dan retensi. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Lakens (2022) serta Campbell dan Stanley (1963) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Rancangan yang dapat diaudit mencantumkan alasan pemilihan fidelitas implementasi, prosedur penanganan unit alokasi, dan kriteria keputusan untuk waktu pengukuran. Tanpa penjelasan itu, pembaca sulit menilai apakah observasi fidelitas benar-benar mendukung

klaim. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Kasus peta konsep pada stoikiometri dapat digunakan untuk memperlihatkan hubungan antara pola individual dan pola kelas. Sebagian miskonsepsi mungkin hanya dimiliki beberapa peserta tetapi sangat kuat, sedangkan pola lain tersebar luas karena bersumber dari analogi atau buku teks yang sama. Analisis perlu menjaga kedua tingkat tersebut.

Langkah operasional yang disarankan:

- Definisikan fidelitas implementasi dalam istilah operasional.
- Tentukan bukti utama dari log aktivitas belajar dan bukti pendukung dari skor keyakinan.
- Periksa analisis attrition sebelum membuat kategori atau model.
- Catat potensi regresi menuju rerata beserta strategi mitigasinya.
- Sajikan perbedaan perubahan bersama batas interpretasi dan ketidakpastian.

Pemeriksaan kualitas dapat dilakukan melalui audit terhadap kepatuhan intervensi, analisis kasus yang bertentangan, dan evaluasi validitas internal. Apabila ditemukan perbedaan instruktur, peneliti perlu menurunkan kekuatan klaim atau menambah bukti. Transparansi seperti ini meningkatkan kegunaan temuan bagi penelitian lanjutan dan praktik kelas.

Kesimpulannya, fidelitas implementasi perlu dirancang sebagai proses argumentatif. Data tidak berbicara sendiri; maknanya bergantung pada definisi konstruk, kualitas prosedur, dan konteks. Oleh sebab itu, klaim akhir sebaiknya menggunakan bahasa yang sesuai dengan kekuatan bukti, misalnya “menunjukkan pola”, “mengindikasikan”, atau “mendukung penjelasan”.

7.5.11 Ancaman Validitas Internal

Kajian yang kuat memperlakukan ancaman validitas internal sebagai bagian dari rantai argumentasi, bukan sebagai unsur pelengkap. Fokus ini berada dalam desain untuk menguji apakah intervensi diagnostik atau remedial menyebabkan perubahan konsepsi, penalaran, keyakinan, transfer, dan retensi. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Cumming (2014) serta Lakens (2022) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Pada tahap perencanaan, estimand efek dan kelompok pembanding harus dibedakan secara eksplisit. Pada tahap pelaksanaan, konsistensi log aktivitas belajar perlu diperiksa; pada tahap interpretasi, analisis attrition menjadi dasar pembatasan klaim. Setiap keputusan dicatat

dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Dalam kelas yang mempelajari tutorial metakognitif pada mol, data agregat mungkin menunjukkan mayoritas jawaban salah. Akan tetapi, wawancara dapat memperlihatkan beberapa mekanisme berbeda: bahasa sehari-hari, generalisasi aturan, salah membaca simbol, atau model partikel yang tidak tepat. Karena itu, satu persentase tidak boleh menggantikan analisis struktur alasan.

Keputusan yang harus didokumentasikan:

- Definisikan estimand efek dalam istilah operasional.
- Tentukan bukti utama dari pretest-posttest diagnostik dan bukti pendukung dari tugas transfer.
- Periksa validitas internal sebelum membuat kategori atau model.
- Catat potensi history dan maturation beserta strategi mitigasinya.
- Sajikan ukuran efek bersama batas interpretasi dan ketidakpastian.

Kriteria keberhasilan tidak cukup berupa banyaknya data. Yang lebih penting adalah apakah analisis attrition dapat ditelusuri, apakah validitas konstruk diperiksa, dan apakah *power* dijelaskan. Peneliti harus menghindari mengandalkan signifikansi saja; jika risiko itu tidak dapat dihilangkan, dampaknya perlu dianalisis dan dicantumkan sebagai batas penelitian.

Pada akhirnya, keputusan tentang ancaman validitas internal menentukan apakah penelitian menghasilkan daftar jawaban salah atau pemahaman yang dapat ditindaklanjuti mengenai cara peserta didik berpikir. Sasaran utamanya ialah bukti yang cukup untuk merancang diagnosis, pembelajaran, dan penelitian lanjutan secara bertanggung jawab.

7.5.12 Ancaman Validitas Eksternal

Subbagian ini menempatkan ancaman validitas eksternal sebagai simpul penting dalam penelitian eksperimen dan kuasi-eksperimen. Fokus ini berada dalam desain untuk menguji apakah intervensi diagnostik atau remedial menyebabkan perubahan konsepsi, penalaran, keyakinan, transfer, dan retensi. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Shadish et al. (2002) serta Campbell dan Stanley (1963) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Secara operasional, peneliti perlu menetapkan unit alokasi, menjelaskan waktu pengukuran, dan mendokumentasikan fidelitas implementasi. Keputusan tersebut kemudian diterjemahkan ke dalam pilihan pretest-posttest diagnostik, tugas transfer, serta observasi

fidelitas. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Sebagai ilustrasi, pada topik teks perubahan konseptual pada kalor, peserta didik dapat memberikan jawaban benar pada tingkat simbolik tetapi alasan submikroskopiknya keliru. Peneliti tidak cukup menghitung benar-salah; ia perlu membandingkan *pretest-posttest* diagnostik dengan tugas transfer, memeriksa tingkat keyakinan, dan menguji konsistensi pada konteks lain. Hasil analisis mungkin menunjukkan bahwa kesalahan hanya muncul pada satu representasi, atau justru merupakan hubungan konseptual yang stabil.

Pertanyaan audit metodologis:

- Definisikan unit alokasi dalam istilah operasional.
- Tentukan bukti utama dari skor keyakinan dan bukti pendukung dari tes retensi.
- Periksa validitas konstruk sebelum membuat kategori atau model.
- Catat potensi kontaminasi kelompok beserta strategi mitigasinya.
- Sajikan interval kepercayaan bersama batas interpretasi dan ketidakpastian.

Mutu inferensi terutama ditentukan oleh validitas internal, validitas eksternal, dan kepatuhan intervensi. Ancaman utama ialah regresi menuju rerata serta *teaching to the test*. Karena itu, laporan perlu menyajikan bukti yang mendukung sekaligus bukti yang membatasi interpretasi. Prinsip ini konsisten dengan penekanan Shadish et al. (2002) dan Campbell dan Stanley (1963) bahwa validitas melekat pada interpretasi dan penggunaan hasil, bukan sekadar pada nama instrumen atau desain.

Dengan demikian, ancaman validitas eksternal berfungsi menghubungkan keputusan desain dengan interpretasi. Produk yang diharapkan bukan hanya perbedaan perubahan, tetapi juga penjelasan mengenai bagaimana produk tersebut diperoleh dan dalam kondisi apa ia dapat digunakan. Penelitian miskonsepsi yang baik selalu menyisakan ruang bagi penjelasan alternatif dan verifikasi lanjutan.

7.5.13 Analisis Ukuran Efek dan Ketidakpastian

Pembahasan mengenai analisis ukuran efek dan ketidakpastian diperlukan agar penelitian eksperimen dan kuasi-eksperimen tidak berhenti pada prosedur administratif. Fokus ini berada dalam desain untuk menguji apakah intervensi diagnostik atau remedial menyebabkan perubahan konsepsi, penalaran, keyakinan, transfer, dan retensi. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Campbell

dan Stanley (1963) serta Lakens (2022) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Penerapan praktis dapat disusun dalam urutan: mendefinisikan kelompok pembanding, mengoperasionalkan kesetaraan awal, mengumpulkan skor keyakinan, memeriksa validitas konstruk, lalu melaporkan ukuran efek. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

$$d = (M_1 - M_2)/SD_{pooled}; g = J \times d$$

Misalnya, pada simulasi partikel untuk gas, perubahan skor sesudah pembelajaran belum otomatis menandakan perubahan konsep. Peserta didik mungkin mempelajari pola jawaban tes. Bukti yang lebih meyakinkan mencakup transfer ke situasi baru, penjelasan kausal, representasi partikel, dan retensi setelah jeda waktu.

Bukti minimum yang perlu tersedia:

- Definisikan kelompok pembanding dalam istilah operasional.
- Tentukan bukti utama dari tugas transfer dan bukti pendukung dari observasi fidelitas.
- Periksa validitas eksternal sebelum membuat kategori atau model.
- Catat potensi teaching to the test beserta strategi mitigasinya.
- Sajikan analisis sensitivitas bersama batas interpretasi dan ketidakpastian.

Temuan yang baik tidak selalu berarti hasil yang besar atau sesuai harapan. Temuan bernilai ketika validitas konstruk terjaga, *power* dijelaskan, dan ketidakpastian disajikan. Kerangka Campbell dan Stanley (1963) serta Lakens (2022) menolong peneliti membedakan kekuatan bukti dari daya tarik narasi.

Inti metodologis subbagian ini adalah keterlacakan. Pembaca harus dapat mengikuti perjalanan dari pertanyaan, data, pengodean, analisis, sampai ukuran efek. Ketika jejak itu jelas, temuan dapat digunakan untuk memperbaiki pembelajaran tanpa mengubah diagnosis konseptual menjadi label bagi peserta didik.

7.5.14 Pelaporan CONSORT dan TREND

Dalam metodologi penelitian miskonsepsi kimia, pelaporan consort dan trend menentukan mutu hubungan antara data dan kesimpulan. Fokus ini berada dalam desain untuk menguji apakah intervensi diagnostik atau remedial menyebabkan perubahan konsepsi, penalaran, keyakinan, transfer, dan retensi. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Schulz et al. (2010) serta Campbell dan

Stanley (1963) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Prosedur minimal mencakup definisi waktu pengukuran, alasan memilih tugas transfer, aturan mengubah respons menjadi kategori, serta mekanisme pemeriksaan validitas eksternal. Semua penyimpangan dari rencana awal perlu dicatat. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Pada penelitian mengenai refutational text pada ikatan, peneliti dapat menyusun beberapa tugas ekuivalen yang berubah konteks, representasi, atau tuntutan penalarannya. Jika pola respons tetap berulang dan disertai keyakinan tinggi, terdapat dasar yang lebih kuat untuk menyebut indikasi miskonsepsi. Jika pola berubah-ubah, interpretasi yang lebih hati-hati adalah pengetahuan parsial atau ketergantungan konteks.

Kriteria pemeriksaan sebelum analisis:

- Definisikan waktu pengukuran dalam istilah operasional.
- Tentukan bukti utama dari tes retensi dan bukti pendukung dari log aktivitas belajar.
- Periksa power sebelum membuat kategori atau model.
- Catat potensi perbedaan instruktur beserta strategi mitigasinya.
- Sajikan profil responder bersama batas interpretasi dan ketidakpastian.

Pemeriksaan mutu hendaknya dilakukan selama penelitian, bukan hanya pada akhir. Catatan tentang validitas eksternal, keputusan terkait kepatuhan intervensi, dan kejadian yang berpotensi menimbulkan kontaminasi kelompok perlu dipelihara dalam *audit trail*. Hal ini memungkinkan koreksi dini dan pelaporan yang jujur.

Kesimpulannya, pelaporan consort dan trend perlu dirancang sebagai proses argumentatif. Data tidak berbicara sendiri; maknanya bergantung pada definisi konstruk, kualitas prosedur, dan konteks. Oleh sebab itu, klaim akhir sebaiknya menggunakan bahasa yang sesuai dengan kekuatan bukti, misalnya “menunjukkan pola”, “mengindikasikan”, atau “mendukung penjelasan”.

7.5.15 Sintesis Desain Eksperimental Miskonsepsi

Keputusan tentang sintesis desain eksperimental miskonsepsi harus dibuat sebelum data utama dikumpulkan, bukan ditambahkan setelah hasil diketahui. Fokus ini berada dalam desain untuk menguji apakah intervensi diagnostik atau remedial menyebabkan perubahan konsepsi, penalaran, keyakinan, transfer, dan retensi. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Des Jarlais

et al. (2004) serta Lakens (2022) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Tahap kerja dimulai dengan penegasan kesetaraan awal. Setelah itu, peneliti menyusun aturan untuk estimand efek dan menentukan bagaimana kelompok pembanding akan diperiksa. Bukti sebaiknya berasal dari lebih dari satu bentuk, misalnya tes retensi dan log aktivitas belajar. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Pada topik POE pada kesetimbangan, peneliti juga perlu memperhatikan bahasa. Istilah yang sama dapat diberi makna berbeda oleh peserta didik. Klarifikasi melalui pertanyaan tindak lanjut, gambar, atau contoh tandingan membantu membedakan masalah terminologi dari kesalahan konseptual.

Langkah operasional yang disarankan:

- Definisikan kesetaraan awal dalam istilah operasional.
- Tentukan bukti utama dari observasi fidelitas dan bukti pendukung dari pretest-posttest diagnostik.
- Periksa kepatuhan intervensi sebelum membuat kategori atau model.
- Catat potensi mengandalkan signifikansi saja beserta strategi mitigasinya.
- Sajikan laporan fidelitas bersama batas interpretasi dan ketidakpastian.

Kualitas penelitian meningkat ketika peneliti menunjukkan rantai bukti dari tes retensi menuju analisis sensitivitas, menyatakan asumsi, dan memeriksa alternatif penjelasan. Sebaliknya, *teaching to the test* dapat menghasilkan klaim yang tampak tegas tetapi tidak stabil. Rekomendasi praktis harus proporsional dengan bukti.

Pada akhirnya, keputusan tentang sintesis desain eksperimental miskonsepsi menentukan apakah penelitian menghasilkan daftar jawaban salah atau pemahaman yang dapat ditindaklanjuti mengenai cara peserta didik berpikir. Sasaran utamanya ialah bukti yang cukup untuk merancang diagnosis, pembelajaran, dan penelitian lanjutan secara bertanggung jawab.

7.6 Penelitian Mixed Methods

Subbab ini membahas penggabungan data kuantitatif dan kualitatif untuk memperoleh keluasan pemetaan sekaligus kedalaman penjelasan mengenai pola miskonsepsi. Fokus utamanya adalah menjaga keselarasan antara pertanyaan penelitian, bentuk bukti, teknik analisis, dan kekuatan klaim. Dalam konteks miskonsepsi kimia, keputusan metodologis harus mampu membedakan pola konseptual dari respons yang timbul karena keterbatasan bahasa, kesalahan hitung, atau kondisi administrasi. Pembahasan berikut dipecah menjadi lima belas bagian agar setiap keputusan dapat ditelaah secara operasional.

7.6.1 Rasional Penggunaan Mixed Methods

Rasional penggunaan *mixed methods* perlu dipahami sebagai keputusan metodologis yang mempunyai konsekuensi pada desain, analisis, dan pelaporan. Fokus ini berada dalam penggabungan data kuantitatif dan kualitatif untuk memperoleh keluasan pemetaan sekaligus kedalaman penjelasan mengenai pola miskonsepsi. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Levitt et al. (2018) serta Johnson dan Onwuegbuzie (2004) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Peneliti sebaiknya menyiapkan matriks yang menghubungkan pertanyaan penelitian dengan unit analisis, gambar partikel, teknik analisis, dan evaluasi intervensi. Matriks ini mencegah pengumpulan data yang banyak tetapi tidak relevan. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Ilustrasi eksperimen disertai studi proses juga menunjukkan bahwa kualitas instrumen memengaruhi temuan. Distraktor yang tidak berasal dari pemikiran peserta didik dapat menghasilkan pilihan acak, sedangkan pertanyaan yang terlalu menuntut matematika dapat mengukur keterampilan hitung alih-alih konsepsi. Peneliti perlu menilai kesesuaian bukti dengan konstruk yang dimaksud.

Keputusan yang harus didokumentasikan:

- Definisikan unit analisis dalam istilah operasional.
- Tentukan bukti utama dari artefak pembelajaran dan bukti pendukung dari skala keyakinan.
- Periksa nilai tambah integrasi sebelum membuat kategori atau model.
- Catat potensi dua studi paralel tanpa integrasi beserta strategi mitigasinya.
- Sajikan joint display bersama batas interpretasi dan ketidakpastian.

Peneliti sebaiknya memisahkan keputusan apriori dari keputusan eksploratori. Pemisahan tersebut membantu mengendalikan *joint display* dekoratif dan mencegah dominasi satu komponen. Selanjutnya, penanganan divergensi dan kualitas tiap komponen dilaporkan bersama ketidakpastian, sehingga pembaca dapat menilai kekuatan temuan tanpa bergantung pada kesimpulan penulis semata.

Dengan demikian, rasional penggunaan mixed methods berfungsi menghubungkan keputusan desain dengan interpretasi. Produk yang diharapkan bukan hanya evaluasi intervensi, tetapi juga penjelasan mengenai bagaimana produk tersebut diperoleh dan dalam kondisi apa ia

dapat digunakan. Penelitian miskonsepsi yang baik selalu menyisakan ruang bagi penjelasan alternatif dan verifikasi lanjutan.

7.6.2 Desain Konvergen

Ketelitian penelitian pada bagian desain konvergen sangat menentukan apakah pola jawaban dapat ditafsirkan sebagai bukti konseptual. Fokus ini berada dalam penggabungan data kuantitatif dan kualitatif untuk memperoleh keluasan pemetaan sekaligus kedalaman penjelasan mengenai pola miskonsepsi. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. AERA, APA, dan NCME (2014) serta Levitt et al. (2018) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Dokumentasi metodologis hendaknya memuat siapa yang melakukan artefak pembelajaran, kapan data diperoleh, bagaimana urutan pengumpulan dikendalikan, dan atas dasar apa meta-inferensi terpadu disusun. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Contoh pada kajian peta konsep dan wawancara memperlihatkan pentingnya pemisahan antara ketidaktahuan, kesalahan prosedural, dan miskonsepsi. Seorang peserta dapat gagal karena lupa rumus, sedangkan peserta lain memilih jawaban yang sama karena menggunakan model mental yang tidak ilmiah. Perbedaan ini baru terlihat ketika artefak pembelajaran dibaca bersama skala keyakinan dan bukti tentang alasan.

Pertanyaan audit metodologis:

- Definisikan bentuk meta-inferensi dalam istilah operasional.
- Tentukan bukti utama dari tes diagnostik dan bukti pendukung dari wawancara.
- Periksa kualitas tiap komponen sebelum membuat kategori atau model.
- Catat potensi dominasi satu komponen beserta strategi mitigasinya.
- Sajikan tipologi peserta bersama batas interpretasi dan ketidakpastian.

Standar metodologis menuntut koherensi antara konstruk, prosedur, analisis, dan klaim. Dalam konteks ini, nilai tambah integrasi tidak dapat digantikan oleh indeks tunggal. Demikian pula, legitimasi integrasi harus dinilai dengan bukti yang sesuai. AERA, APA, dan NCME (2014) dan Levitt et al. (2018) memberikan dasar untuk pelaporan yang lengkap dan dapat ditelaah.

Inti metodologis subbagian ini adalah keterlacakan. Pembaca harus dapat mengikuti perjalanan dari pertanyaan, data, pengodean, analisis, sampai meta-inferensi terpadu. Ketika

jejak itu jelas, temuan dapat digunakan untuk memperbaiki pembelajaran tanpa mengubah diagnosis konseptual menjadi label bagi peserta didik.

7.6.3 Desain Eksplanatori Sekuensial

Peneliti sering memusatkan perhatian pada teknik, padahal desain eksplanatori sekuensial justru mengatur batas inferensi yang dapat dibuat. Fokus ini berada dalam penggabungan data kuantitatif dan kualitatif untuk memperoleh keluasan pemetaan sekaligus kedalaman penjelasan mengenai pola miskonsepsi. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Creswell dan Plano Clark (2018) serta Johnson dan Onwuegbuzie (2004) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Rancangan yang dapat diaudit mencantumkan alasan pemilihan prioritas komponen, prosedur penanganan titik integrasi, dan kriteria keputusan untuk unit analisis. Tanpa penjelasan itu, pembaca sulit menilai apakah tes diagnostik benar-benar mendukung klaim. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Kasus respons gambar dan tes bertingkat dapat digunakan untuk memperlihatkan hubungan antara pola individual dan pola kelas. Sebagian miskonsepsi mungkin hanya dimiliki beberapa peserta tetapi sangat kuat, sedangkan pola lain tersebar luas karena bersumber dari analogi atau buku teks yang sama. Analisis perlu menjaga kedua tingkat tersebut.

Bukti minimum yang perlu tersedia:

- Definisikan prioritas komponen dalam istilah operasional.
- Tentukan bukti utama dari skala keyakinan dan bukti pendukung dari observasi.
- Periksa legitimasi integrasi sebelum membuat kategori atau model.
- Catat potensi sampel tidak terhubung beserta strategi mitigasinya.
- Sajikan penjelasan pola statistik bersama batas interpretasi dan ketidakpastian.

Pemeriksaan kualitas dapat dilakukan melalui audit terhadap kualitas tiap komponen, analisis kasus yang bertentangan, dan evaluasi kesesuaian sampel. Apabila ditemukan dua studi paralel tanpa integrasi, peneliti perlu menurunkan kekuatan klaim atau menambah bukti. Transparansi seperti ini meningkatkan kegunaan temuan bagi penelitian lanjutan dan praktik kelas.

Kesimpulannya, desain eksplanatori sekuensial perlu dirancang sebagai proses argumentatif. Data tidak berbicara sendiri; maknanya bergantung pada definisi konstruk, kualitas prosedur,

dan konteks. Oleh sebab itu, klaim akhir sebaiknya menggunakan bahasa yang sesuai dengan kekuatan bukti, misalnya “menunjukkan pola”, “mengindikasikan”, atau “mendukung penjelasan”.

7.6.4 Desain Eksploratori Sekuensial

Kajian yang kuat memperlakukan desain eksploratori sekuensial sebagai bagian dari rantai argumentasi, bukan sebagai unsur pelengkap. Fokus ini berada dalam penggabungan data kuantitatif dan kualitatif untuk memperoleh keluasan pemetaan sekaligus kedalaman penjelasan mengenai pola miskonsepsi. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Johnson dan Onwuegbuzie (2004) serta Levitt et al. (2018) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Pada tahap perencanaan, urutan pengumpulan dan strategi sampling terhubung harus dibedakan secara eksplisit. Pada tahap pelaksanaan, konsistensi skala keyakinan perlu diperiksa; pada tahap interpretasi, legitimasi integrasi menjadi dasar pembatasan klaim. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Dalam kelas yang mempelajari data kelas dan studi kasus guru, data agregat mungkin menunjukkan mayoritas jawaban salah. Akan tetapi, wawancara dapat memperlihatkan beberapa mekanisme berbeda: bahasa sehari-hari, generalisasi aturan, salah membaca simbol, atau model partikel yang tidak tepat. Karena itu, satu persentase tidak boleh menggantikan analisis struktur alasan.

Kriteria pemeriksaan sebelum analisis:

- Definisikan urutan pengumpulan dalam istilah operasional.
- Tentukan bukti utama dari wawancara dan bukti pendukung dari gambar partikel.
- Periksa kesesuaian sampel sebelum membuat kategori atau model.
- Catat potensi mengabaikan hasil yang bertentangan beserta strategi mitigasinya.
- Sajikan pengembangan instrumen bersama batas interpretasi dan ketidakpastian.

Kriteria keberhasilan tidak cukup berupa banyaknya data. Yang lebih penting adalah apakah legitimasi integrasi dapat ditelusuri, apakah transparansi titik campur diperiksa, dan apakah nilai tambah integrasi dijelaskan. Peneliti harus menghindari dominasi satu komponen; jika risiko itu tidak dapat dihilangkan, dampaknya perlu dianalisis dan dicantumkan sebagai batas penelitian.

Pada akhirnya, keputusan tentang desain eksploratori sekuensial menentukan apakah penelitian menghasilkan daftar jawaban salah atau pemahaman yang dapat ditindaklanjuti mengenai cara peserta didik berpikir. Sasaran utamanya ialah bukti yang cukup untuk merancang diagnosis, pembelajaran, dan penelitian lanjutan secara bertanggung jawab.

7.6.5 Desain Embedded dalam Eksperimen

Subbagian ini menempatkan desain embedded dalam eksperimen sebagai simpul penting dalam penelitian *mixed methods*. Fokus ini berada dalam penggabungan data kuantitatif dan kualitatif untuk memperoleh keluasan pemetaan sekaligus kedalaman penjelasan mengenai pola miskonsepsi. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Fetters et al. (2013) serta Johnson dan Onwuegbuzie (2004) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Secara operasional, peneliti perlu menetapkan titik integrasi, menjelaskan unit analisis, dan mendokumentasikan prioritas komponen. Keputusan tersebut kemudian diterjemahkan ke dalam pilihan wawancara, gambar partikel, serta tes diagnostik. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Sebagai ilustrasi, pada topik skor rendah-keyakinan tinggi pada redoks, peserta didik dapat memberikan jawaban benar pada tingkat simbolik tetapi alasan submikroskopiknya keliru. Peneliti tidak cukup menghitung benar-salah; ia perlu membandingkan wawancara dengan gambar partikel, memeriksa tingkat keyakinan, dan menguji konsistensi pada konteks lain. Hasil analisis mungkin menunjukkan bahwa kesalahan hanya muncul pada satu representasi, atau justru merupakan hubungan konseptual yang stabil.

Langkah operasional yang disarankan:

- Definisikan titik integrasi dalam istilah operasional.
- Tentukan bukti utama dari observasi dan bukti pendukung dari artefak pembelajaran.
- Periksa transparansi titik campur sebelum membuat kategori atau model.
- Catat potensi joint display dekoratif beserta strategi mitigasinya.
- Sajikan evaluasi intervensi bersama batas interpretasi dan ketidakpastian.

Mutu inferensi terutama ditentukan oleh kesesuaian sampel, penanganan divergensi, dan kualitas tiap komponen. Ancaman utama ialah sampel tidak terhubung serta meta-inferensi terlalu luas. Karena itu, laporan perlu menyajikan bukti yang mendukung sekaligus bukti yang membatasi interpretasi. Prinsip ini konsisten dengan penekanan Fetters et al. (2013) dan

Johnson dan Onwuegbuzie (2004) bahwa validitas melekat pada interpretasi dan penggunaan hasil, bukan sekadar pada nama instrumen atau desain.

Dengan demikian, desain embedded dalam eksperimen berfungsi menghubungkan keputusan desain dengan interpretasi. Produk yang diharapkan bukan hanya penjelasan pola statistik, tetapi juga penjelasan mengenai bagaimana produk tersebut diperoleh dan dalam kondisi apa ia dapat digunakan. Penelitian miskonsepsi yang baik selalu menyisakan ruang bagi penjelasan alternatif dan verifikasi lanjutan.

7.6.6 Desain Multitahap

Pembahasan mengenai desain multitahap diperlukan agar penelitian *mixed methods* tidak berhenti pada prosedur administratif. Fokus ini berada dalam penggabungan data kuantitatif dan kualitatif untuk memperoleh keluasan pemetaan sekaligus kedalaman penjelasan mengenai pola miskonsepsi. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Greene et al. (1989) serta Levitt et al. (2018) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Penerapan praktis dapat disusun dalam urutan: mendefinisikan strategi sampling terhubung, mengoperasionalkan bentuk meta-inferensi, mengumpulkan observasi, memeriksa transparansi titik campur, lalu melaporkan pengembangan instrumen. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Misalnya, pada klaster konsepsi kesetimbangan, perubahan skor sesudah pembelajaran belum otomatis menandakan perubahan konsep. Peserta didik mungkin mempelajari pola jawaban tes. Bukti yang lebih meyakinkan mencakup transfer ke situasi baru, penjelasan kausal, representasi partikel, dan retensi setelah jeda waktu.

Keputusan yang harus didokumentasikan:

- Definisikan strategi sampling terhubung dalam istilah operasional.
- Tentukan bukti utama dari gambar partikel dan bukti pendukung dari tes diagnostik.
- Periksa penanganan divergensi sebelum membuat kategori atau model.
- Catat potensi meta-inferensi terlalu luas beserta strategi mitigasinya.
- Sajikan meta-inferensi terpadu bersama batas interpretasi dan ketidakpastian.

Temuan yang baik tidak selalu berarti hasil yang besar atau sesuai harapan. Temuan bernilai ketika transparansi titik campur terjaga, nilai tambah integrasi dijelaskan, dan ketidakpastian

disajikan. Kerangka Greene et al. (1989) serta Levitt et al. (2018) menolong peneliti membedakan kekuatan bukti dari daya tarik narasi.

Inti metodologis subbagian ini adalah keterlacakan. Pembaca harus dapat mengikuti perjalanan dari pertanyaan, data, pengodean, analisis, sampai pengembangan instrumen. Ketika jejak itu jelas, temuan dapat digunakan untuk memperbaiki pembelajaran tanpa mengubah diagnosis konseptual menjadi label bagi peserta didik.

7.6.7 Sampling dalam Mixed Methods

Dalam metodologi penelitian miskonsepsi kimia, sampling dalam *mixed methods* menentukan mutu hubungan antara data dan kesimpulan. Fokus ini berada dalam penggabungan data kuantitatif dan kualitatif untuk memperoleh keluasan pemetaan sekaligus kedalaman penjelasan mengenai pola miskonsepsi. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Levitt et al. (2018) serta Johnson dan Onwuegbuzie (2004) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Prosedur minimal mencakup definisi unit analisis, alasan memilih gambar partikel, aturan mengubah respons menjadi kategori, serta mekanisme pemeriksaan penanganan divergensi. Semua penyimpangan dari rencana awal perlu dicatat. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Pada penelitian mengenai survei luas dilanjutkan wawancara, peneliti dapat menyusun beberapa tugas ekuivalen yang berubah konteks, representasi, atau tuntutan penalarannya. Jika pola respons tetap berulang dan disertai keyakinan tinggi, terdapat dasar yang lebih kuat untuk menyebut indikasi miskonsepsi. Jika pola berubah-ubah, interpretasi yang lebih hati-hati adalah pengetahuan parsial atau ketergantungan konteks.

Pertanyaan audit metodologis:

- Definisikan unit analisis dalam istilah operasional.
- Tentukan bukti utama dari artefak pembelajaran dan bukti pendukung dari skala keyakinan.
- Periksa nilai tambah integrasi sebelum membuat kategori atau model.
- Catat potensi dua studi paralel tanpa integrasi beserta strategi mitigasinya.
- Sajikan joint display bersama batas interpretasi dan ketidakpastian.

Pemeriksaan mutu hendaknya dilakukan selama penelitian, bukan hanya pada akhir. Catatan tentang penanganan divergensi, keputusan terkait kualitas tiap komponen, dan kejadian yang

berpotensi menimbulkan *joint display* dekoratif perlu dipelihara dalam *audit trail*. Hal ini memungkinkan koreksi dini dan pelaporan yang jujur.

Kesimpulannya, sampling dalam mixed methods perlu dirancang sebagai proses argumentatif. Data tidak berbicara sendiri; maknanya bergantung pada definisi konstruk, kualitas prosedur, dan konteks. Oleh sebab itu, klaim akhir sebaiknya menggunakan bahasa yang sesuai dengan kekuatan bukti, misalnya “menunjukkan pola”, “mengindikasikan”, atau “mendukung penjelasan”.

7.6.8 Connecting, Building, Merging, dan Embedding

Keputusan tentang connecting, building, merging, dan embedding harus dibuat sebelum data utama dikumpulkan, bukan ditambahkan setelah hasil diketahui. Fokus ini berada dalam penggabungan data kuantitatif dan kualitatif untuk memperoleh keluasan pemetaan sekaligus kedalaman penjelasan mengenai pola miskonsepsi. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. AERA, APA, dan NCME (2014) serta Levitt et al. (2018) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Tahap kerja dimulai dengan penegasan bentuk meta-inferensi. Setelah itu, peneliti menyusun aturan untuk urutan pengumpulan dan menentukan bagaimana strategi sampling terhubung akan diperiksa. Bukti sebaiknya berasal dari lebih dari satu bentuk, misalnya artefak pembelajaran dan skala keyakinan. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Pada topik wawancara awal untuk menyusun distraktor, peneliti juga perlu memperhatikan bahasa. Istilah yang sama dapat diberi makna berbeda oleh peserta didik. Klarifikasi melalui pertanyaan tindak lanjut, gambar, atau contoh tandingan membantu membedakan masalah terminologi dari kesalahan konseptual.

Bukti minimum yang perlu tersedia:

- Definisikan bentuk meta-inferensi dalam istilah operasional.
- Tentukan bukti utama dari tes diagnostik dan bukti pendukung dari wawancara.
- Periksa kualitas tiap komponen sebelum membuat kategori atau model.
- Catat potensi dominasi satu komponen beserta strategi mitigasinya.
- Sajikan tipologi peserta bersama batas interpretasi dan ketidakpastian.

Kualitas penelitian meningkat ketika peneliti menunjukkan rantai bukti dari artefak pembelajaran menuju meta-inferensi terpadu, menyatakan asumsi, dan memeriksa alternatif

penjelasan. Sebaliknya, meta-inferensi terlalu luas dapat menghasilkan klaim yang tampak tegas tetapi tidak stabil. Rekomendasi praktis harus proporsional dengan bukti.

Pada akhirnya, keputusan tentang connecting, building, merging, dan embedding menentukan apakah penelitian menghasilkan daftar jawaban salah atau pemahaman yang dapat ditindaklanjuti mengenai cara peserta didik berpikir. Sasaran utamanya ialah bukti yang cukup untuk merancang diagnosis, pembelajaran, dan penelitian lanjutan secara bertanggung jawab.

7.6.9 Joint Display untuk Data Miskonsepsi

Joint display untuk data miskonsepsi perlu dipahami sebagai keputusan metodologis yang mempunyai konsekuensi pada desain, analisis, dan pelaporan. Fokus ini berada dalam penggabungan data kuantitatif dan kualitatif untuk memperoleh keluasan pemetaan sekaligus kedalaman penjelasan mengenai pola miskonsepsi. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Creswell dan Plano Clark (2018) serta Johnson dan Onwuegbuzie (2004) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Peneliti sebaiknya menyiapkan matriks yang menghubungkan pertanyaan penelitian dengan prioritas komponen, tes diagnostik, teknik analisis, dan joint display. Matriks ini mencegah pengumpulan data yang banyak tetapi tidak relevan. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Ilustrasi eksperimen disertai studi proses juga menunjukkan bahwa kualitas instrumen memengaruhi temuan. Distraktor yang tidak berasal dari pemikiran peserta didik dapat menghasilkan pilihan acak, sedangkan pertanyaan yang terlalu menuntut matematika dapat mengukur keterampilan hitung alih-alih konsepsi. Peneliti perlu menilai kesesuaian bukti dengan konstruk yang dimaksud.

Kriteria pemeriksaan sebelum analisis:

- Definisikan prioritas komponen dalam istilah operasional.
- Tentukan bukti utama dari skala keyakinan dan bukti pendukung dari observasi.
- Periksa legitimasi integrasi sebelum membuat kategori atau model.
- Catat potensi sampel tidak terhubung beserta strategi mitigasinya.
- Sajikan penjelasan pola statistik bersama batas interpretasi dan ketidakpastian.

Peneliti sebaiknya memisahkan keputusan apriori dari keputusan eksploratori. Pemisahan tersebut membantu mengendalikan dua studi paralel tanpa integrasi dan mencegah

mengabaikan hasil yang bertentangan. Selanjutnya, kualitas tiap komponen dan kesesuaian sampel dilaporkan bersama ketidakpastian, sehingga pembaca dapat menilai kekuatan temuan tanpa bergantung pada kesimpulan penulis semata.

Dengan demikian, joint display untuk data miskonsepsi berfungsi menghubungkan keputusan desain dengan interpretasi. Produk yang diharapkan bukan hanya joint display, tetapi juga penjelasan mengenai bagaimana produk tersebut diperoleh dan dalam kondisi apa ia dapat digunakan. Penelitian miskonsepsi yang baik selalu menyisakan ruang bagi penjelasan alternatif dan verifikasi lanjutan.

7.6.10 Analisis Divergensi dan Kasus Negatif

Ketelitian penelitian pada bagian analisis divergensi dan kasus negatif sangat menentukan apakah pola jawaban dapat ditafsirkan sebagai bukti konseptual. Fokus ini berada dalam penggabungan data kuantitatif dan kualitatif untuk memperoleh keluasan pemetaan sekaligus kedalaman penjelasan mengenai pola miskonsepsi. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Johnson dan Onwuegbuzie (2004) serta Levitt et al. (2018) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Dokumentasi metodologis hendaknya memuat siapa yang melakukan skala keyakinan, kapan data diperoleh, bagaimana strategi sampling terhubung dikendalikan, dan atas dasar apa tipologi peserta disusun. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Contoh pada kajian peta konsep dan wawancara memperlihatkan pentingnya pemisahan antara ketidaktahuan, kesalahan prosedural, dan miskonsepsi. Seorang peserta dapat gagal karena lupa rumus, sedangkan peserta lain memilih jawaban yang sama karena menggunakan model mental yang tidak ilmiah. Perbedaan ini baru terlihat ketika skala keyakinan dibaca bersama observasi dan bukti tentang alasan.

Langkah operasional yang disarankan:

- Definisikan urutan pengumpulan dalam istilah operasional.
- Tentukan bukti utama dari wawancara dan bukti pendukung dari gambar partikel.
- Periksa kesesuaian sampel sebelum membuat kategori atau model.
- Catat potensi mengabaikan hasil yang bertentangan beserta strategi mitigasinya.
- Sajikan pengembangan instrumen bersama batas interpretasi dan ketidakpastian.

Standar metodologis menuntut koherensi antara konstruk, prosedur, analisis, dan klaim. Dalam konteks ini, legitimasi integrasi tidak dapat digantikan oleh indeks tunggal. Demikian pula, transparansi titik campur harus dinilai dengan bukti yang sesuai. Johnson dan Onwuegbuzie (2004) dan Levitt et al. (2018) memberikan dasar untuk pelaporan yang lengkap dan dapat ditelaah.

Inti metodologis subbagian ini adalah keterlacakan. Pembaca harus dapat mengikuti perjalanan dari pertanyaan, data, pengodean, analisis, sampai tipologi peserta. Ketika jejak itu jelas, temuan dapat digunakan untuk memperbaiki pembelajaran tanpa mengubah diagnosis konseptual menjadi label bagi peserta didik.

7.6.11 Meta-Inferensi

Peneliti sering memusatkan perhatian pada teknik, padahal meta-inferensi justru mengatur batas inferensi yang dapat dibuat. Fokus ini berada dalam penggabungan data kuantitatif dan kualitatif untuk memperoleh keluasan pemetaan sekaligus kedalaman penjelasan mengenai pola miskonsepsi. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Fetters et al. (2013) serta Johnson dan Onwuegbuzie (2004) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Rancangan yang dapat diaudit mencantumkan alasan pemilihan titik integrasi, prosedur penanganan unit analisis, dan kriteria keputusan untuk prioritas komponen. Tanpa penjelasan itu, pembaca sulit menilai apakah wawancara benar-benar mendukung klaim. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Kasus respons gambar dan tes bertingkat dapat digunakan untuk memperlihatkan hubungan antara pola individual dan pola kelas. Sebagian miskonsepsi mungkin hanya dimiliki beberapa peserta tetapi sangat kuat, sedangkan pola lain tersebar luas karena bersumber dari analogi atau buku teks yang sama. Analisis perlu menjaga kedua tingkat tersebut.

Keputusan yang harus didokumentasikan:

- Definisikan titik integrasi dalam istilah operasional.
- Tentukan bukti utama dari observasi dan bukti pendukung dari artefak pembelajaran.
- Periksa transparansi titik campur sebelum membuat kategori atau model.
- Catat potensi joint display dekoratif beserta strategi mitigasinya.
- Sajikan evaluasi intervensi bersama batas interpretasi dan ketidakpastian.

Pemeriksaan kualitas dapat dilakukan melalui audit terhadap kesesuaian sampel, analisis kasus yang bertentangan, dan evaluasi penanganan divergensi. Apabila ditemukan sampel tidak terhubung, peneliti perlu menurunkan kekuatan klaim atau menambah bukti. Transparansi seperti ini meningkatkan kegunaan temuan bagi penelitian lanjutan dan praktik kelas.

Kesimpulannya, meta-inferensi perlu dirancang sebagai proses argumentatif. Data tidak berbicara sendiri; maknanya bergantung pada definisi konstruk, kualitas prosedur, dan konteks. Oleh sebab itu, klaim akhir sebaiknya menggunakan bahasa yang sesuai dengan kekuatan bukti, misalnya “menunjukkan pola”, “mengindikasikan”, atau “mendukung penjelasan”.

7.6.12 Validitas dan Legitimasi Integrasi

Kajian yang kuat memperlakukan validitas dan legitimasi integrasi sebagai bagian dari rantai argumentasi, bukan sebagai unsur pelengkap. Fokus ini berada dalam penggabungan data kuantitatif dan kualitatif untuk memperoleh keluasan pemetaan sekaligus kedalaman penjelasan mengenai pola miskonsepsi. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Greene et al. (1989) serta Levitt et al. (2018) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Pada tahap perencanaan, strategi sampling terhubung dan bentuk meta-inferensi harus dibedakan secara eksplisit. Pada tahap pelaksanaan, konsistensi observasi perlu diperiksa; pada tahap interpretasi, transparansi titik campur menjadi dasar pembatasan klaim. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Dalam kelas yang mempelajari data kelas dan studi kasus guru, data agregat mungkin menunjukkan mayoritas jawaban salah. Akan tetapi, wawancara dapat memperlihatkan beberapa mekanisme berbeda: bahasa sehari-hari, generalisasi aturan, salah membaca simbol, atau model partikel yang tidak tepat. Karena itu, satu persentase tidak boleh menggantikan analisis struktur alasan.

Pertanyaan audit metodologis:

- Definisikan strategi sampling terhubung dalam istilah operasional.
- Tentukan bukti utama dari gambar partikel dan bukti pendukung dari tes diagnostik.
- Periksa penanganan divergensi sebelum membuat kategori atau model.
- Catat potensi meta-inferensi terlalu luas beserta strategi mitigasinya.

- Sajikan meta-inferensi terpadu bersama batas interpretasi dan ketidakpastian.

Kriteria keberhasilan tidak cukup berupa banyaknya data. Yang lebih penting adalah apakah transparansi titik campur dapat ditelusuri, apakah nilai tambah integrasi diperiksa, dan apakah legitimasi integrasi dijelaskan. Peneliti harus menghindari mengabaikan hasil yang bertentangan; jika risiko itu tidak dapat dihilangkan, dampaknya perlu dianalisis dan dicantumkan sebagai batas penelitian.

Pada akhirnya, keputusan tentang validitas dan legitimasi integrasi menentukan apakah penelitian menghasilkan daftar jawaban salah atau pemahaman yang dapat ditindaklanjuti mengenai cara peserta didik berpikir. Sasaran utamanya ialah bukti yang cukup untuk merancang diagnosis, pembelajaran, dan penelitian lanjutan secara bertanggung jawab.

7.6.13 Manajemen Tim Peneliti Multimetode

Subbagian ini menempatkan manajemen tim peneliti multimetode sebagai simpul penting dalam penelitian *mixed methods*. Fokus ini berada dalam penggabungan data kuantitatif dan kualitatif untuk memperoleh keluasan pemetaan sekaligus kedalaman penjelasan mengenai pola miskonsepsi. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Levitt et al. (2018) serta Johnson dan Onwuegbuzie (2004) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Secara operasional, peneliti perlu menetapkan unit analisis, menjelaskan prioritas komponen, dan mendokumentasikan titik integrasi. Keputusan tersebut kemudian diterjemahkan ke dalam pilihan gambar partikel, tes diagnostik, serta wawancara. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Sebagai ilustrasi, pada topik skor rendah-keyakinan tinggi pada redoks, peserta didik dapat memberikan jawaban benar pada tingkat simbolik tetapi alasan submikroskopiknya keliru. Peneliti tidak cukup menghitung benar-salah; ia perlu membandingkan gambar partikel dengan tes diagnostik, memeriksa tingkat keyakinan, dan menguji konsistensi pada konteks lain. Hasil analisis mungkin menunjukkan bahwa kesalahan hanya muncul pada satu representasi, atau justru merupakan hubungan konseptual yang stabil.

Bukti minimum yang perlu tersedia:

- Definisikan unit analisis dalam istilah operasional.
- Tentukan bukti utama dari artefak pembelajaran dan bukti pendukung dari skala keyakinan.

- Periksa nilai tambah integrasi sebelum membuat kategori atau model.
- Catat potensi dua studi paralel tanpa integrasi beserta strategi mitigasinya.
- Sajikan joint display bersama batas interpretasi dan ketidakpastian.

Mutu inferensi terutama ditentukan oleh penanganan divergensi, kualitas tiap komponen, dan kesesuaian sampel. Ancaman utama ialah *joint display* dekoratif serta dominasi satu komponen. Karena itu, laporan perlu menyajikan bukti yang mendukung sekaligus bukti yang membatasi interpretasi. Prinsip ini konsisten dengan penekanan Levitt et al. (2018) dan Johnson dan Onwuegbuzie (2004) bahwa validitas melekat pada interpretasi dan penggunaan hasil, bukan sekadar pada nama instrumen atau desain.

Dengan demikian, manajemen tim peneliti multimetode berfungsi menghubungkan keputusan desain dengan interpretasi. Produk yang diharapkan bukan hanya evaluasi intervensi, tetapi juga penjelasan mengenai bagaimana produk tersebut diperoleh dan dalam kondisi apa ia dapat digunakan. Penelitian miskonsepsi yang baik selalu menyisakan ruang bagi penjelasan alternatif dan verifikasi lanjutan.

7.6.14 Pelaporan Mixed Methods

Pembahasan mengenai pelaporan *mixed methods* diperlukan agar penelitian *mixed methods* tidak berhenti pada prosedur administratif. Fokus ini berada dalam penggabungan data kuantitatif dan kualitatif untuk memperoleh keluasan pemetaan sekaligus kedalaman penjelasan mengenai pola miskonsepsi. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. AERA, APA, dan NCME (2014) serta Levitt et al. (2018) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Penerapan praktis dapat disusun dalam urutan: mendefinisikan bentuk meta-inferensi, mengoperasionalkan urutan pengumpulan, mengumpulkan artefak pembelajaran, memeriksa nilai tambah integrasi, lalu melaporkan meta-inferensi terpadu. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Misalnya, pada klaster konsepsi kesetimbangan, perubahan skor sesudah pembelajaran belum otomatis menandakan perubahan konsep. Peserta didik mungkin mempelajari pola jawaban tes. Bukti yang lebih meyakinkan mencakup transfer ke situasi baru, penjelasan kausal, representasi partikel, dan retensi setelah jeda waktu.

Kriteria pemeriksaan sebelum analisis:

- Definisikan bentuk meta-inferensi dalam istilah operasional.

- Tentukan bukti utama dari tes diagnostik dan bukti pendukung dari wawancara.
- Periksa kualitas tiap komponen sebelum membuat kategori atau model.
- Catat potensi dominasi satu komponen beserta strategi mitigasinya.
- Sajikan tipologi peserta bersama batas interpretasi dan ketidakpastian.

Temuan yang baik tidak selalu berarti hasil yang besar atau sesuai harapan. Temuan bernilai ketika nilai tambah integrasi terjaga, legitimasi integrasi dijelaskan, dan ketidakpastian disajikan. Kerangka AERA, APA, dan NCME (2014) serta Levitt et al. (2018) menolong peneliti membedakan kekuatan bukti dari daya tarik narasi.

Inti metodologis subbagian ini adalah keterlacakan. Pembaca harus dapat mengikuti perjalanan dari pertanyaan, data, pengodean, analisis, sampai meta-inferensi terpadu. Ketika jejak itu jelas, temuan dapat digunakan untuk memperbaiki pembelajaran tanpa mengubah diagnosis konseptual menjadi label bagi peserta didik.

7.6.15 Sintesis Desain Mixed Methods Miskonsepsi

Dalam metodologi penelitian miskonsepsi kimia, sintesis desain *mixed methods* miskonsepsi menentukan mutu hubungan antara data dan kesimpulan. Fokus ini berada dalam penggabungan data kuantitatif dan kualitatif untuk memperoleh keluasan pemetaan sekaligus kedalaman penjelasan mengenai pola miskonsepsi. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Creswell dan Plano Clark (2018) serta Johnson dan Onwuegbuzie (2004) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Prosedur minimal mencakup definisi prioritas komponen, alasan memilih tes diagnostik, aturan mengubah respons menjadi kategori, serta mekanisme pemeriksaan kualitas tiap komponen. Semua penyimpangan dari rencana awal perlu dicatat. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Pada penelitian mengenai survei luas dilanjutkan wawancara, peneliti dapat menyusun beberapa tugas ekuivalen yang berubah konteks, representasi, atau tuntutan penalarannya. Jika pola respons tetap berulang dan disertai keyakinan tinggi, terdapat dasar yang lebih kuat untuk menyebut indikasi miskonsepsi. Jika pola berubah-ubah, interpretasi yang lebih hati-hati adalah pengetahuan parsial atau ketergantungan konteks.

Langkah operasional yang disarankan:

- Definisikan prioritas komponen dalam istilah operasional.
- Tentukan bukti utama dari skala keyakinan dan bukti pendukung dari observasi.

- Periksa legitimasi integrasi sebelum membuat kategori atau model.
- Catat potensi sampel tidak terhubung beserta strategi mitigasinya.
- Sajikan penjelasan pola statistik bersama batas interpretasi dan ketidakpastian.

Pemeriksaan mutu hendaknya dilakukan selama penelitian, bukan hanya pada akhir. Catatan tentang kualitas tiap komponen, keputusan terkait kesesuaian sampel, dan kejadian yang berpotensi menimbulkan dua studi paralel tanpa integrasi perlu dipelihara dalam *audit trail*. Hal ini memungkinkan koreksi dini dan pelaporan yang jujur.

Kesimpulannya, sintesis desain mixed methods miskonsepsi perlu dirancang sebagai proses argumentatif. Data tidak berbicara sendiri; maknanya bergantung pada definisi konstruk, kualitas prosedur, dan konteks. Oleh sebab itu, klaim akhir sebaiknya menggunakan bahasa yang sesuai dengan kekuatan bukti, misalnya “menunjukkan pola”, “mengindikasikan”, atau “mendukung penjelasan”.

7.7 Penelitian Pengembangan Instrumen

Subbab ini membahas proses sistematis untuk merancang, menguji, dan memvalidasi alat yang menghasilkan bukti sahih tentang konsepsi, miskonsepsi, alasan, dan keyakinan. Fokus utamanya adalah menjaga keselarasan antara pertanyaan penelitian, bentuk bukti, teknik analisis, dan kekuatan klaim. Dalam konteks miskonsepsi kimia, keputusan metodologis harus mampu membedakan pola konseptual dari respons yang timbul karena keterbatasan bahasa, kesalahan hitung, atau kondisi administrasi. Pembahasan berikut dipecah menjadi lima belas bagian agar setiap keputusan dapat ditelaah secara operasional.

7.7.1 Definisi Konstruk dan Tujuan Penggunaan

Keputusan tentang definisi konstruk dan tujuan penggunaan harus dibuat sebelum data utama dikumpulkan, bukan ditambahkan setelah hasil diketahui. Fokus ini berada dalam proses sistematis untuk merancang, menguji, dan memvalidasi alat yang menghasilkan bukti sahih tentang konsepsi, miskonsepsi, alasan, dan keyakinan. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Messick (1995) serta Boateng et al. (2018) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Tahap kerja dimulai dengan penegasan domain konten. Setelah itu, peneliti menyusun aturan untuk fungsi diagnostik dan menentukan bagaimana interpretasi skor akan diperiksa. Bukti sebaiknya berasal dari lebih dari satu bentuk, misalnya wawancara eksploratori dan penilaian ahli. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti

membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Pada topik tes gambar partikel, peneliti juga perlu memperhatikan bahasa. Istilah yang sama dapat diberi makna berbeda oleh peserta didik. Klarifikasi melalui pertanyaan tindak lanjut, gambar, atau contoh tandingan membantu membedakan masalah terminologi dari kesalahan konseptual.

Keputusan yang harus didokumentasikan:

- Definisikan domain konten dalam istilah operasional.
- Tentukan bukti utama dari distraktor autentik dan bukti pendukung dari uji kognitif.
- Periksa struktur internal sebelum membuat kategori atau model.
- Catat potensi sampel uji terlalu sempit beserta strategi mitigasinya.
- Sajikan argumen validitas bersama batas interpretasi dan ketidakpastian.

Kualitas penelitian meningkat ketika peneliti menunjukkan rantai bukti dari wawancara eksploratori menuju bank butir, menyatakan asumsi, dan memeriksa alternatif penjelasan. Sebaliknya, alpha sebagai satu-satunya bukti dapat menghasilkan klaim yang tampak tegas tetapi tidak stabil. Rekomendasi praktis harus proporsional dengan bukti.

Pada akhirnya, keputusan tentang definisi konstruk dan tujuan penggunaan menentukan apakah penelitian menghasilkan daftar jawaban salah atau pemahaman yang dapat ditindaklanjuti mengenai cara peserta didik berpikir. Sasaran utamanya ialah bukti yang cukup untuk merancang diagnosis, pembelajaran, dan penelitian lanjutan secara bertanggung jawab.

7.7.2 Analisis Domain Konsep Kimia

Analisis domain konsep kimia perlu dipahami sebagai keputusan metodologis yang mempunyai konsekuensi pada desain, analisis, dan pelaporan. Fokus ini berada dalam proses sistematis untuk merancang, menguji, dan memvalidasi alat yang menghasilkan bukti sah tentang konsepsi, miskonsepsi, alasan, dan keyakinan. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Kane (2013) serta Treagust (1988) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Peneliti sebaiknya menyiapkan matriks yang menghubungkan pertanyaan penelitian dengan format respons, distraktor autentik, teknik analisis, dan manual penskoran. Matriks ini mencegah pengumpulan data yang banyak tetapi tidak relevan. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya

berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Ilustrasi tes empat tingkat asam-basa juga menunjukkan bahwa kualitas instrumen memengaruhi temuan. Distraktor yang tidak berasal dari pemikiran peserta didik dapat menghasilkan pilihan acak, sedangkan pertanyaan yang terlalu menuntut matematika dapat mengukur keterampilan hitung alih-alih konsepsi. Peneliti perlu menilai kesesuaian bukti dengan konstruk yang dimaksud.

Pertanyaan audit metodologis:

- Definisikan format respons dalam istilah operasional.
- Tentukan bukti utama dari penilaian ahli dan bukti pendukung dari data uji lapangan.
- Periksa hubungan dengan variabel lain sebelum membuat kategori atau model.
- Catat potensi kategori skor tanpa dasar beserta strategi mitigasinya.
- Sajikan parameter butir bersama batas interpretasi dan ketidakpastian.

Peneliti sebaiknya memisahkan keputusan apriori dari keputusan eksploratori. Pemisahan tersebut membantu mengendalikan validitas disebut sifat tes dan mencegah mengabaikan invariansi. Selanjutnya, struktur internal dan konsekuensi penggunaan dilaporkan bersama ketidakpastian, sehingga pembaca dapat menilai kekuatan temuan tanpa bergantung pada kesimpulan penulis semata.

Dengan demikian, analisis domain konsep kimia berfungsi menghubungkan keputusan desain dengan interpretasi. Produk yang diharapkan bukan hanya manual penskoran, tetapi juga penjelasan mengenai bagaimana produk tersebut diperoleh dan dalam kondisi apa ia dapat digunakan. Penelitian miskonsepsi yang baik selalu menyisakan ruang bagi penjelasan alternatif dan verifikasi lanjutan.

7.7.3 Penyusunan Blueprint Instrumen

Ketelitian penelitian pada bagian penyusunan blueprint instrumen sangat menentukan apakah pola jawaban dapat ditafsirkan sebagai bukti konseptual. Fokus ini berada dalam proses sistematis untuk merancang, menguji, dan memvalidasi alat yang menghasilkan bukti sahih tentang konsepsi, miskonsepsi, alasan, dan keyakinan. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Treagust (1988) serta AERA, APA, dan NCME (2014) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Dokumentasi metodologis hendaknya memuat siapa yang melakukan penilaian ahli, kapan data diperoleh, bagaimana interpretasi skor dikendalikan, dan atas dasar apa argumen validitas disusun. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama

penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Contoh pada kajian peta konsep terstruktur memperlihatkan pentingnya pemisahan antara ketidaktahuan, kesalahan prosedural, dan miskonsepsi. Seorang peserta dapat gagal karena lupa rumus, sedangkan peserta lain memilih jawaban yang sama karena menggunakan model mental yang tidak ilmiah. Perbedaan ini baru terlihat ketika penilaian ahli dibaca bersama data uji lapangan dan bukti tentang alasan.

Bukti minimum yang perlu tersedia:

- Definisikan fungsi diagnostik dalam istilah operasional.
- Tentukan bukti utama dari uji kognitif dan bukti pendukung dari peta konsep.
- Periksa konsekuensi penggunaan sebelum membuat kategori atau model.
- Catat potensi mengabaikan invariansi beserta strategi mitigasinya.
- Sajikan panduan interpretasi bersama batas interpretasi dan ketidakpastian.

Standar metodologis menuntut koherensi antara konstruk, prosedur, analisis, dan klaim. Dalam konteks ini, hubungan dengan variabel lain tidak dapat digantikan oleh indeks tunggal. Demikian pula, reliabilitas yang sesuai harus dinilai dengan bukti yang sesuai. Treagust (1988) dan AERA, APA, dan NCME (2014) memberikan dasar untuk pelaporan yang lengkap dan dapat ditelaah.

Inti metodologis subbagian ini adalah keterlacakan. Pembaca harus dapat mengikuti perjalanan dari pertanyaan, data, pengodean, analisis, sampai argumen validitas. Ketika jejak itu jelas, temuan dapat digunakan untuk memperbaiki pembelajaran tanpa mengubah diagnosis konseptual menjadi label bagi peserta didik.

7.7.4 Eksplorasi Miskonsepsi untuk Distraktor

Peneliti sering memusatkan perhatian pada teknik, padahal eksplorasi miskonsepsi untuk distraktor justru mengatur batas inferensi yang dapat dibuat. Fokus ini berada dalam proses sistematis untuk merancang, menguji, dan memvalidasi alat yang menghasilkan bukti sahih tentang konsepsi, miskonsepsi, alasan, dan keyakinan. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Wren dan Barbera (2014) serta Messick (1995) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Rancangan yang dapat diaudit mencantumkan alasan pemilihan populasi penggunaan, prosedur penanganan konstruk target, dan kriteria keputusan untuk format respons. Tanpa penjelasan itu, pembaca sulit menilai apakah uji kognitif benar-benar mendukung klaim. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean,

penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Kasus wawancara diagnostik dapat digunakan untuk memperlihatkan hubungan antara pola individual dan pola kelas. Sebagian miskonsepsi mungkin hanya dimiliki beberapa peserta tetapi sangat kuat, sedangkan pola lain tersebar luas karena bersumber dari analogi atau buku teks yang sama. Analisis perlu menjaga kedua tingkat tersebut.

Kriteria pemeriksaan sebelum analisis:

- Definisikan populasi penggunaan dalam istilah operasional.
- Tentukan bukti utama dari data uji lapangan dan bukti pendukung dari wawancara eksploratori.
- Periksa reliabilitas yang sesuai sebelum membuat kategori atau model.
- Catat potensi distraktor karangan peneliti beserta strategi mitigasinya.
- Sajikan blueprint bersama batas interpretasi dan ketidakpastian.

Pemeriksaan kualitas dapat dilakukan melalui audit terhadap konsekuensi penggunaan, analisis kasus yang bertentangan, dan evaluasi validitas isi. Apabila ditemukan kategori skor tanpa dasar, peneliti perlu menurunkan kekuatan klaim atau menambah bukti. Transparansi seperti ini meningkatkan kegunaan temuan bagi penelitian lanjutan dan praktik kelas.

Kesimpulannya, eksplorasi miskonsepsi untuk distraktor perlu dirancang sebagai proses argumentatif. Data tidak berbicara sendiri; maknanya bergantung pada definisi konstruk, kualitas prosedur, dan konteks. Oleh sebab itu, klaim akhir sebaiknya menggunakan bahasa yang sesuai dengan kekuatan bukti, misalnya “menunjukkan pola”, “mengindikasikan”, atau “mendukung penjelasan”.

7.7.5 Pengembangan Butir Satu dan Dua Tingkat

Kajian yang kuat memperlakukan pengembangan butir satu dan dua tingkat sebagai bagian dari rantai argumentasi, bukan sebagai unsur pelengkap. Fokus ini berada dalam proses sistematis untuk merancang, menguji, dan memvalidasi alat yang menghasilkan bukti sah tentang konsepsi, miskonsepsi, alasan, dan keyakinan. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Barbera (2013) serta Wren dan Barbera (2014) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Pada tahap perencanaan, interpretasi skor dan domain konten harus dibedakan secara eksplisit. Pada tahap pelaksanaan, konsistensi data uji lapangan perlu diperiksa; pada tahap interpretasi, reliabilitas yang sesuai menjadi dasar pembatasan klaim. Setiap keputusan

dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Dalam kelas yang mempelajari tes adaptif konsep mol, data agregat mungkin menunjukkan mayoritas jawaban salah. Akan tetapi, wawancara dapat memperlihatkan beberapa mekanisme berbeda: bahasa sehari-hari, generalisasi aturan, salah membaca simbol, atau model partikel yang tidak tepat. Karena itu, satu persentase tidak boleh menggantikan analisis struktur alasan.

Langkah operasional yang disarankan:

- Definisikan interpretasi skor dalam istilah operasional.
- Tentukan bukti utama dari peta konsep dan bukti pendukung dari distraktor autentik.
- Periksa validitas isi sebelum membuat kategori atau model.
- Catat potensi alpha sebagai satu-satunya bukti beserta strategi mitigasinya.
- Sajikan bank butir bersama batas interpretasi dan ketidakpastian.

Kriteria keberhasilan tidak cukup berupa banyaknya data. Yang lebih penting adalah apakah reliabilitas yang sesuai dapat ditelusuri, apakah proses respons diperiksa, dan apakah hubungan dengan variabel lain dijelaskan. Peneliti harus menghindari mengabaikan invariansi; jika risiko itu tidak dapat dihilangkan, dampaknya perlu dianalisis dan dicantumkan sebagai batas penelitian.

Pada akhirnya, keputusan tentang pengembangan butir satu dan dua tingkat menentukan apakah penelitian menghasilkan daftar jawaban salah atau pemahaman yang dapat ditindaklanjuti mengenai cara peserta didik berpikir. Sasaran utamanya ialah bukti yang cukup untuk merancang diagnosis, pembelajaran, dan penelitian lanjutan secara bertanggung jawab.

7.7.6 Pengembangan Butir Tiga dan Empat Tingkat

Subbagian ini menempatkan pengembangan butir tiga dan empat tingkat sebagai simpul penting dalam penelitian pengembangan instrumen. Fokus ini berada dalam proses sistematis untuk merancang, menguji, dan memvalidasi alat yang menghasilkan bukti sahih tentang konsepsi, miskonsepsi, alasan, dan keyakinan. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. AERA, APA, dan NCME (2014) serta DeVellis dan Thorpe (2021) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Secara operasional, peneliti perlu menetapkan konstruk target, menjelaskan format respons, dan mendokumentasikan populasi penggunaan. Keputusan tersebut kemudian diterjemahkan

ke dalam pilihan peta konsep, distraktor autentik, serta uji kognitif. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Sebagai ilustrasi, pada topik tes dua tingkat reaksi kimia, peserta didik dapat memberikan jawaban benar pada tingkat simbolik tetapi alasan submikroskopiknya keliru. Peneliti tidak cukup menghitung benar-salah; ia perlu membandingkan peta konsep dengan distraktor autentik, memeriksa tingkat keyakinan, dan menguji konsistensi pada konteks lain. Hasil analisis mungkin menunjukkan bahwa kesalahan hanya muncul pada satu representasi, atau justru merupakan hubungan konseptual yang stabil.

Keputusan yang harus didokumentasikan:

- Definiskan konstruk target dalam istilah operasional.
- Tentukan bukti utama dari wawancara eksploratori dan bukti pendukung dari penilaian ahli.
- Periksa proses respons sebelum membuat kategori atau model.
- Catat potensi validitas disebut sifat tes beserta strategi mitigasinya.
- Sajikan manual penskoran bersama batas interpretasi dan ketidakpastian.

Mutu inferensi terutama ditentukan oleh validitas isi, struktur internal, dan konsekuensi penggunaan. Ancaman utama ialah distraktor karangan peneliti serta sampel uji terlalu sempit. Karena itu, laporan perlu menyajikan bukti yang mendukung sekaligus bukti yang membatasi interpretasi. Prinsip ini konsisten dengan penekanan AERA, APA, dan NCME (2014) dan DeVellis dan Thorpe (2021) bahwa validitas melekat pada interpretasi dan penggunaan hasil, bukan sekadar pada nama instrumen atau desain.

Dengan demikian, pengembangan butir tiga dan empat tingkat berfungsi menghubungkan keputusan desain dengan interpretasi. Produk yang diharapkan bukan hanya blueprint, tetapi juga penjelasan mengenai bagaimana produk tersebut diperoleh dan dalam kondisi apa ia dapat digunakan. Penelitian miskonsepsi yang baik selalu menyisakan ruang bagi penjelasan alternatif dan verifikasi lanjutan.

7.7.7 Instrumen Respons Terbuka dan Gambar

Pembahasan mengenai instrumen respons terbuka dan gambar diperlukan agar penelitian pengembangan instrumen tidak berhenti pada prosedur administratif. Fokus ini berada dalam proses sistematis untuk merancang, menguji, dan memvalidasi alat yang menghasilkan bukti sahih tentang konsepsi, miskonsepsi, alasan, dan keyakinan. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. DeVellis dan

Thorpe (2021) serta Kane (2013) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Penerapan praktis dapat disusun dalam urutan: mendefinisikan domain konten, mengoperasionalkan fungsi diagnostik, mengumpulkan wawancara eksploratori, memeriksa proses respons, lalu melaporkan bank butir. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Misalnya, pada inventori termokimia, perubahan skor sesudah pembelajaran belum otomatis menandakan perubahan konsep. Peserta didik mungkin mempelajari pola jawaban tes. Bukti yang lebih meyakinkan mencakup transfer ke situasi baru, penjelasan kausal, representasi partikel, dan retensi setelah jeda waktu.

Pertanyaan audit metodologis:

- Definisikan domain konten dalam istilah operasional.
- Tentukan bukti utama dari distraktor autentik dan bukti pendukung dari uji kognitif.
- Periksa struktur internal sebelum membuat kategori atau model.
- Catat potensi sampel uji terlalu sempit beserta strategi mitigasinya.
- Sajikan argumen validitas bersama batas interpretasi dan ketidakpastian.

Temuan yang baik tidak selalu berarti hasil yang besar atau sesuai harapan. Temuan bernilai ketika proses respons terjaga, hubungan dengan variabel lain dijelaskan, dan ketidakpastian disajikan. Kerangka DeVellis dan Thorpe (2021) serta Kane (2013) menolong peneliti membedakan kekuatan bukti dari daya tarik narasi.

Inti metodologis subbagian ini adalah keterlacakan. Pembaca harus dapat mengikuti perjalanan dari pertanyaan, data, pengodean, analisis, sampai bank butir. Ketika jejak itu jelas, temuan dapat digunakan untuk memperbaiki pembelajaran tanpa mengubah diagnosis konseptual menjadi label bagi peserta didik.

7.7.8 Validasi Isi oleh Pakar

Dalam metodologi penelitian miskonsepsi kimia, validasi isi oleh pakar menentukan mutu hubungan antara data dan kesimpulan. Fokus ini berada dalam proses sistematis untuk merancang, menguji, dan memvalidasi alat yang menghasilkan bukti sah tentang konsepsi, miskonsepsi, alasan, dan keyakinan. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Boateng et al. (2018) serta Barbera (2013) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Prosedur minimal mencakup definisi format respons, alasan memilih distraktor autentik, aturan mengubah respons menjadi kategori, serta mekanisme pemeriksaan struktur internal. Semua penyimpangan dari rencana awal perlu dicatat. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Pada penelitian mengenai inventori redoks dengan keyakinan, peneliti dapat menyusun beberapa tugas ekuivalen yang berubah konteks, representasi, atau tuntutan penalarannya. Jika pola respons tetap berulang dan disertai keyakinan tinggi, terdapat dasar yang lebih kuat untuk menyebut indikasi miskonsepsi. Jika pola berubah-ubah, interpretasi yang lebih hati-hati adalah pengetahuan parsial atau ketergantungan konteks.

Bukti minimum yang perlu tersedia:

- Definisikan format respons dalam istilah operasional.
- Tentukan bukti utama dari penilaian ahli dan bukti pendukung dari data uji lapangan.
- Periksa hubungan dengan variabel lain sebelum membuat kategori atau model.
- Catat potensi kategori skor tanpa dasar beserta strategi mitigasinya.
- Sajikan parameter butir bersama batas interpretasi dan ketidakpastian.

Pemeriksaan mutu hendaknya dilakukan selama penelitian, bukan hanya pada akhir. Catatan tentang struktur internal, keputusan terkait konsekuensi penggunaan, dan kejadian yang berpotensi menimbulkan validitas disebut sifat tes perlu dipelihara dalam *audit trail*. Hal ini memungkinkan koreksi dini dan pelaporan yang jujur.

Kesimpulannya, validasi isi oleh pakar perlu dirancang sebagai proses argumentatif. Data tidak berbicara sendiri; maknanya bergantung pada definisi konstruk, kualitas prosedur, dan konteks. Oleh sebab itu, klaim akhir sebaiknya menggunakan bahasa yang sesuai dengan kekuatan bukti, misalnya “menunjukkan pola”, “mengindikasikan”, atau “mendukung penjelasan”.

7.7.9 Wawancara Kognitif dan Proses Respons

Keputusan tentang wawancara kognitif dan proses respons harus dibuat sebelum data utama dikumpulkan, bukan ditambahkan setelah hasil diketahui. Fokus ini berada dalam proses sistematis untuk merancang, menguji, dan memvalidasi alat yang menghasilkan bukti sahih tentang konsepsi, miskonsepsi, alasan, dan keyakinan. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Messick (1995) serta Boateng et al. (2018) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Tahap kerja dimulai dengan penegasan fungsi diagnostik. Setelah itu, peneliti menyusun aturan untuk interpretasi skor dan menentukan bagaimana domain konten akan diperiksa. Bukti sebaiknya berasal dari lebih dari satu bentuk, misalnya penilaian ahli dan data uji lapangan. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Pada topik tes gambar partikel, peneliti juga perlu memperhatikan bahasa. Istilah yang sama dapat diberi makna berbeda oleh peserta didik. Klarifikasi melalui pertanyaan tindak lanjut, gambar, atau contoh tandingan membantu membedakan masalah terminologi dari kesalahan konseptual.

Kriteria pemeriksaan sebelum analisis:

- Definisikan fungsi diagnostik dalam istilah operasional.
- Tentukan bukti utama dari uji kognitif dan bukti pendukung dari peta konsep.
- Periksa konsekuensi penggunaan sebelum membuat kategori atau model.
- Catat potensi mengabaikan invariansi beserta strategi mitigasinya.
- Sajikan panduan interpretasi bersama batas interpretasi dan ketidakpastian.

Kualitas penelitian meningkat ketika peneliti menunjukkan rantai bukti dari penilaian ahli menuju argumen validitas, menyatakan asumsi, dan memeriksa alternatif penjelasan. Sebaliknya, sampel uji terlalu sempit dapat menghasilkan klaim yang tampak tegas tetapi tidak stabil. Rekomendasi praktis harus proporsional dengan bukti.

Pada akhirnya, keputusan tentang wawancara kognitif dan proses respons menentukan apakah penelitian menghasilkan daftar jawaban salah atau pemahaman yang dapat ditindaklanjuti mengenai cara peserta didik berpikir. Sasaran utamanya ialah bukti yang cukup untuk merancang diagnosis, pembelajaran, dan penelitian lanjutan secara bertanggung jawab.

7.7.10 Uji Coba Awal dan Revisi

Uji coba awal dan revisi perlu dipahami sebagai keputusan metodologis yang mempunyai konsekuensi pada desain, analisis, dan pelaporan. Fokus ini berada dalam proses sistematis untuk merancang, menguji, dan memvalidasi alat yang menghasilkan bukti sahih tentang konsepsi, miskonsepsi, alasan, dan keyakinan. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Kane (2013) serta Treagust (1988) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Peneliti sebaiknya menyiapkan matriks yang menghubungkan pertanyaan penelitian dengan populasi penggunaan, uji kognitif, teknik analisis, dan parameter butir. Matriks ini mencegah pengumpulan data yang banyak tetapi tidak relevan. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Ilustrasi tes empat tingkat asam-basa juga menunjukkan bahwa kualitas instrumen memengaruhi temuan. Distraktor yang tidak berasal dari pemikiran peserta didik dapat menghasilkan pilihan acak, sedangkan pertanyaan yang terlalu menuntut matematika dapat mengukur keterampilan hitung alih-alih konsepsi. Peneliti perlu menilai kesesuaian bukti dengan konstruk yang dimaksud.

Langkah operasional yang disarankan:

- Definisikan populasi penggunaan dalam istilah operasional.
- Tentukan bukti utama dari data uji lapangan dan bukti pendukung dari wawancara eksploratori.
- Periksa reliabilitas yang sesuai sebelum membuat kategori atau model.
- Catat potensi distraktor karangan peneliti beserta strategi mitigasinya.
- Sajikan blueprint bersama batas interpretasi dan ketidakpastian.

Peneliti sebaiknya memisahkan keputusan apriori dari keputusan eksploratori. Pemisahan tersebut membantu mengendalikan kategori skor tanpa dasar dan mencegah alpha sebagai satu-satunya bukti. Selanjutnya, konsekuensi penggunaan dan validitas isi dilaporkan bersama ketidakpastian, sehingga pembaca dapat menilai kekuatan temuan tanpa bergantung pada kesimpulan penulis semata.

Dengan demikian, uji coba awal dan revisi berfungsi menghubungkan keputusan desain dengan interpretasi. Produk yang diharapkan bukan hanya parameter butir, tetapi juga penjelasan mengenai bagaimana produk tersebut diperoleh dan dalam kondisi apa ia dapat digunakan. Penelitian miskonsepsi yang baik selalu menyisakan ruang bagi penjelasan alternatif dan verifikasi lanjutan.

7.7.11 Analisis Teori Tes Klasik

Ketelitian penelitian pada bagian analisis teori tes klasik sangat menentukan apakah pola jawaban dapat ditafsirkan sebagai bukti konseptual. Fokus ini berada dalam proses sistematis untuk merancang, menguji, dan memvalidasi alat yang menghasilkan bukti sahih tentang konsepsi, miskonsepsi, alasan, dan keyakinan. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Treagust (1988) serta AERA,

APA, dan NCME (2014) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Dokumentasi metodologis hendaknya memuat siapa yang melakukan data uji lapangan, kapan data diperoleh, bagaimana domain konten dikendalikan, dan atas dasar apa panduan interpretasi disusun. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Contoh pada kajian peta konsep terstruktur memperlihatkan pentingnya pemisahan antara ketidaktahuan, kesalahan prosedural, dan miskonsepsi. Seorang peserta dapat gagal karena lupa rumus, sedangkan peserta lain memilih jawaban yang sama karena menggunakan model mental yang tidak ilmiah. Perbedaan ini baru terlihat ketika data uji lapangan dibaca bersama wawancara eksploratori dan bukti tentang alasan.

Keputusan yang harus didokumentasikan:

- Definisikan interpretasi skor dalam istilah operasional.
- Tentukan bukti utama dari peta konsep dan bukti pendukung dari distraktor autentik.
- Periksa validitas isi sebelum membuat kategori atau model.
- Catat potensi alpha sebagai satu-satunya bukti beserta strategi mitigasinya.
- Sajikan bank butir bersama batas interpretasi dan ketidakpastian.

Standar metodologis menuntut koherensi antara konstruk, prosedur, analisis, dan klaim. Dalam konteks ini, reliabilitas yang sesuai tidak dapat digantikan oleh indeks tunggal. Demikian pula, proses respons harus dinilai dengan bukti yang sesuai. Treagust (1988) dan AERA, APA, dan NCME (2014) memberikan dasar untuk pelaporan yang lengkap dan dapat ditelaah.

Inti metodologis subbagian ini adalah keterlacakan. Pembaca harus dapat mengikuti perjalanan dari pertanyaan, data, pengodean, analisis, sampai panduan interpretasi. Ketika jejak itu jelas, temuan dapat digunakan untuk memperbaiki pembelajaran tanpa mengubah diagnosis konseptual menjadi label bagi peserta didik.

7.7.12 Analisis Rasch dan IRT

Peneliti sering memusatkan perhatian pada teknik, padahal analisis rasch dan irt justru mengatur batas inferensi yang dapat dibuat. Fokus ini berada dalam proses sistematis untuk merancang, menguji, dan memvalidasi alat yang menghasilkan bukti sahih tentang konsepsi, miskonsepsi, alasan, dan keyakinan. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Wren dan Barbera (2014) serta Messick

(1995) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Rancangan yang dapat diaudit mencantumkan alasan pemilihan konstruk target, prosedur penanganan format respons, dan kriteria keputusan untuk populasi penggunaan. Tanpa penjelasan itu, pembaca sulit menilai apakah peta konsep benar-benar mendukung klaim. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Kasus wawancara diagnostik dapat digunakan untuk memperlihatkan hubungan antara pola individual dan pola kelas. Sebagian miskonsepsi mungkin hanya dimiliki beberapa peserta tetapi sangat kuat, sedangkan pola lain tersebar luas karena bersumber dari analogi atau buku teks yang sama. Analisis perlu menjaga kedua tingkat tersebut.

Pertanyaan audit metodologis:

- Definisikan konstruk target dalam istilah operasional.
- Tentukan bukti utama dari wawancara eksploratori dan bukti pendukung dari penilaian ahli.
- Periksa proses respons sebelum membuat kategori atau model.
- Catat potensi validitas disebut sifat tes beserta strategi mitigasinya.
- Sajikan manual penskoran bersama batas interpretasi dan ketidakpastian.

Pemeriksaan kualitas dapat dilakukan melalui audit terhadap validitas isi, analisis kasus yang bertentangan, dan evaluasi struktur internal. Apabila ditemukan distraktor karangan peneliti, peneliti perlu menurunkan kekuatan klaim atau menambah bukti. Transparansi seperti ini meningkatkan kegunaan temuan bagi penelitian lanjutan dan praktik kelas.

Kesimpulannya, analisis rasch dan irt perlu dirancang sebagai proses argumentatif. Data tidak berbicara sendiri; maknanya bergantung pada definisi konstruk, kualitas prosedur, dan konteks. Oleh sebab itu, klaim akhir sebaiknya menggunakan bahasa yang sesuai dengan kekuatan bukti, misalnya “menunjukkan pola”, “mengindikasikan”, atau “mendukung penjelasan”.

7.7.13 Reliabilitas dan Ketepatan Pengukuran

Kajian yang kuat memperlakukan reliabilitas dan ketepatan pengukuran sebagai bagian dari rantai argumentasi, bukan sebagai unsur pelengkap. Fokus ini berada dalam proses sistematis untuk merancang, menguji, dan memvalidasi alat yang menghasilkan bukti sahih tentang konsepsi, miskonsepsi, alasan, dan keyakinan. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur

pemahaman, bukan objek yang dapat dilihat secara langsung. Barbera (2013) serta Wren dan Barbera (2014) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Pada tahap perencanaan, domain konten dan fungsi diagnostik harus dibedakan secara eksplisit. Pada tahap pelaksanaan, konsistensi wawancara eksploratori perlu diperiksa; pada tahap interpretasi, proses respons menjadi dasar pembatasan klaim. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Dalam kelas yang mempelajari tes adaptif konsep mol, data agregat mungkin menunjukkan mayoritas jawaban salah. Akan tetapi, wawancara dapat memperlihatkan beberapa mekanisme berbeda: bahasa sehari-hari, generalisasi aturan, salah membaca simbol, atau model partikel yang tidak tepat. Karena itu, satu persentase tidak boleh menggantikan analisis struktur alasan.

Bukti minimum yang perlu tersedia:

- Definisikan domain konten dalam istilah operasional.
- Tentukan bukti utama dari distraktor autentik dan bukti pendukung dari uji kognitif.
- Periksa struktur internal sebelum membuat kategori atau model.
- Catat potensi sampel uji terlalu sempit beserta strategi mitigasinya.
- Sajikan argumen validitas bersama batas interpretasi dan ketidakpastian.

Kriteria keberhasilan tidak cukup berupa banyaknya data. Yang lebih penting adalah apakah proses respons dapat ditelusuri, apakah hubungan dengan variabel lain diperiksa, dan apakah reliabilitas yang sesuai dijelaskan. Peneliti harus menghindari alpha sebagai satu-satunya bukti; jika risiko itu tidak dapat dihilangkan, dampaknya perlu dianalisis dan dicantumkan sebagai batas penelitian.

Pada akhirnya, keputusan tentang reliabilitas dan ketepatan pengukuran menentukan apakah penelitian menghasilkan daftar jawaban salah atau pemahaman yang dapat ditindaklanjuti mengenai cara peserta didik berpikir. Sasaran utamanya ialah bukti yang cukup untuk merancang diagnosis, pembelajaran, dan penelitian lanjutan secara bertanggung jawab.

7.7.14 Argumen Validitas dan Fairness

Subbagian ini menempatkan argumen validitas dan fairness sebagai simpul penting dalam penelitian pengembangan instrumen. Fokus ini berada dalam proses sistematis untuk merancang, menguji, dan memvalidasi alat yang menghasilkan bukti sahih tentang konsepsi, miskonsepsi, alasan, dan keyakinan. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman,

bukan objek yang dapat dilihat secara langsung. AERA, APA, dan NCME (2014) serta DeVellis dan Thorpe (2021) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Secara operasional, peneliti perlu menetapkan format respons, menjelaskan populasi penggunaan, dan mendokumentasikan konstruk target. Keputusan tersebut kemudian diterjemahkan ke dalam pilihan distraktor autentik, uji kognitif, serta peta konsep. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Sebagai ilustrasi, pada topik tes dua tingkat reaksi kimia, peserta didik dapat memberikan jawaban benar pada tingkat simbolik tetapi alasan submikroskopiknya keliru. Peneliti tidak cukup menghitung benar-salah; ia perlu membandingkan distraktor autentik dengan uji kognitif, memeriksa tingkat keyakinan, dan menguji konsistensi pada konteks lain. Hasil analisis mungkin menunjukkan bahwa kesalahan hanya muncul pada satu representasi, atau justru merupakan hubungan konseptual yang stabil.

Kriteria pemeriksaan sebelum analisis:

- Definisikan format respons dalam istilah operasional.
- Tentukan bukti utama dari penilaian ahli dan bukti pendukung dari data uji lapangan.
- Periksa hubungan dengan variabel lain sebelum membuat kategori atau model.
- Catat potensi kategori skor tanpa dasar beserta strategi mitigasinya.
- Sajikan parameter butir bersama batas interpretasi dan ketidakpastian.

Mutu inferensi terutama ditentukan oleh struktur internal, konsekuensi penggunaan, dan validitas isi. Ancaman utama ialah validitas disebut sifat tes serta mengabaikan invariansi. Karena itu, laporan perlu menyajikan bukti yang mendukung sekaligus bukti yang membatasi interpretasi. Prinsip ini konsisten dengan penekanan AERA, APA, dan NCME (2014) dan DeVellis dan Thorpe (2021) bahwa validitas melekat pada interpretasi dan penggunaan hasil, bukan sekadar pada nama instrumen atau desain.

Dengan demikian, argumen validitas dan fairness berfungsi menghubungkan keputusan desain dengan interpretasi. Produk yang diharapkan bukan hanya manual penskoran, tetapi juga penjelasan mengenai bagaimana produk tersebut diperoleh dan dalam kondisi apa ia dapat digunakan. Penelitian miskonsepsi yang baik selalu menyisakan ruang bagi penjelasan alternatif dan verifikasi lanjutan.

7.7.15 Sintesis Siklus Pengembangan Instrumen

Pembahasan mengenai sintesis siklus pengembangan instrumen diperlukan agar penelitian pengembangan instrumen tidak berhenti pada prosedur administratif. Fokus ini berada dalam proses sistematis untuk merancang, menguji, dan memvalidasi alat yang menghasilkan bukti sahih tentang konsepsi, miskonsepsi, alasan, dan keyakinan. Peneliti perlu menyatakan sejak

awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. DeVellis dan Thorpe (2021) serta Kane (2013) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Penerapan praktis dapat disusun dalam urutan: mendefinisikan fungsi diagnostik, mengoperasionalkan interpretasi skor, mengumpulkan penilaian ahli, memeriksa hubungan dengan variabel lain, lalu melaporkan argumen validitas. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Misalnya, pada inventori termokimia, perubahan skor sesudah pembelajaran belum otomatis menandakan perubahan konsep. Peserta didik mungkin mempelajari pola jawaban tes. Bukti yang lebih meyakinkan mencakup transfer ke situasi baru, penjelasan kausal, representasi partikel, dan retensi setelah jeda waktu.

Langkah operasional yang disarankan:

- Definisikan fungsi diagnostik dalam istilah operasional.
- Tentukan bukti utama dari uji kognitif dan bukti pendukung dari peta konsep.
- Periksa konsekuensi penggunaan sebelum membuat kategori atau model.
- Catat potensi mengabaikan invariansi beserta strategi mitigasinya.
- Sajikan panduan interpretasi bersama batas interpretasi dan ketidakpastian.

Temuan yang baik tidak selalu berarti hasil yang besar atau sesuai harapan. Temuan bernilai ketika hubungan dengan variabel lain terjaga, reliabilitas yang sesuai dijelaskan, dan ketidakpastian disajikan. Kerangka DeVellis dan Thorpe (2021) serta Kane (2013) menolong peneliti membedakan kekuatan bukti dari daya tarik narasi.

Inti metodologis subbagian ini adalah keterlacakan. Pembaca harus dapat mengikuti perjalanan dari pertanyaan, data, pengodean, analisis, sampai argumen validitas. Ketika jejak itu jelas, temuan dapat digunakan untuk memperbaiki pembelajaran tanpa mengubah diagnosis konseptual menjadi label bagi peserta didik.

7.8 Penelitian Design-Based Research

Subbab ini membahas penelitian iteratif yang mengembangkan sekaligus menguji desain pembelajaran untuk mendiagnosis dan meremediasi miskonsepsi dalam konteks kelas autentik. Fokus utamanya adalah menjaga keselarasan antara pertanyaan penelitian, bentuk bukti, teknik analisis, dan kekuatan klaim. Dalam konteks miskonsepsi kimia, keputusan metodologis harus mampu membedakan pola konseptual dari respons yang timbul karena

keterbatasan bahasa, kesalahan hitung, atau kondisi administrasi. Pembahasan berikut dipecah menjadi lima belas bagian agar setiap keputusan dapat ditelaah secara operasional.

7.8.1 Karakteristik Design-Based Research

Dalam metodologi penelitian miskonsepsi kimia, karakteristik design-based research menentukan mutu hubungan antara data dan kesimpulan. Fokus ini berada dalam penelitian iteratif yang mengembangkan sekaligus menguji desain pembelajaran untuk mendiagnosis dan meremediasi miskonsepsi dalam konteks kelas autentik. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Treagust (1988) serta Wang dan Hannafin (2005) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Prosedur minimal mencakup definisi indikator keberhasilan, alasan memilih artefak siswa, aturan mengubah respons menjadi kategori, serta mekanisme pemeriksaan kolaborasi pemangku kepentingan. Semua penyimpangan dari rencana awal perlu dicatat. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Pada penelitian mengenai LKPD konflik kognitif, peneliti dapat menyusun beberapa tugas ekuivalen yang berubah konteks, representasi, atau tuntutan penalarannya. Jika pola respons tetap berulang dan disertai keyakinan tinggi, terdapat dasar yang lebih kuat untuk menyebut indikasi miskonsepsi. Jika pola berubah-ubah, interpretasi yang lebih hati-hati adalah pengetahuan parsial atau ketergantungan konteks.

Keputusan yang harus didokumentasikan:

- Definisikan indikator keberhasilan dalam istilah operasional.
- Tentukan bukti utama dari catatan guru dan bukti pendukung dari observasi interaksi.
- Periksa transfer prinsip desain sebelum membuat kategori atau model.
- Catat potensi sekadar pengembangan media beserta strategi mitigasinya.
- Sajikan prototipe bersama batas interpretasi dan ketidakpastian.

Pemeriksaan mutu hendaknya dilakukan selama penelitian, bukan hanya pada akhir. Catatan tentang kolaborasi pemangku kepentingan, keputusan terkait relevansi praktik, dan kejadian yang berpotensi menimbulkan iterasi kosmetik perlu dipelihara dalam *audit trail*. Hal ini memungkinkan koreksi dini dan pelaporan yang jujur.

Kesimpulannya, karakteristik design-based research perlu dirancang sebagai proses argumentatif. Data tidak berbicara sendiri; maknanya bergantung pada definisi konstruk, kualitas prosedur, dan konteks. Oleh sebab itu, klaim akhir sebaiknya menggunakan bahasa

yang sesuai dengan kekuatan bukti, misalnya “menunjukkan pola”, “mengindikasikan”, atau “mendukung penjelasan”.

7.8.2 Analisis Masalah Praktik

Keputusan tentang analisis masalah praktik harus dibuat sebelum data utama dikumpulkan, bukan ditambahkan setelah hasil diketahui. Fokus ini berada dalam penelitian iteratif yang mengembangkan sekaligus menguji desain pembelajaran untuk mendiagnosis dan meremediasi miskonsepsi dalam konteks kelas autentik. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Talanquer (2011) serta Treagust (1988) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Tahap kerja dimulai dengan penegasan produk pengetahuan. Setelah itu, peneliti menyusun aturan untuk teori desain dan menentukan bagaimana siklus iterasi akan diperiksa. Bukti sebaiknya berasal dari lebih dari satu bentuk, misalnya catatan guru dan observasi interaksi. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Pada topik laboratorium virtual kesetimbangan, peneliti juga perlu memperhatikan bahasa. Istilah yang sama dapat diberi makna berbeda oleh peserta didik. Klarifikasi melalui pertanyaan tindak lanjut, gambar, atau contoh tandingan membantu membedakan masalah terminologi dari kesalahan konseptual.

Pertanyaan audit metodologis:

- Definisikan produk pengetahuan dalam istilah operasional.
- Tentukan bukti utama dari tes diagnostik berulang dan bukti pendukung dari log desain.
- Periksa relevansi praktik sebelum membuat kategori atau model.
- Catat potensi perubahan desain tidak terdokumentasi beserta strategi mitigasinya.
- Sajikan prinsip desain bersama batas interpretasi dan ketidakpastian.

Kualitas penelitian meningkat ketika peneliti menunjukkan rantai bukti dari catatan guru menuju panduan implementasi, menyatakan asumsi, dan memeriksa alternatif penjelasan. Sebaliknya, klaim generalisasi produk dapat menghasilkan klaim yang tampak tegas tetapi tidak stabil. Rekomendasi praktis harus proporsional dengan bukti.

Pada akhirnya, keputusan tentang analisis masalah praktik menentukan apakah penelitian menghasilkan daftar jawaban salah atau pemahaman yang dapat ditindaklanjuti mengenai

cara peserta didik berpikir. Sasaran utamanya ialah bukti yang cukup untuk merancang diagnosis, pembelajaran, dan penelitian lanjutan secara bertanggung jawab.

7.8.3 Perumusan Teori dan Dugaan Desain

Perumusan teori dan dugaan desain perlu dipahami sebagai keputusan metodologis yang mempunyai konsekuensi pada desain, analisis, dan pelaporan. Fokus ini berada dalam penelitian iteratif yang mengembangkan sekaligus menguji desain pembelajaran untuk mendiagnosis dan meremediasi miskonsepsi dalam konteks kelas autentik. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Design-Based Research Collective (2003) serta Wang dan Hannafin (2005) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Peneliti sebaiknya menyiapkan matriks yang menghubungkan pertanyaan penelitian dengan masalah praktik, tes diagnostik berulang, teknik analisis, dan prototipe. Matriks ini mencegah pengumpulan data yang banyak tetapi tidak relevan. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Ilustrasi media augmented reality ikatan juga menunjukkan bahwa kualitas instrumen memengaruhi temuan. Distraktor yang tidak berasal dari pemikiran peserta didik dapat menghasilkan pilihan acak, sedangkan pertanyaan yang terlalu menuntut matematika dapat mengukur keterampilan hitung alih-alih konsepsi. Peneliti perlu menilai kesesuaian bukti dengan konstruk yang dimaksud.

Bukti minimum yang perlu tersedia:

- Definisikan masalah praktik dalam istilah operasional.
- Tentukan bukti utama dari observasi interaksi dan bukti pendukung dari wawancara.
- Periksa ketelitian dokumentasi sebelum membuat kategori atau model.
- Catat potensi evaluasi hanya skor akhir beserta strategi mitigasinya.
- Sajikan teori lokal bersama batas interpretasi dan ketidakpastian.

Peneliti sebaiknya memisahkan keputusan apriori dari keputusan eksploratori. Pemisahan tersebut membantu mengendalikan sekadar pengembangan media dan mencegah teori tidak berkembang. Selanjutnya, relevansi praktik dan keterhubungan teori-desain dilaporkan bersama ketidakpastian, sehingga pembaca dapat menilai kekuatan temuan tanpa bergantung pada kesimpulan penulis semata.

Dengan demikian, perumusan teori dan dugaan desain berfungsi menghubungkan keputusan desain dengan interpretasi. Produk yang diharapkan bukan hanya prototipe, tetapi juga penjelasan mengenai bagaimana produk tersebut diperoleh dan dalam kondisi apa ia dapat digunakan. Penelitian miskonsepsi yang baik selalu menyisakan ruang bagi penjelasan alternatif dan verifikasi lanjutan.

7.8.4 Co-Design dengan Guru dan Peserta Didik

Ketelitian penelitian pada bagian co-design dengan guru dan peserta didik sangat menentukan apakah pola jawaban dapat ditafsirkan sebagai bukti konseptual. Fokus ini berada dalam penelitian iteratif yang mengembangkan sekaligus menguji desain pembelajaran untuk mendiagnosis dan meremediasi miskonsepsi dalam konteks kelas autentik. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Wang dan Hannafin (2005) serta Treagust (1988) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Dokumentasi metodologis hendaknya memuat siapa yang melakukan observasi interaksi, kapan data diperoleh, bagaimana siklus iterasi dikendalikan, dan atas dasar apa prinsip desain disusun. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Contoh pada kajian tutorial representasi ganda memperlihatkan pentingnya pemisahan antara ketidaktahuan, kesalahan prosedural, dan miskonsepsi. Seorang peserta dapat gagal karena lupa rumus, sedangkan peserta lain memilih jawaban yang sama karena menggunakan model mental yang tidak ilmiah. Perbedaan ini baru terlihat ketika observasi interaksi dibaca bersama wawancara dan bukti tentang alasan.

Kriteria pemeriksaan sebelum analisis:

- Definisikan teori desain dalam istilah operasional.
- Tentukan bukti utama dari log desain dan bukti pendukung dari artefak siswa.
- Periksa keterhubungan teori-desain sebelum membuat kategori atau model.
- Catat potensi teori tidak berkembang beserta strategi mitigasinya.
- Sajikan lintasan belajar bersama batas interpretasi dan ketidakpastian.

Standar metodologis menuntut koherensi antara konstruk, prosedur, analisis, dan klaim. Dalam konteks ini, ketelitian dokumentasi tidak dapat digantikan oleh indeks tunggal. Demikian pula, analisis iterasi harus dinilai dengan bukti yang sesuai. Wang dan Hannafin

(2005) dan Treagust (1988) memberikan dasar untuk pelaporan yang lengkap dan dapat ditelaah.

Inti metodologis subbagian ini adalah keterlacakan. Pembaca harus dapat mengikuti perjalanan dari pertanyaan, data, pengodean, analisis, sampai prinsip desain. Ketika jejak itu jelas, temuan dapat digunakan untuk memperbaiki pembelajaran tanpa mengubah diagnosis konseptual menjadi label bagi peserta didik.

7.8.5 Pengembangan Prototipe Awal

Peneliti sering memusatkan perhatian pada teknik, padahal pengembangan prototipe awal justru mengatur batas inferensi yang dapat dibuat. Fokus ini berada dalam penelitian iteratif yang mengembangkan sekaligus menguji desain pembelajaran untuk mendiagnosis dan meremediasi miskonsepsi dalam konteks kelas autentik. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Anderson dan Shattuck (2012) serta Wang dan Hannafin (2005) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Rancangan yang dapat diaudit mencantumkan alasan pemilihan prototipe awal, prosedur penanganan indikator keberhasilan, dan kriteria keputusan untuk masalah praktik. Tanpa penjelasan itu, pembaca sulit menilai apakah log desain benar-benar mendukung klaim. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Kasus dashboard miskonsepsi guru dapat digunakan untuk memperlihatkan hubungan antara pola individual dan pola kelas. Sebagian miskonsepsi mungkin hanya dimiliki beberapa peserta tetapi sangat kuat, sedangkan pola lain tersebar luas karena bersumber dari analogi atau buku teks yang sama. Analisis perlu menjaga kedua tingkat tersebut.

Langkah operasional yang disarankan:

- Definisikan prototipe awal dalam istilah operasional.
- Tentukan bukti utama dari wawancara dan bukti pendukung dari catatan guru.
- Periksa analisis iterasi sebelum membuat kategori atau model.
- Catat potensi iterasi kosmetik beserta strategi mitigasinya.
- Sajikan bukti revisi bersama batas interpretasi dan ketidakpastian.

Pemeriksaan kualitas dapat dilakukan melalui audit terhadap keterhubungan teori-desain, analisis kasus yang bertentangan, dan evaluasi kolaborasi pemangku kepentingan. Apabila ditemukan evaluasi hanya skor akhir, peneliti perlu menurunkan kekuatan klaim atau

menambah bukti. Transparansi seperti ini meningkatkan kegunaan temuan bagi penelitian lanjutan dan praktik kelas.

Kesimpulannya, pengembangan prototipe awal perlu dirancang sebagai proses argumentatif. Data tidak berbicara sendiri; maknanya bergantung pada definisi konstruk, kualitas prosedur, dan konteks. Oleh sebab itu, klaim akhir sebaiknya menggunakan bahasa yang sesuai dengan kekuatan bukti, misalnya “menunjukkan pola”, “mengindikasikan”, atau “mendukung penjelasan”.

7.8.6 Siklus Implementasi dan Revisi

Kajian yang kuat memperlakukan siklus implementasi dan revisi sebagai bagian dari rantai argumentasi, bukan sebagai unsur pelengkap. Fokus ini berada dalam penelitian iteratif yang mengembangkan sekaligus menguji desain pembelajaran untuk mendiagnosis dan meremediasi miskonsepsi dalam konteks kelas autentik. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. McKenney dan Reeves (2018) serta Treagust (1988) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Pada tahap perencanaan, siklus iterasi dan produk pengetahuan harus dibedakan secara eksplisit. Pada tahap pelaksanaan, konsistensi wawancara perlu diperiksa; pada tahap interpretasi, analisis iterasi menjadi dasar pembatasan klaim. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Dalam kelas yang mempelajari permainan stoikiometri, data agregat mungkin menunjukkan mayoritas jawaban salah. Akan tetapi, wawancara dapat memperlihatkan beberapa mekanisme berbeda: bahasa sehari-hari, generalisasi aturan, salah membaca simbol, atau model partikel yang tidak tepat. Karena itu, satu persentase tidak boleh menggantikan analisis struktur alasan.

Keputusan yang harus didokumentasikan:

- Definisikan siklus iterasi dalam istilah operasional.
- Tentukan bukti utama dari artefak siswa dan bukti pendukung dari tes diagnostik berulang.
- Periksa kolaborasi pemangku kepentingan sebelum membuat kategori atau model.
- Catat potensi klaim generalisasi produk beserta strategi mitigasinya.
- Sajikan panduan implementasi bersama batas interpretasi dan ketidakpastian.

Kriteria keberhasilan tidak cukup berupa banyaknya data. Yang lebih penting adalah apakah analisis iterasi dapat ditelusuri, apakah transfer prinsip desain diperiksa, dan apakah ketelitian dokumentasi dijelaskan. Peneliti harus menghindari teori tidak berkembang; jika risiko itu tidak dapat dihilangkan, dampaknya perlu dianalisis dan dicantumkan sebagai batas penelitian.

Pada akhirnya, keputusan tentang siklus implementasi dan revisi menentukan apakah penelitian menghasilkan daftar jawaban salah atau pemahaman yang dapat ditindaklanjuti mengenai cara peserta didik berpikir. Sasaran utamanya ialah bukti yang cukup untuk merancang diagnosis, pembelajaran, dan penelitian lanjutan secara bertanggung jawab.

7.8.7 Pengumpulan Data Multisumber

Subbagian ini menempatkan pengumpulan data multisumber sebagai simpul penting dalam penelitian design-based research. Fokus ini berada dalam penelitian iteratif yang mengembangkan sekaligus menguji desain pembelajaran untuk mendiagnosis dan meremediasi miskonsepsi dalam konteks kelas autentik. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Treagust (1988) serta Wang dan Hannafin (2005) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Secara operasional, peneliti perlu menetapkan indikator keberhasilan, menjelaskan masalah praktik, dan mendokumentasikan prototipe awal. Keputusan tersebut kemudian diterjemahkan ke dalam pilihan artefak siswa, tes diagnostik berulang, serta log desain. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Sebagai ilustrasi, pada topik simulasi submikroskopik larutan, peserta didik dapat memberikan jawaban benar pada tingkat simbolik tetapi alasan submikroskopiknya keliru. Peneliti tidak cukup menghitung benar-salah; ia perlu membandingkan artefak siswa dengan tes diagnostik berulang, memeriksa tingkat keyakinan, dan menguji konsistensi pada konteks lain. Hasil analisis mungkin menunjukkan bahwa kesalahan hanya muncul pada satu representasi, atau justru merupakan hubungan konseptual yang stabil.

Pertanyaan audit metodologis:

- Definisikan indikator keberhasilan dalam istilah operasional.
- Tentukan bukti utama dari catatan guru dan bukti pendukung dari observasi interaksi.
- Periksa transfer prinsip desain sebelum membuat kategori atau model.
- Catat potensi sekadar pengembangan media beserta strategi mitigasinya.

- Sajikan prototipe bersama batas interpretasi dan ketidakpastian.

Mutu inferensi terutama ditentukan oleh kolaborasi pemangku kepentingan, relevansi praktik, dan keterhubungan teori-desain. Ancaman utama ialah iterasi kosmetik serta perubahan desain tidak terdokumentasi. Karena itu, laporan perlu menyajikan bukti yang mendukung sekaligus bukti yang membatasi interpretasi. Prinsip ini konsisten dengan penekanan Treagust (1988) dan Wang dan Hannafin (2005) bahwa validitas melekat pada interpretasi dan penggunaan hasil, bukan sekadar pada nama instrumen atau desain.

Dengan demikian, pengumpulan data multisumber berfungsi menghubungkan keputusan desain dengan interpretasi. Produk yang diharapkan bukan hanya bukti revisi, tetapi juga penjelasan mengenai bagaimana produk tersebut diperoleh dan dalam kondisi apa ia dapat digunakan. Penelitian miskonsepsi yang baik selalu menyisakan ruang bagi penjelasan alternatif dan verifikasi lanjutan.

7.8.8 Analisis Proses dan Hasil

Pembahasan mengenai analisis proses dan hasil diperlukan agar penelitian design-based research tidak berhenti pada prosedur administratif. Fokus ini berada dalam penelitian iteratif yang mengembangkan sekaligus menguji desain pembelajaran untuk mendiagnosis dan meremediasi miskonsepsi dalam konteks kelas autentik. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Talanquer (2011) serta Treagust (1988) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Penerapan praktis dapat disusun dalam urutan: mendefinisikan produk pengetahuan, mengoperasionalkan teori desain, mengumpulkan catatan guru, memeriksa transfer prinsip desain, lalu melaporkan panduan implementasi. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Misalnya, pada platform diagnostik adaptif, perubahan skor sesudah pembelajaran belum otomatis menandakan perubahan konsep. Peserta didik mungkin mempelajari pola jawaban tes. Bukti yang lebih meyakinkan mencakup transfer ke situasi baru, penjelasan kausal, representasi partikel, dan retensi setelah jeda waktu.

Bukti minimum yang perlu tersedia:

- Definisikan produk pengetahuan dalam istilah operasional.
- Tentukan bukti utama dari tes diagnostik berulang dan bukti pendukung dari log desain.
- Periksa relevansi praktik sebelum membuat kategori atau model.

- Catat potensi perubahan desain tidak terdokumentasi beserta strategi mitigasinya.
- Sajikan prinsip desain bersama batas interpretasi dan ketidakpastian.

Temuan yang baik tidak selalu berarti hasil yang besar atau sesuai harapan. Temuan bernilai ketika transfer prinsip desain terjaga, ketelitian dokumentasi dijelaskan, dan ketidakpastian disajikan. Kerangka Talanquer (2011) serta Treagust (1988) menolong peneliti membedakan kekuatan bukti dari daya tarik narasi.

Inti metodologis subbagian ini adalah keterlacakan. Pembaca harus dapat mengikuti perjalanan dari pertanyaan, data, pengodean, analisis, sampai panduan implementasi. Ketika jejak itu jelas, temuan dapat digunakan untuk memperbaiki pembelajaran tanpa mengubah diagnosis konseptual menjadi label bagi peserta didik.

7.8.9 Dokumentasi Keputusan Desain

Dalam metodologi penelitian miskonsepsi kimia, dokumentasi keputusan desain menentukan mutu hubungan antara data dan kesimpulan. Fokus ini berada dalam penelitian iteratif yang mengembangkan sekaligus menguji desain pembelajaran untuk mendiagnosis dan meremediasi miskonsepsi dalam konteks kelas autentik. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Design-Based Research Collective (2003) serta Wang dan Hannafin (2005) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Prosedur minimal mencakup definisi masalah praktik, alasan memilih tes diagnostik berulang, aturan mengubah respons menjadi kategori, serta mekanisme pemeriksaan relevansi praktik. Semua penyimpangan dari rencana awal perlu dicatat. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Pada penelitian mengenai LKPD konflik kognitif, peneliti dapat menyusun beberapa tugas ekuivalen yang berubah konteks, representasi, atau tuntutan penalarannya. Jika pola respons tetap berulang dan disertai keyakinan tinggi, terdapat dasar yang lebih kuat untuk menyebut indikasi miskonsepsi. Jika pola berubah-ubah, interpretasi yang lebih hati-hati adalah pengetahuan parsial atau ketergantungan konteks.

Kriteria pemeriksaan sebelum analisis:

- Definisikan masalah praktik dalam istilah operasional.
- Tentukan bukti utama dari observasi interaksi dan bukti pendukung dari wawancara.
- Periksa ketelitian dokumentasi sebelum membuat kategori atau model.

- Catat potensi evaluasi hanya skor akhir beserta strategi mitigasinya.
- Sajikan teori lokal bersama batas interpretasi dan ketidakpastian.

Pemeriksaan mutu hendaknya dilakukan selama penelitian, bukan hanya pada akhir. Catatan tentang relevansi praktik, keputusan terkait keterhubungan teori-desain, dan kejadian yang berpotensi menimbulkan sekadar pengembangan media perlu dipelihara dalam *audit trail*. Hal ini memungkinkan koreksi dini dan pelaporan yang jujur.

Kesimpulannya, dokumentasi keputusan desain perlu dirancang sebagai proses argumentatif. Data tidak berbicara sendiri; maknanya bergantung pada definisi konstruk, kualitas prosedur, dan konteks. Oleh sebab itu, klaim akhir sebaiknya menggunakan bahasa yang sesuai dengan kekuatan bukti, misalnya “menunjukkan pola”, “mengindikasikan”, atau “mendukung penjelasan”.

7.8.10 Fidelitas dan Adaptasi Lokal

Keputusan tentang fidelitas dan adaptasi lokal harus dibuat sebelum data utama dikumpulkan, bukan ditambahkan setelah hasil diketahui. Fokus ini berada dalam penelitian iteratif yang mengembangkan sekaligus menguji desain pembelajaran untuk mendiagnosis dan meremediasi miskonsepsi dalam konteks kelas autentik. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Wang dan Hannafin (2005) serta Treagust (1988) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Tahap kerja dimulai dengan penegasan teori desain. Setelah itu, peneliti menyusun aturan untuk siklus iterasi dan menentukan bagaimana produk pengetahuan akan diperiksa. Bukti sebaiknya berasal dari lebih dari satu bentuk, misalnya observasi interaksi dan wawancara. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Pada topik laboratorium virtual kesetimbangan, peneliti juga perlu memperhatikan bahasa. Istilah yang sama dapat diberi makna berbeda oleh peserta didik. Klarifikasi melalui pertanyaan tindak lanjut, gambar, atau contoh tandingan membantu membedakan masalah terminologi dari kesalahan konseptual.

Langkah operasional yang disarankan:

- Definisikan teori desain dalam istilah operasional.
- Tentukan bukti utama dari log desain dan bukti pendukung dari artefak siswa.
- Periksa keterhubungan teori-desain sebelum membuat kategori atau model.

- Catat potensi teori tidak berkembang beserta strategi mitigasinya.
- Sajikan lintasan belajar bersama batas interpretasi dan ketidakpastian.

Kualitas penelitian meningkat ketika peneliti menunjukkan rantai bukti dari observasi interaksi menuju prinsip desain, menyatakan asumsi, dan memeriksa alternatif penjelasan. Sebaliknya, perubahan desain tidak terdokumentasi dapat menghasilkan klaim yang tampak tegas tetapi tidak stabil. Rekomendasi praktis harus proporsional dengan bukti.

Pada akhirnya, keputusan tentang fidelitas dan adaptasi lokal menentukan apakah penelitian menghasilkan daftar jawaban salah atau pemahaman yang dapat ditindaklanjuti mengenai cara peserta didik berpikir. Sasaran utamanya ialah bukti yang cukup untuk merancang diagnosis, pembelajaran, dan penelitian lanjutan secara bertanggung jawab.

7.8.11 Pembentukan Prinsip Desain

Pembentukan prinsip desain perlu dipahami sebagai keputusan metodologis yang mempunyai konsekuensi pada desain, analisis, dan pelaporan. Fokus ini berada dalam penelitian iteratif yang mengembangkan sekaligus menguji desain pembelajaran untuk mendiagnosis dan meremediasi miskonsepsi dalam konteks kelas autentik. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Anderson dan Shattuck (2012) serta Wang dan Hannafin (2005) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Peneliti sebaiknya menyiapkan matriks yang menghubungkan pertanyaan penelitian dengan prototipe awal, log desain, teknik analisis, dan teori lokal. Matriks ini mencegah pengumpulan data yang banyak tetapi tidak relevan. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Ilustrasi media augmented reality ikatan juga menunjukkan bahwa kualitas instrumen memengaruhi temuan. Distraktor yang tidak berasal dari pemikiran peserta didik dapat menghasilkan pilihan acak, sedangkan pertanyaan yang terlalu menuntut matematika dapat mengukur keterampilan hitung alih-alih konsepsi. Peneliti perlu menilai kesesuaian bukti dengan konstruk yang dimaksud.

Keputusan yang harus didokumentasikan:

- Definisikan prototipe awal dalam istilah operasional.
- Tentukan bukti utama dari wawancara dan bukti pendukung dari catatan guru.
- Periksa analisis iterasi sebelum membuat kategori atau model.
- Catat potensi iterasi kosmetik beserta strategi mitigasinya.

- Sajikan bukti revisi bersama batas interpretasi dan ketidakpastian.

Peneliti sebaiknya memisahkan keputusan apriori dari keputusan eksploratori. Pemisahan tersebut membantu mengendalikan evaluasi hanya skor akhir dan mencegah klaim generalisasi produk. Selanjutnya, keterhubungan teori-desain dan kolaborasi pemangku kepentingan dilaporkan bersama ketidakpastian, sehingga pembaca dapat menilai kekuatan temuan tanpa bergantung pada kesimpulan penulis semata.

Dengan demikian, pembentukan prinsip desain berfungsi menghubungkan keputusan desain dengan interpretasi. Produk yang diharapkan bukan hanya teori lokal, tetapi juga penjelasan mengenai bagaimana produk tersebut diperoleh dan dalam kondisi apa ia dapat digunakan. Penelitian miskonsepsi yang baik selalu menyisakan ruang bagi penjelasan alternatif dan verifikasi lanjutan.

7.8.12 Pengembangan Teori Lokal

Ketelitian penelitian pada bagian pengembangan teori lokal sangat menentukan apakah pola jawaban dapat ditafsirkan sebagai bukti konseptual. Fokus ini berada dalam penelitian iteratif yang mengembangkan sekaligus menguji desain pembelajaran untuk mendiagnosis dan meremediasi miskonsepsi dalam konteks kelas autentik. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. McKenney dan Reeves (2018) serta Treagust (1988) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Dokumentasi metodologis hendaknya memuat siapa yang melakukan wawancara, kapan data diperoleh, bagaimana produk pengetahuan dikendalikan, dan atas dasar apa lintasan belajar disusun. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Contoh pada kajian tutorial representasi ganda memperlihatkan pentingnya pemisahan antara ketidaktahuan, kesalahan prosedural, dan miskonsepsi. Seorang peserta dapat gagal karena lupa rumus, sedangkan peserta lain memilih jawaban yang sama karena menggunakan model mental yang tidak ilmiah. Perbedaan ini baru terlihat ketika wawancara dibaca bersama catatan guru dan bukti tentang alasan.

Pertanyaan audit metodologis:

- Definisikan siklus iterasi dalam istilah operasional.
- Tentukan bukti utama dari artefak siswa dan bukti pendukung dari tes diagnostik berulang.

- Periksa kolaborasi pemangku kepentingan sebelum membuat kategori atau model.
- Catat potensi klaim generalisasi produk beserta strategi mitigasinya.
- Sajikan panduan implementasi bersama batas interpretasi dan ketidakpastian.

Standar metodologis menuntut koherensi antara konstruk, prosedur, analisis, dan klaim. Dalam konteks ini, analisis iterasi tidak dapat digantikan oleh indeks tunggal. Demikian pula, transfer prinsip desain harus dinilai dengan bukti yang sesuai. McKenney dan Reeves (2018) dan Treagust (1988) memberikan dasar untuk pelaporan yang lengkap dan dapat ditelaah.

Inti metodologis subbagian ini adalah keterlacakan. Pembaca harus dapat mengikuti perjalanan dari pertanyaan, data, pengodean, analisis, sampai lintasan belajar. Ketika jejak itu jelas, temuan dapat digunakan untuk memperbaiki pembelajaran tanpa mengubah diagnosis konseptual menjadi label bagi peserta didik.

7.8.13 Skalabilitas dan Transfer

Peneliti sering memusatkan perhatian pada teknik, padahal skalabilitas dan transfer justru mengatur batas inferensi yang dapat dibuat. Fokus ini berada dalam penelitian iteratif yang mengembangkan sekaligus menguji desain pembelajaran untuk mendiagnosis dan meremediasi miskonsepsi dalam konteks kelas autentik. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Treagust (1988) serta Wang dan Hannafin (2005) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Rancangan yang dapat diaudit mencantumkan alasan pemilihan indikator keberhasilan, prosedur penanganan masalah praktik, dan kriteria keputusan untuk prototipe awal. Tanpa penjelasan itu, pembaca sulit menilai apakah artefak siswa benar-benar mendukung klaim. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Kasus dashboard miskonsepsi guru dapat digunakan untuk memperlihatkan hubungan antara pola individual dan pola kelas. Sebagian miskonsepsi mungkin hanya dimiliki beberapa peserta tetapi sangat kuat, sedangkan pola lain tersebar luas karena bersumber dari analogi atau buku teks yang sama. Analisis perlu menjaga kedua tingkat tersebut.

Bukti minimum yang perlu tersedia:

- Definisikan indikator keberhasilan dalam istilah operasional.
- Tentukan bukti utama dari catatan guru dan bukti pendukung dari observasi interaksi.
- Periksa transfer prinsip desain sebelum membuat kategori atau model.

- Catat potensi sekadar pengembangan media beserta strategi mitigasinya.
- Sajikan prototipe bersama batas interpretasi dan ketidakpastian.

Pemeriksaan kualitas dapat dilakukan melalui audit terhadap kolaborasi pemangku kepentingan, analisis kasus yang bertentangan, dan evaluasi relevansi praktik. Apabila ditemukan iterasi kosmetik, peneliti perlu menurunkan kekuatan klaim atau menambah bukti. Transparansi seperti ini meningkatkan kegunaan temuan bagi penelitian lanjutan dan praktik kelas.

Kesimpulannya, skalabilitas dan transfer perlu dirancang sebagai proses argumentatif. Data tidak berbicara sendiri; maknanya bergantung pada definisi konstruk, kualitas prosedur, dan konteks. Oleh sebab itu, klaim akhir sebaiknya menggunakan bahasa yang sesuai dengan kekuatan bukti, misalnya “menunjukkan pola”, “mengindikasikan”, atau “mendukung penjelasan”.

7.8.14 Pelaporan DBR yang Transparan

Kajian yang kuat memperlakukan pelaporan dbr yang transparan sebagai bagian dari rantai argumentasi, bukan sebagai unsur pelengkap. Fokus ini berada dalam penelitian iteratif yang mengembangkan sekaligus menguji desain pembelajaran untuk mendiagnosis dan meremediasi miskonsepsi dalam konteks kelas autentik. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Talanquer (2011) serta Treagust (1988) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Pada tahap perencanaan, produk pengetahuan dan teori desain harus dibedakan secara eksplisit. Pada tahap pelaksanaan, konsistensi catatan guru perlu diperiksa; pada tahap interpretasi, transfer prinsip desain menjadi dasar pembatasan klaim. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Dalam kelas yang mempelajari permainan stoikiometri, data agregat mungkin menunjukkan mayoritas jawaban salah. Akan tetapi, wawancara dapat memperlihatkan beberapa mekanisme berbeda: bahasa sehari-hari, generalisasi aturan, salah membaca simbol, atau model partikel yang tidak tepat. Karena itu, satu persentase tidak boleh menggantikan analisis struktur alasan.

Kriteria pemeriksaan sebelum analisis:

- Definiskan produk pengetahuan dalam istilah operasional.
- Tentukan bukti utama dari tes diagnostik berulang dan bukti pendukung dari log desain.

- Periksa relevansi praktik sebelum membuat kategori atau model.
- Catat potensi perubahan desain tidak terdokumentasi beserta strategi mitigasinya.
- Sajikan prinsip desain bersama batas interpretasi dan ketidakpastian.

Kriteria keberhasilan tidak cukup berupa banyaknya data. Yang lebih penting adalah apakah transfer prinsip desain dapat ditelusuri, apakah ketelitian dokumentasi diperiksa, dan apakah analisis iterasi dijelaskan. Peneliti harus menghindari klaim generalisasi produk; jika risiko itu tidak dapat dihilangkan, dampaknya perlu dianalisis dan dicantumkan sebagai batas penelitian.

Pada akhirnya, keputusan tentang pelaporan dbr yang transparan menentukan apakah penelitian menghasilkan daftar jawaban salah atau pemahaman yang dapat ditindaklanjuti mengenai cara peserta didik berpikir. Sasaran utamanya ialah bukti yang cukup untuk merancang diagnosis, pembelajaran, dan penelitian lanjutan secara bertanggung jawab.

7.8.15 Sintesis Protokol DBR Miskonsepsi

Subbagian ini menempatkan sintesis protokol dbr miskonsepsi sebagai simpul penting dalam penelitian design-based research. Fokus ini berada dalam penelitian iteratif yang mengembangkan sekaligus menguji desain pembelajaran untuk mendiagnosis dan meremediasi miskonsepsi dalam konteks kelas autentik. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Design-Based Research Collective (2003) serta Wang dan Hannafin (2005) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Secara operasional, peneliti perlu menetapkan masalah praktik, menjelaskan prototipe awal, dan mendokumentasikan indikator keberhasilan. Keputusan tersebut kemudian diterjemahkan ke dalam pilihan tes diagnostik berulang, log desain, serta artefak siswa. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Sebagai ilustrasi, pada topik simulasi submikroskopik larutan, peserta didik dapat memberikan jawaban benar pada tingkat simbolik tetapi alasan submikroskopiknya keliru. Peneliti tidak cukup menghitung benar-salah; ia perlu membandingkan tes diagnostik berulang dengan log desain, memeriksa tingkat keyakinan, dan menguji konsistensi pada konteks lain. Hasil analisis mungkin menunjukkan bahwa kesalahan hanya muncul pada satu representasi, atau justru merupakan hubungan konseptual yang stabil.

Langkah operasional yang disarankan:

- Definisikan masalah praktik dalam istilah operasional.
- Tentukan bukti utama dari observasi interaksi dan bukti pendukung dari wawancara.
- Periksa ketelitian dokumentasi sebelum membuat kategori atau model.
- Catat potensi evaluasi hanya skor akhir beserta strategi mitigasinya.
- Sajikan teori lokal bersama batas interpretasi dan ketidakpastian.

Mutu inferensi terutama ditentukan oleh relevansi praktik, keterhubungan teori-desain, dan kolaborasi pemangku kepentingan. Ancaman utama ialah sekadar pengembangan media serta teori tidak berkembang. Karena itu, laporan perlu menyajikan bukti yang mendukung sekaligus bukti yang membatasi interpretasi. Prinsip ini konsisten dengan penekanan Design-Based Research Collective (2003) dan Wang dan Hannafin (2005) bahwa validitas melekat pada interpretasi dan penggunaan hasil, bukan sekadar pada nama instrumen atau desain.

Dengan demikian, sintesis protokol dbr miskonsepsi berfungsi menghubungkan keputusan desain dengan interpretasi. Produk yang diharapkan bukan hanya prototipe, tetapi juga penjelasan mengenai bagaimana produk tersebut diperoleh dan dalam kondisi apa ia dapat digunakan. Penelitian miskonsepsi yang baik selalu menyisakan ruang bagi penjelasan alternatif dan verifikasi lanjutan.

7.9 Penelitian Tindakan Kelas

Subbab ini membahas penelitian reflektif oleh guru untuk memperbaiki diagnosis dan remediasi miskonsepsi melalui siklus perencanaan, tindakan, observasi, dan refleksi. Fokus utamanya adalah menjaga keselarasan antara pertanyaan penelitian, bentuk bukti, teknik analisis, dan kekuatan klaim. Dalam konteks miskonsepsi kimia, keputusan metodologis harus mampu membedakan pola konseptual dari respons yang timbul karena keterbatasan bahasa, kesalahan hitung, atau kondisi administrasi. Pembahasan berikut dipecah menjadi lima belas bagian agar setiap keputusan dapat ditelaah secara operasional.

7.9.1 Hakikat PTK dalam Penanganan Miskonsepsi

Pembahasan mengenai hakikat ptk dalam penanganan miskonsepsi diperlukan agar penelitian tindakan kelas tidak berhenti pada prosedur administratif. Fokus ini berada dalam penelitian reflektif oleh guru untuk memperbaiki diagnosis dan remediasi miskonsepsi melalui siklus perencanaan, tindakan, observasi, dan refleksi. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Mertler (2020) serta Kaltakci Gurel et al. (2015) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Penerapan praktis dapat disusun dalam urutan: mendefinisikan indikator perbaikan, mengoperasionalkan durasi siklus, mengumpulkan jurnal refleksi, memeriksa refleksi kritis,

lalu melaporkan catatan siklus. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Misalnya, pada POE pada laju reaksi, perubahan skor sesudah pembelajaran belum otomatis menandakan perubahan konsep. Peserta didik mungkin mempelajari pola jawaban tes. Bukti yang lebih meyakinkan mencakup transfer ke situasi baru, penjelasan kausal, representasi partikel, dan retensi setelah jeda waktu.

Keputusan yang harus didokumentasikan:

- Definisikan indikator perbaikan dalam istilah operasional.
- Tentukan bukti utama dari observasi kelas dan bukti pendukung dari diskusi kelompok.
- Periksa data sebelum-selama-sesudah sebelum membuat kategori atau model.
- Catat potensi refleksi deskriptif beserta strategi mitigasinya.
- Sajikan revisi pembelajaran bersama batas interpretasi dan ketidakpastian.

Temuan yang baik tidak selalu berarti hasil yang besar atau sesuai harapan. Temuan bernilai ketika refleksi kritis terjaga, kolaborasi dijelaskan, dan ketidakpastian disajikan. Kerangka Mertler (2020) serta Kaltakci Gurel et al. (2015) menolong peneliti membedakan kekuatan bukti dari daya tarik narasi.

Inti metodologis subbagian ini adalah keterlacakan. Pembaca harus dapat mengikuti perjalanan dari pertanyaan, data, pengodean, analisis, sampai catatan siklus. Ketika jejak itu jelas, temuan dapat digunakan untuk memperbaiki pembelajaran tanpa mengubah diagnosis konseptual menjadi label bagi peserta didik.

7.9.2 Identifikasi Masalah Konseptual Kelas

Dalam metodologi penelitian miskonsepsi kimia, identifikasi masalah konseptual kelas menentukan mutu hubungan antara data dan kesimpulan. Fokus ini berada dalam penelitian reflektif oleh guru untuk memperbaiki diagnosis dan remediasi miskonsepsi melalui siklus perencanaan, tindakan, observasi, dan refleksi. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. BERA (2024) serta Mertler (2020) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Prosedur minimal mencakup definisi tindakan pedagogis, alasan memilih observasi kelas, aturan mengubah respons menjadi kategori, serta mekanisme pemeriksaan data sebelum-selama-sesudah. Semua penyimpangan dari rencana awal perlu dicatat. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons

ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Pada penelitian mengenai peta konsep ikatan, peneliti dapat menyusun beberapa tugas ekuivalen yang berubah konteks, representasi, atau tuntutan penalarannya. Jika pola respons tetap berulang dan disertai keyakinan tinggi, terdapat dasar yang lebih kuat untuk menyebut indikasi miskonsepsi. Jika pola berubah-ubah, interpretasi yang lebih hati-hati adalah pengetahuan parsial atau ketergantungan konteks.

Pertanyaan audit metodologis:

- Definisikan tindakan pedagogis dalam istilah operasional.
- Tentukan bukti utama dari hasil kerja siswa dan bukti pendukung dari tes formatif.
- Periksa kolaborasi sebelum membuat kategori atau model.
- Catat potensi siklus kedua mengulang tanpa revisi beserta strategi mitigasinya.
- Sajikan praktik baik lokal bersama batas interpretasi dan ketidakpastian.

Pemeriksaan mutu hendaknya dilakukan selama penelitian, bukan hanya pada akhir. Catatan tentang data sebelum-selama-sesudah, keputusan terkait transparansi perubahan, dan kejadian yang berpotensi menimbulkan indikator kabur perlu dipelihara dalam *audit trail*. Hal ini memungkinkan koreksi dini dan pelaporan yang jujur.

Kesimpulannya, identifikasi masalah konseptual kelas perlu dirancang sebagai proses argumentatif. Data tidak berbicara sendiri; maknanya bergantung pada definisi konstruk, kualitas prosedur, dan konteks. Oleh sebab itu, klaim akhir sebaiknya menggunakan bahasa yang sesuai dengan kekuatan bukti, misalnya “menunjukkan pola”, “mengindikasikan”, atau “mendukung penjelasan”.

7.9.3 Perencanaan Tindakan

Keputusan tentang perencanaan tindakan harus dibuat sebelum data utama dikumpulkan, bukan ditambahkan setelah hasil diketahui. Fokus ini berada dalam penelitian reflektif oleh guru untuk memperbaiki diagnosis dan remediasi miskonsepsi melalui siklus perencanaan, tindakan, observasi, dan refleksi. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Treagust (1988) serta Kaltakci Gurel et al. (2015) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Tahap kerja dimulai dengan penegasan durasi siklus. Setelah itu, peneliti menyusun aturan untuk kriteria penghentian dan menentukan bagaimana indikator perbaikan akan diperiksa. Bukti sebaiknya berasal dari lebih dari satu bentuk, misalnya hasil kerja siswa dan tes formatif. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean,

penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Pada topik analogi termokimia, peneliti juga perlu memperhatikan bahasa. Istilah yang sama dapat diberi makna berbeda oleh peserta didik. Klarifikasi melalui pertanyaan tindak lanjut, gambar, atau contoh tandingan membantu membedakan masalah terminologi dari kesalahan konseptual.

Bukti minimum yang perlu tersedia:

- Definisikan durasi siklus dalam istilah operasional.
- Tentukan bukti utama dari diskusi kelompok dan bukti pendukung dari asesmen diagnostik.
- Periksa transparansi perubahan sebelum membuat kategori atau model.
- Catat potensi konflik peran guru-peneliti beserta strategi mitigasinya.
- Sajikan agenda siklus berikutnya bersama batas interpretasi dan ketidakpastian.

Kualitas penelitian meningkat ketika peneliti menunjukkan rantai bukti dari hasil kerja siswa menuju revisi pembelajaran, menyatakan asumsi, dan memeriksa alternatif penjelasan. Sebaliknya, refleksi deskriptif dapat menghasilkan klaim yang tampak tegas tetapi tidak stabil. Rekomendasi praktis harus proporsional dengan bukti.

Pada akhirnya, keputusan tentang perencanaan tindakan menentukan apakah penelitian menghasilkan daftar jawaban salah atau pemahaman yang dapat ditindaklanjuti mengenai cara peserta didik berpikir. Sasaran utamanya ialah bukti yang cukup untuk merancang diagnosis, pembelajaran, dan penelitian lanjutan secara bertanggung jawab.

7.9.4 Penetapan Indikator Keberhasilan

Penetapan indikator keberhasilan perlu dipahami sebagai keputusan metodologis yang mempunyai konsekuensi pada desain, analisis, dan pelaporan. Fokus ini berada dalam penelitian reflektif oleh guru untuk memperbaiki diagnosis dan remediasi miskonsepsi melalui siklus perencanaan, tindakan, observasi, dan refleksi. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Kaltakci Gurel et al. (2015) serta Mertler (2020) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Peneliti sebaiknya menyiapkan matriks yang menghubungkan pertanyaan penelitian dengan bukti proses, diskusi kelompok, teknik analisis, dan praktik baik lokal. Matriks ini mencegah pengumpulan data yang banyak tetapi tidak relevan. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data

hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Ilustrasi diskusi representasi redoks juga menunjukkan bahwa kualitas instrumen memengaruhi temuan. Distraktor yang tidak berasal dari pemikiran peserta didik dapat menghasilkan pilihan acak, sedangkan pertanyaan yang terlalu menuntut matematika dapat mengukur keterampilan hitung alih-alih konsepsi. Peneliti perlu menilai kesesuaian bukti dengan konstruk yang dimaksud.

Kriteria pemeriksaan sebelum analisis:

- Definisikan bukti proses dalam istilah operasional.
- Tentukan bukti utama dari tes formatif dan bukti pendukung dari jurnal refleksi.
- Periksa keberlanjutan perbaikan sebelum membuat kategori atau model.
- Catat potensi PTK sebagai formalitas beserta strategi mitigasinya.
- Sajikan rencana tindakan bersama batas interpretasi dan ketidakpastian.

Peneliti sebaiknya memisahkan keputusan apriori dari keputusan eksploratori. Pemisahan tersebut membantu mengendalikan siklus kedua mengulang tanpa revisi dan mencegah hanya membandingkan nilai. Selanjutnya, transparansi perubahan dan keterhubungan tindakan dan masalah dilaporkan bersama ketidakpastian, sehingga pembaca dapat menilai kekuatan temuan tanpa bergantung pada kesimpulan penulis semata.

Dengan demikian, penetapan indikator keberhasilan berfungsi menghubungkan keputusan desain dengan interpretasi. Produk yang diharapkan bukan hanya praktik baik lokal, tetapi juga penjelasan mengenai bagaimana produk tersebut diperoleh dan dalam kondisi apa ia dapat digunakan. Penelitian miskonsepsi yang baik selalu menyisakan ruang bagi penjelasan alternatif dan verifikasi lanjutan.

7.9.5 Siklus Diagnosis Awal

Ketelitian penelitian pada bagian siklus diagnosis awal sangat menentukan apakah pola jawaban dapat ditafsirkan sebagai bukti konseptual. Fokus ini berada dalam penelitian reflektif oleh guru untuk memperbaiki diagnosis dan remediasi miskonsepsi melalui siklus perencanaan, tindakan, observasi, dan refleksi. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Braun dan Clarke (2021) serta Kaltakci Gurel et al. (2015) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Dokumentasi metodologis hendaknya memuat siapa yang melakukan tes formatif, kapan data diperoleh, bagaimana indikator perbaikan dikendalikan, dan atas dasar apa agenda siklus berikutnya disusun. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen,

aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Contoh pada kajian praktikum elektrolit memperlihatkan pentingnya pemisahan antara ketidaktahuan, kesalahan prosedural, dan miskonsepsi. Seorang peserta dapat gagal karena lupa rumus, sedangkan peserta lain memilih jawaban yang sama karena menggunakan model mental yang tidak ilmiah. Perbedaan ini baru terlihat ketika tes formatif dibaca bersama jurnal refleksi dan bukti tentang alasan.

Langkah operasional yang disarankan:

- Definisikan kriteria penghentian dalam istilah operasional.
- Tentukan bukti utama dari asesmen diagnostik dan bukti pendukung dari observasi kelas.
- Periksa keterhubungan tindakan dan masalah sebelum membuat kategori atau model.
- Catat potensi hanya membandingkan nilai beserta strategi mitigasinya.
- Sajikan catatan siklus bersama batas interpretasi dan ketidakpastian.

Standar metodologis menuntut koherensi antara konstruk, prosedur, analisis, dan klaim. Dalam konteks ini, keberlanjutan perbaikan tidak dapat digantikan oleh indeks tunggal. Demikian pula, refleksi kritis harus dinilai dengan bukti yang sesuai. Braun dan Clarke (2021) dan Kaltakci Gurel et al. (2015) memberikan dasar untuk pelaporan yang lengkap dan dapat ditelaah.

Inti metodologis subbagian ini adalah keterlacakan. Pembaca harus dapat mengikuti perjalanan dari pertanyaan, data, pengodean, analisis, sampai agenda siklus berikutnya. Ketika jejak itu jelas, temuan dapat digunakan untuk memperbaiki pembelajaran tanpa mengubah diagnosis konseptual menjadi label bagi peserta didik.

7.9.6 Pelaksanaan Intervensi

Peneliti sering memusatkan perhatian pada teknik, padahal pelaksanaan intervensi justru mengatur batas inferensi yang dapat dibuat. Fokus ini berada dalam penelitian reflektif oleh guru untuk memperbaiki diagnosis dan remediasi miskonsepsi melalui siklus perencanaan, tindakan, observasi, dan refleksi. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Kemmis et al. (2014) serta Mertler (2020) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Rancangan yang dapat diaudit mencantumkan alasan pemilihan masalah prioritas kelas, prosedur penanganan tindakan pedagogis, dan kriteria keputusan untuk bukti proses. Tanpa penjelasan itu, pembaca sulit menilai apakah asesmen diagnostik benar-benar mendukung

klaim. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Kasus refutational text atom dapat digunakan untuk memperlihatkan hubungan antara pola individual dan pola kelas. Sebagian miskonsepsi mungkin hanya dimiliki beberapa peserta tetapi sangat kuat, sedangkan pola lain tersebar luas karena bersumber dari analogi atau buku teks yang sama. Analisis perlu menjaga kedua tingkat tersebut.

Keputusan yang harus didokumentasikan:

- Definisikan masalah prioritas kelas dalam istilah operasional.
- Tentukan bukti utama dari jurnal refleksi dan bukti pendukung dari hasil kerja siswa.
- Periksa refleksi kritis sebelum membuat kategori atau model.
- Catat potensi indikator kabur beserta strategi mitigasinya.
- Sajikan profil perubahan konsepsi bersama batas interpretasi dan ketidakpastian.

Pemeriksaan kualitas dapat dilakukan melalui audit terhadap keterhubungan tindakan dan masalah, analisis kasus yang bertentangan, dan evaluasi data sebelum-selama-sesudah. Apabila ditemukan PTK sebagai formalitas, peneliti perlu menurunkan kekuatan klaim atau menambah bukti. Transparansi seperti ini meningkatkan kegunaan temuan bagi penelitian lanjutan dan praktik kelas.

Kesimpulannya, pelaksanaan intervensi perlu dirancang sebagai proses argumentatif. Data tidak berbicara sendiri; maknanya bergantung pada definisi konstruk, kualitas prosedur, dan konteks. Oleh sebab itu, klaim akhir sebaiknya menggunakan bahasa yang sesuai dengan kekuatan bukti, misalnya “menunjukkan pola”, “mengindikasikan”, atau “mendukung penjelasan”.

7.9.7 Observasi Proses Perubahan Konsep

Kajian yang kuat memperlakukan observasi proses perubahan konsep sebagai bagian dari rantai argumentasi, bukan sebagai unsur pelengkap. Fokus ini berada dalam penelitian reflektif oleh guru untuk memperbaiki diagnosis dan remediasi miskonsepsi melalui siklus perencanaan, tindakan, observasi, dan refleksi. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Mertler (2020) serta Kaltakci Gurel et al. (2015) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Pada tahap perencanaan, indikator perbaikan dan durasi siklus harus dibedakan secara eksplisit. Pada tahap pelaksanaan, konsistensi jurnal refleksi perlu diperiksa; pada tahap

interpretasi, refleksi kritis menjadi dasar pembatasan klaim. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Dalam kelas yang mempelajari peer instruction stoikiometri, data agregat mungkin menunjukkan mayoritas jawaban salah. Akan tetapi, wawancara dapat memperlihatkan beberapa mekanisme berbeda: bahasa sehari-hari, generalisasi aturan, salah membaca simbol, atau model partikel yang tidak tepat. Karena itu, satu persentase tidak boleh menggantikan analisis struktur alasan.

Pertanyaan audit metodologis:

- Definisikan indikator perbaikan dalam istilah operasional.
- Tentukan bukti utama dari observasi kelas dan bukti pendukung dari diskusi kelompok.
- Periksa data sebelum-selama-sesudah sebelum membuat kategori atau model.
- Catat potensi refleksi deskriptif beserta strategi mitigasinya.
- Sajikan revisi pembelajaran bersama batas interpretasi dan ketidakpastian.

Kriteria keberhasilan tidak cukup berupa banyaknya data. Yang lebih penting adalah apakah refleksi kritis dapat ditelusuri, apakah kolaborasi diperiksa, dan apakah keberlanjutan perbaikan dijelaskan. Peneliti harus menghindari hanya membandingkan nilai; jika risiko itu tidak dapat dihilangkan, dampaknya perlu dianalisis dan dicantumkan sebagai batas penelitian.

Pada akhirnya, keputusan tentang observasi proses perubahan konsep menentukan apakah penelitian menghasilkan daftar jawaban salah atau pemahaman yang dapat ditindaklanjuti mengenai cara peserta didik berpikir. Sasaran utamanya ialah bukti yang cukup untuk merancang diagnosis, pembelajaran, dan penelitian lanjutan secara bertanggung jawab.

7.9.8 Refleksi Berbasis Bukti

Subbagian ini menempatkan refleksi berbasis bukti sebagai simpul penting dalam penelitian tindakan kelas. Fokus ini berada dalam penelitian reflektif oleh guru untuk memperbaiki diagnosis dan remediasi miskonsepsi melalui siklus perencanaan, tindakan, observasi, dan refleksi. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. BERA (2024) serta Mertler (2020) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Secara operasional, peneliti perlu menetapkan tindakan pedagogis, menjelaskan bukti proses, dan mendokumentasikan masalah prioritas kelas. Keputusan tersebut kemudian diterjemahkan ke dalam pilihan observasi kelas, diskusi kelompok, serta asesmen diagnostik.

Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Sebagai ilustrasi, pada topik konflik kognitif pada asam-basa, peserta didik dapat memberikan jawaban benar pada tingkat simbolik tetapi alasan submikroskopiknya keliru. Peneliti tidak cukup menghitung benar-salah; ia perlu membandingkan observasi kelas dengan diskusi kelompok, memeriksa tingkat keyakinan, dan menguji konsistensi pada konteks lain. Hasil analisis mungkin menunjukkan bahwa kesalahan hanya muncul pada satu representasi, atau justru merupakan hubungan konseptual yang stabil.

Bukti minimum yang perlu tersedia:

- Definisikan tindakan pedagogis dalam istilah operasional.
- Tentukan bukti utama dari hasil kerja siswa dan bukti pendukung dari tes formatif.
- Periksa kolaborasi sebelum membuat kategori atau model.
- Catat potensi siklus kedua mengulang tanpa revisi beserta strategi mitigasinya.
- Sajikan praktik baik lokal bersama batas interpretasi dan ketidakpastian.

Mutu inferensi terutama ditentukan oleh data sebelum-selama-sesudah, transparansi perubahan, dan keterhubungan tindakan dan masalah. Ancaman utama ialah indikator kabur serta konflik peran guru-peneliti. Karena itu, laporan perlu menyajikan bukti yang mendukung sekaligus bukti yang membatasi interpretasi. Prinsip ini konsisten dengan penekanan BERA (2024) dan Mertler (2020) bahwa validitas melekat pada interpretasi dan penggunaan hasil, bukan sekadar pada nama instrumen atau desain.

Dengan demikian, refleksi berbasis bukti berfungsi menghubungkan keputusan desain dengan interpretasi. Produk yang diharapkan bukan hanya profil perubahan konsepsi, tetapi juga penjelasan mengenai bagaimana produk tersebut diperoleh dan dalam kondisi apa ia dapat digunakan. Penelitian miskonsepsi yang baik selalu menyisakan ruang bagi penjelasan alternatif dan verifikasi lanjutan.

7.9.9 Revisi Antar-Siklus

Pembahasan mengenai revisi antar-siklus diperlukan agar penelitian tindakan kelas tidak berhenti pada prosedur administratif. Fokus ini berada dalam penelitian reflektif oleh guru untuk memperbaiki diagnosis dan remediasi miskonsepsi melalui siklus perencanaan, tindakan, observasi, dan refleksi. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Treagust (1988) serta Kaltakci Gurel et al.

(2015) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Penerapan praktis dapat disusun dalam urutan: mendefinisikan durasi siklus, mengoperasionalkan kriteria penghentian, mengumpulkan hasil kerja siswa, memeriksa kolaborasi, lalu melaporkan revisi pembelajaran. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Misalnya, pada POE pada laju reaksi, perubahan skor sesudah pembelajaran belum otomatis menandakan perubahan konsep. Peserta didik mungkin mempelajari pola jawaban tes. Bukti yang lebih meyakinkan mencakup transfer ke situasi baru, penjelasan kausal, representasi partikel, dan retensi setelah jeda waktu.

Kriteria pemeriksaan sebelum analisis:

- Definisikan durasi siklus dalam istilah operasional.
- Tentukan bukti utama dari diskusi kelompok dan bukti pendukung dari asesmen diagnostik.
- Periksa transparansi perubahan sebelum membuat kategori atau model.
- Catat potensi konflik peran guru-peneliti beserta strategi mitigasinya.
- Sajikan agenda siklus berikutnya bersama batas interpretasi dan ketidakpastian.

Temuan yang baik tidak selalu berarti hasil yang besar atau sesuai harapan. Temuan bernilai ketika kolaborasi terjaga, keberlanjutan perbaikan dijelaskan, dan ketidakpastian disajikan. Kerangka Treagust (1988) serta Kaltakci Gurel et al. (2015) menolong peneliti membedakan kekuatan bukti dari daya tarik narasi.

Inti metodologis subbagian ini adalah keterlacakan. Pembaca harus dapat mengikuti perjalanan dari pertanyaan, data, pengodean, analisis, sampai revisi pembelajaran. Ketika jejak itu jelas, temuan dapat digunakan untuk memperbaiki pembelajaran tanpa mengubah diagnosis konseptual menjadi label bagi peserta didik.

7.9.10 Kolaborasi Guru dan Pengamat

Dalam metodologi penelitian miskonsepsi kimia, kolaborasi guru dan pengamat menentukan mutu hubungan antara data dan kesimpulan. Fokus ini berada dalam penelitian reflektif oleh guru untuk memperbaiki diagnosis dan remediasi miskonsepsi melalui siklus perencanaan, tindakan, observasi, dan refleksi. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Kaltakci Gurel et al. (2015) serta Mertler

(2020) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Prosedur minimal mencakup definisi bukti proses, alasan memilih diskusi kelompok, aturan mengubah respons menjadi kategori, serta mekanisme pemeriksaan transparansi perubahan. Semua penyimpangan dari rencana awal perlu dicatat. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Pada penelitian mengenai peta konsep ikatan, peneliti dapat menyusun beberapa tugas ekuivalen yang berubah konteks, representasi, atau tuntutan penalarannya. Jika pola respons tetap berulang dan disertai keyakinan tinggi, terdapat dasar yang lebih kuat untuk menyebut indikasi miskonsepsi. Jika pola berubah-ubah, interpretasi yang lebih hati-hati adalah pengetahuan parsial atau ketergantungan konteks.

Langkah operasional yang disarankan:

- Definisikan bukti proses dalam istilah operasional.
- Tentukan bukti utama dari tes formatif dan bukti pendukung dari jurnal refleksi.
- Periksa keberlanjutan perbaikan sebelum membuat kategori atau model.
- Catat potensi PTK sebagai formalitas beserta strategi mitigasinya.
- Sajikan rencana tindakan bersama batas interpretasi dan ketidakpastian.

Pemeriksaan mutu hendaknya dilakukan selama penelitian, bukan hanya pada akhir. Catatan tentang transparansi perubahan, keputusan terkait keterhubungan tindakan dan masalah, dan kejadian yang berpotensi menimbulkan siklus kedua mengulang tanpa revisi perlu dipelihara dalam *audit trail*. Hal ini memungkinkan koreksi dini dan pelaporan yang jujur.

Kesimpulannya, kolaborasi guru dan pengamat perlu dirancang sebagai proses argumentatif. Data tidak berbicara sendiri; maknanya bergantung pada definisi konstruk, kualitas prosedur, dan konteks. Oleh sebab itu, klaim akhir sebaiknya menggunakan bahasa yang sesuai dengan kekuatan bukti, misalnya “menunjukkan pola”, “mengindikasikan”, atau “mendukung penjelasan”.

7.9.11 Penggunaan Data Kuantitatif dalam PTK

Keputusan tentang penggunaan data kuantitatif dalam ptk harus dibuat sebelum data utama dikumpulkan, bukan ditambahkan setelah hasil diketahui. Fokus ini berada dalam penelitian reflektif oleh guru untuk memperbaiki diagnosis dan remediasi miskonsepsi melalui siklus perencanaan, tindakan, observasi, dan refleksi. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Braun dan Clarke (2021) serta

Kaltakci Gurel et al. (2015) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Tahap kerja dimulai dengan penegasan kriteria penghentian. Setelah itu, peneliti menyusun aturan untuk indikator perbaikan dan menentukan bagaimana durasi siklus akan diperiksa. Bukti sebaiknya berasal dari lebih dari satu bentuk, misalnya tes formatif dan jurnal refleksi. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Pada topik analogi termokimia, peneliti juga perlu memperhatikan bahasa. Istilah yang sama dapat diberi makna berbeda oleh peserta didik. Klarifikasi melalui pertanyaan tindak lanjut, gambar, atau contoh tandingan membantu membedakan masalah terminologi dari kesalahan konseptual.

Keputusan yang harus didokumentasikan:

- Definisikan kriteria penghentian dalam istilah operasional.
- Tentukan bukti utama dari asesmen diagnostik dan bukti pendukung dari observasi kelas.
- Periksa keterhubungan tindakan dan masalah sebelum membuat kategori atau model.
- Catat potensi hanya membandingkan nilai beserta strategi mitigasinya.
- Sajikan catatan siklus bersama batas interpretasi dan ketidakpastian.

Kualitas penelitian meningkat ketika peneliti menunjukkan rantai bukti dari tes formatif menuju agenda siklus berikutnya, menyatakan asumsi, dan memeriksa alternatif penjelasan. Sebaliknya, konflik peran guru-peneliti dapat menghasilkan klaim yang tampak tegas tetapi tidak stabil. Rekomendasi praktis harus proporsional dengan bukti.

Pada akhirnya, keputusan tentang penggunaan data kuantitatif dalam ptk menentukan apakah penelitian menghasilkan daftar jawaban salah atau pemahaman yang dapat ditindaklanjuti mengenai cara peserta didik berpikir. Sasaran utamanya ialah bukti yang cukup untuk merancang diagnosis, pembelajaran, dan penelitian lanjutan secara bertanggung jawab.

7.9.12 Penggunaan Data Kualitatif dalam PTK

Penggunaan data kualitatif dalam ptk perlu dipahami sebagai keputusan metodologis yang mempunyai konsekuensi pada desain, analisis, dan pelaporan. Fokus ini berada dalam penelitian reflektif oleh guru untuk memperbaiki diagnosis dan remediasi miskonsepsi melalui siklus perencanaan, tindakan, observasi, dan refleksi. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Kemmis et

al. (2014) serta Mertler (2020) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Peneliti sebaiknya menyiapkan matriks yang menghubungkan pertanyaan penelitian dengan masalah prioritas kelas, asesmen diagnostik, teknik analisis, dan rencana tindakan. Matriks ini mencegah pengumpulan data yang banyak tetapi tidak relevan. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Ilustrasi diskusi representasi redoks juga menunjukkan bahwa kualitas instrumen memengaruhi temuan. Distraktor yang tidak berasal dari pemikiran peserta didik dapat menghasilkan pilihan acak, sedangkan pertanyaan yang terlalu menuntut matematika dapat mengukur keterampilan hitung alih-alih konsepsi. Peneliti perlu menilai kesesuaian bukti dengan konstruk yang dimaksud.

Pertanyaan audit metodologis:

- Definisikan masalah prioritas kelas dalam istilah operasional.
- Tentukan bukti utama dari jurnal refleksi dan bukti pendukung dari hasil kerja siswa.
- Periksa refleksi kritis sebelum membuat kategori atau model.
- Catat potensi indikator kabur beserta strategi mitigasinya.
- Sajikan profil perubahan konsepsi bersama batas interpretasi dan ketidakpastian.

Peneliti sebaiknya memisahkan keputusan apriori dari keputusan eksploratori. Pemisahan tersebut membantu mengendalikan PTK sebagai formalitas dan mencegah refleksi deskriptif. Selanjutnya, keterhubungan tindakan dan masalah dan data sebelum-selama-sesudah dilaporkan bersama ketidakpastian, sehingga pembaca dapat menilai kekuatan temuan tanpa bergantung pada kesimpulan penulis semata.

Dengan demikian, penggunaan data kualitatif dalam ptk berfungsi menghubungkan keputusan desain dengan interpretasi. Produk yang diharapkan bukan hanya rencana tindakan, tetapi juga penjelasan mengenai bagaimana produk tersebut diperoleh dan dalam kondisi apa ia dapat digunakan. Penelitian miskonsepsi yang baik selalu menyisakan ruang bagi penjelasan alternatif dan verifikasi lanjutan.

7.9.13 Etika Peran Ganda Guru-Peneliti

Ketelitian penelitian pada bagian etika peran ganda guru-peneliti sangat menentukan apakah pola jawaban dapat ditafsirkan sebagai bukti konseptual. Fokus ini berada dalam penelitian reflektif oleh guru untuk memperbaiki diagnosis dan remediasi miskonsepsi melalui siklus perencanaan, tindakan, observasi, dan refleksi. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegakan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur

pemahaman, bukan objek yang dapat dilihat secara langsung. Mertler (2020) serta Kaltakci Gurel et al. (2015) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Dokumentasi metodologis hendaknya memuat siapa yang melakukan jurnal refleksi, kapan data diperoleh, bagaimana durasi siklus dikendalikan, dan atas dasar apa catatan siklus disusun. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Contoh pada kajian praktikum elektrolit memperlihatkan pentingnya pemisahan antara ketidaktahuan, kesalahan prosedural, dan miskonsepsi. Seorang peserta dapat gagal karena lupa rumus, sedangkan peserta lain memilih jawaban yang sama karena menggunakan model mental yang tidak ilmiah. Perbedaan ini baru terlihat ketika jurnal refleksi dibaca bersama hasil kerja siswa dan bukti tentang alasan.

Bukti minimum yang perlu tersedia:

- Definisikan indikator perbaikan dalam istilah operasional.
- Tentukan bukti utama dari observasi kelas dan bukti pendukung dari diskusi kelompok.
- Periksa data sebelum-selama-sesudah sebelum membuat kategori atau model.
- Catat potensi refleksi deskriptif beserta strategi mitigasinya.
- Sajikan revisi pembelajaran bersama batas interpretasi dan ketidakpastian.

Standar metodologis menuntut koherensi antara konstruk, prosedur, analisis, dan klaim. Dalam konteks ini, refleksi kritis tidak dapat digantikan oleh indeks tunggal. Demikian pula, kolaborasi harus dinilai dengan bukti yang sesuai. Mertler (2020) dan Kaltakci Gurel et al. (2015) memberikan dasar untuk pelaporan yang lengkap dan dapat ditelaah.

Inti metodologis subbagian ini adalah keterlacakan. Pembaca harus dapat mengikuti perjalanan dari pertanyaan, data, pengodean, analisis, sampai catatan siklus. Ketika jejak itu jelas, temuan dapat digunakan untuk memperbaiki pembelajaran tanpa mengubah diagnosis konseptual menjadi label bagi peserta didik.

7.9.14 Pelaporan dan Diseminasi PTK

Peneliti sering memusatkan perhatian pada teknik, padahal pelaporan dan diseminasi ptk justru mengatur batas inferensi yang dapat dibuat. Fokus ini berada dalam penelitian reflektif oleh guru untuk memperbaiki diagnosis dan remediasi miskonsepsi melalui siklus perencanaan, tindakan, observasi, dan refleksi. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. BERA (2024) serta Mertler

(2020) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Rancangan yang dapat diaudit mencantumkan alasan pemilihan tindakan pedagogis, prosedur penanganan bukti proses, dan kriteria keputusan untuk masalah prioritas kelas. Tanpa penjelasan itu, pembaca sulit menilai apakah observasi kelas benar-benar mendukung klaim. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Kasus refutational text atom dapat digunakan untuk memperlihatkan hubungan antara pola individual dan pola kelas. Sebagian miskonsepsi mungkin hanya dimiliki beberapa peserta tetapi sangat kuat, sedangkan pola lain tersebar luas karena bersumber dari analogi atau buku teks yang sama. Analisis perlu menjaga kedua tingkat tersebut.

Kriteria pemeriksaan sebelum analisis:

- Definisikan tindakan pedagogis dalam istilah operasional.
- Tentukan bukti utama dari hasil kerja siswa dan bukti pendukung dari tes formatif.
- Periksa kolaborasi sebelum membuat kategori atau model.
- Catat potensi siklus kedua mengulang tanpa revisi beserta strategi mitigasinya.
- Sajikan praktik baik lokal bersama batas interpretasi dan ketidakpastian.

Pemeriksaan kualitas dapat dilakukan melalui audit terhadap data sebelum-selama-sesudah, analisis kasus yang bertentangan, dan evaluasi transparansi perubahan. Apabila ditemukan indikator kabur, peneliti perlu menurunkan kekuatan klaim atau menambah bukti. Transparansi seperti ini meningkatkan kegunaan temuan bagi penelitian lanjutan dan praktik kelas.

Kesimpulannya, pelaporan dan diseminasi ptk perlu dirancang sebagai proses argumentatif. Data tidak berbicara sendiri; maknanya bergantung pada definisi konstruk, kualitas prosedur, dan konteks. Oleh sebab itu, klaim akhir sebaiknya menggunakan bahasa yang sesuai dengan kekuatan bukti, misalnya “menunjukkan pola”, “mengindikasikan”, atau “mendukung penjelasan”.

7.9.15 Sintesis Protokol PTK Miskonsepsi

Kajian yang kuat memperlakukan sintesis protokol ptk miskonsepsi sebagai bagian dari rantai argumentasi, bukan sebagai unsur pelengkap. Fokus ini berada dalam penelitian reflektif oleh guru untuk memperbaiki diagnosis dan remediasi miskonsepsi melalui siklus perencanaan, tindakan, observasi, dan refleksi. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman,

bukan objek yang dapat dilihat secara langsung. Treagust (1988) serta Kaltakci Gurel et al. (2015) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Pada tahap perencanaan, durasi siklus dan kriteria penghentian harus dibedakan secara eksplisit. Pada tahap pelaksanaan, konsistensi hasil kerja siswa perlu diperiksa; pada tahap interpretasi, kolaborasi menjadi dasar pembatasan klaim. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Dalam kelas yang mempelajari peer instruction stoikiometri, data agregat mungkin menunjukkan mayoritas jawaban salah. Akan tetapi, wawancara dapat memperlihatkan beberapa mekanisme berbeda: bahasa sehari-hari, generalisasi aturan, salah membaca simbol, atau model partikel yang tidak tepat. Karena itu, satu persentase tidak boleh menggantikan analisis struktur alasan.

Langkah operasional yang disarankan:

- Definisikan durasi siklus dalam istilah operasional.
- Tentukan bukti utama dari diskusi kelompok dan bukti pendukung dari asesmen diagnostik.
- Periksa transparansi perubahan sebelum membuat kategori atau model.
- Catat potensi konflik peran guru-peneliti beserta strategi mitigasinya.
- Sajikan agenda siklus berikutnya bersama batas interpretasi dan ketidakpastian.

Kriteria keberhasilan tidak cukup berupa banyaknya data. Yang lebih penting adalah apakah kolaborasi dapat ditelusuri, apakah keberlanjutan perbaikan diperiksa, dan apakah refleksi kritis dijelaskan. Peneliti harus menghindari refleksi deskriptif; jika risiko itu tidak dapat dihilangkan, dampaknya perlu dianalisis dan dicantumkan sebagai batas penelitian.

Pada akhirnya, keputusan tentang sintesis protokol ptk miskonsepsi menentukan apakah penelitian menghasilkan daftar jawaban salah atau pemahaman yang dapat ditindaklanjuti mengenai cara peserta didik berpikir. Sasaran utamanya ialah bukti yang cukup untuk merancang diagnosis, pembelajaran, dan penelitian lanjutan secara bertanggung jawab.

7.10 Systematic Literature Review

Subbab ini membahas prosedur transparan dan dapat direplikasi untuk mengidentifikasi, menilai, serta mensintesis penelitian tentang miskonsepsi kimia. Fokus utamanya adalah menjaga keselarasan antara pertanyaan penelitian, bentuk bukti, teknik analisis, dan kekuatan klaim. Dalam konteks miskonsepsi kimia, keputusan metodologis harus mampu membedakan pola konseptual dari respons yang timbul karena keterbatasan bahasa, kesalahan hitung, atau kondisi administrasi. Pembahasan berikut dipecah menjadi lima belas bagian agar setiap keputusan dapat ditelaah secara operasional.

7.10.1 Perumusan Pertanyaan Review

Subbagian ini menempatkan perumusan pertanyaan review sebagai simpul penting dalam systematic literature review. Fokus ini berada dalam prosedur transparan dan dapat direplikasi untuk mengidentifikasi, menilai, serta mensintesis penelitian tentang miskonsepsi kimia. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Kaltakci Gurel et al. (2015) serta Page et al. (2021) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Secara operasional, peneliti perlu menetapkan penilaian mutu, menjelaskan pertanyaan review, dan mendokumentasikan strategi pencarian. Keputusan tersebut kemudian diterjemahkan ke dalam pilihan instrumen, rekaman bibliografis, serta karakteristik studi. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Sebagai ilustrasi, pada topik miskonsepsi kesetimbangan, peserta didik dapat memberikan jawaban benar pada tingkat simbolik tetapi alasan submikroskopiknya keliru. Peneliti tidak cukup menghitung benar-salah; ia perlu membandingkan instrumen dengan rekaman bibliografis, memeriksa tingkat keyakinan, dan menguji konsistensi pada konteks lain. Hasil analisis mungkin menunjukkan bahwa kesalahan hanya muncul pada satu representasi, atau justru merupakan hubungan konseptual yang stabil.

Keputusan yang harus didokumentasikan:

- Definisikan penilaian mutu dalam istilah operasional.
- Tentukan bukti utama dari indikator kualitas dan bukti pendukung dari teks penuh artikel.
- Periksa transparansi sintesis sebelum membuat kategori atau model.
- Catat potensi PRISMA diperlakukan sebagai metode beserta strategi mitigasinya.
- Sajikan diagram alir PRISMA bersama batas interpretasi dan ketidakpastian.

Mutu inferensi terutama ditentukan oleh penilaian bias, protokol apriori, dan seleksi ganda. Ancaman utama ialah vote counting sederhana serta pencarian satu basis data. Karena itu, laporan perlu menyajikan bukti yang mendukung sekaligus bukti yang membatasi interpretasi. Prinsip ini konsisten dengan penekanan Kaltakci Gurel et al. (2015) dan Page et al. (2021) bahwa validitas melekat pada interpretasi dan penggunaan hasil, bukan sekadar pada nama instrumen atau desain.

Dengan demikian, perumusan pertanyaan review berfungsi menghubungkan keputusan desain dengan interpretasi. Produk yang diharapkan bukan hanya meta-analisis bila layak, tetapi juga penjelasan mengenai bagaimana produk tersebut diperoleh dan dalam kondisi apa ia dapat digunakan. Penelitian miskonsepsi yang baik selalu menyisakan ruang bagi penjelasan alternatif dan verifikasi lanjutan.

7.10.2 Penyusunan dan Registrasi Protokol

Pembahasan mengenai penyusunan dan registrasi protokol diperlukan agar systematic literature review tidak berhenti pada prosedur administratif. Fokus ini berada dalam prosedur transparan dan dapat direplikasi untuk mengidentifikasi, menilai, serta mensintesis penelitian tentang miskonsepsi kimia. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Appelbaum et al. (2018) serta Kaltakci Gurel et al. (2015) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Penerapan praktis dapat disusun dalam urutan: mendefinisikan metode sintesis, mengoperasionalkan basis data, mengumpulkan indikator kualitas, memeriksa transparansi sintesis, lalu melaporkan agenda riset. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Misalnya, pada instrumen diagnosis kimia, perubahan skor sesudah pembelajaran belum otomatis menandakan perubahan konsep. Peserta didik mungkin mempelajari pola jawaban tes. Bukti yang lebih meyakinkan mencakup transfer ke situasi baru, penjelasan kausal, representasi partikel, dan retensi setelah jeda waktu.

Pertanyaan audit metodologis:

- Definisikan metode sintesis dalam istilah operasional.
- Tentukan bukti utama dari rekaman bibliografis dan bukti pendukung dari karakteristik studi.
- Periksa protokol apriori sebelum membuat kategori atau model.
- Catat potensi pencarian satu basis data beserta strategi mitigasinya.
- Sajikan tabel karakteristik bersama batas interpretasi dan ketidakpastian.

Temuan yang baik tidak selalu berarti hasil yang besar atau sesuai harapan. Temuan bernilai ketika transparansi sintesis terjaga, cakupan pencarian dijelaskan, dan ketidakpastian disajikan. Kerangka Appelbaum et al. (2018) serta Kaltakci Gurel et al. (2015) menolong peneliti membedakan kekuatan bukti dari daya tarik narasi.

Inti metodologis subbagian ini adalah keterlacakan. Pembaca harus dapat mengikuti perjalanan dari pertanyaan, data, pengodean, analisis, sampai agenda riset. Ketika jejak itu jelas, temuan dapat digunakan untuk memperbaiki pembelajaran tanpa mengubah diagnosis konseptual menjadi label bagi peserta didik.

7.10.3 Pemilihan Basis Data

Dalam metodologi penelitian miskonsepsi kimia, pemilihan basis data menentukan mutu hubungan antara data dan kesimpulan. Fokus ini berada dalam prosedur transparan dan dapat direplikasi untuk mengidentifikasi, menilai, serta mensintesis penelitian tentang miskonsepsi kimia. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Page et al. (2021) serta Page et al. (2021) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Prosedur minimal mencakup definisi pertanyaan review, alasan memilih rekaman bibliografis, aturan mengubah respons menjadi kategori, serta mekanisme pemeriksaan protokol apriori. Semua penyimpangan dari rencana awal perlu dicatat. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Pada penelitian mengenai intervensi perubahan konseptual, peneliti dapat menyusun beberapa tugas ekuivalen yang berubah konteks, representasi, atau tuntutan penalarannya. Jika pola respons tetap berulang dan disertai keyakinan tinggi, terdapat dasar yang lebih kuat untuk menyebut indikasi miskonsepsi. Jika pola berubah-ubah, interpretasi yang lebih hati-hati adalah pengetahuan parsial atau ketergantungan konteks.

Bukti minimum yang perlu tersedia:

- Definisikan pertanyaan review dalam istilah operasional.
- Tentukan bukti utama dari teks penuh artikel dan bukti pendukung dari temuan miskonsepsi.
- Periksa cakupan pencarian sebelum membuat kategori atau model.
- Catat potensi kata kunci terlalu sempit beserta strategi mitigasinya.
- Sajikan peta topik bersama batas interpretasi dan ketidakpastian.

Pemeriksaan mutu hendaknya dilakukan selama penelitian, bukan hanya pada akhir. Catatan tentang protokol apriori, keputusan terkait seleksi ganda, dan kejadian yang berpotensi menimbulkan PRISMA diperlakukan sebagai metode perlu dipelihara dalam *audit trail*. Hal ini memungkinkan koreksi dini dan pelaporan yang jujur.

Kesimpulannya, pemilihan basis data perlu dirancang sebagai proses argumentatif. Data tidak berbicara sendiri; maknanya bergantung pada definisi konstruk, kualitas prosedur, dan konteks. Oleh sebab itu, klaim akhir sebaiknya menggunakan bahasa yang sesuai dengan kekuatan bukti, misalnya “menunjukkan pola”, “mengindikasikan”, atau “mendukung penjelasan”.

7.10.4 Pengembangan Strategi Pencarian

Keputusan tentang pengembangan strategi pencarian harus dibuat sebelum data utama dikumpulkan, bukan ditambahkan setelah hasil diketahui. Fokus ini berada dalam prosedur transparan dan dapat direplikasi untuk mengidentifikasi, menilai, serta mensintesis penelitian tentang miskonsepsi kimia. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Page et al. (2021) serta Kaltakci Gurel et al. (2015) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Tahap kerja dimulai dengan penegasan basis data. Setelah itu, peneliti menyusun aturan untuk kriteria inklusi dan menentukan bagaimana metode sintesis akan diperiksa. Bukti sebaiknya berasal dari lebih dari satu bentuk, misalnya teks penuh artikel dan temuan miskonsepsi. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Pada topik representasi submikroskopik, peneliti juga perlu memperhatikan bahasa. Istilah yang sama dapat diberi makna berbeda oleh peserta didik. Klarifikasi melalui pertanyaan tindak lanjut, gambar, atau contoh tandingan membantu membedakan masalah terminologi dari kesalahan konseptual.

Kriteria pemeriksaan sebelum analisis:

- Definisikan basis data dalam istilah operasional.
- Tentukan bukti utama dari karakteristik studi dan bukti pendukung dari instrumen.
- Periksa seleksi ganda sebelum membuat kategori atau model.
- Catat potensi duplikasi sampel beserta strategi mitigasinya.
- Sajikan sintesis naratif bersama batas interpretasi dan ketidakpastian.

Kualitas penelitian meningkat ketika peneliti menunjukkan rantai bukti dari teks penuh artikel menuju tabel karakteristik, menyatakan asumsi, dan memeriksa alternatif penjelasan. Sebaliknya, pencarian satu basis data dapat menghasilkan klaim yang tampak tegas tetapi tidak stabil. Rekomendasi praktis harus proporsional dengan bukti.

Pada akhirnya, keputusan tentang pengembangan strategi pencarian menentukan apakah penelitian menghasilkan daftar jawaban salah atau pemahaman yang dapat ditindaklanjuti mengenai cara peserta didik berpikir. Sasaran utamanya ialah bukti yang cukup untuk merancang diagnosis, pembelajaran, dan penelitian lanjutan secara bertanggung jawab.

7.10.5 Pencarian Literatur Abu-Abu

Pencarian literatur abu-abu perlu dipahami sebagai keputusan metodologis yang mempunyai konsekuensi pada desain, analisis, dan pelaporan. Fokus ini berada dalam prosedur transparan dan dapat direplikasi untuk mengidentifikasi, menilai, serta mensintesis penelitian tentang miskonsepsi kimia. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Shea et al. (2017) serta Page et al. (2021) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Peneliti sebaiknya menyiapkan matriks yang menghubungkan pertanyaan penelitian dengan strategi pencarian, karakteristik studi, teknik analisis, dan peta topik. Matriks ini mencegah pengumpulan data yang banyak tetapi tidak relevan. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Ilustrasi keyakinan dan miskonsepsi juga menunjukkan bahwa kualitas instrumen memengaruhi temuan. Distraktor yang tidak berasal dari pemikiran peserta didik dapat menghasilkan pilihan acak, sedangkan pertanyaan yang terlalu menuntut matematika dapat mengukur keterampilan hitung alih-alih konsepsi. Peneliti perlu menilai kesesuaian bukti dengan konstruk yang dimaksud.

Langkah operasional yang disarankan:

- Definisikan strategi pencarian dalam istilah operasional.
- Tentukan bukti utama dari temuan miskonsepsi dan bukti pendukung dari indikator kualitas.
- Periksa ekstraksi teruji sebelum membuat kategori atau model.
- Catat potensi vote counting sederhana beserta strategi mitigasinya.
- Sajikan meta-analisis bila layak bersama batas interpretasi dan ketidakpastian.

Peneliti sebaiknya memisahkan keputusan apriori dari keputusan eksploratori. Pemisahan tersebut membantu mengendalikan kata kunci terlalu sempit dan mencegah mengabaikan publikasi non-Inggris. Selanjutnya, seleksi ganda dan penilaian bias dilaporkan bersama

ketidakpastian, sehingga pembaca dapat menilai kekuatan temuan tanpa bergantung pada kesimpulan penulis semata.

Dengan demikian, pencarian literatur abu-abu berfungsi menghubungkan keputusan desain dengan interpretasi. Produk yang diharapkan bukan hanya peta topik, tetapi juga penjelasan mengenai bagaimana produk tersebut diperoleh dan dalam kondisi apa ia dapat digunakan. Penelitian miskonsepsi yang baik selalu menyisakan ruang bagi penjelasan alternatif dan verifikasi lanjutan.

7.10.6 Manajemen Rekaman dan Duplikasi

Ketelitian penelitian pada bagian manajemen rekaman dan duplikasi sangat menentukan apakah pola jawaban dapat ditafsirkan sebagai bukti konseptual. Fokus ini berada dalam prosedur transparan dan dapat direplikasi untuk mengidentifikasi, menilai, serta mensintesis penelitian tentang miskonsepsi kimia. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Whiting et al. (2016) serta Kaltakci Gurel et al. (2015) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Dokumentasi metodologis hendaknya memuat siapa yang melakukan temuan miskonsepsi, kapan data diperoleh, bagaimana metode sintesis dikendalikan, dan atas dasar apa sintesis naratif disusun. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Contoh pada kajian miskonsepsi guru kimia memperlihatkan pentingnya pemisahan antara ketidaktahuan, kesalahan prosedural, dan miskonsepsi. Seorang peserta dapat gagal karena lupa rumus, sedangkan peserta lain memilih jawaban yang sama karena menggunakan model mental yang tidak ilmiah. Perbedaan ini baru terlihat ketika temuan miskonsepsi dibaca bersama indikator kualitas dan bukti tentang alasan.

Keputusan yang harus didokumentasikan:

- Definisikan kriteria inklusi dalam istilah operasional.
- Tentukan bukti utama dari instrumen dan bukti pendukung dari rekaman bibliografis.
- Periksa penilaian bias sebelum membuat kategori atau model.
- Catat potensi mengabaikan publikasi non-Inggris beserta strategi mitigasinya.
- Sajikan agenda riset bersama batas interpretasi dan ketidakpastian.

Standar metodologis menuntut koherensi antara konstruk, prosedur, analisis, dan klaim. Dalam konteks ini, ekstraksi teruji tidak dapat digantikan oleh indeks tunggal. Demikian pula,

transparansi sintesis harus dinilai dengan bukti yang sesuai. Whiting et al. (2016) dan Kaltakci Gurel et al. (2015) memberikan dasar untuk pelaporan yang lengkap dan dapat ditelaah.

Inti metodologis subbagian ini adalah keterlacakan. Pembaca harus dapat mengikuti perjalanan dari pertanyaan, data, pengodean, analisis, sampai sintesis naratif. Ketika jejak itu jelas, temuan dapat digunakan untuk memperbaiki pembelajaran tanpa mengubah diagnosis konseptual menjadi label bagi peserta didik.

7.10.7 Kriteria Inklusi dan Eksklusi

Peneliti sering memusatkan perhatian pada teknik, padahal kriteria inklusi dan eksklusi justru mengatur batas inferensi yang dapat dibuat. Fokus ini berada dalam prosedur transparan dan dapat direplikasi untuk mengidentifikasi, menilai, serta mensintesis penelitian tentang miskonsepsi kimia. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Kaltakci Gurel et al. (2015) serta Page et al. (2021) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Rancangan yang dapat diaudit mencantumkan alasan pemilihan penilaian mutu, prosedur penanganan pertanyaan review, dan kriteria keputusan untuk strategi pencarian. Tanpa penjelasan itu, pembaca sulit menilai apakah instrumen benar-benar mendukung klaim. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Kasus peta konsep kimia dapat digunakan untuk memperlihatkan hubungan antara pola individual dan pola kelas. Sebagian miskonsepsi mungkin hanya dimiliki beberapa peserta tetapi sangat kuat, sedangkan pola lain tersebar luas karena bersumber dari analogi atau buku teks yang sama. Analisis perlu menjaga kedua tingkat tersebut.

Pertanyaan audit metodologis:

- Definisikan penilaian mutu dalam istilah operasional.
- Tentukan bukti utama dari indikator kualitas dan bukti pendukung dari teks penuh artikel.
- Periksa transparansi sintesis sebelum membuat kategori atau model.
- Catat potensi PRISMA diperlakukan sebagai metode beserta strategi mitigasinya.
- Sajikan diagram alir PRISMA bersama batas interpretasi dan ketidakpastian.

Pemeriksaan kualitas dapat dilakukan melalui audit terhadap penilaian bias, analisis kasus yang bertentangan, dan evaluasi protokol apriori. Apabila ditemukan vote counting

sederhana, peneliti perlu menurunkan kekuatan klaim atau menambah bukti. Transparansi seperti ini meningkatkan kegunaan temuan bagi penelitian lanjutan dan praktik kelas.

Kesimpulannya, kriteria inklusi dan eksklusi perlu dirancang sebagai proses argumentatif. Data tidak berbicara sendiri; maknanya bergantung pada definisi konstruk, kualitas prosedur, dan konteks. Oleh sebab itu, klaim akhir sebaiknya menggunakan bahasa yang sesuai dengan kekuatan bukti, misalnya “menunjukkan pola”, “mengindikasikan”, atau “mendukung penjelasan”.

7.10.8 Proses Screening

Kajian yang kuat memperlakukan proses screening sebagai bagian dari rantai argumentasi, bukan sebagai unsur pelengkap. Fokus ini berada dalam prosedur transparan dan dapat direplikasi untuk mengidentifikasi, menilai, serta mensintesis penelitian tentang miskonsepsi kimia. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Appelbaum et al. (2018) serta Kaltakci Gurel et al. (2015) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Pada tahap perencanaan, metode sintesis dan basis data harus dibedakan secara eksplisit. Pada tahap pelaksanaan, konsistensi indikator kualitas perlu diperiksa; pada tahap interpretasi, transparansi sintesis menjadi dasar pembatasan klaim. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Dalam kelas yang mempelajari remediasi berbasis teknologi, data agregat mungkin menunjukkan mayoritas jawaban salah. Akan tetapi, wawancara dapat memperlihatkan beberapa mekanisme berbeda: bahasa sehari-hari, generalisasi aturan, salah membaca simbol, atau model partikel yang tidak tepat. Karena itu, satu persentase tidak boleh menggantikan analisis struktur alasan.

Bukti minimum yang perlu tersedia:

- Definisikan metode sintesis dalam istilah operasional.
- Tentukan bukti utama dari rekaman bibliografis dan bukti pendukung dari karakteristik studi.
- Periksa protokol apriori sebelum membuat kategori atau model.
- Catat potensi pencarian satu basis data beserta strategi mitigasinya.
- Sajikan tabel karakteristik bersama batas interpretasi dan ketidakpastian.

Kriteria keberhasilan tidak cukup berupa banyaknya data. Yang lebih penting adalah apakah transparansi sintesis dapat ditelusuri, apakah cakupan pencarian diperiksa, dan apakah ekstraksi teruji dijelaskan. Peneliti harus menghindari mengabaikan publikasi non-Inggris; jika risiko itu tidak dapat dihilangkan, dampaknya perlu dianalisis dan dicantumkan sebagai batas penelitian.

Pada akhirnya, keputusan tentang proses screening menentukan apakah penelitian menghasilkan daftar jawaban salah atau pemahaman yang dapat ditindaklanjuti mengenai cara peserta didik berpikir. Sasaran utamanya ialah bukti yang cukup untuk merancang diagnosis, pembelajaran, dan penelitian lanjutan secara bertanggung jawab.

7.10.9 Ekstraksi Data

Subbagian ini menempatkan ekstraksi data sebagai simpul penting dalam systematic literature review. Fokus ini berada dalam prosedur transparan dan dapat direplikasi untuk mengidentifikasi, menilai, serta mensintesis penelitian tentang miskonsepsi kimia. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Page et al. (2021) serta Page et al. (2021) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Secara operasional, peneliti perlu menetapkan pertanyaan review, menjelaskan strategi pencarian, dan mendokumentasikan penilaian mutu. Keputusan tersebut kemudian diterjemahkan ke dalam pilihan rekaman bibliografis, karakteristik studi, serta instrumen. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Sebagai ilustrasi, pada topik miskonsepsi kesetimbangan, peserta didik dapat memberikan jawaban benar pada tingkat simbolik tetapi alasan submikroskopiknya keliru. Peneliti tidak cukup menghitung benar-salah; ia perlu membandingkan rekaman bibliografis dengan karakteristik studi, memeriksa tingkat keyakinan, dan menguji konsistensi pada konteks lain. Hasil analisis mungkin menunjukkan bahwa kesalahan hanya muncul pada satu representasi, atau justru merupakan hubungan konseptual yang stabil.

Kriteria pemeriksaan sebelum analisis:

- Definisikan pertanyaan review dalam istilah operasional.
- Tentukan bukti utama dari teks penuh artikel dan bukti pendukung dari temuan miskonsepsi.
- Periksa cakupan pencarian sebelum membuat kategori atau model.

- Catat potensi kata kunci terlalu sempit beserta strategi mitigasinya.
- Sajikan peta topik bersama batas interpretasi dan ketidakpastian.

Mutu inferensi terutama ditentukan oleh protokol apriori, seleksi ganda, dan penilaian bias. Ancaman utama ialah PRISMA diperlakukan sebagai metode serta duplikasi sampel. Karena itu, laporan perlu menyajikan bukti yang mendukung sekaligus bukti yang membatasi interpretasi. Prinsip ini konsisten dengan penekanan Page et al. (2021) dan Page et al. (2021) bahwa validitas melekat pada interpretasi dan penggunaan hasil, bukan sekadar pada nama instrumen atau desain.

Dengan demikian, ekstraksi data berfungsi menghubungkan keputusan desain dengan interpretasi. Produk yang diharapkan bukan hanya diagram alir PRISMA, tetapi juga penjelasan mengenai bagaimana produk tersebut diperoleh dan dalam kondisi apa ia dapat digunakan. Penelitian miskonsepsi yang baik selalu menyisakan ruang bagi penjelasan alternatif dan verifikasi lanjutan.

7.10.10 Penilaian Kualitas dan Risiko Bias

Pembahasan mengenai penilaian kualitas dan risiko bias diperlukan agar systematic literature review tidak berhenti pada prosedur administratif. Fokus ini berada dalam prosedur transparan dan dapat direplikasi untuk mengidentifikasi, menilai, serta mensintesis penelitian tentang miskonsepsi kimia. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Page et al. (2021) serta Kaltakci Gurel et al. (2015) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Penerapan praktis dapat disusun dalam urutan: mendefinisikan basis data, mengoperasionalkan kriteria inklusi, mengumpulkan teks penuh artikel, memeriksa cakupan pencarian, lalu melaporkan tabel karakteristik. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Misalnya, pada instrumen diagnosis kimia, perubahan skor sesudah pembelajaran belum otomatis menandakan perubahan konsep. Peserta didik mungkin mempelajari pola jawaban tes. Bukti yang lebih meyakinkan mencakup transfer ke situasi baru, penjelasan kausal, representasi partikel, dan retensi setelah jeda waktu.

Langkah operasional yang disarankan:

- Definisikan basis data dalam istilah operasional.
- Tentukan bukti utama dari karakteristik studi dan bukti pendukung dari instrumen.

- Periksa seleksi ganda sebelum membuat kategori atau model.
- Catat potensi duplikasi sampel beserta strategi mitigasinya.
- Sajikan sintesis naratif bersama batas interpretasi dan ketidakpastian.

Temuan yang baik tidak selalu berarti hasil yang besar atau sesuai harapan. Temuan bernilai ketika cakupan pencarian terjaga, ekstraksi teruji dijelaskan, dan ketidakpastian disajikan. Kerangka Page et al. (2021) serta Kaltakci Gurel et al. (2015) menolong peneliti membedakan kekuatan bukti dari daya tarik narasi.

Inti metodologis subbagian ini adalah keterlacakan. Pembaca harus dapat mengikuti perjalanan dari pertanyaan, data, pengodean, analisis, sampai tabel karakteristik. Ketika jejak itu jelas, temuan dapat digunakan untuk memperbaiki pembelajaran tanpa mengubah diagnosis konseptual menjadi label bagi peserta didik.

7.10.11 Sintesis Naratif

Dalam metodologi penelitian miskonsepsi kimia, sintesis naratif menentukan mutu hubungan antara data dan kesimpulan. Fokus ini berada dalam prosedur transparan dan dapat direplikasi untuk mengidentifikasi, menilai, serta mensintesis penelitian tentang miskonsepsi kimia. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Shea et al. (2017) serta Page et al. (2021) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Prosedur minimal mencakup definisi strategi pencarian, alasan memilih karakteristik studi, aturan mengubah respons menjadi kategori, serta mekanisme pemeriksaan seleksi ganda. Semua penyimpangan dari rencana awal perlu dicatat. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Pada penelitian mengenai intervensi perubahan konseptual, peneliti dapat menyusun beberapa tugas ekuivalen yang berubah konteks, representasi, atau tuntutan penalarannya. Jika pola respons tetap berulang dan disertai keyakinan tinggi, terdapat dasar yang lebih kuat untuk menyebut indikasi miskonsepsi. Jika pola berubah-ubah, interpretasi yang lebih hati-hati adalah pengetahuan parsial atau ketergantungan konteks.

Keputusan yang harus didokumentasikan:

- Definisikan strategi pencarian dalam istilah operasional.
- Tentukan bukti utama dari temuan miskonsepsi dan bukti pendukung dari indikator kualitas.

- Periksa ekstraksi teruji sebelum membuat kategori atau model.
- Catat potensi vote counting sederhana beserta strategi mitigasinya.
- Sajikan meta-analisis bila layak bersama batas interpretasi dan ketidakpastian.

Pemeriksaan mutu hendaknya dilakukan selama penelitian, bukan hanya pada akhir. Catatan tentang seleksi ganda, keputusan terkait penilaian bias, dan kejadian yang berpotensi menimbulkan kata kunci terlalu sempit perlu dipelihara dalam *audit trail*. Hal ini memungkinkan koreksi dini dan pelaporan yang jujur.

Kesimpulannya, sintesis naratif perlu dirancang sebagai proses argumentatif. Data tidak berbicara sendiri; maknanya bergantung pada definisi konstruk, kualitas prosedur, dan konteks. Oleh sebab itu, klaim akhir sebaiknya menggunakan bahasa yang sesuai dengan kekuatan bukti, misalnya “menunjukkan pola”, “mengindikasikan”, atau “mendukung penjelasan”.

7.10.12 Meta-Analisis Bila Memenuhi Syarat

Keputusan tentang meta-analisis bila memenuhi syarat harus dibuat sebelum data utama dikumpulkan, bukan ditambahkan setelah hasil diketahui. Fokus ini berada dalam prosedur transparan dan dapat direplikasi untuk mengidentifikasi, menilai, serta mensintesis penelitian tentang miskonsepsi kimia. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Whiting et al. (2016) serta Kaltakci Gurel et al. (2015) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Tahap kerja dimulai dengan penegasan kriteria inklusi. Setelah itu, peneliti menyusun aturan untuk metode sintesis dan menentukan bagaimana basis data akan diperiksa. Bukti sebaiknya berasal dari lebih dari satu bentuk, misalnya temuan miskonsepsi dan indikator kualitas. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

$$Q = \sum w_i(\theta_i - \bar{\theta})^2; I^2 = \max[0, (Q - df)/Q] \times 100\%$$

Pada topik representasi submikroskopik, peneliti juga perlu memperhatikan bahasa. Istilah yang sama dapat diberi makna berbeda oleh peserta didik. Klarifikasi melalui pertanyaan tindak lanjut, gambar, atau contoh tandingan membantu membedakan masalah terminologi dari kesalahan konseptual.

Pertanyaan audit metodologis:

- Definisikan kriteria inklusi dalam istilah operasional.

- Tentukan bukti utama dari instrumen dan bukti pendukung dari rekaman bibliografis.
- Periksa penilaian bias sebelum membuat kategori atau model.
- Catat potensi mengabaikan publikasi non-Inggris beserta strategi mitigasinya.
- Sajikan agenda riset bersama batas interpretasi dan ketidakpastian.

Kualitas penelitian meningkat ketika peneliti menunjukkan rantai bukti dari temuan miskonsepsi menuju sintesis naratif, menyatakan asumsi, dan memeriksa alternatif penjelasan. Sebaliknya, duplikasi sampel dapat menghasilkan klaim yang tampak tegas tetapi tidak stabil. Rekomendasi praktis harus proporsional dengan bukti.

Pada akhirnya, keputusan tentang meta-analisis bila memenuhi syarat menentukan apakah penelitian menghasilkan daftar jawaban salah atau pemahaman yang dapat ditindaklanjuti mengenai cara peserta didik berpikir. Sasaran utamanya ialah bukti yang cukup untuk merancang diagnosis, pembelajaran, dan penelitian lanjutan secara bertanggung jawab.

7.10.13 Analisis Heterogenitas

Analisis heterogenitas perlu dipahami sebagai keputusan metodologis yang mempunyai konsekuensi pada desain, analisis, dan pelaporan. Fokus ini berada dalam prosedur transparan dan dapat direplikasi untuk mengidentifikasi, menilai, serta mensintesis penelitian tentang miskonsepsi kimia. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Kaltakci Gurel et al. (2015) serta Page et al. (2021) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Peneliti sebaiknya menyiapkan matriks yang menghubungkan pertanyaan penelitian dengan penilaian mutu, instrumen, teknik analisis, dan meta-analisis bila layak. Matriks ini mencegah pengumpulan data yang banyak tetapi tidak relevan. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Ilustrasi keyakinan dan miskonsepsi juga menunjukkan bahwa kualitas instrumen memengaruhi temuan. Distraktor yang tidak berasal dari pemikiran peserta didik dapat menghasilkan pilihan acak, sedangkan pertanyaan yang terlalu menuntut matematika dapat mengukur keterampilan hitung alih-alih konsepsi. Peneliti perlu menilai kesesuaian bukti dengan konstruk yang dimaksud.

Bukti minimum yang perlu tersedia:

- Definisikan penilaian mutu dalam istilah operasional.
- Tentukan bukti utama dari indikator kualitas dan bukti pendukung dari teks penuh artikel.

- Periksa transparansi sintesis sebelum membuat kategori atau model.
- Catat potensi PRISMA diperlakukan sebagai metode beserta strategi mitigasinya.
- Sajikan diagram alir PRISMA bersama batas interpretasi dan ketidakpastian.

Peneliti sebaiknya memisahkan keputusan apriori dari keputusan eksploratori. Pemisahan tersebut membantu mengendalikan vote counting sederhana dan mencegah pencarian satu basis data. Selanjutnya, penilaian bias dan protokol apriori dilaporkan bersama ketidakpastian, sehingga pembaca dapat menilai kekuatan temuan tanpa bergantung pada kesimpulan penulis semata.

Dengan demikian, analisis heterogenitas berfungsi menghubungkan keputusan desain dengan interpretasi. Produk yang diharapkan bukan hanya meta-analisis bila layak, tetapi juga penjelasan mengenai bagaimana produk tersebut diperoleh dan dalam kondisi apa ia dapat digunakan. Penelitian miskonsepsi yang baik selalu menyisakan ruang bagi penjelasan alternatif dan verifikasi lanjutan.

7.10.14 Pelaporan PRISMA 2020

Ketelitian penelitian pada bagian pelaporan prisma 2020 sangat menentukan apakah pola jawaban dapat ditafsirkan sebagai bukti konseptual. Fokus ini berada dalam prosedur transparan dan dapat direplikasi untuk mengidentifikasi, menilai, serta mensintesis penelitian tentang miskonsepsi kimia. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Appelbaum et al. (2018) serta Kaltakci Gurel et al. (2015) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Dokumentasi metodologis hendaknya memuat siapa yang melakukan indikator kualitas, kapan data diperoleh, bagaimana basis data dikendalikan, dan atas dasar apa agenda riset disusun. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Contoh pada kajian miskonsepsi guru kimia memperlihatkan pentingnya pemisahan antara ketidaktahuan, kesalahan prosedural, dan miskonsepsi. Seorang peserta dapat gagal karena lupa rumus, sedangkan peserta lain memilih jawaban yang sama karena menggunakan model mental yang tidak ilmiah. Perbedaan ini baru terlihat ketika indikator kualitas dibaca bersama teks penuh artikel dan bukti tentang alasan.

Kriteria pemeriksaan sebelum analisis:

- Definisikan metode sintesis dalam istilah operasional.

- Tentukan bukti utama dari rekaman bibliografis dan bukti pendukung dari karakteristik studi.
- Periksa protokol apriori sebelum membuat kategori atau model.
- Catat potensi pencarian satu basis data beserta strategi mitigasinya.
- Sajikan tabel karakteristik bersama batas interpretasi dan ketidakpastian.

Standar metodologis menuntut koherensi antara konstruk, prosedur, analisis, dan klaim. Dalam konteks ini, transparansi sintesis tidak dapat digantikan oleh indeks tunggal. Demikian pula, cakupan pencarian harus dinilai dengan bukti yang sesuai. Appelbaum et al. (2018) dan Kaltakci Gurel et al. (2015) memberikan dasar untuk pelaporan yang lengkap dan dapat ditelaah.

Inti metodologis subbagian ini adalah keterlacakan. Pembaca harus dapat mengikuti perjalanan dari pertanyaan, data, pengodean, analisis, sampai agenda riset. Ketika jejak itu jelas, temuan dapat digunakan untuk memperbaiki pembelajaran tanpa mengubah diagnosis konseptual menjadi label bagi peserta didik.

7.10.15 Sintesis SLR Miskonsepsi Kimia

Peneliti sering memusatkan perhatian pada teknik, padahal sintesis slr miskonsepsi kimia justru mengatur batas inferensi yang dapat dibuat. Fokus ini berada dalam prosedur transparan dan dapat direplikasi untuk mengidentifikasi, menilai, serta mensintesis penelitian tentang miskonsepsi kimia. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Page et al. (2021) serta Page et al. (2021) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Rancangan yang dapat diaudit mencantumkan alasan pemilihan pertanyaan review, prosedur penanganan strategi pencarian, dan kriteria keputusan untuk penilaian mutu. Tanpa penjelasan itu, pembaca sulit menilai apakah rekaman bibliografis benar-benar mendukung klaim. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Kasus peta konsep kimia dapat digunakan untuk memperlihatkan hubungan antara pola individual dan pola kelas. Sebagian miskonsepsi mungkin hanya dimiliki beberapa peserta tetapi sangat kuat, sedangkan pola lain tersebar luas karena bersumber dari analogi atau buku teks yang sama. Analisis perlu menjaga kedua tingkat tersebut.

Langkah operasional yang disarankan:

- Definisikan pertanyaan review dalam istilah operasional.
- Tentukan bukti utama dari teks penuh artikel dan bukti pendukung dari temuan miskonsepsi.
- Periksa cakupan pencarian sebelum membuat kategori atau model.
- Catat potensi kata kunci terlalu sempit beserta strategi mitigasinya.
- Sajikan peta topik bersama batas interpretasi dan ketidakpastian.

Pemeriksaan kualitas dapat dilakukan melalui audit terhadap protokol apriori, analisis kasus yang bertentangan, dan evaluasi seleksi ganda. Apabila ditemukan PRISMA diperlakukan sebagai metode, peneliti perlu menurunkan kekuatan klaim atau menambah bukti. Transparansi seperti ini meningkatkan kegunaan temuan bagi penelitian lanjutan dan praktik kelas.

Kesimpulannya, sintesis slr miskonsepsi kimia perlu dirancang sebagai proses argumentatif. Data tidak berbicara sendiri; maknanya bergantung pada definisi konstruk, kualitas prosedur, dan konteks. Oleh sebab itu, klaim akhir sebaiknya menggunakan bahasa yang sesuai dengan kekuatan bukti, misalnya “menunjukkan pola”, “mengindikasikan”, atau “mendukung penjelasan”.

7.11 Pemilihan Subjek dan Sampel

Subbab ini membahas penetapan siapa, kelas mana, atau dokumen apa yang diteliti agar bukti sesuai dengan pertanyaan dan populasi sasaran. Fokus utamanya adalah menjaga keselarasan antara pertanyaan penelitian, bentuk bukti, teknik analisis, dan kekuatan klaim. Dalam konteks miskonsepsi kimia, keputusan metodologis harus mampu membedakan pola konseptual dari respons yang timbul karena keterbatasan bahasa, kesalahan hitung, atau kondisi administrasi. Pembahasan berikut dipecah menjadi lima belas bagian agar setiap keputusan dapat ditelaah secara operasional.

7.11.1 Populasi Target dan Terjangkau

Kajian yang kuat memperlakukan populasi target dan terjangkau sebagai bagian dari rantai argumentasi, bukan sebagai unsur pelengkap. Fokus ini berada dalam penetapan siapa, kelas mana, atau dokumen apa yang diteliti agar bukti sesuai dengan pertanyaan dan populasi sasaran. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Lohr (2021) serta Robinson (2014) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Pada tahap perencanaan, populasi terjangkau dan kriteria inklusi harus dibedakan secara eksplisit. Pada tahap pelaksanaan, konsistensi data sekolah perlu diperiksa; pada tahap interpretasi, representativitas menjadi dasar pembatasan klaim. Setiap keputusan dicatat

dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Dalam kelas yang mempelajari subsample kasus miskonsepsi kuat, data agregat mungkin menunjukkan mayoritas jawaban salah. Akan tetapi, wawancara dapat memperlihatkan beberapa mekanisme berbeda: bahasa sehari-hari, generalisasi aturan, salah membaca simbol, atau model partikel yang tidak tepat. Karena itu, satu persentase tidak boleh menggantikan analisis struktur alasan.

Keputusan yang harus didokumentasikan:

- Definisikan populasi terjangkau dalam istilah operasional.
- Tentukan bukti utama dari profil peserta dan bukti pendukung dari hasil prates.
- Periksa kecukupan informasi sebelum membuat kategori atau model.
- Catat potensi kriteria berubah setelah data beserta strategi mitigasinya.
- Sajikan analisis kesetaraan bersama batas interpretasi dan ketidakpastian.

Kriteria keberhasilan tidak cukup berupa banyaknya data. Yang lebih penting adalah apakah representativitas dapat ditelusuri, apakah *power* atau presisi diperiksa, dan apakah analisis attrition dijelaskan. Peneliti harus menghindari sampel terlalu homogen; jika risiko itu tidak dapat dihilangkan, dampaknya perlu dianalisis dan dicantumkan sebagai batas penelitian.

Pada akhirnya, keputusan tentang populasi target dan terjangkau menentukan apakah penelitian menghasilkan daftar jawaban salah atau pemahaman yang dapat ditindaklanjuti mengenai cara peserta didik berpikir. Sasaran utamanya ialah bukti yang cukup untuk merancang diagnosis, pembelajaran, dan penelitian lanjutan secara bertanggung jawab.

7.11.2 Unit Analisis dan Unit Sampling

Subbagian ini menempatkan unit analisis dan unit sampling sebagai simpul penting dalam pemilihan subjek dan sampel. Fokus ini berada dalam penetapan siapa, kelas mana, atau dokumen apa yang diteliti agar bukti sesuai dengan pertanyaan dan populasi sasaran. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Lakens (2022) serta Lohr (2021) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Secara operasional, peneliti perlu menetapkan unit sampling, menjelaskan ukuran sampel, dan mendokumentasikan populasi target. Keputusan tersebut kemudian diterjemahkan ke dalam pilihan profil peserta, hasil prates, serta daftar kelas. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna

untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Sebagai ilustrasi, pada topik siswa kelas XI kesetimbangan, peserta didik dapat memberikan jawaban benar pada tingkat simbolik tetapi alasan submikroskopiknya keliru. Peneliti tidak cukup menghitung benar-salah; ia perlu membandingkan profil peserta dengan hasil prates, memeriksa tingkat keyakinan, dan menguji konsistensi pada konteks lain. Hasil analisis mungkin menunjukkan bahwa kesalahan hanya muncul pada satu representasi, atau justru merupakan hubungan konseptual yang stabil.

Pertanyaan audit metodologis:

- Definisikan unit sampling dalam istilah operasional.
- Tentukan bukti utama dari riwayat belajar dan bukti pendukung dari kesediaan partisipasi.
- Periksa power atau presisi sebelum membuat kategori atau model.
- Catat potensi overrekrutmen tanpa alasan beserta strategi mitigasinya.
- Sajikan catatan nonpartisipasi bersama batas interpretasi dan ketidakpastian.

Mutu inferensi terutama ditentukan oleh kecukupan informasi, transparansi rekrutmen, dan coverage. Ancaman utama ialah mengabaikan clustering serta menganggap besar selalu lebih baik. Karena itu, laporan perlu menyajikan bukti yang mendukung sekaligus bukti yang membatasi interpretasi. Prinsip ini konsisten dengan penekanan Lakens (2022) dan Lohr (2021) bahwa validitas melekat pada interpretasi dan penggunaan hasil, bukan sekadar pada nama instrumen atau desain.

Dengan demikian, unit analisis dan unit sampling berfungsi menghubungkan keputusan desain dengan interpretasi. Produk yang diharapkan bukan hanya tabel karakteristik, tetapi juga penjelasan mengenai bagaimana produk tersebut diperoleh dan dalam kondisi apa ia dapat digunakan. Penelitian miskonsepsi yang baik selalu menyisakan ruang bagi penjelasan alternatif dan verifikasi lanjutan.

7.11.3 Probability Sampling

Pembahasan mengenai probability sampling diperlukan agar pemilihan subjek dan sampel tidak berhenti pada prosedur administratif. Fokus ini berada dalam penetapan siapa, kelas mana, atau dokumen apa yang diteliti agar bukti sesuai dengan pertanyaan dan populasi sasaran. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Palinkas et al. (2015) serta Robinson (2014) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Penerapan praktis dapat disusun dalam urutan: mendefinisikan kriteria inklusi, mengoperasionalkan strategi rekrutmen, mengumpulkan riwayat belajar, memeriksa power

atau presisi, lalu melaporkan analisis kesetaraan. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Misalnya, pada mahasiswa kimia dasar, perubahan skor sesudah pembelajaran belum otomatis menandakan perubahan konsep. Peserta didik mungkin mempelajari pola jawaban tes. Bukti yang lebih meyakinkan mencakup transfer ke situasi baru, penjelasan kausal, representasi partikel, dan retensi setelah jeda waktu.

Bukti minimum yang perlu tersedia:

- Definisikan kriteria inklusi dalam istilah operasional.
- Tentukan bukti utama dari hasil pretes dan bukti pendukung dari daftar kelas.
- Periksa transparansi rekrutmen sebelum membuat kategori atau model.
- Catat potensi menganggap besar selalu lebih baik beserta strategi mitigasinya.
- Sajikan batas generalisasi bersama batas interpretasi dan ketidakpastian.

Temuan yang baik tidak selalu berarti hasil yang besar atau sesuai harapan. Temuan bernilai ketika *power* atau presisi terjaga, analisis attrition dijelaskan, dan ketidakpastian disajikan. Kerangka Palinkas et al. (2015) serta Robinson (2014) menolong peneliti membedakan kekuatan bukti dari daya tarik narasi.

Inti metodologis subbagian ini adalah keterlacakan. Pembaca harus dapat mengikuti perjalanan dari pertanyaan, data, pengodean, analisis, sampai analisis kesetaraan. Ketika jejak itu jelas, temuan dapat digunakan untuk memperbaiki pembelajaran tanpa mengubah diagnosis konseptual menjadi label bagi peserta didik.

7.11.4 Simple Random dan Systematic Sampling

Dalam metodologi penelitian miskonsepsi kimia, simple random dan systematic sampling menentukan mutu hubungan antara data dan kesimpulan. Fokus ini berada dalam penetapan siapa, kelas mana, atau dokumen apa yang diteliti agar bukti sesuai dengan pertanyaan dan populasi sasaran. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Robinson (2014) serta Lohr (2021) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Prosedur minimal mencakup definisi ukuran sampel, alasan memilih hasil pretes, aturan mengubah respons menjadi kategori, serta mekanisme pemeriksaan transparansi rekrutmen. Semua penyimpangan dari rencana awal perlu dicatat. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur

data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Pada penelitian mengenai guru kimia SMA, peneliti dapat menyusun beberapa tugas ekuivalen yang berubah konteks, representasi, atau tuntutan penalarannya. Jika pola respons tetap berulang dan disertai keyakinan tinggi, terdapat dasar yang lebih kuat untuk menyebut indikasi miskonsepsi. Jika pola berubah-ubah, interpretasi yang lebih hati-hati adalah pengetahuan parsial atau ketergantungan konteks.

Kriteria pemeriksaan sebelum analisis:

- Definisikan ukuran sampel dalam istilah operasional.
- Tentukan bukti utama dari kesediaan partisipasi dan bukti pendukung dari data sekolah.
- Periksa analisis attrition sebelum membuat kategori atau model.
- Catat potensi convenience sample tanpa batas klaim beserta strategi mitigasinya.
- Sajikan diagram alir sampel bersama batas interpretasi dan ketidakpastian.

Pemeriksaan mutu hendaknya dilakukan selama penelitian, bukan hanya pada akhir. Catatan tentang transparansi rekrutmen, keputusan terkait coverage, dan kejadian yang berpotensi menimbulkan overrekrutmen tanpa alasan perlu dipelihara dalam *audit trail*. Hal ini memungkinkan koreksi dini dan pelaporan yang jujur.

Kesimpulannya, simple random dan systematic sampling perlu dirancang sebagai proses argumentatif. Data tidak berbicara sendiri; maknanya bergantung pada definisi konstruk, kualitas prosedur, dan konteks. Oleh sebab itu, klaim akhir sebaiknya menggunakan bahasa yang sesuai dengan kekuatan bukti, misalnya “menunjukkan pola”, “mengindikasikan”, atau “mendukung penjelasan”.

7.11.5 Stratified Sampling

Keputusan tentang stratified sampling harus dibuat sebelum data utama dikumpulkan, bukan ditambahkan setelah hasil diketahui. Fokus ini berada dalam penetapan siapa, kelas mana, atau dokumen apa yang diteliti agar bukti sesuai dengan pertanyaan dan populasi sasaran. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Malterud et al. (2016) serta Robinson (2014) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Tahap kerja dimulai dengan penegasan strategi rekrutmen. Setelah itu, peneliti menyusun aturan untuk populasi terjangkau dan menentukan bagaimana kriteria inklusi akan diperiksa. Bukti sebaiknya berasal dari lebih dari satu bentuk, misalnya kesediaan partisipasi dan data sekolah. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan

pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Pada topik kelas dengan capaian beragam, peneliti juga perlu memperhatikan bahasa. Istilah yang sama dapat diberi makna berbeda oleh peserta didik. Klarifikasi melalui pertanyaan tindak lanjut, gambar, atau contoh tandingan membantu membedakan masalah terminologi dari kesalahan konseptual.

Langkah operasional yang disarankan:

- Definisikan strategi rekrutmen dalam istilah operasional.
- Tentukan bukti utama dari daftar kelas dan bukti pendukung dari profil peserta.
- Periksa coverage sebelum membuat kategori atau model.
- Catat potensi sampel terlalu homogen beserta strategi mitigasinya.
- Sajikan justifikasi ukuran bersama batas interpretasi dan ketidakpastian.

Kualitas penelitian meningkat ketika peneliti menunjukkan rantai bukti dari kesediaan partisipasi menuju batas generalisasi, menyatakan asumsi, dan memeriksa alternatif penjelasan. Sebaliknya, menganggap besar selalu lebih baik dapat menghasilkan klaim yang tampak tegas tetapi tidak stabil. Rekomendasi praktis harus proporsional dengan bukti.

Pada akhirnya, keputusan tentang stratified sampling menentukan apakah penelitian menghasilkan daftar jawaban salah atau pemahaman yang dapat ditindaklanjuti mengenai cara peserta didik berpikir. Sasaran utamanya ialah bukti yang cukup untuk merancang diagnosis, pembelajaran, dan penelitian lanjutan secara bertanggung jawab.

7.11.6 Cluster dan Multistage Sampling

Cluster dan multistage sampling perlu dipahami sebagai keputusan metodologis yang mempunyai konsekuensi pada desain, analisis, dan pelaporan. Fokus ini berada dalam penetapan siapa, kelas mana, atau dokumen apa yang diteliti agar bukti sesuai dengan pertanyaan dan populasi sasaran. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Cochran (1977) serta Lohr (2021) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Peneliti sebaiknya menyiapkan matriks yang menghubungkan pertanyaan penelitian dengan populasi target, daftar kelas, teknik analisis, dan diagram alir sampel. Matriks ini mencegah pengumpulan data yang banyak tetapi tidak relevan. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk

replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Ilustrasi sekolah urban dan rural juga menunjukkan bahwa kualitas instrumen memengaruhi temuan. Distraktor yang tidak berasal dari pemikiran peserta didik dapat menghasilkan pilihan acak, sedangkan pertanyaan yang terlalu menuntut matematika dapat mengukur keterampilan hitung alih-alih konsepsi. Peneliti perlu menilai kesesuaian bukti dengan konstruk yang dimaksud.

Keputusan yang harus didokumentasikan:

- Definisikan populasi target dalam istilah operasional.
- Tentukan bukti utama dari data sekolah dan bukti pendukung dari riwayat belajar.
- Periksa representativitas sebelum membuat kategori atau model.
- Catat potensi mengabaikan clustering beserta strategi mitigasinya.
- Sajikan tabel karakteristik bersama batas interpretasi dan ketidakpastian.

Peneliti sebaiknya memisahkan keputusan apriori dari keputusan eksploratori. Pemisahan tersebut membantu mengendalikan convenience sample tanpa batas klaim dan mencegah kriteria berubah setelah data. Selanjutnya, coverage dan kecukupan informasi dilaporkan bersama ketidakpastian, sehingga pembaca dapat menilai kekuatan temuan tanpa bergantung pada kesimpulan penulis semata.

Dengan demikian, cluster dan multistage sampling berfungsi menghubungkan keputusan desain dengan interpretasi. Produk yang diharapkan bukan hanya diagram alir sampel, tetapi juga penjelasan mengenai bagaimana produk tersebut diperoleh dan dalam kondisi apa ia dapat digunakan. Penelitian miskonsepsi yang baik selalu menyisakan ruang bagi penjelasan alternatif dan verifikasi lanjutan.

7.11.7 Purposive Sampling

Ketelitian penelitian pada bagian purposive sampling sangat menentukan apakah pola jawaban dapat ditafsirkan sebagai bukti konseptual. Fokus ini berada dalam penetapan siapa, kelas mana, atau dokumen apa yang diteliti agar bukti sesuai dengan pertanyaan dan populasi sasaran. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Lohr (2021) serta Robinson (2014) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Dokumentasi metodologis hendaknya memuat siapa yang melakukan data sekolah, kapan data diperoleh, bagaimana kriteria inklusi dikendalikan, dan atas dasar apa justifikasi ukuran disusun. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu

peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Contoh pada kajian peserta berkebutuhan khusus memperlihatkan pentingnya pemisahan antara ketidaktahuan, kesalahan prosedural, dan miskonsepsi. Seorang peserta dapat gagal karena lupa rumus, sedangkan peserta lain memilih jawaban yang sama karena menggunakan model mental yang tidak ilmiah. Perbedaan ini baru terlihat ketika data sekolah dibaca bersama riwayat belajar dan bukti tentang alasan.

Pertanyaan audit metodologis:

- Definisikan populasi terjangkau dalam istilah operasional.
- Tentukan bukti utama dari profil peserta dan bukti pendukung dari hasil prates.
- Periksa kecukupan informasi sebelum membuat kategori atau model.
- Catat potensi kriteria berubah setelah data beserta strategi mitigasinya.
- Sajikan analisis kesetaraan bersama batas interpretasi dan ketidakpastian.

Standar metodologis menuntut koherensi antara konstruk, prosedur, analisis, dan klaim. Dalam konteks ini, representativitas tidak dapat digantikan oleh indeks tunggal. Demikian pula, *power* atau presisi harus dinilai dengan bukti yang sesuai. Lohr (2021) dan Robinson (2014) memberikan dasar untuk pelaporan yang lengkap dan dapat ditelaah.

Inti metodologis subbagian ini adalah keterlacakan. Pembaca harus dapat mengikuti perjalanan dari pertanyaan, data, pengodean, analisis, sampai justifikasi ukuran. Ketika jejak itu jelas, temuan dapat digunakan untuk memperbaiki pembelajaran tanpa mengubah diagnosis konseptual menjadi label bagi peserta didik.

7.11.8 Maximum Variation dan Criterion Sampling

Peneliti sering memusatkan perhatian pada teknik, padahal maximum variation dan criterion sampling justru mengatur batas inferensi yang dapat dibuat. Fokus ini berada dalam penetapan siapa, kelas mana, atau dokumen apa yang diteliti agar bukti sesuai dengan pertanyaan dan populasi sasaran. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Lakens (2022) serta Lohr (2021) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Rancangan yang dapat diaudit mencantumkan alasan pemilihan unit sampling, prosedur penanganan ukuran sampel, dan kriteria keputusan untuk populasi target. Tanpa penjelasan itu, pembaca sulit menilai apakah profil peserta benar-benar mendukung klaim. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi

ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Kasus sampel wawancara maksimum variasi dapat digunakan untuk memperlihatkan hubungan antara pola individual dan pola kelas. Sebagian miskonsepsi mungkin hanya dimiliki beberapa peserta tetapi sangat kuat, sedangkan pola lain tersebar luas karena bersumber dari analogi atau buku teks yang sama. Analisis perlu menjaga kedua tingkat tersebut.

Bukti minimum yang perlu tersedia:

- Definisikan unit sampling dalam istilah operasional.
- Tentukan bukti utama dari riwayat belajar dan bukti pendukung dari kesediaan partisipasi.
- Periksa power atau presisi sebelum membuat kategori atau model.
- Catat potensi overrekrutmen tanpa alasan beserta strategi mitigasinya.
- Sajikan catatan nonpartisipasi bersama batas interpretasi dan ketidakpastian.

Pemeriksaan kualitas dapat dilakukan melalui audit terhadap kecukupan informasi, analisis kasus yang bertentangan, dan evaluasi transparansi rekrutmen. Apabila ditemukan mengabaikan clustering, peneliti perlu menurunkan kekuatan klaim atau menambah bukti. Transparansi seperti ini meningkatkan kegunaan temuan bagi penelitian lanjutan dan praktik kelas.

Kesimpulannya, maximum variation dan criterion sampling perlu dirancang sebagai proses argumentatif. Data tidak berbicara sendiri; maknanya bergantung pada definisi konstruk, kualitas prosedur, dan konteks. Oleh sebab itu, klaim akhir sebaiknya menggunakan bahasa yang sesuai dengan kekuatan bukti, misalnya “menunjukkan pola”, “mengindikasikan”, atau “mendukung penjelasan”.

7.11.9 Sampling Kasus Negatif dan Ekstrem

Kajian yang kuat memperlakukan sampling kasus negatif dan ekstrem sebagai bagian dari rantai argumentasi, bukan sebagai unsur pelengkap. Fokus ini berada dalam penetapan siapa, kelas mana, atau dokumen apa yang diteliti agar bukti sesuai dengan pertanyaan dan populasi sasaran. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Palinkas et al. (2015) serta Robinson (2014) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Pada tahap perencanaan, kriteria inklusi dan strategi rekrutmen harus dibedakan secara eksplisit. Pada tahap pelaksanaan, konsistensi riwayat belajar perlu diperiksa; pada tahap interpretasi, power atau presisi menjadi dasar pembatasan klaim. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya

berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Dalam kelas yang mempelajari subsample kasus miskonsepsi kuat, data agregat mungkin menunjukkan mayoritas jawaban salah. Akan tetapi, wawancara dapat memperlihatkan beberapa mekanisme berbeda: bahasa sehari-hari, generalisasi aturan, salah membaca simbol, atau model partikel yang tidak tepat. Karena itu, satu persentase tidak boleh menggantikan analisis struktur alasan.

Kriteria pemeriksaan sebelum analisis:

- Definisikan kriteria inklusi dalam istilah operasional.
- Tentukan bukti utama dari hasil prates dan bukti pendukung dari daftar kelas.
- Periksa transparansi rekrutmen sebelum membuat kategori atau model.
- Catat potensi menganggap besar selalu lebih baik beserta strategi mitigasinya.
- Sajikan batas generalisasi bersama batas interpretasi dan ketidakpastian.

Kriteria keberhasilan tidak cukup berupa banyaknya data. Yang lebih penting adalah apakah *power* atau presisi dapat ditelusuri, apakah analisis attrition diperiksa, dan apakah representativitas dijelaskan. Peneliti harus menghindari kriteria berubah setelah data; jika risiko itu tidak dapat dihilangkan, dampaknya perlu dianalisis dan dicantumkan sebagai batas penelitian.

Pada akhirnya, keputusan tentang sampling kasus negatif dan ekstrem menentukan apakah penelitian menghasilkan daftar jawaban salah atau pemahaman yang dapat ditindaklanjuti mengenai cara peserta didik berpikir. Sasaran utamanya ialah bukti yang cukup untuk merancang diagnosis, pembelajaran, dan penelitian lanjutan secara bertanggung jawab.

7.11.10 Ukuran Sampel Kuantitatif

Subbagian ini menempatkan ukuran sampel kuantitatif sebagai simpul penting dalam pemilihan subjek dan sampel. Fokus ini berada dalam penetapan siapa, kelas mana, atau dokumen apa yang diteliti agar bukti sesuai dengan pertanyaan dan populasi sasaran. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Robinson (2014) serta Lohr (2021) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Secara operasional, peneliti perlu menetapkan ukuran sampel, menjelaskan populasi target, dan mendokumentasikan unit sampling. Keputusan tersebut kemudian diterjemahkan ke dalam pilihan hasil prates, daftar kelas, serta profil peserta. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna

untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

$$n_0 = z^2 p(1 - p)/e^2$$

Sebagai ilustrasi, pada topik siswa kelas XI kesetimbangan, peserta didik dapat memberikan jawaban benar pada tingkat simbolik tetapi alasan submikroskopiknya keliru. Peneliti tidak cukup menghitung benar-salah; ia perlu membandingkan hasil prates dengan daftar kelas, memeriksa tingkat keyakinan, dan menguji konsistensi pada konteks lain. Hasil analisis mungkin menunjukkan bahwa kesalahan hanya muncul pada satu representasi, atau justru merupakan hubungan konseptual yang stabil.

Langkah operasional yang disarankan:

- Definisikan ukuran sampel dalam istilah operasional.
- Tentukan bukti utama dari kesediaan partisipasi dan bukti pendukung dari data sekolah.
- Periksa analisis attrition sebelum membuat kategori atau model.
- Catat potensi convenience sample tanpa batas klaim beserta strategi mitigasinya.
- Sajikan diagram alir sampel bersama batas interpretasi dan ketidakpastian.

Mutu inferensi terutama ditentukan oleh transparansi rekrutmen, coverage, dan kecukupan informasi. Ancaman utama ialah overrekrutmen tanpa alasan serta sampel terlalu homogen. Karena itu, laporan perlu menyajikan bukti yang mendukung sekaligus bukti yang membatasi interpretasi. Prinsip ini konsisten dengan penekanan Robinson (2014) dan Lohr (2021) bahwa validitas melekat pada interpretasi dan penggunaan hasil, bukan sekadar pada nama instrumen atau desain.

Dengan demikian, ukuran sampel kuantitatif berfungsi menghubungkan keputusan desain dengan interpretasi. Produk yang diharapkan bukan hanya catatan nonpartisipasi, tetapi juga penjelasan mengenai bagaimana produk tersebut diperoleh dan dalam kondisi apa ia dapat digunakan. Penelitian miskonsepsi yang baik selalu menyisakan ruang bagi penjelasan alternatif dan verifikasi lanjutan.

7.11.11 Ukuran Sampel Kualitatif

Pembahasan mengenai ukuran sampel kualitatif diperlukan agar pemilihan subjek dan sampel tidak berhenti pada prosedur administratif. Fokus ini berada dalam penetapan siapa, kelas mana, atau dokumen apa yang diteliti agar bukti sesuai dengan pertanyaan dan populasi sasaran. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Malterud et al. (2016) serta Robinson (2014) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Penerapan praktis dapat disusun dalam urutan: mendefinisikan strategi rekrutmen, mengoperasionalkan populasi terjangkau, mengumpulkan kesediaan partisipasi, memeriksa analisis attrition, lalu melaporkan batas generalisasi. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Misalnya, pada mahasiswa kimia dasar, perubahan skor sesudah pembelajaran belum otomatis menandakan perubahan konsep. Peserta didik mungkin mempelajari pola jawaban tes. Bukti yang lebih meyakinkan mencakup transfer ke situasi baru, penjelasan kausal, representasi partikel, dan retensi setelah jeda waktu.

Keputusan yang harus didokumentasikan:

- Definisikan strategi rekrutmen dalam istilah operasional.
- Tentukan bukti utama dari daftar kelas dan bukti pendukung dari profil peserta.
- Periksa coverage sebelum membuat kategori atau model.
- Catat potensi sampel terlalu homogen beserta strategi mitigasinya.
- Sajikan justifikasi ukuran bersama batas interpretasi dan ketidakpastian.

Temuan yang baik tidak selalu berarti hasil yang besar atau sesuai harapan. Temuan bernilai ketika analisis attrition terjaga, representativitas dijelaskan, dan ketidakpastian disajikan. Kerangka Malterud et al. (2016) serta Robinson (2014) menolong peneliti membedakan kekuatan bukti dari daya tarik narasi.

Inti metodologis subbagian ini adalah keterlacakan. Pembaca harus dapat mengikuti perjalanan dari pertanyaan, data, pengodean, analisis, sampai batas generalisasi. Ketika jejak itu jelas, temuan dapat digunakan untuk memperbaiki pembelajaran tanpa mengubah diagnosis konseptual menjadi label bagi peserta didik.

7.11.12 Sampling untuk Mixed Methods

Dalam metodologi penelitian miskonsepsi kimia, sampling untuk *mixed methods* menentukan mutu hubungan antara data dan kesimpulan. Fokus ini berada dalam penetapan siapa, kelas mana, atau dokumen apa yang diteliti agar bukti sesuai dengan pertanyaan dan populasi sasaran. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Cochran (1977) serta Lohr (2021) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Prosedur minimal mencakup definisi populasi target, alasan memilih daftar kelas, aturan mengubah respons menjadi kategori, serta mekanisme pemeriksaan coverage. Semua penyimpangan dari rencana awal perlu dicatat. Setiap keputusan dicatat dalam protokol,

termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Pada penelitian mengenai guru kimia SMA, peneliti dapat menyusun beberapa tugas ekuivalen yang berubah konteks, representasi, atau tuntutan penalarannya. Jika pola respons tetap berulang dan disertai keyakinan tinggi, terdapat dasar yang lebih kuat untuk menyebut indikasi miskonsepsi. Jika pola berubah-ubah, interpretasi yang lebih hati-hati adalah pengetahuan parsial atau ketergantungan konteks.

Pertanyaan audit metodologis:

- Definisikan populasi target dalam istilah operasional.
- Tentukan bukti utama dari data sekolah dan bukti pendukung dari riwayat belajar.
- Periksa representativitas sebelum membuat kategori atau model.
- Catat potensi mengabaikan clustering beserta strategi mitigasinya.
- Sajikan tabel karakteristik bersama batas interpretasi dan ketidakpastian.

Pemeriksaan mutu hendaknya dilakukan selama penelitian, bukan hanya pada akhir. Catatan tentang coverage, keputusan terkait kecukupan informasi, dan kejadian yang berpotensi menimbulkan convenience sample tanpa batas klaim perlu dipelihara dalam *audit trail*. Hal ini memungkinkan koreksi dini dan pelaporan yang jujur.

Kesimpulannya, sampling untuk mixed methods perlu dirancang sebagai proses argumentatif. Data tidak berbicara sendiri; maknanya bergantung pada definisi konstruk, kualitas prosedur, dan konteks. Oleh sebab itu, klaim akhir sebaiknya menggunakan bahasa yang sesuai dengan kekuatan bukti, misalnya “menunjukkan pola”, “mengindikasikan”, atau “mendukung penjelasan”.

7.11.13 Rekrutmen dan Retensi Partisipan

Keputusan tentang rekrutmen dan retensi partisipan harus dibuat sebelum data utama dikumpulkan, bukan ditambahkan setelah hasil diketahui. Fokus ini berada dalam penetapan siapa, kelas mana, atau dokumen apa yang diteliti agar bukti sesuai dengan pertanyaan dan populasi sasaran. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Lohr (2021) serta Robinson (2014) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Tahap kerja dimulai dengan penegasan populasi terjangkau. Setelah itu, peneliti menyusun aturan untuk kriteria inklusi dan menentukan bagaimana strategi rekrutmen akan diperiksa. Bukti sebaiknya berasal dari lebih dari satu bentuk, misalnya data sekolah dan riwayat belajar.

Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Pada topik kelas dengan capaian beragam, peneliti juga perlu memperhatikan bahasa. Istilah yang sama dapat diberi makna berbeda oleh peserta didik. Klarifikasi melalui pertanyaan tindak lanjut, gambar, atau contoh tandingan membantu membedakan masalah terminologi dari kesalahan konseptual.

Bukti minimum yang perlu tersedia:

- Definisikan populasi terjangkau dalam istilah operasional.
- Tentukan bukti utama dari profil peserta dan bukti pendukung dari hasil pretes.
- Periksa kecukupan informasi sebelum membuat kategori atau model.
- Catat potensi kriteria berubah setelah data beserta strategi mitigasinya.
- Sajikan analisis kesetaraan bersama batas interpretasi dan ketidakpastian.

Kualitas penelitian meningkat ketika peneliti menunjukkan rantai bukti dari data sekolah menuju justifikasi ukuran, menyatakan asumsi, dan memeriksa alternatif penjelasan. Sebaliknya, sampel terlalu homogen dapat menghasilkan klaim yang tampak tegas tetapi tidak stabil. Rekomendasi praktis harus proporsional dengan bukti.

Pada akhirnya, keputusan tentang rekrutmen dan retensi partisipan menentukan apakah penelitian menghasilkan daftar jawaban salah atau pemahaman yang dapat ditindaklanjuti mengenai cara peserta didik berpikir. Sasaran utamanya ialah bukti yang cukup untuk merancang diagnosis, pembelajaran, dan penelitian lanjutan secara bertanggung jawab.

7.11.14 Analisis Bias Seleksi

Analisis bias seleksi perlu dipahami sebagai keputusan metodologis yang mempunyai konsekuensi pada desain, analisis, dan pelaporan. Fokus ini berada dalam penetapan siapa, kelas mana, atau dokumen apa yang diteliti agar bukti sesuai dengan pertanyaan dan populasi sasaran. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Lakens (2022) serta Lohr (2021) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Peneliti sebaiknya menyiapkan matriks yang menghubungkan pertanyaan penelitian dengan unit sampling, profil peserta, teknik analisis, dan tabel karakteristik. Matriks ini mencegah pengumpulan data yang banyak tetapi tidak relevan. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk

replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Ilustrasi sekolah urban dan rural juga menunjukkan bahwa kualitas instrumen memengaruhi temuan. Distraktor yang tidak berasal dari pemikiran peserta didik dapat menghasilkan pilihan acak, sedangkan pertanyaan yang terlalu menuntut matematika dapat mengukur keterampilan hitung alih-alih konsepsi. Peneliti perlu menilai kesesuaian bukti dengan konstruk yang dimaksud.

Kriteria pemeriksaan sebelum analisis:

- Definisikan unit sampling dalam istilah operasional.
- Tentukan bukti utama dari riwayat belajar dan bukti pendukung dari kesediaan partisipasi.
- Periksa power atau presisi sebelum membuat kategori atau model.
- Catat potensi overrekrutmen tanpa alasan beserta strategi mitigasinya.
- Sajikan catatan nonpartisipasi bersama batas interpretasi dan ketidakpastian.

Peneliti sebaiknya memisahkan keputusan apriori dari keputusan eksploratori. Pemisahan tersebut membantu mengendalikan mengabaikan clustering dan mencegah menganggap besar selalu lebih baik. Selanjutnya, kecukupan informasi dan transparansi rekrutmen dilaporkan bersama ketidakpastian, sehingga pembaca dapat menilai kekuatan temuan tanpa bergantung pada kesimpulan penulis semata.

Dengan demikian, analisis bias seleksi berfungsi menghubungkan keputusan desain dengan interpretasi. Produk yang diharapkan bukan hanya tabel karakteristik, tetapi juga penjelasan mengenai bagaimana produk tersebut diperoleh dan dalam kondisi apa ia dapat digunakan. Penelitian miskonsepsi yang baik selalu menyisakan ruang bagi penjelasan alternatif dan verifikasi lanjutan.

7.11.15 Sintesis Strategi Pemilihan Sampel

Ketelitian penelitian pada bagian sintesis strategi pemilihan sampel sangat menentukan apakah pola jawaban dapat ditafsirkan sebagai bukti konseptual. Fokus ini berada dalam penetapan siapa, kelas mana, atau dokumen apa yang diteliti agar bukti sesuai dengan pertanyaan dan populasi sasaran. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Palinkas et al. (2015) serta Robinson (2014) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Dokumentasi metodologis hendaknya memuat siapa yang melakukan riwayat belajar, kapan data diperoleh, bagaimana strategi rekrutmen dikendalikan, dan atas dasar apa analisis kesetaraan disusun. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan

selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Contoh pada kajian peserta berkebutuhan khusus memperlihatkan pentingnya pemisahan antara ketidaktahuan, kesalahan prosedural, dan miskonsepsi. Seorang peserta dapat gagal karena lupa rumus, sedangkan peserta lain memilih jawaban yang sama karena menggunakan model mental yang tidak ilmiah. Perbedaan ini baru terlihat ketika riwayat belajar dibaca bersama kesediaan partisipasi dan bukti tentang alasan.

Langkah operasional yang disarankan:

- Definisikan kriteria inklusi dalam istilah operasional.
- Tentukan bukti utama dari hasil prates dan bukti pendukung dari daftar kelas.
- Periksa transparansi rekrutmen sebelum membuat kategori atau model.
- Catat potensi menganggap besar selalu lebih baik beserta strategi mitigasinya.
- Sajikan batas generalisasi bersama batas interpretasi dan ketidakpastian.

Standar metodologis menuntut koherensi antara konstruk, prosedur, analisis, dan klaim. Dalam konteks ini, *power* atau presisi tidak dapat digantikan oleh indeks tunggal. Demikian pula, analisis attrition harus dinilai dengan bukti yang sesuai. Palinkas et al. (2015) dan Robinson (2014) memberikan dasar untuk pelaporan yang lengkap dan dapat ditelaah.

Inti metodologis subbagian ini adalah keterlacakan. Pembaca harus dapat mengikuti perjalanan dari pertanyaan, data, pengodean, analisis, sampai analisis kesetaraan. Ketika jejak itu jelas, temuan dapat digunakan untuk memperbaiki pembelajaran tanpa mengubah diagnosis konseptual menjadi label bagi peserta didik.

7.12 Teknik Pengumpulan Data

Subbab ini membahas pemilihan dan pelaksanaan teknik yang mampu menangkap jawaban, alasan, model mental, keyakinan, interaksi, dan perubahan konsepsi secara sah. Fokus utamanya adalah menjaga keselarasan antara pertanyaan penelitian, bentuk bukti, teknik analisis, dan kekuatan klaim. Dalam konteks miskonsepsi kimia, keputusan metodologis harus mampu membedakan pola konseptual dari respons yang timbul karena keterbatasan bahasa, kesalahan hitung, atau kondisi administrasi. Pembahasan berikut dipecah menjadi lima belas bagian agar setiap keputusan dapat ditelaah secara operasional.

7.12.1 Pemilihan Teknik Berdasarkan Konstruk

Peneliti sering memusatkan perhatian pada teknik, padahal pemilihan teknik berdasarkan konstruk justru mengatur batas inferensi yang dapat dibuat. Fokus ini berada dalam pemilihan dan pelaksanaan teknik yang mampu menangkap jawaban, alasan, model mental, keyakinan, interaksi, dan perubahan konsepsi secara sah. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah

membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. O'Brien et al. (2014) serta Kaltakci Gurel et al. (2015) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Rancangan yang dapat diaudit mencantumkan alasan pemilihan rekaman data, prosedur penanganan konstruk yang dicari, dan kriteria keputusan untuk urutan teknik. Tanpa penjelasan itu, pembaca sulit menilai apakah gambar benar-benar mendukung klaim. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Kasus peta konsep termokimia dapat digunakan untuk memperlihatkan hubungan antara pola individual dan pola kelas. Sebagian miskonsepsi mungkin hanya dimiliki beberapa peserta tetapi sangat kuat, sedangkan pola lain tersebar luas karena bersumber dari analogi atau buku teks yang sama. Analisis perlu menjaga kedua tingkat tersebut.

Keputusan yang harus didokumentasikan:

- Definisikan rekaman data dalam istilah operasional.
- Tentukan bukti utama dari observasi dan video dan bukti pendukung dari pertanyaan terbuka.
- Periksa dokumentasi konteks sebelum membuat kategori atau model.
- Catat potensi leading question beserta strategi mitigasinya.
- Sajikan protokol pengumpulan bersama batas interpretasi dan ketidakpastian.

Pemeriksaan kualitas dapat dilakukan melalui audit terhadap keamanan penyimpanan, analisis kasus yang bertentangan, dan evaluasi kesesuaian teknik-konstruk. Apabila ditemukan mengabaikan bahasa peserta, peneliti perlu menurunkan kekuatan klaim atau menambah bukti. Transparansi seperti ini meningkatkan kegunaan temuan bagi penelitian lanjutan dan praktik kelas.

Kesimpulannya, pemilihan teknik berdasarkan konstruk perlu dirancang sebagai proses argumentatif. Data tidak berbicara sendiri; maknanya bergantung pada definisi konstruk, kualitas prosedur, dan konteks. Oleh sebab itu, klaim akhir sebaiknya menggunakan bahasa yang sesuai dengan kekuatan bukti, misalnya “menunjukkan pola”, “mengindikasikan”, atau “mendukung penjelasan”.

7.12.2 Tes Pilihan Ganda Diagnostik

Kajian yang kuat memperlakukan tes pilihan ganda diagnostik sebagai bagian dari rantai argumentasi, bukan sebagai unsur pelengkap. Fokus ini berada dalam pemilihan dan pelaksanaan teknik yang mampu menangkap jawaban, alasan, model mental, keyakinan, interaksi, dan perubahan konsepsi secara sah. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut

dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. AERA, APA, dan NCME (2014) serta O'Brien et al. (2014) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Pada tahap perencanaan, waktu pengukuran dan tingkat keterbukaan harus dibedakan secara eksplisit. Pada tahap pelaksanaan, konsistensi observasi dan video perlu diperiksa; pada tahap interpretasi, dokumentasi konteks menjadi dasar pembatasan klaim. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Dalam kelas yang mempelajari log simulasi atom, data agregat mungkin menunjukkan mayoritas jawaban salah. Akan tetapi, wawancara dapat memperlihatkan beberapa mekanisme berbeda: bahasa sehari-hari, generalisasi aturan, salah membaca simbol, atau model partikel yang tidak tepat. Karena itu, satu persentase tidak boleh menggantikan analisis struktur alasan.

Pertanyaan audit metodologis:

- Definisikan waktu pengukuran dalam istilah operasional.
- Tentukan bukti utama dari tes pilihan bertingkat dan bukti pendukung dari wawancara.
- Periksa kesesuaian teknik-konstruk sebelum membuat kategori atau model.
- Catat potensi administrator memberi petunjuk beserta strategi mitigasinya.
- Sajikan jadwal bersama batas interpretasi dan ketidakpastian.

Kriteria keberhasilan tidak cukup berupa banyaknya data. Yang lebih penting adalah apakah dokumentasi konteks dapat ditelusuri, apakah prosedur baku diperiksa, dan apakah pilot testing dijelaskan. Peneliti harus menghindari instrumen terlalu membebani; jika risiko itu tidak dapat dihilangkan, dampaknya perlu dianalisis dan dicantumkan sebagai batas penelitian.

Pada akhirnya, keputusan tentang tes pilihan ganda diagnostik menentukan apakah penelitian menghasilkan daftar jawaban salah atau pemahaman yang dapat ditindaklanjuti mengenai cara peserta didik berpikir. Sasaran utamanya ialah bukti yang cukup untuk merancang diagnosis, pembelajaran, dan penelitian lanjutan secara bertanggung jawab.

7.12.3 Tes Dua, Tiga, dan Empat Tingkat

Subbagian ini menempatkan tes dua, tiga, dan empat tingkat sebagai simpul penting dalam teknik pengumpulan data. Fokus ini berada dalam pemilihan dan pelaksanaan teknik yang mampu menangkap jawaban, alasan, model mental, keyakinan, interaksi, dan perubahan konsepsi secara sah. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung,

ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Treagust (1988) serta Kaltakci Gurel et al. (2015) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Secara operasional, peneliti perlu menetapkan konstruk yang dicari, menjelaskan urutan teknik, dan mendokumentasikan rekaman data. Keputusan tersebut kemudian diterjemahkan ke dalam pilihan tes pilihan bertingkat, wawancara, serta gambar. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Sebagai ilustrasi, pada topik wawancara tentang ikatan, peserta didik dapat memberikan jawaban benar pada tingkat simbolik tetapi alasan submikroskopiknya keliru. Peneliti tidak cukup menghitung benar-salah; ia perlu membandingkan tes pilihan bertingkat dengan wawancara, memeriksa tingkat keyakinan, dan menguji konsistensi pada konteks lain. Hasil analisis mungkin menunjukkan bahwa kesalahan hanya muncul pada satu representasi, atau justru merupakan hubungan konseptual yang stabil.

Bukti minimum yang perlu tersedia:

- Definisikan konstruk yang dicari dalam istilah operasional.
- Tentukan bukti utama dari pertanyaan terbuka dan bukti pendukung dari think-aloud.
- Periksa prosedur baku sebelum membuat kategori atau model.
- Catat potensi rekaman tidak lengkap beserta strategi mitigasinya.
- Sajikan skrip instruksi bersama batas interpretasi dan ketidakpastian.

Mutu inferensi terutama ditentukan oleh kesesuaian teknik-konstruk, pelatihan pengumpul data, dan keamanan penyimpanan. Ancaman utama ialah leading question serta transkripsi selektif. Karena itu, laporan perlu menyajikan bukti yang mendukung sekaligus bukti yang membatasi interpretasi. Prinsip ini konsisten dengan penekanan Treagust (1988) dan Kaltakci Gurel et al. (2015) bahwa validitas melekat pada interpretasi dan penggunaan hasil, bukan sekadar pada nama instrumen atau desain.

Dengan demikian, tes dua, tiga, dan empat tingkat berfungsi menghubungkan keputusan desain dengan interpretasi. Produk yang diharapkan bukan hanya protokol pengumpulan, tetapi juga penjelasan mengenai bagaimana produk tersebut diperoleh dan dalam kondisi apa ia dapat digunakan. Penelitian miskonsepsi yang baik selalu menyisakan ruang bagi penjelasan alternatif dan verifikasi lanjutan.

7.12.4 Pertanyaan Terbuka

Pembahasan mengenai pertanyaan terbuka diperlukan agar teknik pengumpulan data tidak berhenti pada prosedur administratif. Fokus ini berada dalam pemilihan dan pelaksanaan

teknik yang mampu menangkap jawaban, alasan, model mental, keyakinan, interaksi, dan perubahan konsepsi secara sah. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Kaltakci Gurel et al. (2015) serta O'Brien et al. (2014) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Penerapan praktis dapat disusun dalam urutan: mendefinisikan tingkat keterbukaan, mengoperasionalkan standarisasi administrator, mengumpulkan pertanyaan terbuka, memeriksa prosedur baku, lalu melaporkan jadwal. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Misalnya, pada gambar partikel larutan, perubahan skor sesudah pembelajaran belum otomatis menandakan perubahan konsep. Peserta didik mungkin mempelajari pola jawaban tes. Bukti yang lebih meyakinkan mencakup transfer ke situasi baru, penjelasan kausal, representasi partikel, dan retensi setelah jeda waktu.

Kriteria pemeriksaan sebelum analisis:

- Definisikan tingkat keterbukaan dalam istilah operasional.
- Tentukan bukti utama dari wawancara dan bukti pendukung dari gambar.
- Periksa pelatihan pengumpul data sebelum membuat kategori atau model.
- Catat potensi transkripsi selektif beserta strategi mitigasinya.
- Sajikan format observasi bersama batas interpretasi dan ketidakpastian.

Temuan yang baik tidak selalu berarti hasil yang besar atau sesuai harapan. Temuan bernilai ketika prosedur baku terjaga, pilot testing dijelaskan, dan ketidakpastian disajikan. Kerangka Kaltakci Gurel et al. (2015) serta O'Brien et al. (2014) menolong peneliti membedakan kekuatan bukti dari daya tarik narasi.

Inti metodologis subbagian ini adalah keterlacakan. Pembaca harus dapat mengikuti perjalanan dari pertanyaan, data, pengodean, analisis, sampai jadwal. Ketika jejak itu jelas, temuan dapat digunakan untuk memperbaiki pembelajaran tanpa mengubah diagnosis konseptual menjadi label bagi peserta didik.

7.12.5 Wawancara Klinis dan Semi-Terstruktur

Dalam metodologi penelitian miskonsepsi kimia, wawancara klinis dan semi-terstruktur menentukan mutu hubungan antara data dan kesimpulan. Fokus ini berada dalam pemilihan dan pelaksanaan teknik yang mampu menangkap jawaban, alasan, model mental, keyakinan, interaksi, dan perubahan konsepsi secara sah. Peneliti perlu menyatakan sejak awal apa

yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Nyachwaya et al. (2011) serta Kaltakci Gurel et al. (2015) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Prosedur minimal mencakup definisi urutan teknik, alasan memilih wawancara, aturan mengubah respons menjadi kategori, serta mekanisme pemeriksaan pelatihan pengumpul data. Semua penyimpangan dari rencana awal perlu dicatat. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Pada penelitian mengenai think-aloud stoikiometri, peneliti dapat menyusun beberapa tugas ekuivalen yang berubah konteks, representasi, atau tuntutan penalarannya. Jika pola respons tetap berulang dan disertai keyakinan tinggi, terdapat dasar yang lebih kuat untuk menyebut indikasi miskonsepsi. Jika pola berubah-ubah, interpretasi yang lebih hati-hati adalah pengetahuan parsial atau ketergantungan konteks.

Langkah operasional yang disarankan:

- Definisikan urutan teknik dalam istilah operasional.
- Tentukan bukti utama dari think-aloud dan bukti pendukung dari observasi dan video.
- Periksa pilot testing sebelum membuat kategori atau model.
- Catat potensi mengabaikan bahasa peserta beserta strategi mitigasinya.
- Sajikan sistem penamaan data bersama batas interpretasi dan ketidakpastian.

Pemeriksaan mutu hendaknya dilakukan selama penelitian, bukan hanya pada akhir. Catatan tentang pelatihan pengumpul data, keputusan terkait keamanan penyimpanan, dan kejadian yang berpotensi menimbulkan rekaman tidak lengkap perlu dipelihara dalam *audit trail*. Hal ini memungkinkan koreksi dini dan pelaporan yang jujur.

Kesimpulannya, wawancara klinis dan semi-terstruktur perlu dirancang sebagai proses argumentatif. Data tidak berbicara sendiri; maknanya bergantung pada definisi konstruk, kualitas prosedur, dan konteks. Oleh sebab itu, klaim akhir sebaiknya menggunakan bahasa yang sesuai dengan kekuatan bukti, misalnya “menunjukkan pola”, “mengindikasikan”, atau “mendukung penjelasan”.

7.12.6 Think-Aloud Protocol

Keputusan tentang think-aloud protocol harus dibuat sebelum data utama dikumpulkan, bukan ditambahkan setelah hasil diketahui. Fokus ini berada dalam pemilihan dan pelaksanaan teknik yang mampu menangkap jawaban, alasan, model mental, keyakinan, interaksi, dan perubahan konsepsi secara sah. Peneliti perlu menyatakan sejak awal apa

yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Tong et al. (2007) serta O'Brien et al. (2014) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Tahap kerja dimulai dengan penegasan standarisasi administrator. Setelah itu, peneliti menyusun aturan untuk waktu pengukuran dan menentukan bagaimana tingkat keterbukaan akan diperiksa. Bukti sebaiknya berasal dari lebih dari satu bentuk, misalnya think-aloud dan observasi dan video. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Pada topik observasi diskusi kesetimbangan, peneliti juga perlu memperhatikan bahasa. Istilah yang sama dapat diberi makna berbeda oleh peserta didik. Klarifikasi melalui pertanyaan tindak lanjut, gambar, atau contoh tandingan membantu membedakan masalah terminologi dari kesalahan konseptual.

Keputusan yang harus didokumentasikan:

- Definisikan standarisasi administrator dalam istilah operasional.
- Tentukan bukti utama dari gambar dan bukti pendukung dari tes pilihan bertingkat.
- Periksa keamanan penyimpanan sebelum membuat kategori atau model.
- Catat potensi instrumen terlalu membebani beserta strategi mitigasinya.
- Sajikan log penyimpangan bersama batas interpretasi dan ketidakpastian.

Kualitas penelitian meningkat ketika peneliti menunjukkan rantai bukti dari think-aloud menuju format observasi, menyatakan asumsi, dan memeriksa alternatif penjelasan. Sebaliknya, transkripsi selektif dapat menghasilkan klaim yang tampak tegas tetapi tidak stabil. Rekomendasi praktis harus proporsional dengan bukti.

Pada akhirnya, keputusan tentang think-aloud protocol menentukan apakah penelitian menghasilkan daftar jawaban salah atau pemahaman yang dapat ditindaklanjuti mengenai cara peserta didik berpikir. Sasaran utamanya ialah bukti yang cukup untuk merancang diagnosis, pembelajaran, dan penelitian lanjutan secara bertanggung jawab.

7.12.7 Tugas Menggambar dan Model Partikel

Tugas menggambar dan model partikel perlu dipahami sebagai keputusan metodologis yang mempunyai konsekuensi pada desain, analisis, dan pelaporan. Fokus ini berada dalam pemilihan dan pelaksanaan teknik yang mampu menangkap jawaban, alasan, model mental, keyakinan, interaksi, dan perubahan konsepsi secara sah. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut

dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. O'Brien et al. (2014) serta Kaltakci Gurel et al. (2015) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Peneliti sebaiknya menyiapkan matriks yang menghubungkan pertanyaan penelitian dengan rekaman data, gambar, teknik analisis, dan sistem penamaan data. Matriks ini mencegah pengumpulan data yang banyak tetapi tidak relevan. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Ilustrasi tes CRI redoks juga menunjukkan bahwa kualitas instrumen memengaruhi temuan. Distraktor yang tidak berasal dari pemikiran peserta didik dapat menghasilkan pilihan acak, sedangkan pertanyaan yang terlalu menuntut matematika dapat mengukur keterampilan hitung alih-alih konsepsi. Peneliti perlu menilai kesesuaian bukti dengan konstruk yang dimaksud.

Pertanyaan audit metodologis:

- Definisikan rekaman data dalam istilah operasional.
- Tentukan bukti utama dari observasi dan video dan bukti pendukung dari pertanyaan terbuka.
- Periksa dokumentasi konteks sebelum membuat kategori atau model.
- Catat potensi leading question beserta strategi mitigasinya.
- Sajikan protokol pengumpulan bersama batas interpretasi dan ketidakpastian.

Peneliti sebaiknya memisahkan keputusan apriori dari keputusan eksploratori. Pemisahan tersebut membantu mengendalikan mengabaikan bahasa peserta dan mencegah administrator memberi petunjuk. Selanjutnya, keamanan penyimpanan dan kesesuaian teknik-konstruk dilaporkan bersama ketidakpastian, sehingga pembaca dapat menilai kekuatan temuan tanpa bergantung pada kesimpulan penulis semata.

Dengan demikian, tugas menggambar dan model partikel berfungsi menghubungkan keputusan desain dengan interpretasi. Produk yang diharapkan bukan hanya sistem penamaan data, tetapi juga penjelasan mengenai bagaimana produk tersebut diperoleh dan dalam kondisi apa ia dapat digunakan. Penelitian miskonsepsi yang baik selalu menyisakan ruang bagi penjelasan alternatif dan verifikasi lanjutan.

7.12.8 Peta Konsep

Ketelitian penelitian pada bagian peta konsep sangat menentukan apakah pola jawaban dapat ditafsirkan sebagai bukti konseptual. Fokus ini berada dalam pemilihan dan pelaksanaan teknik yang mampu menangkap jawaban, alasan, model mental, keyakinan, interaksi, dan

perubahan konsepsi secara sah. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. AERA, APA, dan NCME (2014) serta O'Brien et al. (2014) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Dokumentasi metodologis hendaknya memuat siapa yang melakukan observasi dan video, kapan data diperoleh, bagaimana tingkat keterbukaan dikendalikan, dan atas dasar apa log penyimpangan disusun. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Contoh pada kajian video praktikum asam-basa memperlihatkan pentingnya pemisahan antara ketidaktahuan, kesalahan prosedural, dan miskonsepsi. Seorang peserta dapat gagal karena lupa rumus, sedangkan peserta lain memilih jawaban yang sama karena menggunakan model mental yang tidak ilmiah. Perbedaan ini baru terlihat ketika observasi dan video dibaca bersama pertanyaan terbuka dan bukti tentang alasan.

Bukti minimum yang perlu tersedia:

- Definisikan waktu pengukuran dalam istilah operasional.
- Tentukan bukti utama dari tes pilihan bertingkat dan bukti pendukung dari wawancara.
- Periksa kesesuaian teknik-konstruk sebelum membuat kategori atau model.
- Catat potensi administrator memberi petunjuk beserta strategi mitigasinya.
- Sajikan jadwal bersama batas interpretasi dan ketidakpastian.

Standar metodologis menuntut koherensi antara konstruk, prosedur, analisis, dan klaim. Dalam konteks ini, dokumentasi konteks tidak dapat digantikan oleh indeks tunggal. Demikian pula, prosedur baku harus dinilai dengan bukti yang sesuai. AERA, APA, dan NCME (2014) dan O'Brien et al. (2014) memberikan dasar untuk pelaporan yang lengkap dan dapat ditelaah.

Inti metodologis subbagian ini adalah keterlacakan. Pembaca harus dapat mengikuti perjalanan dari pertanyaan, data, pengodean, analisis, sampai log penyimpangan. Ketika jejak itu jelas, temuan dapat digunakan untuk memperbaiki pembelajaran tanpa mengubah diagnosis konseptual menjadi label bagi peserta didik.

7.12.9 Observasi Kelas

Peneliti sering memusatkan perhatian pada teknik, padahal observasi kelas justru mengatur batas inferensi yang dapat dibuat. Fokus ini berada dalam pemilihan dan pelaksanaan teknik yang mampu menangkap jawaban, alasan, model mental, keyakinan, interaksi, dan perubahan konsepsi secara sah. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa

yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Treagust (1988) serta Kaltakci Gurel et al. (2015) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Rancangan yang dapat diaudit mencantumkan alasan pemilihan konstruk yang dicari, prosedur penanganan urutan teknik, dan kriteria keputusan untuk rekaman data. Tanpa penjelasan itu, pembaca sulit menilai apakah tes pilihan bertingkat benar-benar mendukung klaim. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Kasus peta konsep termokimia dapat digunakan untuk memperlihatkan hubungan antara pola individual dan pola kelas. Sebagian miskonsepsi mungkin hanya dimiliki beberapa peserta tetapi sangat kuat, sedangkan pola lain tersebar luas karena bersumber dari analogi atau buku teks yang sama. Analisis perlu menjaga kedua tingkat tersebut.

Kriteria pemeriksaan sebelum analisis:

- Definisikan konstruk yang dicari dalam istilah operasional.
- Tentukan bukti utama dari pertanyaan terbuka dan bukti pendukung dari think-aloud.
- Periksa prosedur baku sebelum membuat kategori atau model.
- Catat potensi rekaman tidak lengkap beserta strategi mitigasinya.
- Sajikan skrip instruksi bersama batas interpretasi dan ketidakpastian.

Pemeriksaan kualitas dapat dilakukan melalui audit terhadap kesesuaian teknik-konstruk, analisis kasus yang bertentangan, dan evaluasi pelatihan pengumpul data. Apabila ditemukan leading question, peneliti perlu menurunkan kekuatan klaim atau menambah bukti. Transparansi seperti ini meningkatkan kegunaan temuan bagi penelitian lanjutan dan praktik kelas.

Kesimpulannya, observasi kelas perlu dirancang sebagai proses argumentatif. Data tidak berbicara sendiri; maknanya bergantung pada definisi konstruk, kualitas prosedur, dan konteks. Oleh sebab itu, klaim akhir sebaiknya menggunakan bahasa yang sesuai dengan kekuatan bukti, misalnya “menunjukkan pola”, “mengindikasikan”, atau “mendukung penjelasan”.

7.12.10 Diskusi Kelompok Terarah

Kajian yang kuat memperlakukan diskusi kelompok terarah sebagai bagian dari rantai argumentasi, bukan sebagai unsur pelengkap. Fokus ini berada dalam pemilihan dan pelaksanaan teknik yang mampu menangkap jawaban, alasan, model mental, keyakinan,

interaksi, dan perubahan konsepsi secara sahih. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Kaltakci Gurel et al. (2015) serta O'Brien et al. (2014) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Pada tahap perencanaan, tingkat keterbukaan dan standarisasi administrator harus dibedakan secara eksplisit. Pada tahap pelaksanaan, konsistensi pertanyaan terbuka perlu diperiksa; pada tahap interpretasi, prosedur baku menjadi dasar pembatasan klaim. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Dalam kelas yang mempelajari log simulasi atom, data agregat mungkin menunjukkan mayoritas jawaban salah. Akan tetapi, wawancara dapat memperlihatkan beberapa mekanisme berbeda: bahasa sehari-hari, generalisasi aturan, salah membaca simbol, atau model partikel yang tidak tepat. Karena itu, satu persentase tidak boleh menggantikan analisis struktur alasan.

Langkah operasional yang disarankan:

- Definisikan tingkat keterbukaan dalam istilah operasional.
- Tentukan bukti utama dari wawancara dan bukti pendukung dari gambar.
- Periksa pelatihan pengumpul data sebelum membuat kategori atau model.
- Catat potensi transkripsi selektif beserta strategi mitigasinya.
- Sajikan format observasi bersama batas interpretasi dan ketidakpastian.

Kriteria keberhasilan tidak cukup berupa banyaknya data. Yang lebih penting adalah apakah prosedur baku dapat ditelusuri, apakah pilot testing diperiksa, dan apakah dokumentasi konteks dijelaskan. Peneliti harus menghindari administrator memberi petunjuk; jika risiko itu tidak dapat dihilangkan, dampaknya perlu dianalisis dan dicantumkan sebagai batas penelitian.

Pada akhirnya, keputusan tentang diskusi kelompok terarah menentukan apakah penelitian menghasilkan daftar jawaban salah atau pemahaman yang dapat ditindaklanjuti mengenai cara peserta didik berpikir. Sasaran utamanya ialah bukti yang cukup untuk merancang diagnosis, pembelajaran, dan penelitian lanjutan secara bertanggung jawab.

7.12.11 Dokumentasi Artefak Belajar

Subbagian ini menempatkan dokumentasi artefak belajar sebagai simpul penting dalam teknik pengumpulan data. Fokus ini berada dalam pemilihan dan pelaksanaan teknik yang mampu menangkap jawaban, alasan, model mental, keyakinan, interaksi, dan perubahan konsepsi

secara sahih. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Nyachwaya et al. (2011) serta Kaltakci Gurel et al. (2015) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Secara operasional, peneliti perlu menetapkan urutan teknik, menjelaskan rekaman data, dan mendokumentasikan konstruk yang dicari. Keputusan tersebut kemudian diterjemahkan ke dalam pilihan wawancara, gambar, serta tes pilihan bertingkat. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Sebagai ilustrasi, pada topik wawancara tentang ikatan, peserta didik dapat memberikan jawaban benar pada tingkat simbolik tetapi alasan submikroskopiknya keliru. Peneliti tidak cukup menghitung benar-salah; ia perlu membandingkan wawancara dengan gambar, memeriksa tingkat keyakinan, dan menguji konsistensi pada konteks lain. Hasil analisis mungkin menunjukkan bahwa kesalahan hanya muncul pada satu representasi, atau justru merupakan hubungan konseptual yang stabil.

Keputusan yang harus didokumentasikan:

- Definisikan urutan teknik dalam istilah operasional.
- Tentukan bukti utama dari think-aloud dan bukti pendukung dari observasi dan video.
- Periksa pilot testing sebelum membuat kategori atau model.
- Catat potensi mengabaikan bahasa peserta beserta strategi mitigasinya.
- Sajikan sistem penamaan data bersama batas interpretasi dan ketidakpastian.

Mutu inferensi terutama ditentukan oleh pelatihan pengumpul data, keamanan penyimpanan, dan kesesuaian teknik-konstruk. Ancaman utama ialah rekaman tidak lengkap serta instrumen terlalu membebani. Karena itu, laporan perlu menyajikan bukti yang mendukung sekaligus bukti yang membatasi interpretasi. Prinsip ini konsisten dengan penekanan Nyachwaya et al. (2011) dan Kaltakci Gurel et al. (2015) bahwa validitas melekat pada interpretasi dan penggunaan hasil, bukan sekadar pada nama instrumen atau desain.

Dengan demikian, dokumentasi artefak belajar berfungsi menghubungkan keputusan desain dengan interpretasi. Produk yang diharapkan bukan hanya skrip instruksi, tetapi juga penjelasan mengenai bagaimana produk tersebut diperoleh dan dalam kondisi apa ia dapat digunakan. Penelitian miskonsepsi yang baik selalu menyisakan ruang bagi penjelasan alternatif dan verifikasi lanjutan.

7.12.12 Data Digital dan Log Aktivitas

Pembahasan mengenai data digital dan log aktivitas diperlukan agar teknik pengumpulan data tidak berhenti pada prosedur administratif. Fokus ini berada dalam pemilihan dan pelaksanaan teknik yang mampu menangkap jawaban, alasan, model mental, keyakinan, interaksi, dan perubahan konsepsi secara sah. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Tong et al. (2007) serta O'Brien et al. (2014) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Penerapan praktis dapat disusun dalam urutan: mendefinisikan standarisasi administrator, mengoperasionalkan waktu pengukuran, mengumpulkan think-aloud, memeriksa pilot testing, lalu melaporkan format observasi. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Misalnya, pada gambar partikel larutan, perubahan skor sesudah pembelajaran belum otomatis menandakan perubahan konsep. Peserta didik mungkin mempelajari pola jawaban tes. Bukti yang lebih meyakinkan mencakup transfer ke situasi baru, penjelasan kausal, representasi partikel, dan retensi setelah jeda waktu.

Pertanyaan audit metodologis:

- Definisikan standarisasi administrator dalam istilah operasional.
- Tentukan bukti utama dari gambar dan bukti pendukung dari tes pilihan bertingkat.
- Periksa keamanan penyimpanan sebelum membuat kategori atau model.
- Catat potensi instrumen terlalu membebani beserta strategi mitigasinya.
- Sajikan log penyimpangan bersama batas interpretasi dan ketidakpastian.

Temuan yang baik tidak selalu berarti hasil yang besar atau sesuai harapan. Temuan bernilai ketika pilot testing terjaga, dokumentasi konteks dijelaskan, dan ketidakpastian disajikan. Kerangka Tong et al. (2007) serta O'Brien et al. (2014) menolong peneliti membedakan kekuatan bukti dari daya tarik narasi.

Inti metodologis subbagian ini adalah keterlacakan. Pembaca harus dapat mengikuti perjalanan dari pertanyaan, data, pengodean, analisis, sampai format observasi. Ketika jejak itu jelas, temuan dapat digunakan untuk memperbaiki pembelajaran tanpa mengubah diagnosis konseptual menjadi label bagi peserta didik.

7.12.13 Prosedur Pilot Pengumpulan Data

Dalam metodologi penelitian miskonsepsi kimia, prosedur pilot pengumpulan data menentukan mutu hubungan antara data dan kesimpulan. Fokus ini berada dalam pemilihan dan pelaksanaan teknik yang mampu menangkap jawaban, alasan, model mental, keyakinan, interaksi, dan perubahan konsepsi secara sah. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. O'Brien et al. (2014) serta Kaltakci Gurel et al. (2015) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Prosedur minimal mencakup definisi rekaman data, alasan memilih gambar, aturan mengubah respons menjadi kategori, serta mekanisme pemeriksaan keamanan penyimpanan. Semua penyimpangan dari rencana awal perlu dicatat. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Pada penelitian mengenai think-aloud stoikiometri, peneliti dapat menyusun beberapa tugas ekuivalen yang berubah konteks, representasi, atau tuntutan penalarannya. Jika pola respons tetap berulang dan disertai keyakinan tinggi, terdapat dasar yang lebih kuat untuk menyebut indikasi miskonsepsi. Jika pola berubah-ubah, interpretasi yang lebih hati-hati adalah pengetahuan parsial atau ketergantungan konteks.

Bukti minimum yang perlu tersedia:

- Definisikan rekaman data dalam istilah operasional.
- Tentukan bukti utama dari observasi dan video dan bukti pendukung dari pertanyaan terbuka.
- Periksa dokumentasi konteks sebelum membuat kategori atau model.
- Catat potensi leading question beserta strategi mitigasinya.
- Sajikan protokol pengumpulan bersama batas interpretasi dan ketidakpastian.

Pemeriksaan mutu hendaknya dilakukan selama penelitian, bukan hanya pada akhir. Catatan tentang keamanan penyimpanan, keputusan terkait kesesuaian teknik-konstruk, dan kejadian yang berpotensi menimbulkan mengabaikan bahasa peserta perlu dipelihara dalam *audit trail*. Hal ini memungkinkan koreksi dini dan pelaporan yang jujur.

Kesimpulannya, prosedur pilot pengumpulan data perlu dirancang sebagai proses argumentatif. Data tidak berbicara sendiri; maknanya bergantung pada definisi konstruk, kualitas prosedur, dan konteks. Oleh sebab itu, klaim akhir sebaiknya menggunakan bahasa

yang sesuai dengan kekuatan bukti, misalnya “menunjukkan pola”, “mengindikasikan”, atau “mendukung penjelasan”.

7.12.14 Manajemen dan Keamanan Data

Keputusan tentang manajemen dan keamanan data harus dibuat sebelum data utama dikumpulkan, bukan ditambahkan setelah hasil diketahui. Fokus ini berada dalam pemilihan dan pelaksanaan teknik yang mampu menangkap jawaban, alasan, model mental, keyakinan, interaksi, dan perubahan konsepsi secara sah. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. AERA, APA, dan NCME (2014) serta O’Brien et al. (2014) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Tahap kerja dimulai dengan penegasan waktu pengukuran. Setelah itu, peneliti menyusun aturan untuk tingkat keterbukaan dan menentukan bagaimana standarisasi administrator akan diperiksa. Bukti sebaiknya berasal dari lebih dari satu bentuk, misalnya observasi dan video dan pertanyaan terbuka. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Pada topik observasi diskusi kesetimbangan, peneliti juga perlu memperhatikan bahasa. Istilah yang sama dapat diberi makna berbeda oleh peserta didik. Klarifikasi melalui pertanyaan tindak lanjut, gambar, atau contoh tandingan membantu membedakan masalah terminologi dari kesalahan konseptual.

Kriteria pemeriksaan sebelum analisis:

- Definisikan waktu pengukuran dalam istilah operasional.
- Tentukan bukti utama dari tes pilihan bertingkat dan bukti pendukung dari wawancara.
- Periksa kesesuaian teknik-konstruk sebelum membuat kategori atau model.
- Catat potensi administrator memberi petunjuk beserta strategi mitigasinya.
- Sajikan jadwal bersama batas interpretasi dan ketidakpastian.

Kualitas penelitian meningkat ketika peneliti menunjukkan rantai bukti dari observasi dan video menuju log penyimpangan, menyatakan asumsi, dan memeriksa alternatif penjelasan. Sebaliknya, instrumen terlalu membebani dapat menghasilkan klaim yang tampak tegas tetapi tidak stabil. Rekomendasi praktis harus proporsional dengan bukti.

Pada akhirnya, keputusan tentang manajemen dan keamanan data menentukan apakah penelitian menghasilkan daftar jawaban salah atau pemahaman yang dapat ditindaklanjuti

mengenai cara peserta didik berpikir. Sasaran utamanya ialah bukti yang cukup untuk merancang diagnosis, pembelajaran, dan penelitian lanjutan secara bertanggung jawab.

7.12.15 Sintesis Strategi Pengumpulan Data

Sintesis strategi pengumpulan data perlu dipahami sebagai keputusan metodologis yang mempunyai konsekuensi pada desain, analisis, dan pelaporan. Fokus ini berada dalam pemilihan dan pelaksanaan teknik yang mampu menangkap jawaban, alasan, model mental, keyakinan, interaksi, dan perubahan konsepsi secara sah. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Treagust (1988) serta Kaltakci Gurel et al. (2015) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Peneliti sebaiknya menyiapkan matriks yang menghubungkan pertanyaan penelitian dengan konstruk yang dicari, tes pilihan bertingkat, teknik analisis, dan protokol pengumpulan. Matriks ini mencegah pengumpulan data yang banyak tetapi tidak relevan. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Ilustrasi tes CRI redoks juga menunjukkan bahwa kualitas instrumen memengaruhi temuan. Distraktor yang tidak berasal dari pemikiran peserta didik dapat menghasilkan pilihan acak, sedangkan pertanyaan yang terlalu menuntut matematika dapat mengukur keterampilan hitung alih-alih konsepsi. Peneliti perlu menilai kesesuaian bukti dengan konstruk yang dimaksud.

Langkah operasional yang disarankan:

- Definisikan konstruk yang dicari dalam istilah operasional.
- Tentukan bukti utama dari pertanyaan terbuka dan bukti pendukung dari think-aloud.
- Periksa prosedur baku sebelum membuat kategori atau model.
- Catat potensi rekaman tidak lengkap beserta strategi mitigasinya.
- Sajikan skrip instruksi bersama batas interpretasi dan ketidakpastian.

Peneliti sebaiknya memisahkan keputusan apriori dari keputusan eksploratori. Pemisahan tersebut membantu mengendalikan leading question dan mencegah transkripsi selektif. Selanjutnya, kesesuaian teknik-konstruk dan pelatihan pengumpul data dilaporkan bersama ketidakpastian, sehingga pembaca dapat menilai kekuatan temuan tanpa bergantung pada kesimpulan penulis semata.

Dengan demikian, sintesis strategi pengumpulan data berfungsi menghubungkan keputusan desain dengan interpretasi. Produk yang diharapkan bukan hanya protokol pengumpulan,

tetapi juga penjelasan mengenai bagaimana produk tersebut diperoleh dan dalam kondisi apa ia dapat digunakan. Penelitian miskonsepsi yang baik selalu menyisakan ruang bagi penjelasan alternatif dan verifikasi lanjutan.

7.13 Analisis Data Kuantitatif

Subbab ini membahas pengolahan numerik untuk mengestimasi prevalensi, kualitas instrumen, asosiasi, perubahan, heterogenitas, dan ketidakpastian temuan miskonsepsi. Fokus utamanya adalah menjaga keselarasan antara pertanyaan penelitian, bentuk bukti, teknik analisis, dan kekuatan klaim. Dalam konteks miskonsepsi kimia, keputusan metodologis harus mampu membedakan pola konseptual dari respons yang timbul karena keterbatasan bahasa, kesalahan hitung, atau kondisi administrasi. Pembahasan berikut dipecah menjadi lima belas bagian agar setiap keputusan dapat ditelaah secara operasional.

7.13.1 Persiapan dan Pembersihan Data

Ketelitian penelitian pada bagian persiapan dan pembersihan data sangat menentukan apakah pola jawaban dapat ditafsirkan sebagai bukti konseptual. Fokus ini berada dalam pengolahan numerik untuk mengestimasi prevalensi, kualitas instrumen, asosiasi, perubahan, heterogenitas, dan ketidakpastian temuan miskonsepsi. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Embretson dan Reise (2000) serta Appelbaum et al. (2018) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Dokumentasi metodologis hendaknya memuat siapa yang melakukan kategori konsepsi, kapan data diperoleh, bagaimana estimand dikendalikan, dan atas dasar apa interval estimasi disusun. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Contoh pada kajian Rasch pada CCI memperlihatkan pentingnya pemisahan antara ketidaktahuan, kesalahan prosedural, dan miskonsepsi. Seorang peserta dapat gagal karena lupa rumus, sedangkan peserta lain memilih jawaban yang sama karena menggunakan model mental yang tidak ilmiah. Perbedaan ini baru terlihat ketika kategori konsepsi dibaca bersama skor pre-post dan bukti tentang alasan.

Keputusan yang harus didokumentasikan:

- Definisikan unit analisis dalam istilah operasional.
- Tentukan bukti utama dari tingkat keyakinan dan bukti pendukung dari data kelas.

- Periksa interval kepercayaan sebelum membuat kategori atau model.
- Catat potensi menghapus data hilang otomatis beserta strategi mitigasinya.
- Sajikan ukuran efek bersama batas interpretasi dan ketidakpastian.

Standar metodologis menuntut koherensi antara konstruk, prosedur, analisis, dan klaim. Dalam konteks ini, robustness tidak dapat digantikan oleh indeks tunggal. Demikian pula, ukuran efek harus dinilai dengan bukti yang sesuai. Embretson dan Reise (2000) dan Appelbaum et al. (2018) memberikan dasar untuk pelaporan yang lengkap dan dapat ditelaah.

Inti metodologis subbagian ini adalah keterlacakan. Pembaca harus dapat mengikuti perjalanan dari pertanyaan, data, pengodean, analisis, sampai interval estimasi. Ketika jejak itu jelas, temuan dapat digunakan untuk memperbaiki pembelajaran tanpa mengubah diagnosis konseptual menjadi label bagi peserta didik.

7.13.2 Statistik Deskriptif

Peneliti sering memusatkan perhatian pada teknik, padahal statistik deskriptif justru mengatur batas inferensi yang dapat dibuat. Fokus ini berada dalam pengolahan numerik untuk mengestimasi prevalensi, kualitas instrumen, asosiasi, perubahan, heterogenitas, dan ketidakpastian temuan miskonsepsi. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Little dan Rubin (2019) serta Lakens (2022) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Rancangan yang dapat diaudit mencantumkan alasan pemilihan penanganan data hilang, prosedur penanganan model statistik, dan kriteria keputusan untuk skala pengukuran. Tanpa penjelasan itu, pembaca sulit menilai apakah tingkat keyakinan benar-benar mendukung klaim. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Kasus klaster profil asam-basa dapat digunakan untuk memperlihatkan hubungan antara pola individual dan pola kelas. Sebagian miskonsepsi mungkin hanya dimiliki beberapa peserta tetapi sangat kuat, sedangkan pola lain tersebar luas karena bersumber dari analogi atau buku teks yang sama. Analisis perlu menjaga kedua tingkat tersebut.

Pertanyaan audit metodologis:

- Definisikan penanganan data hilang dalam istilah operasional.
- Tentukan bukti utama dari skor pre-post dan bukti pendukung dari parameter instrumen.
- Periksa ukuran efek sebelum membuat kategori atau model.

- Catat potensi menginterpretasi alpha keliru beserta strategi mitigasinya.
- Sajikan parameter butir bersama batas interpretasi dan ketidakpastian.

Pemeriksaan kualitas dapat dilakukan melalui audit terhadap interval kepercayaan, analisis kasus yang bertentangan, dan evaluasi kontrol multipel. Apabila ditemukan mengabaikan clustering, peneliti perlu menurunkan kekuatan klaim atau menambah bukti. Transparansi seperti ini meningkatkan kegunaan temuan bagi penelitian lanjutan dan praktik kelas.

Kesimpulannya, statistik deskriptif perlu dirancang sebagai proses argumentatif. Data tidak berbicara sendiri; maknanya bergantung pada definisi konstruk, kualitas prosedur, dan konteks. Oleh sebab itu, klaim akhir sebaiknya menggunakan bahasa yang sesuai dengan kekuatan bukti, misalnya “menunjukkan pola”, “mengindikasikan”, atau “mendukung penjelasan”.

7.13.3 Estimasi Prevalensi dan Interval Kepercayaan

Kajian yang kuat memperlakukan estimasi prevalensi dan interval kepercayaan sebagai bagian dari rantai argumentasi, bukan sebagai unsur pelengkap. Fokus ini berada dalam pengolahan numerik untuk mengestimasi prevalensi, kualitas instrumen, asosiasi, perubahan, heterogenitas, dan ketidakpastian temuan miskonsepsi. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. van Buuren (2018) serta Little dan Rubin (2019) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Pada tahap perencanaan, estimand dan pelaporan ketidakpastian harus dibedakan secara eksplisit. Pada tahap pelaksanaan, konsistensi skor pre-post perlu diperiksa; pada tahap interpretasi, ukuran efek menjadi dasar pembatasan klaim. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

$$\hat{p} = x/n; SE(\hat{p}) = \sqrt{[\hat{p}(1 - \hat{p})/n]}$$

Dalam kelas yang mempelajari meta-analisis intervensi, data agregat mungkin menunjukkan mayoritas jawaban salah. Akan tetapi, wawancara dapat memperlihatkan beberapa mekanisme berbeda: bahasa sehari-hari, generalisasi aturan, salah membaca simbol, atau model partikel yang tidak tepat. Karena itu, satu persentase tidak boleh menggantikan analisis struktur alasan.

Bukti minimum yang perlu tersedia:

- Definisikan estimand dalam istilah operasional.

- Tentukan bukti utama dari data kelas dan bukti pendukung dari skor butir.
- Periksa kontrol multipel sebelum membuat kategori atau model.
- Catat potensi menyembunyikan hasil nol beserta strategi mitigasinya.
- Sajikan analisis sensitivitas bersama batas interpretasi dan ketidakpastian.

Kriteria keberhasilan tidak cukup berupa banyaknya data. Yang lebih penting adalah apakah ukuran efek dapat ditelusuri, apakah reproducibility diperiksa, dan apakah robustness dijelaskan. Peneliti harus menghindari menghapus data hilang otomatis; jika risiko itu tidak dapat dihilangkan, dampaknya perlu dianalisis dan dicantumkan sebagai batas penelitian.

Pada akhirnya, keputusan tentang estimasi prevalensi dan interval kepercayaan menentukan apakah penelitian menghasilkan daftar jawaban salah atau pemahaman yang dapat ditindaklanjuti mengenai cara peserta didik berpikir. Sasaran utamanya ialah bukti yang cukup untuk merancang diagnosis, pembelajaran, dan penelitian lanjutan secara bertanggung jawab.

7.13.4 Analisis Tabulasi Silang

Subbagian ini menempatkan analisis tabulasi silang sebagai simpul penting dalam analisis data kuantitatif. Fokus ini berada dalam pengolahan numerik untuk mengestimasi prevalensi, kualitas instrumen, asosiasi, perubahan, heterogenitas, dan ketidakpastian temuan miskonsepsi. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Appelbaum et al. (2018) serta Cumming (2014) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Secara operasional, peneliti perlu menetapkan model statistik, menjelaskan skala pengukuran, dan mendokumentasikan penanganan data hilang. Keputusan tersebut kemudian diterjemahkan ke dalam pilihan data kelas, skor butir, serta tingkat keyakinan. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Sebagai ilustrasi, pada topik prevalensi miskonsepsi mol, peserta didik dapat memberikan jawaban benar pada tingkat simbolik tetapi alasan submikroskopiknya keliru. Peneliti tidak cukup menghitung benar-salah; ia perlu membandingkan data kelas dengan skor butir, memeriksa tingkat keyakinan, dan menguji konsistensi pada konteks lain. Hasil analisis mungkin menunjukkan bahwa kesalahan hanya muncul pada satu representasi, atau justru merupakan hubungan konseptual yang stabil.

Kriteria pemeriksaan sebelum analisis:

- Definisikan model statistik dalam istilah operasional.
- Tentukan bukti utama dari parameter instrumen dan bukti pendukung dari kategori konsepsi.
- Periksa reproducibility sebelum membuat kategori atau model.
- Catat potensi p-hacking beserta strategi mitigasinya.
- Sajikan statistik deskriptif bersama batas interpretasi dan ketidakpastian.

Mutu inferensi terutama ditentukan oleh kontrol multipel, diagnostik asumsi, dan interval kepercayaan. Ancaman utama ialah menginterpretasi alpha keliru serta dichotomization berlebihan. Karena itu, laporan perlu menyajikan bukti yang mendukung sekaligus bukti yang membatasi interpretasi. Prinsip ini konsisten dengan penekanan Appelbaum et al. (2018) dan Cumming (2014) bahwa validitas melekat pada interpretasi dan penggunaan hasil, bukan sekadar pada nama instrumen atau desain.

Dengan demikian, analisis tabulasi silang berfungsi menghubungkan keputusan desain dengan interpretasi. Produk yang diharapkan bukan hanya parameter butir, tetapi juga penjelasan mengenai bagaimana produk tersebut diperoleh dan dalam kondisi apa ia dapat digunakan. Penelitian miskonsepsi yang baik selalu menyisakan ruang bagi penjelasan alternatif dan verifikasi lanjutan.

7.13.5 Uji Perbedaan Kelompok

Pembahasan mengenai uji perbedaan kelompok diperlukan agar analisis data kuantitatif tidak berhenti pada prosedur administratif. Fokus ini berada dalam pengolahan numerik untuk mengestimasi prevalensi, kualitas instrumen, asosiasi, perubahan, heterogenitas, dan ketidakpastian temuan miskonsepsi. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Cumming (2014) serta Taber (2018) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Penerapan praktis dapat disusun dalam urutan: mendefinisikan pelaporan ketidakpastian, mengoperasionalkan unit analisis, mengumpulkan parameter instrumen, memeriksa reproducibility, lalu melaporkan analisis sensitivitas. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Misalnya, pada perubahan skor kesetimbangan, perubahan skor sesudah pembelajaran belum otomatis menandakan perubahan konsep. Peserta didik mungkin mempelajari pola jawaban

tes. Bukti yang lebih meyakinkan mencakup transfer ke situasi baru, penjelasan kausal, representasi partikel, dan retensi setelah jeda waktu.

Langkah operasional yang disarankan:

- Definisikan pelaporan ketidakpastian dalam istilah operasional.
- Tentukan bukti utama dari skor butir dan bukti pendukung dari tingkat keyakinan.
- Periksa diagnostik asumsi sebelum membuat kategori atau model.
- Catat potensi dichotomization berlebihan beserta strategi mitigasinya.
- Sajikan interval estimasi bersama batas interpretasi dan ketidakpastian.

Temuan yang baik tidak selalu berarti hasil yang besar atau sesuai harapan. Temuan bernilai ketika reproducibility terjaga, robustness dijelaskan, dan ketidakpastian disajikan. Kerangka Cumming (2014) serta Taber (2018) menolong peneliti membedakan kekuatan bukti dari daya tarik narasi.

Inti metodologis subbagian ini adalah keterlacakan. Pembaca harus dapat mengikuti perjalanan dari pertanyaan, data, pengodean, analisis, sampai analisis sensitivitas. Ketika jejak itu jelas, temuan dapat digunakan untuk memperbaiki pembelajaran tanpa mengubah diagnosis konseptual menjadi label bagi peserta didik.

7.13.6 Korelasi dan Regresi

Dalam metodologi penelitian miskonsepsi kimia, korelasi dan regresi menentukan mutu hubungan antara data dan kesimpulan. Fokus ini berada dalam pengolahan numerik untuk mengestimasi prevalensi, kualitas instrumen, asosiasi, perubahan, heterogenitas, dan ketidakpastian temuan miskonsepsi. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Wasserstein dan Lazar (2016) serta van Buuren (2018) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Prosedur minimal mencakup definisi skala pengukuran, alasan memilih skor butir, aturan mengubah respons menjadi kategori, serta mekanisme pemeriksaan diagnostik asumsi. Semua penyimpangan dari rencana awal perlu dicatat. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Pada penelitian mengenai hubungan keyakinan dan akurasi, peneliti dapat menyusun beberapa tugas ekuivalen yang berubah konteks, representasi, atau tuntutan penalarannya. Jika pola respons tetap berulang dan disertai keyakinan tinggi, terdapat dasar yang lebih kuat

untuk menyebut indikasi miskonsepsi. Jika pola berubah-ubah, interpretasi yang lebih hati-hati adalah pengetahuan parsial atau ketergantungan konteks.

Keputusan yang harus didokumentasikan:

- Definisikan skala pengukuran dalam istilah operasional.
- Tentukan bukti utama dari kategori konsepsi dan bukti pendukung dari skor pre-post.
- Periksa robustness sebelum membuat kategori atau model.
- Catat potensi mengabaikan clustering beserta strategi mitigasinya.
- Sajikan model regresi bersama batas interpretasi dan ketidakpastian.

Pemeriksaan mutu hendaknya dilakukan selama penelitian, bukan hanya pada akhir. Catatan tentang diagnostik asumsi, keputusan terkait interval kepercayaan, dan kejadian yang berpotensi menimbulkan p-hacking perlu dipelihara dalam *audit trail*. Hal ini memungkinkan koreksi dini dan pelaporan yang jujur.

Kesimpulannya, korelasi dan regresi perlu dirancang sebagai proses argumentatif. Data tidak berbicara sendiri; maknanya bergantung pada definisi konstruk, kualitas prosedur, dan konteks. Oleh sebab itu, klaim akhir sebaiknya menggunakan bahasa yang sesuai dengan kekuatan bukti, misalnya “menunjukkan pola”, “mengindikasikan”, atau “mendukung penjelasan”.

7.13.7 Analisis Data Pretest-Posttest

Keputusan tentang analisis data pretest-posttest harus dibuat sebelum data utama dikumpulkan, bukan ditambahkan setelah hasil diketahui. Fokus ini berada dalam pengolahan numerik untuk mengestimasi prevalensi, kualitas instrumen, asosiasi, perubahan, heterogenitas, dan ketidakpastian temuan miskonsepsi. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Lakens (2022) serta Wasserstein dan Lazar (2016) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Tahap kerja dimulai dengan penegasan unit analisis. Setelah itu, peneliti menyusun aturan untuk estimand dan menentukan bagaimana pelaporan ketidakpastian akan diperiksa. Bukti sebaiknya berasal dari lebih dari satu bentuk, misalnya kategori konsepsi dan skor pre-post. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Pada topik perbedaan kelas redoks, peneliti juga perlu memperhatikan bahasa. Istilah yang sama dapat diberi makna berbeda oleh peserta didik. Klarifikasi melalui pertanyaan tindak

lanjut, gambar, atau contoh tandingan membantu membedakan masalah terminologi dari kesalahan konseptual.

Pertanyaan audit metodologis:

- Definisikan unit analisis dalam istilah operasional.
- Tentukan bukti utama dari tingkat keyakinan dan bukti pendukung dari data kelas.
- Periksa interval kepercayaan sebelum membuat kategori atau model.
- Catat potensi menghapus data hilang otomatis beserta strategi mitigasinya.
- Sajikan ukuran efek bersama batas interpretasi dan ketidakpastian.

Kualitas penelitian meningkat ketika peneliti menunjukkan rantai bukti dari kategori konsepsi menuju interval estimasi, menyatakan asumsi, dan memeriksa alternatif penjelasan. Sebaliknya, dichotomization berlebihan dapat menghasilkan klaim yang tampak tegas tetapi tidak stabil. Rekomendasi praktis harus proporsional dengan bukti.

Pada akhirnya, keputusan tentang analisis data pretest-posttest menentukan apakah penelitian menghasilkan daftar jawaban salah atau pemahaman yang dapat ditindaklanjuti mengenai cara peserta didik berpikir. Sasaran utamanya ialah bukti yang cukup untuk merancang diagnosis, pembelajaran, dan penelitian lanjutan secara bertanggung jawab.

7.13.8 Ukuran Efek

Ukuran efek perlu dipahami sebagai keputusan metodologis yang mempunyai konsekuensi pada desain, analisis, dan pelaporan. Fokus ini berada dalam pengolahan numerik untuk mengestimasi prevalensi, kualitas instrumen, asosiasi, perubahan, heterogenitas, dan ketidakpastian temuan miskonsepsi. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Taber (2018) serta Embretson dan Reise (2000) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Peneliti sebaiknya menyiapkan matriks yang menghubungkan pertanyaan penelitian dengan penanganan data hilang, tingkat keyakinan, teknik analisis, dan model regresi. Matriks ini mencegah pengumpulan data yang banyak tetapi tidak relevan. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

$$d = (M_1 - M_2)/SD_{pooled}; \quad g = J \times d$$

Ilustrasi reliabilitas tes termokimia juga menunjukkan bahwa kualitas instrumen memengaruhi temuan. Distraktor yang tidak berasal dari pemikiran peserta didik dapat

menghasilkan pilihan acak, sedangkan pertanyaan yang terlalu menuntut matematika dapat mengukur keterampilan hitung alih-alih konsepsi. Peneliti perlu menilai kesesuaian bukti dengan konstruk yang dimaksud.

Bukti minimum yang perlu tersedia:

- Definisikan penanganan data hilang dalam istilah operasional.
- Tentukan bukti utama dari skor pre-post dan bukti pendukung dari parameter instrumen.
- Periksa ukuran efek sebelum membuat kategori atau model.
- Catat potensi menginterpretasi alpha keliru beserta strategi mitigasinya.
- Sajikan parameter butir bersama batas interpretasi dan ketidakpastian.

Peneliti sebaiknya memisahkan keputusan apriori dari keputusan eksploratori. Pemisahan tersebut membantu mengendalikan mengabaikan clustering dan mencegah menyembunyikan hasil nol. Selanjutnya, interval kepercayaan dan kontrol multipel dilaporkan bersama ketidakpastian, sehingga pembaca dapat menilai kekuatan temuan tanpa bergantung pada kesimpulan penulis semata.

Dengan demikian, ukuran efek berfungsi menghubungkan keputusan desain dengan interpretasi. Produk yang diharapkan bukan hanya model regresi, tetapi juga penjelasan mengenai bagaimana produk tersebut diperoleh dan dalam kondisi apa ia dapat digunakan. Penelitian miskonsepsi yang baik selalu menyisakan ruang bagi penjelasan alternatif dan verifikasi lanjutan.

7.13.9 Analisis Data Berjenjang

Ketelitian penelitian pada bagian analisis data berjenjang sangat menentukan apakah pola jawaban dapat ditafsirkan sebagai bukti konseptual. Fokus ini berada dalam pengolahan numerik untuk mengestimasi prevalensi, kualitas instrumen, asosiasi, perubahan, heterogenitas, dan ketidakpastian temuan miskonsepsi. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Embretson dan Reise (2000) serta Appelbaum et al. (2018) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Dokumentasi metodologis hendaknya memuat siapa yang melakukan skor pre-post, kapan data diperoleh, bagaimana pelaporan ketidakpastian dikendalikan, dan atas dasar apa ukuran efek disusun. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Contoh pada kajian Rasch pada CCI memperlihatkan pentingnya pemisahan antara ketidaktahuan, kesalahan prosedural, dan miskonsepsi. Seorang peserta dapat gagal karena lupa rumus, sedangkan peserta lain memilih jawaban yang sama karena menggunakan model mental yang tidak ilmiah. Perbedaan ini baru terlihat ketika skor pre-post dibaca bersama parameter instrumen dan bukti tentang alasan.

Kriteria pemeriksaan sebelum analisis:

- Definisikan estimand dalam istilah operasional.
- Tentukan bukti utama dari data kelas dan bukti pendukung dari skor butir.
- Periksa kontrol multipel sebelum membuat kategori atau model.
- Catat potensi menyembunyikan hasil nol beserta strategi mitigasinya.
- Sajikan analisis sensitivitas bersama batas interpretasi dan ketidakpastian.

Standar metodologis menuntut koherensi antara konstruk, prosedur, analisis, dan klaim. Dalam konteks ini, ukuran efek tidak dapat digantikan oleh indeks tunggal. Demikian pula, reproducibility harus dinilai dengan bukti yang sesuai. Embretson dan Reise (2000) dan Appelbaum et al. (2018) memberikan dasar untuk pelaporan yang lengkap dan dapat ditelaah.

Inti metodologis subbagian ini adalah keterlacakan. Pembaca harus dapat mengikuti perjalanan dari pertanyaan, data, pengodean, analisis, sampai ukuran efek. Ketika jejak itu jelas, temuan dapat digunakan untuk memperbaiki pembelajaran tanpa mengubah diagnosis konseptual menjadi label bagi peserta didik.

7.13.10 Penanganan Data Hilang

Peneliti sering memusatkan perhatian pada teknik, padahal penanganan data hilang justru mengatur batas inferensi yang dapat dibuat. Fokus ini berada dalam pengolahan numerik untuk mengestimasi prevalensi, kualitas instrumen, asosiasi, perubahan, heterogenitas, dan ketidakpastian temuan miskonsepsi. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Little dan Rubin (2019) serta Lakens (2022) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Rancangan yang dapat diaudit mencantumkan alasan pemilihan model statistik, prosedur penanganan skala pengukuran, dan kriteria keputusan untuk penanganan data hilang. Tanpa penjelasan itu, pembaca sulit menilai apakah data kelas benar-benar mendukung klaim. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Kasus kluster profil asam-basa dapat digunakan untuk memperlihatkan hubungan antara pola individual dan pola kelas. Sebagian miskonsepsi mungkin hanya dimiliki beberapa peserta tetapi sangat kuat, sedangkan pola lain tersebar luas karena bersumber dari analogi atau buku teks yang sama. Analisis perlu menjaga kedua tingkat tersebut.

Langkah operasional yang disarankan:

- Definisikan model statistik dalam istilah operasional.
- Tentukan bukti utama dari parameter instrumen dan bukti pendukung dari kategori konsepsi.
- Periksa reproducibility sebelum membuat kategori atau model.
- Catat potensi p-hacking beserta strategi mitigasinya.
- Sajikan statistik deskriptif bersama batas interpretasi dan ketidakpastian.

Pemeriksaan kualitas dapat dilakukan melalui audit terhadap kontrol multipel, analisis kasus yang bertentangan, dan evaluasi diagnostik asumsi. Apabila ditemukan menginterpretasi alpha keliru, peneliti perlu menurunkan kekuatan klaim atau menambah bukti. Transparansi seperti ini meningkatkan kegunaan temuan bagi penelitian lanjutan dan praktik kelas.

Kesimpulannya, penanganan data hilang perlu dirancang sebagai proses argumentatif. Data tidak berbicara sendiri; maknanya bergantung pada definisi konstruk, kualitas prosedur, dan konteks. Oleh sebab itu, klaim akhir sebaiknya menggunakan bahasa yang sesuai dengan kekuatan bukti, misalnya “menunjukkan pola”, “mengindikasikan”, atau “mendukung penjelasan”.

7.13.11 Reliabilitas Skor

Kajian yang kuat memperlakukan reliabilitas skor sebagai bagian dari rantai argumentasi, bukan sebagai unsur pelengkap. Fokus ini berada dalam pengolahan numerik untuk mengestimasi prevalensi, kualitas instrumen, asosiasi, perubahan, heterogenitas, dan ketidakpastian temuan miskonsepsi. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. van Buuren (2018) serta Little dan Rubin (2019) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Pada tahap perencanaan, pelaporan ketidakpastian dan unit analisis harus dibedakan secara eksplisit. Pada tahap pelaksanaan, konsistensi parameter instrumen perlu diperiksa; pada tahap interpretasi, reproducibility menjadi dasar pembatasan klaim. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

$$\omega = (\Sigma \lambda_i)^2 / [(\Sigma \lambda_i)^2 + \Sigma \theta_i]$$

Dalam kelas yang mempelajari meta-analisis intervensi, data agregat mungkin menunjukkan mayoritas jawaban salah. Akan tetapi, wawancara dapat memperlihatkan beberapa mekanisme berbeda: bahasa sehari-hari, generalisasi aturan, salah membaca simbol, atau model partikel yang tidak tepat. Karena itu, satu persentase tidak boleh menggantikan analisis struktur alasan.

Keputusan yang harus didokumentasikan:

- Definisikan pelaporan ketidakpastian dalam istilah operasional.
- Tentukan bukti utama dari skor butir dan bukti pendukung dari tingkat keyakinan.
- Periksa diagnostik asumsi sebelum membuat kategori atau model.
- Catat potensi dichotomization berlebihan beserta strategi mitigasinya.
- Sajikan interval estimasi bersama batas interpretasi dan ketidakpastian.

Kriteria keberhasilan tidak cukup berupa banyaknya data. Yang lebih penting adalah apakah reproducibility dapat ditelusuri, apakah robustness diperiksa, dan apakah ukuran efek dijelaskan. Peneliti harus menghindari menyembunyikan hasil nol; jika risiko itu tidak dapat dihilangkan, dampaknya perlu dianalisis dan dicantumkan sebagai batas penelitian.

Pada akhirnya, keputusan tentang reliabilitas skor menentukan apakah penelitian menghasilkan daftar jawaban salah atau pemahaman yang dapat ditindaklanjuti mengenai cara peserta didik berpikir. Sasaran utamanya ialah bukti yang cukup untuk merancang diagnosis, pembelajaran, dan penelitian lanjutan secara bertanggung jawab.

7.13.12 Analisis Butir Klasik

Subbagian ini menempatkan analisis butir klasik sebagai simpul penting dalam analisis data kuantitatif. Fokus ini berada dalam pengolahan numerik untuk mengestimasi prevalensi, kualitas instrumen, asosiasi, perubahan, heterogenitas, dan ketidakpastian temuan miskonsepsi. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Appelbaum et al. (2018) serta Cumming (2014) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Secara operasional, peneliti perlu menetapkan skala pengukuran, menjelaskan penanganan data hilang, dan mendokumentasikan model statistik. Keputusan tersebut kemudian diterjemahkan ke dalam pilihan skor butir, tingkat keyakinan, serta data kelas. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

$$p_i = R_i/N; D_i = p_i(\text{atas}) - p_i(\text{bawah})$$

Sebagai ilustrasi, pada topik prevalensi miskonsepsi mol, peserta didik dapat memberikan jawaban benar pada tingkat simbolik tetapi alasan submikroskopiknya keliru. Peneliti tidak cukup menghitung benar-salah; ia perlu membandingkan skor butir dengan tingkat keyakinan, memeriksa tingkat keyakinan, dan menguji konsistensi pada konteks lain. Hasil analisis mungkin menunjukkan bahwa kesalahan hanya muncul pada satu representasi, atau justru merupakan hubungan konseptual yang stabil.

Pertanyaan audit metodologis:

- Definisikan skala pengukuran dalam istilah operasional.
- Tentukan bukti utama dari kategori konsepsi dan bukti pendukung dari skor pre-post.
- Periksa robustness sebelum membuat kategori atau model.
- Catat potensi mengabaikan clustering beserta strategi mitigasinya.
- Sajikan model regresi bersama batas interpretasi dan ketidakpastian.

Mutu inferensi terutama ditentukan oleh diagnostik asumsi, interval kepercayaan, dan kontrol multipel. Ancaman utama ialah p-hacking serta menghapus data hilang otomatis. Karena itu, laporan perlu menyajikan bukti yang mendukung sekaligus bukti yang membatasi interpretasi. Prinsip ini konsisten dengan penekanan Appelbaum et al. (2018) dan Cumming (2014) bahwa validitas melekat pada interpretasi dan penggunaan hasil, bukan sekadar pada nama instrumen atau desain.

Dengan demikian, analisis butir klasik berfungsi menghubungkan keputusan desain dengan interpretasi. Produk yang diharapkan bukan hanya statistik deskriptif, tetapi juga penjelasan mengenai bagaimana produk tersebut diperoleh dan dalam kondisi apa ia dapat digunakan. Penelitian miskonsepsi yang baik selalu menyisakan ruang bagi penjelasan alternatif dan verifikasi lanjutan.

7.13.13 Rasch dan Item Response Theory

Pembahasan mengenai rasch dan item response theory diperlukan agar analisis data kuantitatif tidak berhenti pada prosedur administratif. Fokus ini berada dalam pengolahan numerik untuk mengestimasi prevalensi, kualitas instrumen, asosiasi, perubahan, heterogenitas, dan ketidakpastian temuan miskonsepsi. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Cumming (2014) serta Taber (2018) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Penerapan praktis dapat disusun dalam urutan: mendefinisikan unit analisis, mengoperasionalkan estimand, mengumpulkan kategori konsepsi, memeriksa robustness, lalu melaporkan interval estimasi. Setiap keputusan dicatat dalam protokol, termasuk versi

instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Misalnya, pada perubahan skor kesetimbangan, perubahan skor sesudah pembelajaran belum otomatis menandakan perubahan konsep. Peserta didik mungkin mempelajari pola jawaban tes. Bukti yang lebih meyakinkan mencakup transfer ke situasi baru, penjelasan kausal, representasi partikel, dan retensi setelah jeda waktu.

Bukti minimum yang perlu tersedia:

- Definisikan unit analisis dalam istilah operasional.
- Tentukan bukti utama dari tingkat keyakinan dan bukti pendukung dari data kelas.
- Periksa interval kepercayaan sebelum membuat kategori atau model.
- Catat potensi menghapus data hilang otomatis beserta strategi mitigasinya.
- Sajikan ukuran efek bersama batas interpretasi dan ketidakpastian.

Temuan yang baik tidak selalu berarti hasil yang besar atau sesuai harapan. Temuan bernilai ketika robustness terjaga, ukuran efek dijelaskan, dan ketidakpastian disajikan. Kerangka Cumming (2014) serta Taber (2018) menolong peneliti membedakan kekuatan bukti dari daya tarik narasi.

Inti metodologis subbagian ini adalah keterlacakan. Pembaca harus dapat mengikuti perjalanan dari pertanyaan, data, pengodean, analisis, sampai interval estimasi. Ketika jejak itu jelas, temuan dapat digunakan untuk memperbaiki pembelajaran tanpa mengubah diagnosis konseptual menjadi label bagi peserta didik.

7.13.14 Kontrol Kesalahan Multipel dan Sensitivitas

Dalam metodologi penelitian miskonsepsi kimia, kontrol kesalahan multipel dan sensitivitas menentukan mutu hubungan antara data dan kesimpulan. Fokus ini berada dalam pengolahan numerik untuk mengestimasi prevalensi, kualitas instrumen, asosiasi, perubahan, heterogenitas, dan ketidakpastian temuan miskonsepsi. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Wasserstein dan Lazar (2016) serta van Buuren (2018) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Prosedur minimal mencakup definisi penanganan data hilang, alasan memilih tingkat keyakinan, aturan mengubah respons menjadi kategori, serta mekanisme pemeriksaan interval kepercayaan. Semua penyimpangan dari rencana awal perlu dicatat. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak

hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

$p(k) \leq (k/m)q$ untuk prosedur Benjamini–Hochberg

Pada penelitian mengenai hubungan keyakinan dan akurasi, peneliti dapat menyusun beberapa tugas ekuivalen yang berubah konteks, representasi, atau tuntutan penalarannya. Jika pola respons tetap berulang dan disertai keyakinan tinggi, terdapat dasar yang lebih kuat untuk menyebut indikasi miskonsepsi. Jika pola berubah-ubah, interpretasi yang lebih hati-hati adalah pengetahuan parsial atau ketergantungan konteks.

Kriteria pemeriksaan sebelum analisis:

- Definisikan penanganan data hilang dalam istilah operasional.
- Tentukan bukti utama dari skor pre-post dan bukti pendukung dari parameter instrumen.
- Periksa ukuran efek sebelum membuat kategori atau model.
- Catat potensi menginterpretasi alpha keliru beserta strategi mitigasinya.
- Sajikan parameter butir bersama batas interpretasi dan ketidakpastian.

Pemeriksaan mutu hendaknya dilakukan selama penelitian, bukan hanya pada akhir. Catatan tentang interval kepercayaan, keputusan terkait kontrol multipel, dan kejadian yang berpotensi menimbulkan mengabaikan clustering perlu dipelihara dalam *audit trail*. Hal ini memungkinkan koreksi dini dan pelaporan yang jujur.

Kesimpulannya, kontrol kesalahan multipel dan sensitivitas perlu dirancang sebagai proses argumentatif. Data tidak berbicara sendiri; maknanya bergantung pada definisi konstruk, kualitas prosedur, dan konteks. Oleh sebab itu, klaim akhir sebaiknya menggunakan bahasa yang sesuai dengan kekuatan bukti, misalnya “menunjukkan pola”, “mengindikasikan”, atau “mendukung penjelasan”.

7.13.15 Sintesis Analisis Kuantitatif Miskonsepsi

Keputusan tentang sintesis analisis kuantitatif miskonsepsi harus dibuat sebelum data utama dikumpulkan, bukan ditambahkan setelah hasil diketahui. Fokus ini berada dalam pengolahan numerik untuk mengestimasi prevalensi, kualitas instrumen, asosiasi, perubahan, heterogenitas, dan ketidakpastian temuan miskonsepsi. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Lakens (2022) serta Wasserstein dan Lazar (2016) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Tahap kerja dimulai dengan penegasan estimand. Setelah itu, peneliti menyusun aturan untuk pelaporan ketidakpastian dan menentukan bagaimana unit analisis akan diperiksa. Bukti sebaiknya berasal dari lebih dari satu bentuk, misalnya skor pre-post dan parameter

instrumen. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Pada topik perbedaan kelas redoks, peneliti juga perlu memperhatikan bahasa. Istilah yang sama dapat diberi makna berbeda oleh peserta didik. Klarifikasi melalui pertanyaan tindak lanjut, gambar, atau contoh tandingan membantu membedakan masalah terminologi dari kesalahan konseptual.

Langkah operasional yang disarankan:

- Definisikan estimand dalam istilah operasional.
- Tentukan bukti utama dari data kelas dan bukti pendukung dari skor butir.
- Periksa kontrol multipel sebelum membuat kategori atau model.
- Catat potensi menyembunyikan hasil nol beserta strategi mitigasinya.
- Sajikan analisis sensitivitas bersama batas interpretasi dan ketidakpastian.

Kualitas penelitian meningkat ketika peneliti menunjukkan rantai bukti dari skor pre-post menuju ukuran efek, menyatakan asumsi, dan memeriksa alternatif penjelasan. Sebaliknya, menghapus data hilang otomatis dapat menghasilkan klaim yang tampak tegas tetapi tidak stabil. Rekomendasi praktis harus proporsional dengan bukti.

Pada akhirnya, keputusan tentang sintesis analisis kuantitatif miskonsepsi menentukan apakah penelitian menghasilkan daftar jawaban salah atau pemahaman yang dapat ditindaklanjuti mengenai cara peserta didik berpikir. Sasaran utamanya ialah bukti yang cukup untuk merancang diagnosis, pembelajaran, dan penelitian lanjutan secara bertanggung jawab.

7.14 Analisis Data Kualitatif

Subbab ini membahas proses sistematis untuk mengubah tuturan, tulisan, gambar, observasi, dan artefak menjadi tema, model mental, mekanisme, atau tipologi konsepsi. Fokus utamanya adalah menjaga keselarasan antara pertanyaan penelitian, bentuk bukti, teknik analisis, dan kekuatan klaim. Dalam konteks miskonsepsi kimia, keputusan metodologis harus mampu membedakan pola konseptual dari respons yang timbul karena keterbatasan bahasa, kesalahan hitung, atau kondisi administrasi. Pembahasan berikut dipecah menjadi lima belas bagian agar setiap keputusan dapat ditelaah secara operasional.

7.14.1 Orientasi Analisis Kualitatif

Orientasi analisis kualitatif perlu dipahami sebagai keputusan metodologis yang mempunyai konsekuensi pada desain, analisis, dan pelaporan. Fokus ini berada dalam proses sistematis untuk mengubah tuturan, tulisan, gambar, observasi, dan artefak menjadi tema, model mental, mekanisme, atau tipologi konsepsi. Peneliti perlu menyatakan sejak awal apa yang

dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Braun dan Clarke (2006) serta Braun dan Clarke (2021) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Peneliti sebaiknya menyiapkan matriks yang menghubungkan pertanyaan penelitian dengan peran teori, video kelas, teknik analisis, dan matriks kasus. Matriks ini mencegah pengumpulan data yang banyak tetapi tidak relevan. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Ilustrasi gambar gas juga menunjukkan bahwa kualitas instrumen memengaruhi temuan. Distraktor yang tidak berasal dari pemikiran peserta didik dapat menghasilkan pilihan acak, sedangkan pertanyaan yang terlalu menuntut matematika dapat mengukur keterampilan hitung alih-alih konsepsi. Peneliti perlu menilai kesesuaian bukti dengan konstruk yang dimaksud.

Keputusan yang harus didokumentasikan:

- Definisikan peran teori dalam istilah operasional.
- Tentukan bukti utama dari artefak tugas dan bukti pendukung dari jawaban terbuka.
- Periksa audit trail sebelum membuat kategori atau model.
- Catat potensi coding tanpa pertanyaan beserta strategi mitigasinya.
- Sajikan codebook bersama batas interpretasi dan ketidakpastian.

Peneliti sebaiknya memisahkan keputusan apriori dari keputusan eksploratori. Pemisahan tersebut membantu mengendalikan kutipan tanpa analisis dan mencegah tema sebagai topik semata. Selanjutnya, kecukupan kutipan dan kredibilitas dilaporkan bersama ketidakpastian, sehingga pembaca dapat menilai kekuatan temuan tanpa bergantung pada kesimpulan penulis semata.

Dengan demikian, orientasi analisis kualitatif berfungsi menghubungkan keputusan desain dengan interpretasi. Produk yang diharapkan bukan hanya matriks kasus, tetapi juga penjelasan mengenai bagaimana produk tersebut diperoleh dan dalam kondisi apa ia dapat digunakan. Penelitian miskonsepsi yang baik selalu menyisakan ruang bagi penjelasan alternatif dan verifikasi lanjutan.

7.14.2 Transkripsi dan Familiarisasi

Ketelitian penelitian pada bagian transkripsi dan familiarisasi sangat menentukan apakah pola jawaban dapat ditafsirkan sebagai bukti konseptual. Fokus ini berada dalam proses sistematis untuk mengubah tuturan, tulisan, gambar, observasi, dan artefak menjadi tema, model

mental, mekanisme, atau tipologi konsepsi. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Braun dan Clarke (2021) serta Lincoln dan Guba (1985) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Dokumentasi metodologis hendaknya memuat siapa yang melakukan artefak tugas, kapan data diperoleh, bagaimana unit makna dikendalikan, dan atas dasar apa narasi analitik disusun. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Contoh pada kajian penjelasan redoks memperlihatkan pentingnya pemisahan antara ketidaktahuan, kesalahan prosedural, dan miskonsepsi. Seorang peserta dapat gagal karena lupa rumus, sedangkan peserta lain memilih jawaban yang sama karena menggunakan model mental yang tidak ilmiah. Perbedaan ini baru terlihat ketika artefak tugas dibaca bersama jawaban terbuka dan bukti tentang alasan.

Pertanyaan audit metodologis:

- Definisikan strategi penutupan analisis dalam istilah operasional.
- Tentukan bukti utama dari transkrip wawancara dan bukti pendukung dari gambar partikel.
- Periksa kredibilitas sebelum membuat kategori atau model.
- Catat potensi tema sebagai topik semata beserta strategi mitigasinya.
- Sajikan tema bersama batas interpretasi dan ketidakpastian.

Standar metodologis menuntut koherensi antara konstruk, prosedur, analisis, dan klaim. Dalam konteks ini, *audit trail* tidak dapat digantikan oleh indeks tunggal. Demikian pula, dependabilitas harus dinilai dengan bukti yang sesuai. Braun dan Clarke (2021) dan Lincoln dan Guba (1985) memberikan dasar untuk pelaporan yang lengkap dan dapat ditelaah.

Inti metodologis subbagian ini adalah keterlacakan. Pembaca harus dapat mengikuti perjalanan dari pertanyaan, data, pengodean, analisis, sampai narasi analitik. Ketika jejak itu jelas, temuan dapat digunakan untuk memperbaiki pembelajaran tanpa mengubah diagnosis konseptual menjadi label bagi peserta didik.

7.14.3 Penentuan Unit Makna

Peneliti sering memusatkan perhatian pada teknik, padahal penentuan unit makna justru mengatur batas inferensi yang dapat dibuat. Fokus ini berada dalam proses sistematis untuk mengubah tuturan, tulisan, gambar, observasi, dan artefak menjadi tema, model mental,

mekanisme, atau tipologi konsepsi. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Miles et al. (2014) serta Braun dan Clarke (2006) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Rancangan yang dapat diaudit mencantumkan alasan pemilihan tradisi analisis, prosedur penanganan tingkat induksi, dan kriteria keputusan untuk peran teori. Tanpa penjelasan itu, pembaca sulit menilai apakah transkrip wawancara benar-benar mendukung klaim. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Kasus argumentasi laju dapat digunakan untuk memperlihatkan hubungan antara pola individual dan pola kelas. Sebagian miskonsepsi mungkin hanya dimiliki beberapa peserta tetapi sangat kuat, sedangkan pola lain tersebar luas karena bersumber dari analogi atau buku teks yang sama. Analisis perlu menjaga kedua tingkat tersebut.

Bukti minimum yang perlu tersedia:

- Definisikan tradisi analisis dalam istilah operasional.
- Tentukan bukti utama dari jawaban terbuka dan bukti pendukung dari catatan observasi.
- Periksa dependabilitas sebelum membuat kategori atau model.
- Catat potensi menghitung kode sebagai makna beserta strategi mitigasinya.
- Sajikan tipologi konsepsi bersama batas interpretasi dan ketidakpastian.

Pemeriksaan kualitas dapat dilakukan melalui audit terhadap kredibilitas, analisis kasus yang bertentangan, dan evaluasi confirmabilitas. Apabila ditemukan coding tanpa pertanyaan, peneliti perlu menurunkan kekuatan klaim atau menambah bukti. Transparansi seperti ini meningkatkan kegunaan temuan bagi penelitian lanjutan dan praktik kelas.

Kesimpulannya, penentuan unit makna perlu dirancang sebagai proses argumentatif. Data tidak berbicara sendiri; maknanya bergantung pada definisi konstruk, kualitas prosedur, dan konteks. Oleh sebab itu, klaim akhir sebaiknya menggunakan bahasa yang sesuai dengan kekuatan bukti, misalnya “menunjukkan pola”, “mengindikasikan”, atau “mendukung penjelasan”.

7.14.4 Coding Awal

Kajian yang kuat memperlakukan coding awal sebagai bagian dari rantai argumentasi, bukan sebagai unsur pelengkap. Fokus ini berada dalam proses sistematis untuk mengubah tuturan, tulisan, gambar, observasi, dan artefak menjadi tema, model mental, mekanisme, atau tipologi konsepsi. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang

dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Saldaña (2021) serta Lincoln dan Guba (1985) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Pada tahap perencanaan, unit makna dan kedalaman interpretasi harus dibedakan secara eksplisit. Pada tahap pelaksanaan, konsistensi jawaban terbuka perlu diperiksa; pada tahap interpretasi, dependabilitas menjadi dasar pembatasan klaim. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Dalam kelas yang mempelajari diskusi elektrolit, data agregat mungkin menunjukkan mayoritas jawaban salah. Akan tetapi, wawancara dapat memperlihatkan beberapa mekanisme berbeda: bahasa sehari-hari, generalisasi aturan, salah membaca simbol, atau model partikel yang tidak tepat. Karena itu, satu persentase tidak boleh menggantikan analisis struktur alasan.

Kriteria pemeriksaan sebelum analisis:

- Definisikan unit makna dalam istilah operasional.
- Tentukan bukti utama dari gambar partikel dan bukti pendukung dari video kelas.
- Periksa konfirmabilitas sebelum membuat kategori atau model.
- Catat potensi menghapus kasus negatif beserta strategi mitigasinya.
- Sajikan model mekanisme bersama batas interpretasi dan ketidakpastian.

Kriteria keberhasilan tidak cukup berupa banyaknya data. Yang lebih penting adalah apakah dependabilitas dapat ditelusuri, apakah refleksivitas diperiksa, dan apakah *audit trail* dijelaskan. Peneliti harus menghindari tema sebagai topik semata; jika risiko itu tidak dapat dihilangkan, dampaknya perlu dianalisis dan dicantumkan sebagai batas penelitian.

Pada akhirnya, keputusan tentang coding awal menentukan apakah penelitian menghasilkan daftar jawaban salah atau pemahaman yang dapat ditindaklanjuti mengenai cara peserta didik berpikir. Sasaran utamanya ialah bukti yang cukup untuk merancang diagnosis, pembelajaran, dan penelitian lanjutan secara bertanggung jawab.

7.14.5 Pengembangan Codebook

Subbagian ini menempatkan pengembangan codebook sebagai simpul penting dalam analisis data kualitatif. Fokus ini berada dalam proses sistematis untuk mengubah tuturan, tulisan, gambar, observasi, dan artefak menjadi tema, model mental, mekanisme, atau tipologi konsepsi. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan,

atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Lincoln dan Guba (1985) serta O'Brien et al. (2014) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Secara operasional, peneliti perlu menetapkan tingkat induksi, menjelaskan peran teori, dan mendokumentasikan tradisi analisis. Keputusan tersebut kemudian diterjemahkan ke dalam pilihan gambar partikel, video kelas, serta transkrip wawancara. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Sebagai ilustrasi, pada topik model mental ikatan, peserta didik dapat memberikan jawaban benar pada tingkat simbolik tetapi alasan submikroskopiknya keliru. Peneliti tidak cukup menghitung benar-salah; ia perlu membandingkan gambar partikel dengan video kelas, memeriksa tingkat keyakinan, dan menguji konsistensi pada konteks lain. Hasil analisis mungkin menunjukkan bahwa kesalahan hanya muncul pada satu representasi, atau justru merupakan hubungan konseptual yang stabil.

Langkah operasional yang disarankan:

- Definisikan tingkat induksi dalam istilah operasional.
- Tentukan bukti utama dari catatan observasi dan bukti pendukung dari artefak tugas.
- Periksa refleksivitas sebelum membuat kategori atau model.
- Catat potensi kutipan tanpa analisis beserta strategi mitigasinya.
- Sajikan matriks kasus bersama batas interpretasi dan ketidakpastian.

Mutu inferensi terutama ditentukan oleh konfirmabilitas, kecukupan kutipan, dan kredibilitas. Ancaman utama ialah menghitung kode sebagai makna serta klaim saturasi otomatis. Karena itu, laporan perlu menyajikan bukti yang mendukung sekaligus bukti yang membatasi interpretasi. Prinsip ini konsisten dengan penekanan Lincoln dan Guba (1985) dan O'Brien et al. (2014) bahwa validitas melekat pada interpretasi dan penggunaan hasil, bukan sekadar pada nama instrumen atau desain.

Dengan demikian, pengembangan codebook berfungsi menghubungkan keputusan desain dengan interpretasi. Produk yang diharapkan bukan hanya tipologi konsepsi, tetapi juga penjelasan mengenai bagaimana produk tersebut diperoleh dan dalam kondisi apa ia dapat digunakan. Penelitian miskonsepsi yang baik selalu menyisakan ruang bagi penjelasan alternatif dan verifikasi lanjutan.

7.14.6 Analisis Tematik Refleksif

Pembahasan mengenai analisis tematik refleksif diperlukan agar analisis data kualitatif tidak berhenti pada prosedur administratif. Fokus ini berada dalam proses sistematis untuk

mengubah tuturan, tulisan, gambar, observasi, dan artefak menjadi tema, model mental, mekanisme, atau tipologi konsepsi. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Nowell et al. (2017) serta Miles et al. (2014) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Penerapan praktis dapat disusun dalam urutan: mendefinisikan kedalaman interpretasi, mengoperasionalkan strategi penutupan analisis, mengumpulkan catatan observasi, memeriksa reflektivitas, lalu melaporkan model mekanisme. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Misalnya, pada penalaran asam kuat, perubahan skor sesudah pembelajaran belum otomatis menandakan perubahan konsep. Peserta didik mungkin mempelajari pola jawaban tes. Bukti yang lebih meyakinkan mencakup transfer ke situasi baru, penjelasan kausal, representasi partikel, dan retensi setelah jeda waktu.

Keputusan yang harus didokumentasikan:

- Definisikan kedalaman interpretasi dalam istilah operasional.
- Tentukan bukti utama dari video kelas dan bukti pendukung dari transkrip wawancara.
- Periksa kecukupan kutipan sebelum membuat kategori atau model.
- Catat potensi klaim saturasi otomatis beserta strategi mitigasinya.
- Sajikan narasi analitik bersama batas interpretasi dan ketidakpastian.

Temuan yang baik tidak selalu berarti hasil yang besar atau sesuai harapan. Temuan bernilai ketika reflektivitas terjaga, *audit trail* dijelaskan, dan ketidakpastian disajikan. Kerangka Nowell et al. (2017) serta Miles et al. (2014) menolong peneliti membedakan kekuatan bukti dari daya tarik narasi.

Inti metodologis subbagian ini adalah keterlacakan. Pembaca harus dapat mengikuti perjalanan dari pertanyaan, data, pengodean, analisis, sampai model mekanisme. Ketika jejak itu jelas, temuan dapat digunakan untuk memperbaiki pembelajaran tanpa mengubah diagnosis konseptual menjadi label bagi peserta didik.

7.14.7 Analisis Isi Kualitatif

Dalam metodologi penelitian miskonsepsi kimia, analisis isi kualitatif menentukan mutu hubungan antara data dan kesimpulan. Fokus ini berada dalam proses sistematis untuk mengubah tuturan, tulisan, gambar, observasi, dan artefak menjadi tema, model mental, mekanisme, atau tipologi konsepsi. Peneliti perlu menyatakan sejak awal apa yang dimaksud,

indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. O'Brien et al. (2014) serta Nowell et al. (2017) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Prosedur minimal mencakup definisi peran teori, alasan memilih video kelas, aturan mengubah respons menjadi kategori, serta mekanisme pemeriksaan kecukupan kutipan. Semua penyimpangan dari rencana awal perlu dicatat. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Pada penelitian mengenai analogi kalor, peneliti dapat menyusun beberapa tugas ekuivalen yang berubah konteks, representasi, atau tuntutan penalarannya. Jika pola respons tetap berulang dan disertai keyakinan tinggi, terdapat dasar yang lebih kuat untuk menyebut indikasi miskonsepsi. Jika pola berubah-ubah, interpretasi yang lebih hati-hati adalah pengetahuan parsial atau ketergantungan konteks.

Pertanyaan audit metodologis:

- Definisikan peran teori dalam istilah operasional.
- Tentukan bukti utama dari artefak tugas dan bukti pendukung dari jawaban terbuka.
- Periksa audit trail sebelum membuat kategori atau model.
- Catat potensi coding tanpa pertanyaan beserta strategi mitigasinya.
- Sajikan codebook bersama batas interpretasi dan ketidakpastian.

Pemeriksaan mutu hendaknya dilakukan selama penelitian, bukan hanya pada akhir. Catatan tentang kecukupan kutipan, keputusan terkait kredibilitas, dan kejadian yang berpotensi menimbulkan kutipan tanpa analisis perlu dipelihara dalam *audit trail*. Hal ini memungkinkan koreksi dini dan pelaporan yang jujur.

Kesimpulannya, analisis isi kualitatif perlu dirancang sebagai proses argumentatif. Data tidak berbicara sendiri; maknanya bergantung pada definisi konstruk, kualitas prosedur, dan konteks. Oleh sebab itu, klaim akhir sebaiknya menggunakan bahasa yang sesuai dengan kekuatan bukti, misalnya “menunjukkan pola”, “mengindikasikan”, atau “mendukung penjelasan”.

7.14.8 Analisis Fenomenografis Konsepsi

Keputusan tentang analisis fenomenografis konsepsi harus dibuat sebelum data utama dikumpulkan, bukan ditambahkan setelah hasil diketahui. Fokus ini berada dalam proses sistematis untuk mengubah tuturan, tulisan, gambar, observasi, dan artefak menjadi tema, model mental, mekanisme, atau tipologi konsepsi. Peneliti perlu menyatakan sejak awal apa

yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Braun dan Clarke (2006) serta Braun dan Clarke (2021) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Tahap kerja dimulai dengan penegasan strategi penutupan analisis. Setelah itu, peneliti menyusun aturan untuk unit makna dan menentukan bagaimana kedalaman interpretasi akan diperiksa. Bukti sebaiknya berasal dari lebih dari satu bentuk, misalnya artefak tugas dan jawaban terbuka. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Pada topik bahasa keseimbangan, peneliti juga perlu memperhatikan bahasa. Istilah yang sama dapat diberi makna berbeda oleh peserta didik. Klarifikasi melalui pertanyaan tindak lanjut, gambar, atau contoh tandingan membantu membedakan masalah terminologi dari kesalahan konseptual.

Bukti minimum yang perlu tersedia:

- Definisikan strategi penutupan analisis dalam istilah operasional.
- Tentukan bukti utama dari transkrip wawancara dan bukti pendukung dari gambar partikel.
- Periksa kredibilitas sebelum membuat kategori atau model.
- Catat potensi tema sebagai topik semata beserta strategi mitigasinya.
- Sajikan tema bersama batas interpretasi dan ketidakpastian.

Kualitas penelitian meningkat ketika peneliti menunjukkan rantai bukti dari artefak tugas menuju narasi analitik, menyatakan asumsi, dan memeriksa alternatif penjelasan. Sebaliknya, klaim saturasi otomatis dapat menghasilkan klaim yang tampak tegas tetapi tidak stabil. Rekomendasi praktis harus proporsional dengan bukti.

Pada akhirnya, keputusan tentang analisis fenomenografis konsepsi menentukan apakah penelitian menghasilkan daftar jawaban salah atau pemahaman yang dapat ditindaklanjuti mengenai cara peserta didik berpikir. Sasaran utamanya ialah bukti yang cukup untuk merancang diagnosis, pembelajaran, dan penelitian lanjutan secara bertanggung jawab.

7.14.9 Analisis Model Mental

Analisis model mental perlu dipahami sebagai keputusan metodologis yang mempunyai konsekuensi pada desain, analisis, dan pelaporan. Fokus ini berada dalam proses sistematis untuk mengubah tuturan, tulisan, gambar, observasi, dan artefak menjadi tema, model mental, mekanisme, atau tipologi konsepsi. Peneliti perlu menyatakan sejak awal apa yang

dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Braun dan Clarke (2021) serta Lincoln dan Guba (1985) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Peneliti sebaiknya menyiapkan matriks yang menghubungkan pertanyaan penelitian dengan tradisi analisis, transkrip wawancara, teknik analisis, dan codebook. Matriks ini mencegah pengumpulan data yang banyak tetapi tidak relevan. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Ilustrasi gambar gas juga menunjukkan bahwa kualitas instrumen memengaruhi temuan. Distraktor yang tidak berasal dari pemikiran peserta didik dapat menghasilkan pilihan acak, sedangkan pertanyaan yang terlalu menuntut matematika dapat mengukur keterampilan hitung alih-alih konsepsi. Peneliti perlu menilai kesesuaian bukti dengan konstruk yang dimaksud.

Kriteria pemeriksaan sebelum analisis:

- Definisikan tradisi analisis dalam istilah operasional.
- Tentukan bukti utama dari jawaban terbuka dan bukti pendukung dari catatan observasi.
- Periksa dependabilitas sebelum membuat kategori atau model.
- Catat potensi menghitung kode sebagai makna beserta strategi mitigasinya.
- Sajikan tipologi konsepsi bersama batas interpretasi dan ketidakpastian.

Peneliti sebaiknya memisahkan keputusan apriori dari keputusan eksploratori. Pemisahan tersebut membantu mengendalikan coding tanpa pertanyaan dan mencegah menghapus kasus negatif. Selanjutnya, kredibilitas dan konfirmabilitas dilaporkan bersama ketidakpastian, sehingga pembaca dapat menilai kekuatan temuan tanpa bergantung pada kesimpulan penulis semata.

Dengan demikian, analisis model mental berfungsi menghubungkan keputusan desain dengan interpretasi. Produk yang diharapkan bukan hanya codebook, tetapi juga penjelasan mengenai bagaimana produk tersebut diperoleh dan dalam kondisi apa ia dapat digunakan. Penelitian miskonsepsi yang baik selalu menyisakan ruang bagi penjelasan alternatif dan verifikasi lanjutan.

7.14.10 Analisis Gambar dan Representasi

Ketelitian penelitian pada bagian analisis gambar dan representasi sangat menentukan apakah pola jawaban dapat ditafsirkan sebagai bukti konseptual. Fokus ini berada dalam proses sistematis untuk mengubah tuturan, tulisan, gambar, observasi, dan artefak menjadi

tema, model mental, mekanisme, atau tipologi konsepsi. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Miles et al. (2014) serta Braun dan Clarke (2006) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Dokumentasi metodologis hendaknya memuat siapa yang melakukan jawaban terbuka, kapan data diperoleh, bagaimana kedalaman interpretasi dikendalikan, dan atas dasar apa tema disusun. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Contoh pada kajian penjelasan redoks memperlihatkan pentingnya pemisahan antara ketidaktahuan, kesalahan prosedural, dan miskonsepsi. Seorang peserta dapat gagal karena lupa rumus, sedangkan peserta lain memilih jawaban yang sama karena menggunakan model mental yang tidak ilmiah. Perbedaan ini baru terlihat ketika jawaban terbuka dibaca bersama catatan observasi dan bukti tentang alasan.

Langkah operasional yang disarankan:

- Definisikan unit makna dalam istilah operasional.
- Tentukan bukti utama dari gambar partikel dan bukti pendukung dari video kelas.
- Periksa konfirmabilitas sebelum membuat kategori atau model.
- Catat potensi menghapus kasus negatif beserta strategi mitigasinya.
- Sajikan model mekanisme bersama batas interpretasi dan ketidakpastian.

Standar metodologis menuntut koherensi antara konstruk, prosedur, analisis, dan klaim. Dalam konteks ini, dependabilitas tidak dapat digantikan oleh indeks tunggal. Demikian pula, reflektivitas harus dinilai dengan bukti yang sesuai. Miles et al. (2014) dan Braun dan Clarke (2006) memberikan dasar untuk pelaporan yang lengkap dan dapat ditelaah.

Inti metodologis subbagian ini adalah keterlacakan. Pembaca harus dapat mengikuti perjalanan dari pertanyaan, data, pengodean, analisis, sampai tema. Ketika jejak itu jelas, temuan dapat digunakan untuk memperbaiki pembelajaran tanpa mengubah diagnosis konseptual menjadi label bagi peserta didik.

7.14.11 Analisis Interaksi Kelas

Peneliti sering memusatkan perhatian pada teknik, padahal analisis interaksi kelas justru mengatur batas inferensi yang dapat dibuat. Fokus ini berada dalam proses sistematis untuk mengubah tuturan, tulisan, gambar, observasi, dan artefak menjadi tema, model mental, mekanisme, atau tipologi konsepsi. Peneliti perlu menyatakan sejak awal apa yang dimaksud,

indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Saldaña (2021) serta Lincoln dan Guba (1985) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Rancangan yang dapat diaudit mencantumkan alasan pemilihan tingkat induksi, prosedur penanganan peran teori, dan kriteria keputusan untuk tradisi analisis. Tanpa penjelasan itu, pembaca sulit menilai apakah gambar partikel benar-benar mendukung klaim. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Kasus argumentasi laju dapat digunakan untuk memperlihatkan hubungan antara pola individual dan pola kelas. Sebagian miskonsepsi mungkin hanya dimiliki beberapa peserta tetapi sangat kuat, sedangkan pola lain tersebar luas karena bersumber dari analogi atau buku teks yang sama. Analisis perlu menjaga kedua tingkat tersebut.

Keputusan yang harus didokumentasikan:

- Definisikan tingkat induksi dalam istilah operasional.
- Tentukan bukti utama dari catatan observasi dan bukti pendukung dari artefak tugas.
- Periksa reflektivitas sebelum membuat kategori atau model.
- Catat potensi kutipan tanpa analisis beserta strategi mitigasinya.
- Sajikan matriks kasus bersama batas interpretasi dan ketidakpastian.

Pemeriksaan kualitas dapat dilakukan melalui audit terhadap konfirmabilitas, analisis kasus yang bertentangan, dan evaluasi kecukupan kutipan. Apabila ditemukan menghitung kode sebagai makna, peneliti perlu menurunkan kekuatan klaim atau menambah bukti. Transparansi seperti ini meningkatkan kegunaan temuan bagi penelitian lanjutan dan praktik kelas.

Kesimpulannya, analisis interaksi kelas perlu dirancang sebagai proses argumentatif. Data tidak berbicara sendiri; maknanya bergantung pada definisi konstruk, kualitas prosedur, dan konteks. Oleh sebab itu, klaim akhir sebaiknya menggunakan bahasa yang sesuai dengan kekuatan bukti, misalnya “menunjukkan pola”, “mengindikasikan”, atau “mendukung penjelasan”.

7.14.12 Kasus Negatif dan Disconfirming Evidence

Kajian yang kuat memperlakukan kasus negatif dan disconfirming evidence sebagai bagian dari rantai argumentasi, bukan sebagai unsur pelengkap. Fokus ini berada dalam proses sistematis untuk mengubah tuturan, tulisan, gambar, observasi, dan artefak menjadi tema, model mental, mekanisme, atau tipologi konsepsi. Peneliti perlu menyatakan sejak awal apa

yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Lincoln dan Guba (1985) serta O'Brien et al. (2014) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Pada tahap perencanaan, kedalaman interpretasi dan strategi penutupan analisis harus dibedakan secara eksplisit. Pada tahap pelaksanaan, konsistensi catatan observasi perlu diperiksa; pada tahap interpretasi, reflektivitas menjadi dasar pembatasan klaim. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Dalam kelas yang mempelajari diskusi elektrolit, data agregat mungkin menunjukkan mayoritas jawaban salah. Akan tetapi, wawancara dapat memperlihatkan beberapa mekanisme berbeda: bahasa sehari-hari, generalisasi aturan, salah membaca simbol, atau model partikel yang tidak tepat. Karena itu, satu persentase tidak boleh menggantikan analisis struktur alasan.

Pertanyaan audit metodologis:

- Definisikan kedalaman interpretasi dalam istilah operasional.
- Tentukan bukti utama dari video kelas dan bukti pendukung dari transkrip wawancara.
- Periksa kecukupan kutipan sebelum membuat kategori atau model.
- Catat potensi klaim saturasi otomatis beserta strategi mitigasinya.
- Sajikan narasi analitik bersama batas interpretasi dan ketidakpastian.

Kriteria keberhasilan tidak cukup berupa banyaknya data. Yang lebih penting adalah apakah reflektivitas dapat ditelusuri, apakah *audit trail* diperiksa, dan apakah dependabilitas dijelaskan. Peneliti harus menghindari menghapus kasus negatif; jika risiko itu tidak dapat dihilangkan, dampaknya perlu dianalisis dan dicantumkan sebagai batas penelitian.

Pada akhirnya, keputusan tentang kasus negatif dan disconfirming evidence menentukan apakah penelitian menghasilkan daftar jawaban salah atau pemahaman yang dapat ditindaklanjuti mengenai cara peserta didik berpikir. Sasaran utamanya ialah bukti yang cukup untuk merancang diagnosis, pembelajaran, dan penelitian lanjutan secara bertanggung jawab.

7.14.13 Kredibilitas serta Audit Trail

Subbagian ini menempatkan kredibilitas serta *audit trail* sebagai simpul penting dalam analisis data kualitatif. Fokus ini berada dalam proses sistematis untuk mengubah tuturan, tulisan, gambar, observasi, dan artefak menjadi tema, model mental, mekanisme, atau tipologi konsepsi. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat

diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Nowell et al. (2017) serta Miles et al. (2014) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Secara operasional, peneliti perlu menetapkan peran teori, menjelaskan tradisi analisis, dan mendokumentasikan tingkat induksi. Keputusan tersebut kemudian diterjemahkan ke dalam pilihan video kelas, transkrip wawancara, serta gambar partikel. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Sebagai ilustrasi, pada topik model mental ikatan, peserta didik dapat memberikan jawaban benar pada tingkat simbolik tetapi alasan submikroskopiknya keliru. Peneliti tidak cukup menghitung benar-salah; ia perlu membandingkan video kelas dengan transkrip wawancara, memeriksa tingkat keyakinan, dan menguji konsistensi pada konteks lain. Hasil analisis mungkin menunjukkan bahwa kesalahan hanya muncul pada satu representasi, atau justru merupakan hubungan konseptual yang stabil.

Bukti minimum yang perlu tersedia:

- Definisikan peran teori dalam istilah operasional.
- Tentukan bukti utama dari artefak tugas dan bukti pendukung dari jawaban terbuka.
- Periksa audit trail sebelum membuat kategori atau model.
- Catat potensi coding tanpa pertanyaan beserta strategi mitigasinya.
- Sajikan codebook bersama batas interpretasi dan ketidakpastian.

Mutu inferensi terutama ditentukan oleh kecukupan kutipan, kredibilitas, dan konfirmabilitas. Ancaman utama ialah kutipan tanpa analisis serta tema sebagai topik semata. Karena itu, laporan perlu menyajikan bukti yang mendukung sekaligus bukti yang membatasi interpretasi. Prinsip ini konsisten dengan penekanan Nowell et al. (2017) dan Miles et al. (2014) bahwa validitas melekat pada interpretasi dan penggunaan hasil, bukan sekadar pada nama instrumen atau desain.

Dengan demikian, kredibilitas serta audit trail berfungsi menghubungkan keputusan desain dengan interpretasi. Produk yang diharapkan bukan hanya matriks kasus, tetapi juga penjelasan mengenai bagaimana produk tersebut diperoleh dan dalam kondisi apa ia dapat digunakan. Penelitian miskonsepsi yang baik selalu menyisakan ruang bagi penjelasan alternatif dan verifikasi lanjutan.

7.14.14 Penyajian Kutipan dan Interpretasi

Pembahasan mengenai penyajian kutipan dan interpretasi diperlukan agar analisis data kualitatif tidak berhenti pada prosedur administratif. Fokus ini berada dalam proses sistematis untuk mengubah tuturan, tulisan, gambar, observasi, dan artefak menjadi tema, model mental, mekanisme, atau tipologi konsepsi. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. O'Brien et al. (2014) serta Nowell et al. (2017) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Penerapan praktis dapat disusun dalam urutan: mendefinisikan strategi penutupan analisis, mengoperasionalkan unit makna, mengumpulkan artefak tugas, memeriksa *audit trail*, lalu melaporkan narasi analitik. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Misalnya, pada penalaran asam kuat, perubahan skor sesudah pembelajaran belum otomatis menandakan perubahan konsep. Peserta didik mungkin mempelajari pola jawaban tes. Bukti yang lebih meyakinkan mencakup transfer ke situasi baru, penjelasan kausal, representasi partikel, dan retensi setelah jeda waktu.

Kriteria pemeriksaan sebelum analisis:

- Definisikan strategi penutupan analisis dalam istilah operasional.
- Tentukan bukti utama dari transkrip wawancara dan bukti pendukung dari gambar partikel.
- Periksa kredibilitas sebelum membuat kategori atau model.
- Catat potensi tema sebagai topik semata beserta strategi mitigasinya.
- Sajikan tema bersama batas interpretasi dan ketidakpastian.

Temuan yang baik tidak selalu berarti hasil yang besar atau sesuai harapan. Temuan bernilai ketika *audit trail* terjaga, dependabilitas dijelaskan, dan ketidakpastian disajikan. Kerangka O'Brien et al. (2014) serta Nowell et al. (2017) menolong peneliti membedakan kekuatan bukti dari daya tarik narasi.

Inti metodologis subbagian ini adalah keterlacakan. Pembaca harus dapat mengikuti perjalanan dari pertanyaan, data, pengodean, analisis, sampai narasi analitik. Ketika jejak itu jelas, temuan dapat digunakan untuk memperbaiki pembelajaran tanpa mengubah diagnosis konseptual menjadi label bagi peserta didik.

7.14.15 Sintesis Analisis Kualitatif Miskonsepsi

Dalam metodologi penelitian miskonsepsi kimia, sintesis analisis kualitatif miskonsepsi menentukan mutu hubungan antara data dan kesimpulan. Fokus ini berada dalam proses sistematis untuk mengubah tuturan, tulisan, gambar, observasi, dan artefak menjadi tema, model mental, mekanisme, atau tipologi konsepsi. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Braun dan Clarke (2006) serta Braun dan Clarke (2021) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Prosedur minimal mencakup definisi tradisi analisis, alasan memilih transkrip wawancara, aturan mengubah respons menjadi kategori, serta mekanisme pemeriksaan kredibilitas. Semua penyimpangan dari rencana awal perlu dicatat. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Pada penelitian mengenai analogi kalor, peneliti dapat menyusun beberapa tugas ekuivalen yang berubah konteks, representasi, atau tuntutan penalarannya. Jika pola respons tetap berulang dan disertai keyakinan tinggi, terdapat dasar yang lebih kuat untuk menyebut indikasi miskonsepsi. Jika pola berubah-ubah, interpretasi yang lebih hati-hati adalah pengetahuan parsial atau ketergantungan konteks.

Langkah operasional yang disarankan:

- Definisikan tradisi analisis dalam istilah operasional.
- Tentukan bukti utama dari jawaban terbuka dan bukti pendukung dari catatan observasi.
- Periksa dependabilitas sebelum membuat kategori atau model.
- Catat potensi menghitung kode sebagai makna beserta strategi mitigasinya.
- Sajikan tipologi konsepsi bersama batas interpretasi dan ketidakpastian.

Pemeriksaan mutu hendaknya dilakukan selama penelitian, bukan hanya pada akhir. Catatan tentang kredibilitas, keputusan terkait konfirmabilitas, dan kejadian yang berpotensi menimbulkan coding tanpa pertanyaan perlu dipelihara dalam *audit trail*. Hal ini memungkinkan koreksi dini dan pelaporan yang jujur.

Kesimpulannya, sintesis analisis kualitatif miskonsepsi perlu dirancang sebagai proses argumentatif. Data tidak berbicara sendiri; maknanya bergantung pada definisi konstruk, kualitas prosedur, dan konteks. Oleh sebab itu, klaim akhir sebaiknya menggunakan bahasa yang sesuai dengan kekuatan bukti, misalnya “menunjukkan pola”, “mengindikasikan”, atau “mendukung penjelasan”.

7.15 Triangulasi Data

Subbab ini membahas penggunaan perbandingan terencana antar-sumber, metode, peneliti, teori, atau waktu untuk menguji kekuatan dan batas interpretasi miskonsepsi. Fokus utamanya adalah menjaga keselarasan antara pertanyaan penelitian, bentuk bukti, teknik analisis, dan kekuatan klaim. Dalam konteks miskonsepsi kimia, keputusan metodologis harus mampu membedakan pola konseptual dari respons yang timbul karena keterbatasan bahasa, kesalahan hitung, atau kondisi administrasi. Pembahasan berikut dipecah menjadi lima belas bagian agar setiap keputusan dapat ditelaah secara operasional.

7.15.1 Pengertian dan Tujuan Triangulasi

Keputusan tentang pengertian dan tujuan triangulasi harus dibuat sebelum data utama dikumpulkan, bukan ditambahkan setelah hasil diketahui. Fokus ini berada dalam penggunaan perbandingan terencana antar-sumber, metode, peneliti, teori, atau waktu untuk menguji kekuatan dan batas interpretasi miskonsepsi. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Carter et al. (2014) serta AERA, APA, dan NCME (2014) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Tahap kerja dimulai dengan penegasan jenis sumber. Setelah itu, peneliti menyusun aturan untuk penanganan perbedaan dan menentukan bagaimana waktu integrasi akan diperiksa. Bukti sebaiknya berasal dari lebih dari satu bentuk, misalnya wawancara dan observasi. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Pada topik diskusi kelompok mengubah respons individu, peneliti juga perlu memperhatikan bahasa. Istilah yang sama dapat diberi makna berbeda oleh peserta didik. Klarifikasi melalui pertanyaan tindak lanjut, gambar, atau contoh tandingan membantu membedakan masalah terminologi dari kesalahan konseptual.

Keputusan yang harus didokumentasikan:

- Definisikan jenis sumber dalam istilah operasional.
- Tentukan bukti utama dari gambar dan bukti pendukung dari dokumen.
- Periksa transparansi protokol sebelum membuat kategori atau model.
- Catat potensi triangulasi setelah hasil diketahui beserta strategi mitigasinya.
- Sajikan status bukti bersama batas interpretasi dan ketidakpastian.

Kualitas penelitian meningkat ketika peneliti menunjukkan rantai bukti dari wawancara menuju peta divergensi, menyatakan asumsi, dan memeriksa alternatif penjelasan.

Sebaliknya, memaksa konsensus dapat menghasilkan klaim yang tampak tegas tetapi tidak stabil. Rekomendasi praktis harus proporsional dengan bukti.

Pada akhirnya, keputusan tentang pengertian dan tujuan triangulasi menentukan apakah penelitian menghasilkan daftar jawaban salah atau pemahaman yang dapat ditindaklanjuti mengenai cara peserta didik berpikir. Sasaran utamanya ialah bukti yang cukup untuk merancang diagnosis, pembelajaran, dan penelitian lanjutan secara bertanggung jawab.

7.15.2 Triangulasi Sumber

Triangulasi sumber perlu dipahami sebagai keputusan metodologis yang mempunyai konsekuensi pada desain, analisis, dan pelaporan. Fokus ini berada dalam penggunaan perbandingan terencana antar-sumber, metode, peneliti, teori, atau waktu untuk menguji kekuatan dan batas interpretasi miskonsepsi. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Lincoln dan Guba (1985) serta Carter et al. (2014) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Peneliti sebaiknya menyiapkan matriks yang menghubungkan pertanyaan penelitian dengan aturan konvergensi, gambar, teknik analisis, dan narasi penjelas. Matriks ini mencegah pengumpulan data yang banyak tetapi tidak relevan. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Ilustrasi hasil *pretest* kembali pada tes retensi juga menunjukkan bahwa kualitas instrumen memengaruhi temuan. Distraktor yang tidak berasal dari pemikiran peserta didik dapat menghasilkan pilihan acak, sedangkan pertanyaan yang terlalu menuntut matematika dapat mengukur keterampilan hitung alih-alih konsepsi. Peneliti perlu menilai kesesuaian bukti dengan konstruk yang dimaksud.

Pertanyaan audit metodologis:

- Definisikan aturan konvergensi dalam istilah operasional.
- Tentukan bukti utama dari observasi dan bukti pendukung dari pengukuran berulang.
- Periksa analisis divergensi sebelum membuat kategori atau model.
- Catat potensi mengabaikan ketergantungan sumber beserta strategi mitigasinya.
- Sajikan revisi kategori bersama batas interpretasi dan ketidakpastian.

Peneliti sebaiknya memisahkan keputusan apriori dari keputusan eksploratori. Pemisahan tersebut membantu mengendalikan menghitung jumlah metode dan mencegah memakai istilah tanpa prosedur. Selanjutnya, transparansi protokol dan jejak keputusan dilaporkan

bersama ketidakpastian, sehingga pembaca dapat menilai kekuatan temuan tanpa bergantung pada kesimpulan penulis semata.

Dengan demikian, triangulasi sumber berfungsi menghubungkan keputusan desain dengan interpretasi. Produk yang diharapkan bukan hanya narasi jelas, tetapi juga penjelasan mengenai bagaimana produk tersebut diperoleh dan dalam kondisi apa ia dapat digunakan. Penelitian miskonsepsi yang baik selalu menyisakan ruang bagi penjelasan alternatif dan verifikasi lanjutan.

7.15.3 Triangulasi Metode

Ketelitian penelitian pada bagian triangulasi metode sangat menentukan apakah pola jawaban dapat ditafsirkan sebagai bukti konseptual. Fokus ini berada dalam penggunaan perbandingan terencana antar-sumber, metode, peneliti, teori, atau waktu untuk menguji kekuatan dan batas interpretasi miskonsepsi. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Fetters et al. (2013) serta AERA, APA, dan NCME (2014) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Dokumentasi metodologis hendaknya memuat siapa yang melakukan observasi, kapan data diperoleh, bagaimana waktu integrasi dikendalikan, dan atas dasar apa status bukti disusun. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Contoh pada kajian perbedaan guru dan siswa memperlihatkan pentingnya pemisahan antara ketidaktahuan, kesalahan prosedural, dan miskonsepsi. Seorang peserta dapat gagal karena lupa rumus, sedangkan peserta lain memilih jawaban yang sama karena menggunakan model mental yang tidak ilmiah. Perbedaan ini baru terlihat ketika observasi dibaca bersama pengukuran berulang dan bukti tentang alasan.

Bukti minimum yang perlu tersedia:

- Definisikan penanganan perbedaan dalam istilah operasional.
- Tentukan bukti utama dari dokumen dan bukti pendukung dari tes.
- Periksa jejak keputusan sebelum membuat kategori atau model.
- Catat potensi memakai istilah tanpa prosedur beserta strategi mitigasinya.
- Sajikan batas interpretasi bersama batas interpretasi dan ketidakpastian.

Standar metodologis menuntut koherensi antara konstruk, prosedur, analisis, dan klaim. Dalam konteks ini, analisis divergensi tidak dapat digantikan oleh indeks tunggal. Demikian

pula, proporsionalitas klaim harus dinilai dengan bukti yang sesuai. Feters et al. (2013) dan AERA, APA, dan NCME (2014) memberikan dasar untuk pelaporan yang lengkap dan dapat ditelaah.

Inti metodologis subbagian ini adalah keterlacakan. Pembaca harus dapat mengikuti perjalanan dari pertanyaan, data, pengodean, analisis, sampai status bukti. Ketika jejak itu jelas, temuan dapat digunakan untuk memperbaiki pembelajaran tanpa mengubah diagnosis konseptual menjadi label bagi peserta didik.

7.15.4 Triangulasi Peneliti

Peneliti sering memusatkan perhatian pada teknik, padahal triangulasi peneliti justru mengatur batas inferensi yang dapat dibuat. Fokus ini berada dalam penggunaan perbandingan terencana antar-sumber, metode, peneliti, teori, atau waktu untuk menguji kekuatan dan batas interpretasi miskonsepsi. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. AERA, APA, dan NCME (2014) serta Carter et al. (2014) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Rancangan yang dapat diaudit mencantumkan alasan pemilihan tim penafsir, prosedur penanganan objek yang ditriangulasi, dan kriteria keputusan untuk aturan konvergensi. Tanpa penjelasan itu, pembaca sulit menilai apakah dokumen benar-benar mendukung klaim. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Kasus data daring dan kertas dapat digunakan untuk memperlihatkan hubungan antara pola individual dan pola kelas. Sebagian miskonsepsi mungkin hanya dimiliki beberapa peserta tetapi sangat kuat, sedangkan pola lain tersebar luas karena bersumber dari analogi atau buku teks yang sama. Analisis perlu menjaga kedua tingkat tersebut.

Kriteria pemeriksaan sebelum analisis:

- Definisikan tim penafsir dalam istilah operasional.
- Tentukan bukti utama dari pengukuran berulang dan bukti pendukung dari wawancara.
- Periksa proporsionalitas klaim sebelum membuat kategori atau model.
- Catat potensi menganggap kesamaan selalu benar beserta strategi mitigasinya.
- Sajikan matriks konvergensi bersama batas interpretasi dan ketidakpastian.

Pemeriksaan kualitas dapat dilakukan melalui audit terhadap jejak keputusan, analisis kasus yang bertentangan, dan evaluasi komplementaritas. Apabila ditemukan mengabaikan ketergantungan sumber, peneliti perlu menurunkan kekuatan klaim atau menambah bukti.

Transparansi seperti ini meningkatkan kegunaan temuan bagi penelitian lanjutan dan praktik kelas.

Kesimpulannya, triangulasi peneliti perlu dirancang sebagai proses argumentatif. Data tidak berbicara sendiri; maknanya bergantung pada definisi konstruk, kualitas prosedur, dan konteks. Oleh sebab itu, klaim akhir sebaiknya menggunakan bahasa yang sesuai dengan kekuatan bukti, misalnya “menunjukkan pola”, “mengindikasikan”, atau “mendukung penjelasan”.

7.15.5 Triangulasi Teori

Kajian yang kuat memperlakukan triangulasi teori sebagai bagian dari rantai argumentasi, bukan sebagai unsur pelengkap. Fokus ini berada dalam penggunaan perbandingan terencana antar-sumber, metode, peneliti, teori, atau waktu untuk menguji kekuatan dan batas interpretasi miskonsepsi. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Kaltakci Gurel et al. (2015) serta AERA, APA, dan NCME (2014) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Pada tahap perencanaan, waktu integrasi dan jenis sumber harus dibedakan secara eksplisit. Pada tahap pelaksanaan, konsistensi pengukuran berulang perlu diperiksa; pada tahap interpretasi, proporsionalitas klaim menjadi dasar pembatasan klaim. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Dalam kelas yang mempelajari dua peneliti berbeda kode, data agregat mungkin menunjukkan mayoritas jawaban salah. Akan tetapi, wawancara dapat memperlihatkan beberapa mekanisme berbeda: bahasa sehari-hari, generalisasi aturan, salah membaca simbol, atau model partikel yang tidak tepat. Karena itu, satu persentase tidak boleh menggantikan analisis struktur alasan.

Langkah operasional yang disarankan:

- Definisikan waktu integrasi dalam istilah operasional.
- Tentukan bukti utama dari tes dan bukti pendukung dari gambar.
- Periksa komplementaritas sebelum membuat kategori atau model.
- Catat potensi memaksa konsensus beserta strategi mitigasinya.
- Sajikan peta divergensi bersama batas interpretasi dan ketidakpastian.

Kriteria keberhasilan tidak cukup berupa banyaknya data. Yang lebih penting adalah apakah proporsionalitas klaim dapat ditelusuri, apakah independensi relatif sumber diperiksa, dan

apakah analisis divergensi dijelaskan. Peneliti harus menghindari memakai istilah tanpa prosedur; jika risiko itu tidak dapat dihilangkan, dampaknya perlu dianalisis dan dicantumkan sebagai batas penelitian.

Pada akhirnya, keputusan tentang triangulasi teori menentukan apakah penelitian menghasilkan daftar jawaban salah atau pemahaman yang dapat ditindaklanjuti mengenai cara peserta didik berpikir. Sasaran utamanya ialah bukti yang cukup untuk merancang diagnosis, pembelajaran, dan penelitian lanjutan secara bertanggung jawab.

7.15.6 Triangulasi Waktu

Subbagian ini menempatkan triangulasi waktu sebagai simpul penting dalam triangulasi data. Fokus ini berada dalam penggunaan perbandingan terencana antar-sumber, metode, peneliti, teori, atau waktu untuk menguji kekuatan dan batas interpretasi miskonsepsi. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Farmer et al. (2006) serta Carter et al. (2014) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Secara operasional, peneliti perlu menetapkan objek yang ditriangulasi, menjelaskan aturan konvergensi, dan mendokumentasikan tim penafsir. Keputusan tersebut kemudian diterjemahkan ke dalam pilihan tes, gambar, serta dokumen. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Sebagai ilustrasi, pada topik jawaban benar tetapi gambar salah, peserta didik dapat memberikan jawaban benar pada tingkat simbolik tetapi alasan submikroskopiknya keliru. Peneliti tidak cukup menghitung benar-salah; ia perlu membandingkan tes dengan gambar, memeriksa tingkat keyakinan, dan menguji konsistensi pada konteks lain. Hasil analisis mungkin menunjukkan bahwa kesalahan hanya muncul pada satu representasi, atau justru merupakan hubungan konseptual yang stabil.

Keputusan yang harus didokumentasikan:

- Definisikan objek yang ditriangulasi dalam istilah operasional.
- Tentukan bukti utama dari wawancara dan bukti pendukung dari observasi.
- Periksa independensi relatif sumber sebelum membuat kategori atau model.
- Catat potensi menghitung jumlah metode beserta strategi mitigasinya.
- Sajikan narasi penjelas bersama batas interpretasi dan ketidakpastian.

Mutu inferensi terutama ditentukan oleh komplementaritas, transparansi protokol, dan jejak keputusan. Ancaman utama ialah menganggap kesamaan selalu benar serta triangulasi

setelah hasil diketahui. Karena itu, laporan perlu menyajikan bukti yang mendukung sekaligus bukti yang membatasi interpretasi. Prinsip ini konsisten dengan penekanan Farmer et al. (2006) dan Carter et al. (2014) bahwa validitas melekat pada interpretasi dan penggunaan hasil, bukan sekadar pada nama instrumen atau desain.

Dengan demikian, triangulasi waktu berfungsi menghubungkan keputusan desain dengan interpretasi. Produk yang diharapkan bukan hanya matriks konvergensi, tetapi juga penjelasan mengenai bagaimana produk tersebut diperoleh dan dalam kondisi apa ia dapat digunakan. Penelitian miskonsepsi yang baik selalu menyisakan ruang bagi penjelasan alternatif dan verifikasi lanjutan.

7.15.7 Penyusunan Matriks Konvergensi

Pembahasan mengenai penyusunan matriks konvergensi diperlukan agar triangulasi data tidak berhenti pada prosedur administratif. Fokus ini berada dalam penggunaan perbandingan terencana antar-sumber, metode, peneliti, teori, atau waktu untuk menguji kekuatan dan batas interpretasi miskonsepsi. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Carter et al. (2014) serta AERA, APA, dan NCME (2014) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Penerapan praktis dapat disusun dalam urutan: mendefinisikan jenis sumber, mengoperasionalkan penanganan perbedaan, mengumpulkan wawancara, memeriksa independensi relatif sumber, lalu melaporkan peta divergensi. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Misalnya, pada wawancara berbeda dari tes, perubahan skor sesudah pembelajaran belum otomatis menandakan perubahan konsep. Peserta didik mungkin mempelajari pola jawaban tes. Bukti yang lebih meyakinkan mencakup transfer ke situasi baru, penjelasan kausal, representasi partikel, dan retensi setelah jeda waktu.

Pertanyaan audit metodologis:

- Definisikan jenis sumber dalam istilah operasional.
- Tentukan bukti utama dari gambar dan bukti pendukung dari dokumen.
- Periksa transparansi protokol sebelum membuat kategori atau model.
- Catat potensi triangulasi setelah hasil diketahui beserta strategi mitigasinya.
- Sajikan status bukti bersama batas interpretasi dan ketidakpastian.

Temuan yang baik tidak selalu berarti hasil yang besar atau sesuai harapan. Temuan bernilai ketika independensi relatif sumber terjaga, analisis divergensi dijelaskan, dan ketidakpastian disajikan. Kerangka Carter et al. (2014) serta AERA, APA, dan NCME (2014) menolong peneliti membedakan kekuatan bukti dari daya tarik narasi.

Inti metodologis subbagian ini adalah keterlacakan. Pembaca harus dapat mengikuti perjalanan dari pertanyaan, data, pengodean, analisis, sampai peta divergensi. Ketika jejak itu jelas, temuan dapat digunakan untuk memperbaiki pembelajaran tanpa mengubah diagnosis konseptual menjadi label bagi peserta didik.

7.15.8 Analisis Konvergensi

Dalam metodologi penelitian miskonsepsi kimia, analisis konvergensi menentukan mutu hubungan antara data dan kesimpulan. Fokus ini berada dalam penggunaan perbandingan terencana antar-sumber, metode, peneliti, teori, atau waktu untuk menguji kekuatan dan batas interpretasi miskonsepsi. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Lincoln dan Guba (1985) serta Carter et al. (2014) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Prosedur minimal mencakup definisi aturan konvergensi, alasan memilih gambar, aturan mengubah respons menjadi kategori, serta mekanisme pemeriksaan transparansi protokol. Semua penyimpangan dari rencana awal perlu dicatat. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Pada penelitian mengenai keyakinan tinggi dengan alasan tidak ilmiah, peneliti dapat menyusun beberapa tugas ekuivalen yang berubah konteks, representasi, atau tuntutan penalarannya. Jika pola respons tetap berulang dan disertai keyakinan tinggi, terdapat dasar yang lebih kuat untuk menyebut indikasi miskonsepsi. Jika pola berubah-ubah, interpretasi yang lebih hati-hati adalah pengetahuan parsial atau ketergantungan konteks.

Bukti minimum yang perlu tersedia:

- Definisikan aturan konvergensi dalam istilah operasional.
- Tentukan bukti utama dari observasi dan bukti pendukung dari pengukuran berulang.
- Periksa analisis divergensi sebelum membuat kategori atau model.
- Catat potensi mengabaikan ketergantungan sumber beserta strategi mitigasinya.
- Sajikan revisi kategori bersama batas interpretasi dan ketidakpastian.

Pemeriksaan mutu hendaknya dilakukan selama penelitian, bukan hanya pada akhir. Catatan tentang transparansi protokol, keputusan terkait jejak keputusan, dan kejadian yang berpotensi menimbulkan menghitung jumlah metode perlu dipelihara dalam *audit trail*. Hal ini memungkinkan koreksi dini dan pelaporan yang jujur.

Kesimpulannya, analisis konvergensi perlu dirancang sebagai proses argumentatif. Data tidak berbicara sendiri; maknanya bergantung pada definisi konstruk, kualitas prosedur, dan konteks. Oleh sebab itu, klaim akhir sebaiknya menggunakan bahasa yang sesuai dengan kekuatan bukti, misalnya “menunjukkan pola”, “mengindikasikan”, atau “mendukung penjelasan”.

7.15.9 Analisis Komplementaritas

Keputusan tentang analisis komplementaritas harus dibuat sebelum data utama dikumpulkan, bukan ditambahkan setelah hasil diketahui. Fokus ini berada dalam penggunaan perbandingan terencana antar-sumber, metode, peneliti, teori, atau waktu untuk menguji kekuatan dan batas interpretasi miskonsepsi. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Fetters et al. (2013) serta AERA, APA, dan NCME (2014) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Tahap kerja dimulai dengan penegasan penanganan perbedaan. Setelah itu, peneliti menyusun aturan untuk waktu integrasi dan menentukan bagaimana jenis sumber akan diperiksa. Bukti sebaiknya berasal dari lebih dari satu bentuk, misalnya observasi dan pengukuran berulang. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Pada topik diskusi kelompok mengubah respons individu, peneliti juga perlu memperhatikan bahasa. Istilah yang sama dapat diberi makna berbeda oleh peserta didik. Klarifikasi melalui pertanyaan tindak lanjut, gambar, atau contoh tandingan membantu membedakan masalah terminologi dari kesalahan konseptual.

Kriteria pemeriksaan sebelum analisis:

- Definisikan penanganan perbedaan dalam istilah operasional.
- Tentukan bukti utama dari dokumen dan bukti pendukung dari tes.
- Periksa jejak keputusan sebelum membuat kategori atau model.
- Catat potensi memakai istilah tanpa prosedur beserta strategi mitigasinya.
- Sajikan batas interpretasi bersama batas interpretasi dan ketidakpastian.

Kualitas penelitian meningkat ketika peneliti menunjukkan rantai bukti dari observasi menuju status bukti, menyatakan asumsi, dan memeriksa alternatif penjelasan. Sebaliknya, triangulasi setelah hasil diketahui dapat menghasilkan klaim yang tampak tegas tetapi tidak stabil. Rekomendasi praktis harus proporsional dengan bukti.

Pada akhirnya, keputusan tentang analisis komplementaritas menentukan apakah penelitian menghasilkan daftar jawaban salah atau pemahaman yang dapat ditindaklanjuti mengenai cara peserta didik berpikir. Sasaran utamanya ialah bukti yang cukup untuk merancang diagnosis, pembelajaran, dan penelitian lanjutan secara bertanggung jawab.

7.15.10 Analisis Divergensi

Analisis divergensi perlu dipahami sebagai keputusan metodologis yang mempunyai konsekuensi pada desain, analisis, dan pelaporan. Fokus ini berada dalam penggunaan perbandingan terencana antar-sumber, metode, peneliti, teori, atau waktu untuk menguji kekuatan dan batas interpretasi miskonsepsi. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. AERA, APA, dan NCME (2014) serta Carter et al. (2014) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Peneliti sebaiknya menyiapkan matriks yang menghubungkan pertanyaan penelitian dengan tim penafsir, dokumen, teknik analisis, dan revisi kategori. Matriks ini mencegah pengumpulan data yang banyak tetapi tidak relevan. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Ilustrasi hasil *pretest* kembali pada tes retensi juga menunjukkan bahwa kualitas instrumen memengaruhi temuan. Distraktor yang tidak berasal dari pemikiran peserta didik dapat menghasilkan pilihan acak, sedangkan pertanyaan yang terlalu menuntut matematika dapat mengukur keterampilan hitung alih-alih konsepsi. Peneliti perlu menilai kesesuaian bukti dengan konstruk yang dimaksud.

Langkah operasional yang disarankan:

- Definisikan tim penafsir dalam istilah operasional.
- Tentukan bukti utama dari pengukuran berulang dan bukti pendukung dari wawancara.
- Periksa proporsionalitas klaim sebelum membuat kategori atau model.
- Catat potensi menganggap kesamaan selalu benar beserta strategi mitigasinya.
- Sajikan matriks konvergensi bersama batas interpretasi dan ketidakpastian.

Peneliti sebaiknya memisahkan keputusan apriori dari keputusan eksploratori. Pemisahan tersebut membantu mengendalikan mengabaikan ketergantungan sumber dan mencegah memaksa konsensus. Selanjutnya, jejak keputusan dan komplementaritas dilaporkan bersama ketidakpastian, sehingga pembaca dapat menilai kekuatan temuan tanpa bergantung pada kesimpulan penulis semata.

Dengan demikian, analisis divergensi berfungsi menghubungkan keputusan desain dengan interpretasi. Produk yang diharapkan bukan hanya revisi kategori, tetapi juga penjelasan mengenai bagaimana produk tersebut diperoleh dan dalam kondisi apa ia dapat digunakan. Penelitian miskonsepsi yang baik selalu menyisakan ruang bagi penjelasan alternatif dan verifikasi lanjutan.

7.15.11 Protokol Penyelesaian Ketidaksesuaian

Ketelitian penelitian pada bagian protokol penyelesaian ketidaksesuaian sangat menentukan apakah pola jawaban dapat ditafsirkan sebagai bukti konseptual. Fokus ini berada dalam penggunaan perbandingan terencana antar-sumber, metode, peneliti, teori, atau waktu untuk menguji kekuatan dan batas interpretasi miskonsepsi. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Kaltakci Gurel et al. (2015) serta AERA, APA, dan NCME (2014) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Dokumentasi metodologis hendaknya memuat siapa yang melakukan pengukuran berulang, kapan data diperoleh, bagaimana jenis sumber dikendalikan, dan atas dasar apa batas interpretasi disusun. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Contoh pada kajian perbedaan guru dan siswa memperlihatkan pentingnya pemisahan antara ketidaktahuan, kesalahan prosedural, dan miskonsepsi. Seorang peserta dapat gagal karena lupa rumus, sedangkan peserta lain memilih jawaban yang sama karena menggunakan model mental yang tidak ilmiah. Perbedaan ini baru terlihat ketika pengukuran berulang dibaca bersama wawancara dan bukti tentang alasan.

Keputusan yang harus didokumentasikan:

- Definisikan waktu integrasi dalam istilah operasional.
- Tentukan bukti utama dari tes dan bukti pendukung dari gambar.
- Periksa komplementaritas sebelum membuat kategori atau model.
- Catat potensi memaksa konsensus beserta strategi mitigasinya.

- Sajikan peta divergensi bersama batas interpretasi dan ketidakpastian.

Standar metodologis menuntut koherensi antara konstruk, prosedur, analisis, dan klaim. Dalam konteks ini, proporsionalitas klaim tidak dapat digantikan oleh indeks tunggal. Demikian pula, independensi relatif sumber harus dinilai dengan bukti yang sesuai. Kaltakci Gurel et al. (2015) dan AERA, APA, dan NCME (2014) memberikan dasar untuk pelaporan yang lengkap dan dapat ditelaah.

Inti metodologis subbagian ini adalah keterlacakan. Pembaca harus dapat mengikuti perjalanan dari pertanyaan, data, pengodean, analisis, sampai batas interpretasi. Ketika jejak itu jelas, temuan dapat digunakan untuk memperbaiki pembelajaran tanpa mengubah diagnosis konseptual menjadi label bagi peserta didik.

7.15.12 Triangulasi dalam Mixed Methods

Peneliti sering memusatkan perhatian pada teknik, padahal triangulasi dalam *mixed methods* justru mengatur batas inferensi yang dapat dibuat. Fokus ini berada dalam penggunaan perbandingan terencana antar-sumber, metode, peneliti, teori, atau waktu untuk menguji kekuatan dan batas interpretasi miskonsepsi. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegakan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Farmer et al. (2006) serta Carter et al. (2014) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Rancangan yang dapat diaudit mencantumkan alasan pemilihan objek yang ditriangulasi, prosedur penanganan aturan konvergensi, dan kriteria keputusan untuk tim penafsir. Tanpa penjelasan itu, pembaca sulit menilai apakah tes benar-benar mendukung klaim. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Kasus data daring dan kertas dapat digunakan untuk memperlihatkan hubungan antara pola individual dan pola kelas. Sebagian miskonsepsi mungkin hanya dimiliki beberapa peserta tetapi sangat kuat, sedangkan pola lain tersebar luas karena bersumber dari analogi atau buku teks yang sama. Analisis perlu menjaga kedua tingkat tersebut.

Pertanyaan audit metodologis:

- Definiskan objek yang ditriangulasi dalam istilah operasional.
- Tentukan bukti utama dari wawancara dan bukti pendukung dari observasi.
- Periksa independensi relatif sumber sebelum membuat kategori atau model.
- Catat potensi menghitung jumlah metode beserta strategi mitigasinya.
- Sajikan narasi penjelasan bersama batas interpretasi dan ketidakpastian.

Pemeriksaan kualitas dapat dilakukan melalui audit terhadap komplementaritas, analisis kasus yang bertentangan, dan evaluasi transparansi protokol. Apabila ditemukan menganggap kesamaan selalu benar, peneliti perlu menurunkan kekuatan klaim atau menambah bukti. Transparansi seperti ini meningkatkan kegunaan temuan bagi penelitian lanjutan dan praktik kelas.

Kesimpulannya, triangulasi dalam mixed methods perlu dirancang sebagai proses argumentatif. Data tidak berbicara sendiri; maknanya bergantung pada definisi konstruk, kualitas prosedur, dan konteks. Oleh sebab itu, klaim akhir sebaiknya menggunakan bahasa yang sesuai dengan kekuatan bukti, misalnya “menunjukkan pola”, “mengindikasikan”, atau “mendukung penjelasan”.

7.15.13 Triangulasi dan Validitas Konstruk

Kajian yang kuat memperlakukan triangulasi dan validitas konstruk sebagai bagian dari rantai argumentasi, bukan sebagai unsur pelengkap. Fokus ini berada dalam penggunaan perbandingan terencana antar-sumber, metode, peneliti, teori, atau waktu untuk menguji kekuatan dan batas interpretasi miskonsepsi. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Carter et al. (2014) serta AERA, APA, dan NCME (2014) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Pada tahap perencanaan, jenis sumber dan penanganan perbedaan harus dibedakan secara eksplisit. Pada tahap pelaksanaan, konsistensi wawancara perlu diperiksa; pada tahap interpretasi, independensi relatif sumber menjadi dasar pembatasan klaim. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Dalam kelas yang mempelajari dua peneliti berbeda kode, data agregat mungkin menunjukkan mayoritas jawaban salah. Akan tetapi, wawancara dapat memperlihatkan beberapa mekanisme berbeda: bahasa sehari-hari, generalisasi aturan, salah membaca simbol, atau model partikel yang tidak tepat. Karena itu, satu persentase tidak boleh menggantikan analisis struktur alasan.

Bukti minimum yang perlu tersedia:

- Definisikan jenis sumber dalam istilah operasional.
- Tentukan bukti utama dari gambar dan bukti pendukung dari dokumen.
- Periksa transparansi protokol sebelum membuat kategori atau model.
- Catat potensi triangulasi setelah hasil diketahui beserta strategi mitigasinya.

- Sajikan status bukti bersama batas interpretasi dan ketidakpastian.

Kriteria keberhasilan tidak cukup berupa banyaknya data. Yang lebih penting adalah apakah independensi relatif sumber dapat ditelusuri, apakah analisis divergensi diperiksa, dan apakah proporsionalitas klaim dijelaskan. Peneliti harus menghindari memaksa konsensus; jika risiko itu tidak dapat dihilangkan, dampaknya perlu dianalisis dan dicantumkan sebagai batas penelitian.

Pada akhirnya, keputusan tentang triangulasi dan validitas konstruk menentukan apakah penelitian menghasilkan daftar jawaban salah atau pemahaman yang dapat ditindaklanjuti mengenai cara peserta didik berpikir. Sasaran utamanya ialah bukti yang cukup untuk merancang diagnosis, pembelajaran, dan penelitian lanjutan secara bertanggung jawab.

7.15.14 Pelaporan Hasil Triangulasi

Subbagian ini menempatkan pelaporan hasil triangulasi sebagai simpul penting dalam triangulasi data. Fokus ini berada dalam penggunaan perbandingan terencana antar-sumber, metode, peneliti, teori, atau waktu untuk menguji kekuatan dan batas interpretasi miskonsepsi. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Lincoln dan Guba (1985) serta Carter et al. (2014) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Secara operasional, peneliti perlu menetapkan aturan konvergensi, menjelaskan tim penafsir, dan mendokumentasikan objek yang ditriangulasi. Keputusan tersebut kemudian diterjemahkan ke dalam pilihan gambar, dokumen, serta tes. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Sebagai ilustrasi, pada topik jawaban benar tetapi gambar salah, peserta didik dapat memberikan jawaban benar pada tingkat simbolik tetapi alasan submikroskopiknya keliru. Peneliti tidak cukup menghitung benar-salah; ia perlu membandingkan gambar dengan dokumen, memeriksa tingkat keyakinan, dan menguji konsistensi pada konteks lain. Hasil analisis mungkin menunjukkan bahwa kesalahan hanya muncul pada satu representasi, atau justru merupakan hubungan konseptual yang stabil.

Kriteria pemeriksaan sebelum analisis:

- Definisikan aturan konvergensi dalam istilah operasional.
- Tentukan bukti utama dari observasi dan bukti pendukung dari pengukuran berulang.
- Periksa analisis divergensi sebelum membuat kategori atau model.

- Catat potensi mengabaikan ketergantungan sumber beserta strategi mitigasinya.
- Sajikan revisi kategori bersama batas interpretasi dan ketidakpastian.

Mutu inferensi terutama ditentukan oleh transparansi protokol, jejak keputusan, dan komplementaritas. Ancaman utama ialah menghitung jumlah metode serta memakai istilah tanpa prosedur. Karena itu, laporan perlu menyajikan bukti yang mendukung sekaligus bukti yang membatasi interpretasi. Prinsip ini konsisten dengan penekanan Lincoln dan Guba (1985) dan Carter et al. (2014) bahwa validitas melekat pada interpretasi dan penggunaan hasil, bukan sekadar pada nama instrumen atau desain.

Dengan demikian, pelaporan hasil triangulasi berfungsi menghubungkan keputusan desain dengan interpretasi. Produk yang diharapkan bukan hanya narasi penjas, tetapi juga penjelasan mengenai bagaimana produk tersebut diperoleh dan dalam kondisi apa ia dapat digunakan. Penelitian miskonsepsi yang baik selalu menyisakan ruang bagi penjelasan alternatif dan verifikasi lanjutan.

7.15.15 Sintesis Protokol Triangulasi Miskonsepsi

Pembahasan mengenai sintesis protokol triangulasi miskonsepsi diperlukan agar triangulasi data tidak berhenti pada prosedur administratif. Fokus ini berada dalam penggunaan perbandingan terencana antar-sumber, metode, peneliti, teori, atau waktu untuk menguji kekuatan dan batas interpretasi miskonsepsi. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Fetters et al. (2013) serta AERA, APA, dan NCME (2014) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Penerapan praktis dapat disusun dalam urutan: mendefinisikan penanganan perbedaan, mengoperasionalkan waktu integrasi, mengumpulkan observasi, memeriksa analisis divergensi, lalu melaporkan status bukti. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Misalnya, pada wawancara berbeda dari tes, perubahan skor sesudah pembelajaran belum otomatis menandakan perubahan konsep. Peserta didik mungkin mempelajari pola jawaban tes. Bukti yang lebih meyakinkan mencakup transfer ke situasi baru, penjelasan kausal, representasi partikel, dan retensi setelah jeda waktu.

Langkah operasional yang disarankan:

- Definisikan penanganan perbedaan dalam istilah operasional.
- Tentukan bukti utama dari dokumen dan bukti pendukung dari tes.

- Periksa jejak keputusan sebelum membuat kategori atau model.
- Catat potensi memakai istilah tanpa prosedur beserta strategi mitigasinya.
- Sajikan batas interpretasi bersama batas interpretasi dan ketidakpastian.

Temuan yang baik tidak selalu berarti hasil yang besar atau sesuai harapan. Temuan bernilai ketika analisis divergensi terjaga, proporsionalitas klaim dijelaskan, dan ketidakpastian disajikan. Kerangka Fetters et al. (2013) serta AERA, APA, dan NCME (2014) menolong peneliti membedakan kekuatan bukti dari daya tarik narasi.

Inti metodologis subbagian ini adalah keterlacakan. Pembaca harus dapat mengikuti perjalanan dari pertanyaan, data, pengodean, analisis, sampai status bukti. Ketika jejak itu jelas, temuan dapat digunakan untuk memperbaiki pembelajaran tanpa mengubah diagnosis konseptual menjadi label bagi peserta didik.

7.16 Analisis Pola Respons

Subbab ini membahas analisis konfigurasi jawaban, alasan, keyakinan, dan konsistensi lintas butir untuk mengenali struktur miskonsepsi yang tidak tampak dari skor total. Fokus utamanya adalah menjaga keselarasan antara pertanyaan penelitian, bentuk bukti, teknik analisis, dan kekuatan klaim. Dalam konteks miskonsepsi kimia, keputusan metodologis harus mampu membedakan pola konseptual dari respons yang timbul karena keterbatasan bahasa, kesalahan hitung, atau kondisi administrasi. Pembahasan berikut dipecah menjadi lima belas bagian agar setiap keputusan dapat ditelaah secara operasional.

7.16.1 Rasional Analisis Pola Respons

Dalam metodologi penelitian miskonsepsi kimia, rasional analisis pola respons menentukan mutu hubungan antara data dan kesimpulan. Fokus ini berada dalam analisis konfigurasi jawaban, alasan, keyakinan, dan konsistensi lintas butir untuk mengenali struktur miskonsepsi yang tidak tampak dari skor total. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Wren dan Barbera (2014) serta Bock (1972) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Prosedur minimal mencakup definisi model klasifikasi, alasan memilih respons lintas konteks, aturan mengubah respons menjadi kategori, serta mekanisme pemeriksaan ketidakpastian klasifikasi. Semua penyimpangan dari rencana awal perlu dicatat. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Pada penelitian mengenai pola distraktor ikatan, peneliti dapat menyusun beberapa tugas ekuivalen yang berubah konteks, representasi, atau tuntutan penalarannya. Jika pola respons tetap berulang dan disertai keyakinan tinggi, terdapat dasar yang lebih kuat untuk menyebut indikasi miskonsepsi. Jika pola berubah-ubah, interpretasi yang lebih hati-hati adalah pengetahuan parsial atau ketergantungan konteks.

Keputusan yang harus didokumentasikan:

- Definisikan model klasifikasi dalam istilah operasional.
- Tentukan bukti utama dari data longitudinal dan bukti pendukung dari pilihan alasan.
- Periksa hubungan dengan bukti kualitatif sebelum membuat kategori atau model.
- Catat potensi cluster fishing beserta strategi mitigasinya.
- Sajikan matriks respons bersama batas interpretasi dan ketidakpastian.

Pemeriksaan mutu hendaknya dilakukan selama penelitian, bukan hanya pada akhir. Catatan tentang ketidakpastian klasifikasi, keputusan terkait stabilitas profil, dan kejadian yang berpotensi menimbulkan overfitting perlu dipelihara dalam *audit trail*. Hal ini memungkinkan koreksi dini dan pelaporan yang jujur.

Kesimpulannya, rasional analisis pola respons perlu dirancang sebagai proses argumentatif. Data tidak berbicara sendiri; maknanya bergantung pada definisi konstruk, kualitas prosedur, dan konteks. Oleh sebab itu, klaim akhir sebaiknya menggunakan bahasa yang sesuai dengan kekuatan bukti, misalnya “menunjukkan pola”, “mengindikasikan”, atau “mendukung penjelasan”.

7.16.2 Matriks Individu × Butir

Keputusan tentang matriks individu × butir harus dibuat sebelum data utama dikumpulkan, bukan ditambahkan setelah hasil diketahui. Fokus ini berada dalam analisis konfigurasi jawaban, alasan, keyakinan, dan konsistensi lintas butir untuk mengenali struktur miskonsepsi yang tidak tampak dari skor total. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Hasan et al. (1999) serta Wren dan Barbera (2014) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Tahap kerja dimulai dengan penegasan validasi profil. Setelah itu, peneliti menyusun aturan untuk fitur respons dan menentukan bagaimana kesetaraan butir akan diperiksa. Bukti sebaiknya berasal dari lebih dari satu bentuk, misalnya data longitudinal dan pilihan alasan. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti

membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Pada topik transisi konsepsi setelah intervensi, peneliti juga perlu memperhatikan bahasa. Istilah yang sama dapat diberi makna berbeda oleh peserta didik. Klarifikasi melalui pertanyaan tindak lanjut, gambar, atau contoh tandingan membantu membedakan masalah terminologi dari kesalahan konseptual.

Pertanyaan audit metodologis:

- Definisikan validasi profil dalam istilah operasional.
- Tentukan bukti utama dari vektor jawaban dan bukti pendukung dari skala keyakinan.
- Periksa stabilitas profil sebelum membuat kategori atau model.
- Catat potensi label yang merendahkan beserta strategi mitigasinya.
- Sajikan profil konsepsi bersama batas interpretasi dan ketidakpastian.

Kualitas penelitian meningkat ketika peneliti menunjukkan rantai bukti dari data longitudinal menuju peta konsistensi, menyatakan asumsi, dan memeriksa alternatif penjelasan. Sebaliknya, menganggap pola sebagai tipe tetap dapat menghasilkan klaim yang tampak tegas tetapi tidak stabil. Rekomendasi praktis harus proporsional dengan bukti.

Pada akhirnya, keputusan tentang matriks individu $\times$ butir menentukan apakah penelitian menghasilkan daftar jawaban salah atau pemahaman yang dapat ditindaklanjuti mengenai cara peserta didik berpikir. Sasaran utamanya ialah bukti yang cukup untuk merancang diagnosis, pembelajaran, dan penelitian lanjutan secara bertanggung jawab.

7.16.3 Analisis Distraktor

Analisis distraktor perlu dipahami sebagai keputusan metodologis yang mempunyai konsekuensi pada desain, analisis, dan pelaporan. Fokus ini berada dalam analisis konfigurasi jawaban, alasan, keyakinan, dan konsistensi lintas butir untuk mengenali struktur miskonsepsi yang tidak tampak dari skor total. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Brandriet dan Bretz (2014) serta Bock (1972) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Peneliti sebaiknya menyiapkan matriks yang menghubungkan pertanyaan penelitian dengan unit pola, vektor jawaban, teknik analisis, dan matriks respons. Matriks ini mencegah pengumpulan data yang banyak tetapi tidak relevan. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Ilustrasi respons tidak konsisten pada asam-basa juga menunjukkan bahwa kualitas instrumen memengaruhi temuan. Distraktor yang tidak berasal dari pemikiran peserta didik dapat menghasilkan pilihan acak, sedangkan pertanyaan yang terlalu menuntut matematika dapat mengukur keterampilan hitung alih-alih konsepsi. Peneliti perlu menilai kesesuaian bukti dengan konstruk yang dimaksud.

Bukti minimum yang perlu tersedia:

- Definisikan unit pola dalam istilah operasional.
- Tentukan bukti utama dari pilihan alasan dan bukti pendukung dari waktu respons.
- Periksa interpretabilitas sebelum membuat kategori atau model.
- Catat potensi mengabaikan probabilitas kelas beserta strategi mitigasinya.
- Sajikan latent class bersama batas interpretasi dan ketidakpastian.

Peneliti sebaiknya memisahkan keputusan apriori dari keputusan eksploratori. Pemisahan tersebut membantu mengendalikan cluster fishing dan mencegah profil bergantung skala. Selanjutnya, stabilitas profil dan ukuran kelompok dilaporkan bersama ketidakpastian, sehingga pembaca dapat menilai kekuatan temuan tanpa bergantung pada kesimpulan penulis semata.

Dengan demikian, analisis distraktor berfungsi menghubungkan keputusan desain dengan interpretasi. Produk yang diharapkan bukan hanya matriks respons, tetapi juga penjelasan mengenai bagaimana produk tersebut diperoleh dan dalam kondisi apa ia dapat digunakan. Penelitian miskonsepsi yang baik selalu menyisakan ruang bagi penjelasan alternatif dan verifikasi lanjutan.

7.16.4 Kombinasi Jawaban dan Alasan

Ketelitian penelitian pada bagian kombinasi jawaban dan alasan sangat menentukan apakah pola jawaban dapat ditafsirkan sebagai bukti konseptual. Fokus ini berada dalam analisis konfigurasi jawaban, alasan, keyakinan, dan konsistensi lintas butir untuk mengenali struktur miskonsepsi yang tidak tampak dari skor total. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Bock (1972) serta Wren dan Barbera (2014) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Dokumentasi metodologis hendaknya memuat siapa yang melakukan pilihan alasan, kapan data diperoleh, bagaimana kesetaraan butir dikendalikan, dan atas dasar apa profil konsepsi disusun. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu

peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Kategori respons = f(jawaban, alasan, keyakinan)

Contoh pada kajian cluster redoks memperlihatkan pentingnya pemisahan antara ketidaktahuan, kesalahan prosedural, dan miskonsepsi. Seorang peserta dapat gagal karena lupa rumus, sedangkan peserta lain memilih jawaban yang sama karena menggunakan model mental yang tidak ilmiah. Perbedaan ini baru terlihat ketika pilihan alasan dibaca bersama waktu respons dan bukti tentang alasan.

Kriteria pemeriksaan sebelum analisis:

- Definisikan fitur respons dalam istilah operasional.
- Tentukan bukti utama dari skala keyakinan dan bukti pendukung dari respons lintas konteks.
- Periksa ukuran kelompok sebelum membuat kategori atau model.
- Catat potensi profil bergantung skala beserta strategi mitigasinya.
- Sajikan transisi profil bersama batas interpretasi dan ketidakpastian.

Standar metodologis menuntut koherensi antara konstruk, prosedur, analisis, dan klaim. Dalam konteks ini, interpretabilitas tidak dapat digantikan oleh indeks tunggal. Demikian pula, validasi silang harus dinilai dengan bukti yang sesuai. Bock (1972) dan Wren dan Barbera (2014) memberikan dasar untuk pelaporan yang lengkap dan dapat ditelaah.

Inti metodologis subbagian ini adalah keterlacakan. Pembaca harus dapat mengikuti perjalanan dari pertanyaan, data, pengodean, analisis, sampai profil konsepsi. Ketika jejak itu jelas, temuan dapat digunakan untuk memperbaiki pembelajaran tanpa mengubah diagnosis konseptual menjadi label bagi peserta didik.

7.16.5 Integrasi Tingkat Keyakinan

Peneliti sering memusatkan perhatian pada teknik, padahal integrasi tingkat keyakinan justru mengatur batas inferensi yang dapat dibuat. Fokus ini berada dalam analisis konfigurasi jawaban, alasan, keyakinan, dan konsistensi lintas butir untuk mengenali struktur miskonsepsi yang tidak tampak dari skor total. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Collins dan Lanza (2010) serta Bock (1972) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Rancangan yang dapat diaudit mencantumkan alasan pemilihan ambang keyakinan, prosedur penanganan model klasifikasi, dan kriteria keputusan untuk unit pola. Tanpa penjelasan itu, pembaca sulit menilai apakah skala keyakinan benar-benar mendukung klaim. Setiap

keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Kasus pola waktu respons dapat digunakan untuk memperlihatkan hubungan antara pola individual dan pola kelas. Sebagian miskonsepsi mungkin hanya dimiliki beberapa peserta tetapi sangat kuat, sedangkan pola lain tersebar luas karena bersumber dari analogi atau buku teks yang sama. Analisis perlu menjaga kedua tingkat tersebut.

Langkah operasional yang disarankan:

- Definisikan ambang keyakinan dalam istilah operasional.
- Tentukan bukti utama dari waktu respons dan bukti pendukung dari data longitudinal.
- Periksa validasi silang sebelum membuat kategori atau model.
- Catat potensi overfitting beserta strategi mitigasinya.
- Sajikan analisis distraktor bersama batas interpretasi dan ketidakpastian.

Pemeriksaan kualitas dapat dilakukan melalui audit terhadap ukuran kelompok, analisis kasus yang bertentangan, dan evaluasi ketidakpastian klasifikasi. Apabila ditemukan mengabaikan probabilitas kelas, peneliti perlu menurunkan kekuatan klaim atau menambah bukti. Transparansi seperti ini meningkatkan kegunaan temuan bagi penelitian lanjutan dan praktik kelas.

Kesimpulannya, integrasi tingkat keyakinan perlu dirancang sebagai proses argumentatif. Data tidak berbicara sendiri; maknanya bergantung pada definisi konstruk, kualitas prosedur, dan konteks. Oleh sebab itu, klaim akhir sebaiknya menggunakan bahasa yang sesuai dengan kekuatan bukti, misalnya “menunjukkan pola”, “mengindikasikan”, atau “mendukung penjelasan”.

7.16.6 Indeks Konsistensi Lintas Butir

Kajian yang kuat memperlakukan indeks konsistensi lintas butir sebagai bagian dari rantai argumentasi, bukan sebagai unsur pelengkap. Fokus ini berada dalam analisis konfigurasi jawaban, alasan, keyakinan, dan konsistensi lintas butir untuk mengenali struktur miskonsepsi yang tidak tampak dari skor total. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Barbera (2013) serta Wren dan Barbera (2014) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Pada tahap perencanaan, kesetaraan butir dan validasi profil harus dibedakan secara eksplisit. Pada tahap pelaksanaan, konsistensi waktu respons perlu diperiksa; pada tahap interpretasi, validasi silang menjadi dasar pembatasan klaim. Setiap keputusan dicatat dalam protokol,

termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Konsistensi = jumlah pasangan respons selaras / seluruh pasangan yang diuji

Dalam kelas yang mempelajari profil gambar dan alasan, data agregat mungkin menunjukkan mayoritas jawaban salah. Akan tetapi, wawancara dapat memperlihatkan beberapa mekanisme berbeda: bahasa sehari-hari, generalisasi aturan, salah membaca simbol, atau model partikel yang tidak tepat. Karena itu, satu persentase tidak boleh menggantikan analisis struktur alasan.

Keputusan yang harus didokumentasikan:

- Definisikan kesetaraan butir dalam istilah operasional.
- Tentukan bukti utama dari respons lintas konteks dan bukti pendukung dari vektor jawaban.
- Periksa ketidakpastian klasifikasi sebelum membuat kategori atau model.
- Catat potensi menganggap pola sebagai tipe tetap beserta strategi mitigasinya.
- Sajikan peta konsistensi bersama batas interpretasi dan ketidakpastian.

Kriteria keberhasilan tidak cukup berupa banyaknya data. Yang lebih penting adalah apakah validasi silang dapat ditelusuri, apakah hubungan dengan bukti kualitatif diperiksa, dan apakah interpretabilitas dijelaskan. Peneliti harus menghindari profil bergantung skala; jika risiko itu tidak dapat dihilangkan, dampaknya perlu dianalisis dan dicantumkan sebagai batas penelitian.

Pada akhirnya, keputusan tentang indeks konsistensi lintas butir menentukan apakah penelitian menghasilkan daftar jawaban salah atau pemahaman yang dapat ditindaklanjuti mengenai cara peserta didik berpikir. Sasaran utamanya ialah bukti yang cukup untuk merancang diagnosis, pembelajaran, dan penelitian lanjutan secara bertanggung jawab.

7.16.7 Pola Lintas Representasi

Subbagian ini menempatkan pola lintas representasi sebagai simpul penting dalam analisis pola respons. Fokus ini berada dalam analisis konfigurasi jawaban, alasan, keyakinan, dan konsistensi lintas butir untuk mengenali struktur miskonsepsi yang tidak tampak dari skor total. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Wren dan Barbera (2014) serta Bock (1972) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Secara operasional, peneliti perlu menetapkan model klasifikasi, menjelaskan unit pola, dan mendokumentasikan ambang keyakinan. Keputusan tersebut kemudian diterjemahkan ke dalam pilihan respons lintas konteks, vektor jawaban, serta skala keyakinan. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Sebagai ilustrasi, pada topik profil skor rendah-keyakinan tinggi, peserta didik dapat memberikan jawaban benar pada tingkat simbolik tetapi alasan submikroskopiknya keliru. Peneliti tidak cukup menghitung benar-salah; ia perlu membandingkan respons lintas konteks dengan vektor jawaban, memeriksa tingkat keyakinan, dan menguji konsistensi pada konteks lain. Hasil analisis mungkin menunjukkan bahwa kesalahan hanya muncul pada satu representasi, atau justru merupakan hubungan konseptual yang stabil.

Pertanyaan audit metodologis:

- Definisikan model klasifikasi dalam istilah operasional.
- Tentukan bukti utama dari data longitudinal dan bukti pendukung dari pilihan alasan.
- Periksa hubungan dengan bukti kualitatif sebelum membuat kategori atau model.
- Catat potensi cluster fishing beserta strategi mitigasinya.
- Sajikan matriks respons bersama batas interpretasi dan ketidakpastian.

Mutu inferensi terutama ditentukan oleh ketidakpastian klasifikasi, stabilitas profil, dan ukuran kelompok. Ancaman utama ialah overfitting serta label yang merendahkan. Karena itu, laporan perlu menyajikan bukti yang mendukung sekaligus bukti yang membatasi interpretasi. Prinsip ini konsisten dengan penekanan Wren dan Barbera (2014) dan Bock (1972) bahwa validitas melekat pada interpretasi dan penggunaan hasil, bukan sekadar pada nama instrumen atau desain.

Dengan demikian, pola lintas representasi berfungsi menghubungkan keputusan desain dengan interpretasi. Produk yang diharapkan bukan hanya analisis distraktor, tetapi juga penjelasan mengenai bagaimana produk tersebut diperoleh dan dalam kondisi apa ia dapat digunakan. Penelitian miskonsepsi yang baik selalu menyisakan ruang bagi penjelasan alternatif dan verifikasi lanjutan.

7.16.8 Cluster Analysis

Pembahasan mengenai cluster analysis diperlukan agar analisis pola respons tidak berhenti pada prosedur administratif. Fokus ini berada dalam analisis konfigurasi jawaban, alasan, keyakinan, dan konsistensi lintas butir untuk mengenali struktur miskonsepsi yang tidak tampak dari skor total. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan

objek yang dapat dilihat secara langsung. Hasan et al. (1999) serta Wren dan Barbera (2014) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Penerapan praktis dapat disusun dalam urutan: mendefinisikan validasi profil, mengoperasionalkan fitur respons, mengumpulkan data longitudinal, memeriksa hubungan dengan bukti kualitatif, lalu melaporkan peta konsistensi. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Misalnya, pada kombinasi makro benar-submikro salah, perubahan skor sesudah pembelajaran belum otomatis menandakan perubahan konsep. Peserta didik mungkin mempelajari pola jawaban tes. Bukti yang lebih meyakinkan mencakup transfer ke situasi baru, penjelasan kausal, representasi partikel, dan retensi setelah jeda waktu.

Bukti minimum yang perlu tersedia:

- Definisikan validasi profil dalam istilah operasional.
- Tentukan bukti utama dari vektor jawaban dan bukti pendukung dari skala keyakinan.
- Periksa stabilitas profil sebelum membuat kategori atau model.
- Catat potensi label yang merendahkan beserta strategi mitigasinya.
- Sajikan profil konsepsi bersama batas interpretasi dan ketidakpastian.

Temuan yang baik tidak selalu berarti hasil yang besar atau sesuai harapan. Temuan bernilai ketika hubungan dengan bukti kualitatif terjaga, interpretabilitas dijelaskan, dan ketidakpastian disajikan. Kerangka Hasan et al. (1999) serta Wren dan Barbera (2014) menolong peneliti membedakan kekuatan bukti dari daya tarik narasi.

Inti metodologis subbagian ini adalah keterlacakan. Pembaca harus dapat mengikuti perjalanan dari pertanyaan, data, pengodean, analisis, sampai peta konsistensi. Ketika jejak itu jelas, temuan dapat digunakan untuk memperbaiki pembelajaran tanpa mengubah diagnosis konseptual menjadi label bagi peserta didik.

7.16.9 Latent Class Analysis

Dalam metodologi penelitian miskonsepsi kimia, latent class analysis menentukan mutu hubungan antara data dan kesimpulan. Fokus ini berada dalam analisis konfigurasi jawaban, alasan, keyakinan, dan konsistensi lintas butir untuk mengenali struktur miskonsepsi yang tidak tampak dari skor total. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Brandriet dan Bretz (2014) serta Bock (1972)

sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Prosedur minimal mencakup definisi unit pola, alasan memilih vektor jawaban, aturan mengubah respons menjadi kategori, serta mekanisme pemeriksaan stabilitas profil. Semua penyimpangan dari rencana awal perlu dicatat. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Pada penelitian mengenai pola distraktor ikatan, peneliti dapat menyusun beberapa tugas ekuivalen yang berubah konteks, representasi, atau tuntutan penalarannya. Jika pola respons tetap berulang dan disertai keyakinan tinggi, terdapat dasar yang lebih kuat untuk menyebut indikasi miskonsepsi. Jika pola berubah-ubah, interpretasi yang lebih hati-hati adalah pengetahuan parsial atau ketergantungan konteks.

Kriteria pemeriksaan sebelum analisis:

- Definisikan unit pola dalam istilah operasional.
- Tentukan bukti utama dari pilihan alasan dan bukti pendukung dari waktu respons.
- Periksa interpretabilitas sebelum membuat kategori atau model.
- Catat potensi mengabaikan probabilitas kelas beserta strategi mitigasinya.
- Sajikan latent class bersama batas interpretasi dan ketidakpastian.

Pemeriksaan mutu hendaknya dilakukan selama penelitian, bukan hanya pada akhir. Catatan tentang stabilitas profil, keputusan terkait ukuran kelompok, dan kejadian yang berpotensi menimbulkan cluster fishing perlu dipelihara dalam *audit trail*. Hal ini memungkinkan koreksi dini dan pelaporan yang jujur.

Kesimpulannya, latent class analysis perlu dirancang sebagai proses argumentatif. Data tidak berbicara sendiri; maknanya bergantung pada definisi konstruk, kualitas prosedur, dan konteks. Oleh sebab itu, klaim akhir sebaiknya menggunakan bahasa yang sesuai dengan kekuatan bukti, misalnya “menunjukkan pola”, “mengindikasikan”, atau “mendukung penjelasan”.

7.16.10 Latent Transition Analysis

Keputusan tentang latent transition analysis harus dibuat sebelum data utama dikumpulkan, bukan ditambahkan setelah hasil diketahui. Fokus ini berada dalam analisis konfigurasi jawaban, alasan, keyakinan, dan konsistensi lintas butir untuk mengenali struktur miskonsepsi yang tidak tampak dari skor total. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Bock (1972) serta Wren dan Barbera (2014)

sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Tahap kerja dimulai dengan penegasan fitur respons. Setelah itu, peneliti menyusun aturan untuk kesetaraan butir dan menentukan bagaimana validasi profil akan diperiksa. Bukti sebaiknya berasal dari lebih dari satu bentuk, misalnya pilihan alasan dan waktu respons. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Pada topik transisi konsepsi setelah intervensi, peneliti juga perlu memperhatikan bahasa. Istilah yang sama dapat diberi makna berbeda oleh peserta didik. Klarifikasi melalui pertanyaan tindak lanjut, gambar, atau contoh tandingan membantu membedakan masalah terminologi dari kesalahan konseptual.

Langkah operasional yang disarankan:

- Definisikan fitur respons dalam istilah operasional.
- Tentukan bukti utama dari skala keyakinan dan bukti pendukung dari respons lintas konteks.
- Periksa ukuran kelompok sebelum membuat kategori atau model.
- Catat potensi profil bergantung skala beserta strategi mitigasinya.
- Sajikan transisi profil bersama batas interpretasi dan ketidakpastian.

Kualitas penelitian meningkat ketika peneliti menunjukkan rantai bukti dari pilihan alasan menuju profil konsepsi, menyatakan asumsi, dan memeriksa alternatif penjelasan. Sebaliknya, label yang merendahkan dapat menghasilkan klaim yang tampak tegas tetapi tidak stabil. Rekomendasi praktis harus proporsional dengan bukti.

Pada akhirnya, keputusan tentang latent transition analysis menentukan apakah penelitian menghasilkan daftar jawaban salah atau pemahaman yang dapat ditindaklanjuti mengenai cara peserta didik berpikir. Sasaran utamanya ialah bukti yang cukup untuk merancang diagnosis, pembelajaran, dan penelitian lanjutan secara bertanggung jawab.

7.16.11 Model Respons Nominal

Model respons nominal perlu dipahami sebagai keputusan metodologis yang mempunyai konsekuensi pada desain, analisis, dan pelaporan. Fokus ini berada dalam analisis konfigurasi jawaban, alasan, keyakinan, dan konsistensi lintas butir untuk mengenali struktur miskonsepsi yang tidak tampak dari skor total. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Collins dan Lanza (2010) serta Bock (1972)

sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Peneliti sebaiknya menyiapkan matriks yang menghubungkan pertanyaan penelitian dengan ambang keyakinan, skala keyakinan, teknik analisis, dan latent class. Matriks ini mencegah pengumpulan data yang banyak tetapi tidak relevan. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Ilustrasi respons tidak konsisten pada asam-basa juga menunjukkan bahwa kualitas instrumen memengaruhi temuan. Distraktor yang tidak berasal dari pemikiran peserta didik dapat menghasilkan pilihan acak, sedangkan pertanyaan yang terlalu menuntut matematika dapat mengukur keterampilan hitung alih-alih konsepsi. Peneliti perlu menilai kesesuaian bukti dengan konstruk yang dimaksud.

Keputusan yang harus didokumentasikan:

- Definisikan ambang keyakinan dalam istilah operasional.
- Tentukan bukti utama dari waktu respons dan bukti pendukung dari data longitudinal.
- Periksa validasi silang sebelum membuat kategori atau model.
- Catat potensi overfitting beserta strategi mitigasinya.
- Sajikan analisis distraktor bersama batas interpretasi dan ketidakpastian.

Peneliti sebaiknya memisahkan keputusan apriori dari keputusan eksploratori. Pemisahan tersebut membantu mengendalikan mengabaikan probabilitas kelas dan mencegah menganggap pola sebagai tipe tetap. Selanjutnya, ukuran kelompok dan ketidakpastian klasifikasi dilaporkan bersama ketidakpastian, sehingga pembaca dapat menilai kekuatan temuan tanpa bergantung pada kesimpulan penulis semata.

Dengan demikian, model respons nominal berfungsi menghubungkan keputusan desain dengan interpretasi. Produk yang diharapkan bukan hanya latent class, tetapi juga penjelasan mengenai bagaimana produk tersebut diperoleh dan dalam kondisi apa ia dapat digunakan. Penelitian miskonsepsi yang baik selalu menyisakan ruang bagi penjelasan alternatif dan verifikasi lanjutan.

7.16.12 Analisis Waktu Respons

Ketelitian penelitian pada bagian analisis waktu respons sangat menentukan apakah pola jawaban dapat ditafsirkan sebagai bukti konseptual. Fokus ini berada dalam analisis konfigurasi jawaban, alasan, keyakinan, dan konsistensi lintas butir untuk mengenali struktur miskonsepsi yang tidak tampak dari skor total. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur

pemahaman, bukan objek yang dapat dilihat secara langsung. Barbera (2013) serta Wren dan Barbera (2014) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Dokumentasi metodologis hendaknya memuat siapa yang melakukan waktu respons, kapan data diperoleh, bagaimana validasi profil dikendalikan, dan atas dasar apa transisi profil disusun. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Contoh pada kajian cluster redoks memperlihatkan pentingnya pemisahan antara ketidaktahuan, kesalahan prosedural, dan miskonsepsi. Seorang peserta dapat gagal karena lupa rumus, sedangkan peserta lain memilih jawaban yang sama karena menggunakan model mental yang tidak ilmiah. Perbedaan ini baru terlihat ketika waktu respons dibaca bersama data longitudinal dan bukti tentang alasan.

Pertanyaan audit metodologis:

- Definisikan kesetaraan butir dalam istilah operasional.
- Tentukan bukti utama dari respons lintas konteks dan bukti pendukung dari vektor jawaban.
- Periksa ketidakpastian klasifikasi sebelum membuat kategori atau model.
- Catat potensi menganggap pola sebagai tipe tetap beserta strategi mitigasinya.
- Sajikan peta konsistensi bersama batas interpretasi dan ketidakpastian.

Standar metodologis menuntut koherensi antara konstruk, prosedur, analisis, dan klaim. Dalam konteks ini, validasi silang tidak dapat digantikan oleh indeks tunggal. Demikian pula, hubungan dengan bukti kualitatif harus dinilai dengan bukti yang sesuai. Barbera (2013) dan Wren dan Barbera (2014) memberikan dasar untuk pelaporan yang lengkap dan dapat ditelaah.

Inti metodologis subbagian ini adalah keterlacakan. Pembaca harus dapat mengikuti perjalanan dari pertanyaan, data, pengodean, analisis, sampai transisi profil. Ketika jejak itu jelas, temuan dapat digunakan untuk memperbaiki pembelajaran tanpa mengubah diagnosis konseptual menjadi label bagi peserta didik.

7.16.13 Validasi Profil dengan Wawancara

Peneliti sering memusatkan perhatian pada teknik, padahal validasi profil dengan wawancara justru mengatur batas inferensi yang dapat dibuat. Fokus ini berada dalam analisis konfigurasi jawaban, alasan, keyakinan, dan konsistensi lintas butir untuk mengenali struktur miskonsepsi yang tidak tampak dari skor total. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan

tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Wren dan Barbera (2014) serta Bock (1972) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Rancangan yang dapat diaudit mencantumkan alasan pemilihan model klasifikasi, prosedur penanganan unit pola, dan kriteria keputusan untuk ambang keyakinan. Tanpa penjelasan itu, pembaca sulit menilai apakah respons lintas konteks benar-benar mendukung klaim. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Kasus pola waktu respons dapat digunakan untuk memperlihatkan hubungan antara pola individual dan pola kelas. Sebagian miskonsepsi mungkin hanya dimiliki beberapa peserta tetapi sangat kuat, sedangkan pola lain tersebar luas karena bersumber dari analogi atau buku teks yang sama. Analisis perlu menjaga kedua tingkat tersebut.

Bukti minimum yang perlu tersedia:

- Definisikan model klasifikasi dalam istilah operasional.
- Tentukan bukti utama dari data longitudinal dan bukti pendukung dari pilihan alasan.
- Periksa hubungan dengan bukti kualitatif sebelum membuat kategori atau model.
- Catat potensi cluster fishing beserta strategi mitigasinya.
- Sajikan matriks respons bersama batas interpretasi dan ketidakpastian.

Pemeriksaan kualitas dapat dilakukan melalui audit terhadap ketidakpastian klasifikasi, analisis kasus yang bertentangan, dan evaluasi stabilitas profil. Apabila ditemukan overfitting, peneliti perlu menurunkan kekuatan klaim atau menambah bukti. Transparansi seperti ini meningkatkan kegunaan temuan bagi penelitian lanjutan dan praktik kelas.

Kesimpulannya, validasi profil dengan wawancara perlu dirancang sebagai proses argumentatif. Data tidak berbicara sendiri; maknanya bergantung pada definisi konstruk, kualitas prosedur, dan konteks. Oleh sebab itu, klaim akhir sebaiknya menggunakan bahasa yang sesuai dengan kekuatan bukti, misalnya “menunjukkan pola”, “mengindikasikan”, atau “mendukung penjelasan”.

7.16.14 Visualisasi Pola Respons

Kajian yang kuat memperlakukan visualisasi pola respons sebagai bagian dari rantai argumentasi, bukan sebagai unsur pelengkap. Fokus ini berada dalam analisis konfigurasi jawaban, alasan, keyakinan, dan konsistensi lintas butir untuk mengenali struktur miskonsepsi yang tidak tampak dari skor total. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman,

bukan objek yang dapat dilihat secara langsung. Hasan et al. (1999) serta Wren dan Barbera (2014) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Pada tahap perencanaan, validasi profil dan fitur respons harus dibedakan secara eksplisit. Pada tahap pelaksanaan, konsistensi data longitudinal perlu diperiksa; pada tahap interpretasi, hubungan dengan bukti kualitatif menjadi dasar pembatasan klaim. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Dalam kelas yang mempelajari profil gambar dan alasan, data agregat mungkin menunjukkan mayoritas jawaban salah. Akan tetapi, wawancara dapat memperlihatkan beberapa mekanisme berbeda: bahasa sehari-hari, generalisasi aturan, salah membaca simbol, atau model partikel yang tidak tepat. Karena itu, satu persentase tidak boleh menggantikan analisis struktur alasan.

Kriteria pemeriksaan sebelum analisis:

- Definisikan validasi profil dalam istilah operasional.
- Tentukan bukti utama dari vektor jawaban dan bukti pendukung dari skala keyakinan.
- Periksa stabilitas profil sebelum membuat kategori atau model.
- Catat potensi label yang merendahkan beserta strategi mitigasinya.
- Sajikan profil konsepsi bersama batas interpretasi dan ketidakpastian.

Kriteria keberhasilan tidak cukup berupa banyaknya data. Yang lebih penting adalah apakah hubungan dengan bukti kualitatif dapat ditelusuri, apakah interpretabilitas diperiksa, dan apakah validasi silang dijelaskan. Peneliti harus menghindari menganggap pola sebagai tipe tetap; jika risiko itu tidak dapat dihilangkan, dampaknya perlu dianalisis dan dicantumkan sebagai batas penelitian.

Pada akhirnya, keputusan tentang visualisasi pola respons menentukan apakah penelitian menghasilkan daftar jawaban salah atau pemahaman yang dapat ditindaklanjuti mengenai cara peserta didik berpikir. Sasaran utamanya ialah bukti yang cukup untuk merancang diagnosis, pembelajaran, dan penelitian lanjutan secara bertanggung jawab.

7.16.15 Sintesis Analisis Pola Respons

Subbagian ini menempatkan sintesis analisis pola respons sebagai simpul penting dalam analisis pola respons. Fokus ini berada dalam analisis konfigurasi jawaban, alasan, keyakinan, dan konsistensi lintas butir untuk mengenali struktur miskonsepsi yang tidak tampak dari skor total. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat

dilihat secara langsung. Brandriet dan Bretz (2014) serta Bock (1972) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Secara operasional, peneliti perlu menetapkan unit pola, menjelaskan ambang keyakinan, dan mendokumentasikan model klasifikasi. Keputusan tersebut kemudian diterjemahkan ke dalam pilihan vektor jawaban, skala keyakinan, serta respons lintas konteks. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Sebagai ilustrasi, pada topik profil skor rendah-keyakinan tinggi, peserta didik dapat memberikan jawaban benar pada tingkat simbolik tetapi alasan submikroskopiknya keliru. Peneliti tidak cukup menghitung benar-salah; ia perlu membandingkan vektor jawaban dengan skala keyakinan, memeriksa tingkat keyakinan, dan menguji konsistensi pada konteks lain. Hasil analisis mungkin menunjukkan bahwa kesalahan hanya muncul pada satu representasi, atau justru merupakan hubungan konseptual yang stabil.

Langkah operasional yang disarankan:

- Definisikan unit pola dalam istilah operasional.
- Tentukan bukti utama dari pilihan alasan dan bukti pendukung dari waktu respons.
- Periksa interpretabilitas sebelum membuat kategori atau model.
- Catat potensi mengabaikan probabilitas kelas beserta strategi mitigasinya.
- Sajikan latent class bersama batas interpretasi dan ketidakpastian.

Mutu inferensi terutama ditentukan oleh stabilitas profil, ukuran kelompok, dan ketidakpastian klasifikasi. Ancaman utama ialah cluster fishing serta profil bergantung skala. Karena itu, laporan perlu menyajikan bukti yang mendukung sekaligus bukti yang membatasi interpretasi. Prinsip ini konsisten dengan penekanan Brandriet dan Bretz (2014) dan Bock (1972) bahwa validitas melekat pada interpretasi dan penggunaan hasil, bukan sekadar pada nama instrumen atau desain.

Dengan demikian, sintesis analisis pola respons berfungsi menghubungkan keputusan desain dengan interpretasi. Produk yang diharapkan bukan hanya matriks respons, tetapi juga penjelasan mengenai bagaimana produk tersebut diperoleh dan dalam kondisi apa ia dapat digunakan. Penelitian miskonsepsi yang baik selalu menyisakan ruang bagi penjelasan alternatif dan verifikasi lanjutan.

7.17 Pemetaan Jaringan Konsep

Subbab ini membahas representasi dan analisis hubungan antarkonsep untuk mengungkap organisasi pengetahuan, proposisi keliru, konsep pusat, dan perubahan struktur konseptual. Fokus utamanya adalah menjaga keselarasan antara pertanyaan penelitian, bentuk bukti, teknik analisis, dan kekuatan klaim. Dalam konteks miskonsepsi kimia, keputusan metodologis

harus mampu membedakan pola konseptual dari respons yang timbul karena keterbatasan bahasa, kesalahan hitung, atau kondisi administrasi. Pembahasan berikut dipecah menjadi lima belas bagian agar setiap keputusan dapat ditelaah secara operasional.

7.17.1 Hakikat Jaringan Konsep

Pembahasan mengenai hakikat jaringan konsep diperlukan agar pemetaan jaringan konsep tidak berhenti pada prosedur administratif. Fokus ini berada dalam representasi dan analisis hubungan antarkonsep untuk mengungkap organisasi pengetahuan, proposisi keliru, konsep pusat, dan perubahan struktur konseptual. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Novak dan Cañas (2008) serta AERA, APA, dan NCME (2014) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Penerapan praktis dapat disusun dalam urutan: mendefinisikan daftar konsep, mengoperasionalkan kata penghubung, mengumpulkan proposisi, memeriksa validitas kriteria skor, lalu melaporkan indeks struktural. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Misalnya, pada asam-basa-pH, perubahan skor sesudah pembelajaran belum otomatis menandakan perubahan konsep. Peserta didik mungkin mempelajari pola jawaban tes. Bukti yang lebih meyakinkan mencakup transfer ke situasi baru, penjelasan kausal, representasi partikel, dan retensi setelah jeda waktu.

Keputusan yang harus didokumentasikan:

- Definisikan daftar konsep dalam istilah operasional.
- Tentukan bukti utama dari urutan konstruksi dan bukti pendukung dari wawancara penjas.
- Periksa reliabilitas pengodean sebelum membuat kategori atau model.
- Catat potensi membandingkan peta tidak ekuivalen beserta strategi mitigasinya.
- Sajikan simpul miskonsepsi bersama batas interpretasi dan ketidakpastian.

Temuan yang baik tidak selalu berarti hasil yang besar atau sesuai harapan. Temuan bernilai ketika validitas kriteria skor terjaga, kesetaraan kesempatan dijelaskan, dan ketidakpastian disajikan. Kerangka Novak dan Cañas (2008) serta AERA, APA, dan NCME (2014) menolong peneliti membedakan kekuatan bukti dari daya tarik narasi.

Inti metodologis subbagian ini adalah keterlacakan. Pembaca harus dapat mengikuti perjalanan dari pertanyaan, data, pengodean, analisis, sampai indeks struktural. Ketika jejak

itu jelas, temuan dapat digunakan untuk memperbaiki pembelajaran tanpa mengubah diagnosis konseptual menjadi label bagi peserta didik.

7.17.2 Peta Konsep Terbuka dan Terarah

Dalam metodologi penelitian miskonsepsi kimia, peta konsep terbuka dan terarah menentukan mutu hubungan antara data dan kesimpulan. Fokus ini berada dalam representasi dan analisis hubungan antarkonsep untuk mengungkap organisasi pengetahuan, proposisi keliru, konsep pusat, dan perubahan struktur konseptual. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Ruiz-Primo dan Shavelson (1996) serta Novak dan Cañas (2008) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Prosedur minimal mencakup definisi arah hubungan, alasan memilih urutan konstruksi, aturan mengubah respons menjadi kategori, serta mekanisme pemeriksaan reliabilitas pengodean. Semua penyimpangan dari rencana awal perlu dicatat. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Pada penelitian mengenai energi-ikatan-entalpi, peneliti dapat menyusun beberapa tugas ekuivalen yang berubah konteks, representasi, atau tuntutan penalarannya. Jika pola respons tetap berulang dan disertai keyakinan tinggi, terdapat dasar yang lebih kuat untuk menyebut indikasi miskonsepsi. Jika pola berubah-ubah, interpretasi yang lebih hati-hati adalah pengetahuan parsial atau ketergantungan konteks.

Pertanyaan audit metodologis:

- Definisikan arah hubungan dalam istilah operasional.
- Tentukan bukti utama dari peta ahli dan bukti pendukung dari peta sebelum-sesudah.
- Periksa kesetaraan kesempatan sebelum membuat kategori atau model.
- Catat potensi jaringan terlalu padat beserta strategi mitigasinya.
- Sajikan perbandingan dengan peta ahli bersama batas interpretasi dan ketidakpastian.

Pemeriksaan mutu hendaknya dilakukan selama penelitian, bukan hanya pada akhir. Catatan tentang reliabilitas pengodean, keputusan terkait interpretasi struktur, dan kejadian yang berpotensi menimbulkan mengabaikan label hubungan perlu dipelihara dalam *audit trail*. Hal ini memungkinkan koreksi dini dan pelaporan yang jujur.

Kesimpulannya, peta konsep terbuka dan terarah perlu dirancang sebagai proses argumentatif. Data tidak berbicara sendiri; maknanya bergantung pada definisi konstruk, kualitas prosedur, dan konteks. Oleh sebab itu, klaim akhir sebaiknya menggunakan bahasa

yang sesuai dengan kekuatan bukti, misalnya “menunjukkan pola”, “mengindikasikan”, atau “mendukung penjelasan”.

7.17.3 Pemilihan Konsep Kunci

Keputusan tentang pemilihan konsep kunci harus dibuat sebelum data utama dikumpulkan, bukan ditambahkan setelah hasil diketahui. Fokus ini berada dalam representasi dan analisis hubungan antarkonsep untuk mengungkap organisasi pengetahuan, proposisi keliru, konsep pusat, dan perubahan struktur konseptual. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Talanquer (2011) serta AERA, APA, dan NCME (2014) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Tahap kerja dimulai dengan penegasan kata penghubung. Setelah itu, peneliti menyusun aturan untuk model jaringan dan menentukan bagaimana daftar konsep akan diperiksa. Bukti sebaiknya berasal dari lebih dari satu bentuk, misalnya peta ahli dan peta sebelum-sesudah. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Pada topik kesetimbangan-laju, peneliti juga perlu memperhatikan bahasa. Istilah yang sama dapat diberi makna berbeda oleh peserta didik. Klarifikasi melalui pertanyaan tindak lanjut, gambar, atau contoh tandingan membantu membedakan masalah terminologi dari kesalahan konseptual.

Bukti minimum yang perlu tersedia:

- Definisikan kata penghubung dalam istilah operasional.
- Tentukan bukti utama dari wawancara penjabar dan bukti pendukung dari peta konsep siswa.
- Periksa interpretasi struktur sebelum membuat kategori atau model.
- Catat potensi menganggap sentralitas sebagai sebab beserta strategi mitigasinya.
- Sajikan perubahan struktur bersama batas interpretasi dan ketidakpastian.

Kualitas penelitian meningkat ketika peneliti menunjukkan rantai bukti dari peta ahli menuju simpul miskonsepsi, menyatakan asumsi, dan memeriksa alternatif penjelasan. Sebaliknya, membandingkan peta tidak ekuivalen dapat menghasilkan klaim yang tampak tegas tetapi tidak stabil. Rekomendasi praktis harus proporsional dengan bukti.

Pada akhirnya, keputusan tentang pemilihan konsep kunci menentukan apakah penelitian menghasilkan daftar jawaban salah atau pemahaman yang dapat ditindaklanjuti mengenai

cara peserta didik berpikir. Sasaran utamanya ialah bukti yang cukup untuk merancang diagnosis, pembelajaran, dan penelitian lanjutan secara bertanggung jawab.

7.17.4 Proposisi dan Kata Penghubung

Proposisi dan kata penghubung perlu dipahami sebagai keputusan metodologis yang mempunyai konsekuensi pada desain, analisis, dan pelaporan. Fokus ini berada dalam representasi dan analisis hubungan antarkonsep untuk mengungkap organisasi pengetahuan, proposisi keliru, konsep pusat, dan perubahan struktur konseptual. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. AERA, APA, dan NCME (2014) serta Novak dan Cañas (2008) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Peneliti sebaiknya menyiapkan matriks yang menghubungkan pertanyaan penelitian dengan skema skor, wawancara penjas, teknik analisis, dan perbandingan dengan peta ahli. Matriks ini mencegah pengumpulan data yang banyak tetapi tidak relevan. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Ilustrasi atom-ion-molekul juga menunjukkan bahwa kualitas instrumen memengaruhi temuan. Distraktor yang tidak berasal dari pemikiran peserta didik dapat menghasilkan pilihan acak, sedangkan pertanyaan yang terlalu menuntut matematika dapat mengukur keterampilan hitung alih-alih konsepsi. Peneliti perlu menilai kesesuaian bukti dengan konstruk yang dimaksud.

Kriteria pemeriksaan sebelum analisis:

- Definisikan skema skor dalam istilah operasional.
- Tentukan bukti utama dari peta sebelum-sesudah dan bukti pendukung dari proposisi.
- Periksa triangulasi makna hubungan sebelum membuat kategori atau model.
- Catat potensi menilai estetika beserta strategi mitigasinya.
- Sajikan peta proposisi bersama batas interpretasi dan ketidakpastian.

Peneliti sebaiknya memisahkan keputusan apriori dari keputusan eksploratori. Pemisahan tersebut membantu mengendalikan jaringan terlalu padat dan mencegah jumlah tautan sebagai pemahaman. Selanjutnya, interpretasi struktur dan kejelasan instruksi dilaporkan bersama ketidakpastian, sehingga pembaca dapat menilai kekuatan temuan tanpa bergantung pada kesimpulan penulis semata.

Dengan demikian, proposisi dan kata penghubung berfungsi menghubungkan keputusan desain dengan interpretasi. Produk yang diharapkan bukan hanya perbandingan dengan peta

ahli, tetapi juga penjelasan mengenai bagaimana produk tersebut diperoleh dan dalam kondisi apa ia dapat digunakan. Penelitian miskonsepsi yang baik selalu menyisakan ruang bagi penjelasan alternatif dan verifikasi lanjutan.

7.17.5 Peta Ahli sebagai Pembanding

Ketelitian penelitian pada bagian peta ahli sebagai pembanding sangat menentukan apakah pola jawaban dapat ditafsirkan sebagai bukti konseptual. Fokus ini berada dalam representasi dan analisis hubungan antarkonsep untuk mengungkap organisasi pengetahuan, proposisi keliru, konsep pusat, dan perubahan struktur konseptual. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Nakhleh (1992) serta AERA, APA, dan NCME (2014) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Dokumentasi metodologis hendaknya memuat siapa yang melakukan peta sebelum-sesudah, kapan data diperoleh, bagaimana daftar konsep dikendalikan, dan atas dasar apa perubahan struktur disusun. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Contoh pada kajian redoks-muatan-elektron memperlihatkan pentingnya pemisahan antara ketidaktahuan, kesalahan prosedural, dan miskonsepsi. Seorang peserta dapat gagal karena lupa rumus, sedangkan peserta lain memilih jawaban yang sama karena menggunakan model mental yang tidak ilmiah. Perbedaan ini baru terlihat ketika peta sebelum-sesudah dibaca bersama proposisi dan bukti tentang alasan.

Langkah operasional yang disarankan:

- Definisikan model jaringan dalam istilah operasional.
- Tentukan bukti utama dari peta konsep siswa dan bukti pendukung dari urutan konstruksi.
- Periksa kejelasan instruksi sebelum membuat kategori atau model.
- Catat potensi jumlah tautan sebagai pemahaman beserta strategi mitigasinya.
- Sajikan indeks struktural bersama batas interpretasi dan ketidakpastian.

Standar metodologis menuntut koherensi antara konstruk, prosedur, analisis, dan klaim. Dalam konteks ini, triangulasi makna hubungan tidak dapat digantikan oleh indeks tunggal. Demikian pula, validitas kriteria skor harus dinilai dengan bukti yang sesuai. Nakhleh (1992) dan AERA, APA, dan NCME (2014) memberikan dasar untuk pelaporan yang lengkap dan dapat ditelaah.

Inti metodologis subbagian ini adalah keterlacakan. Pembaca harus dapat mengikuti perjalanan dari pertanyaan, data, pengodean, analisis, sampai perubahan struktur. Ketika jejak itu jelas, temuan dapat digunakan untuk memperbaiki pembelajaran tanpa mengubah diagnosis konseptual menjadi label bagi peserta didik.

7.17.6 Penskoran Struktural

Peneliti sering memusatkan perhatian pada teknik, padahal penskoran struktural justru mengatur batas inferensi yang dapat dibuat. Fokus ini berada dalam representasi dan analisis hubungan antarkonsep untuk mengungkap organisasi pengetahuan, proposisi keliru, konsep pusat, dan perubahan struktur konseptual. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Novak dan Gowin (1984) serta Novak dan Cañas (2008) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Rancangan yang dapat diaudit mencantumkan alasan pemilihan jenis peta, prosedur penanganan arah hubungan, dan kriteria keputusan untuk skema skor. Tanpa penjelasan itu, pembaca sulit menilai apakah peta konsep siswa benar-benar mendukung klaim. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Kasus gas-tekanan-suhu dapat digunakan untuk memperlihatkan hubungan antara pola individual dan pola kelas. Sebagian miskonsepsi mungkin hanya dimiliki beberapa peserta tetapi sangat kuat, sedangkan pola lain tersebar luas karena bersumber dari analogi atau buku teks yang sama. Analisis perlu menjaga kedua tingkat tersebut.

Keputusan yang harus didokumentasikan:

- Definisikan jenis peta dalam istilah operasional.
- Tentukan bukti utama dari proposisi dan bukti pendukung dari peta ahli.
- Periksa validitas kriteria skor sebelum membuat kategori atau model.
- Catat potensi mengabaikan label hubungan beserta strategi mitigasinya.
- Sajikan jaringan konsep bersama batas interpretasi dan ketidakpastian.

Pemeriksaan kualitas dapat dilakukan melalui audit terhadap kejelasan instruksi, analisis kasus yang bertentangan, dan evaluasi reliabilitas pengodean. Apabila ditemukan menilai estetika, peneliti perlu menurunkan kekuatan klaim atau menambah bukti. Transparansi seperti ini meningkatkan kegunaan temuan bagi penelitian lanjutan dan praktik kelas.

Kesimpulannya, penskoran struktural perlu dirancang sebagai proses argumentatif. Data tidak berbicara sendiri; maknanya bergantung pada definisi konstruk, kualitas prosedur, dan

konteks. Oleh sebab itu, klaim akhir sebaiknya menggunakan bahasa yang sesuai dengan kekuatan bukti, misalnya “menunjukkan pola”, “mengindikasikan”, atau “mendukung penjelasan”.

7.17.7 Penskoran Proposisional

Kajian yang kuat memperlakukan penskoran proposisional sebagai bagian dari rantai argumentasi, bukan sebagai unsur pelengkap. Fokus ini berada dalam representasi dan analisis hubungan antarkonsep untuk mengungkap organisasi pengetahuan, proposisi keliru, konsep pusat, dan perubahan struktur konseptual. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Novak dan Cañas (2008) serta AERA, APA, dan NCME (2014) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Pada tahap perencanaan, daftar konsep dan kata penghubung harus dibedakan secara eksplisit. Pada tahap pelaksanaan, konsistensi proposisi perlu diperiksa; pada tahap interpretasi, validitas kriteria skor menjadi dasar pembatasan klaim. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Dalam kelas yang mempelajari larutan-konsentrasi-kelarutan, data agregat mungkin menunjukkan mayoritas jawaban salah. Akan tetapi, wawancara dapat memperlihatkan beberapa mekanisme berbeda: bahasa sehari-hari, generalisasi aturan, salah membaca simbol, atau model partikel yang tidak tepat. Karena itu, satu persentase tidak boleh menggantikan analisis struktur alasan.

Pertanyaan audit metodologis:

- Definisikan daftar konsep dalam istilah operasional.
- Tentukan bukti utama dari urutan konstruksi dan bukti pendukung dari wawancara penjas.
- Periksa reliabilitas pengodean sebelum membuat kategori atau model.
- Catat potensi membandingkan peta tidak ekuivalen beserta strategi mitigasinya.
- Sajikan simpul miskonsepsi bersama batas interpretasi dan ketidakpastian.

Kriteria keberhasilan tidak cukup berupa banyaknya data. Yang lebih penting adalah apakah validitas kriteria skor dapat ditelusuri, apakah kesetaraan kesempatan diperiksa, dan apakah triangulasi makna hubungan dijelaskan. Peneliti harus menghindari jumlah tautan sebagai pemahaman; jika risiko itu tidak dapat dihilangkan, dampaknya perlu dianalisis dan dicantumkan sebagai batas penelitian.

Pada akhirnya, keputusan tentang penskoran proposisional menentukan apakah penelitian menghasilkan daftar jawaban salah atau pemahaman yang dapat ditindaklanjuti mengenai cara peserta didik berpikir. Sasaran utamanya ialah bukti yang cukup untuk merancang diagnosis, pembelajaran, dan penelitian lanjutan secara bertanggung jawab.

7.17.8 Reliabilitas Pengodean Peta

Subbagian ini menempatkan reliabilitas pengodean peta sebagai simpul penting dalam pemetaan jaringan konsep. Fokus ini berada dalam representasi dan analisis hubungan antarkonsep untuk mengungkap organisasi pengetahuan, proposisi keliru, konsep pusat, dan perubahan struktur konseptual. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Ruiz-Primo dan Shavelson (1996) serta Novak dan Cañas (2008) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Secara operasional, peneliti perlu menetapkan arah hubungan, menjelaskan skema skor, dan mendokumentasikan jenis peta. Keputusan tersebut kemudian diterjemahkan ke dalam pilihan urutan konstruksi, wawancara penjelas, serta peta konsep siswa. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

$$\kappa = (P_o - P_e)/(1 - P_e)$$

Sebagai ilustrasi, pada topik jaringan mol-stoikiometri, peserta didik dapat memberikan jawaban benar pada tingkat simbolik tetapi alasan submikroskopiknya keliru. Peneliti tidak cukup menghitung benar-salah; ia perlu membandingkan urutan konstruksi dengan wawancara penjelas, memeriksa tingkat keyakinan, dan menguji konsistensi pada konteks lain. Hasil analisis mungkin menunjukkan bahwa kesalahan hanya muncul pada satu representasi, atau justru merupakan hubungan konseptual yang stabil.

Bukti minimum yang perlu tersedia:

- Definisikan arah hubungan dalam istilah operasional.
- Tentukan bukti utama dari peta ahli dan bukti pendukung dari peta sebelum-sesudah.
- Periksa kesetaraan kesempatan sebelum membuat kategori atau model.
- Catat potensi jaringan terlalu padat beserta strategi mitigasinya.
- Sajikan perbandingan dengan peta ahli bersama batas interpretasi dan ketidakpastian.

Mutu inferensi terutama ditentukan oleh reliabilitas pengodean, interpretasi struktur, dan kejelasan instruksi. Ancaman utama ialah mengabaikan label hubungan serta menganggap sentralitas sebagai sebab. Karena itu, laporan perlu menyajikan bukti yang mendukung

sekaligus bukti yang membatasi interpretasi. Prinsip ini konsisten dengan penekanan Ruiz-Primo dan Shavelson (1996) dan Novak dan Cañas (2008) bahwa validitas melekat pada interpretasi dan penggunaan hasil, bukan sekadar pada nama instrumen atau desain.

Dengan demikian, reliabilitas pengodean peta berfungsi menghubungkan keputusan desain dengan interpretasi. Produk yang diharapkan bukan hanya jaringan konsep, tetapi juga penjelasan mengenai bagaimana produk tersebut diperoleh dan dalam kondisi apa ia dapat digunakan. Penelitian miskonsepsi yang baik selalu menyisakan ruang bagi penjelasan alternatif dan verifikasi lanjutan.

7.17.9 Konversi Peta menjadi Graf

Pembahasan mengenai konversi peta menjadi graf diperlukan agar pemetaan jaringan konsep tidak berhenti pada prosedur administratif. Fokus ini berada dalam representasi dan analisis hubungan antarkonsep untuk mengungkap organisasi pengetahuan, proposisi keliru, konsep pusat, dan perubahan struktur konseptual. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Talanquer (2011) serta AERA, APA, dan NCME (2014) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Penerapan praktis dapat disusun dalam urutan: mendefinisikan kata penghubung, mengoperasionalkan model jaringan, mengumpulkan peta ahli, memeriksa kesetaraan kesempatan, lalu melaporkan simpul miskonsepsi. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Misalnya, pada asam-basa-pH, perubahan skor sesudah pembelajaran belum otomatis menandakan perubahan konsep. Peserta didik mungkin mempelajari pola jawaban tes. Bukti yang lebih meyakinkan mencakup transfer ke situasi baru, penjelasan kausal, representasi partikel, dan retensi setelah jeda waktu.

Kriteria pemeriksaan sebelum analisis:

- Definisikan kata penghubung dalam istilah operasional.
- Tentukan bukti utama dari wawancara penjas dan bukti pendukung dari peta konsep siswa.
- Periksa interpretasi struktur sebelum membuat kategori atau model.
- Catat potensi menganggap sentralitas sebagai sebab beserta strategi mitigasinya.
- Sajikan perubahan struktur bersama batas interpretasi dan ketidakpastian.

Temuan yang baik tidak selalu berarti hasil yang besar atau sesuai harapan. Temuan bernilai ketika kesetaraan kesempatan terjaga, triangulasi makna hubungan dijelaskan, dan ketidakpastian disajikan. Kerangka Talanquer (2011) serta AERA, APA, dan NCME (2014) menolong peneliti membedakan kekuatan bukti dari daya tarik narasi.

Inti metodologis subbagian ini adalah keterlacakan. Pembaca harus dapat mengikuti perjalanan dari pertanyaan, data, pengodean, analisis, sampai simpul miskonsepsi. Ketika jejak itu jelas, temuan dapat digunakan untuk memperbaiki pembelajaran tanpa mengubah diagnosis konseptual menjadi label bagi peserta didik.

7.17.10 Analisis Derajat dan Sentralitas

Dalam metodologi penelitian miskonsepsi kimia, analisis derajat dan sentralitas menentukan mutu hubungan antara data dan kesimpulan. Fokus ini berada dalam representasi dan analisis hubungan antarkonsep untuk mengungkap organisasi pengetahuan, proposisi keliru, konsep pusat, dan perubahan struktur konseptual. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. AERA, APA, dan NCME (2014) serta Novak dan Cañas (2008) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Prosedur minimal mencakup definisi skema skor, alasan memilih wawancara penjelas, aturan mengubah respons menjadi kategori, serta mekanisme pemeriksaan interpretasi struktur. Semua penyimpangan dari rencana awal perlu dicatat. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

degree(v) = jumlah hubungan yang terhubung dengan simpul v

Pada penelitian mengenai energi-ikatan-entalpi, peneliti dapat menyusun beberapa tugas ekuivalen yang berubah konteks, representasi, atau tuntutan penalarannya. Jika pola respons tetap berulang dan disertai keyakinan tinggi, terdapat dasar yang lebih kuat untuk menyebut indikasi miskonsepsi. Jika pola berubah-ubah, interpretasi yang lebih hati-hati adalah pengetahuan parsial atau ketergantungan konteks.

Langkah operasional yang disarankan:

- Definisikan skema skor dalam istilah operasional.
- Tentukan bukti utama dari peta sebelum-sesudah dan bukti pendukung dari proposisi.
- Periksa triangulasi makna hubungan sebelum membuat kategori atau model.
- Catat potensi menilai estetika beserta strategi mitigasinya.

- Sajikan peta proposisi bersama batas interpretasi dan ketidakpastian.

Pemeriksaan mutu hendaknya dilakukan selama penelitian, bukan hanya pada akhir. Catatan tentang interpretasi struktur, keputusan terkait kejelasan instruksi, dan kejadian yang berpotensi menimbulkan jaringan terlalu padat perlu dipelihara dalam *audit trail*. Hal ini memungkinkan koreksi dini dan pelaporan yang jujur.

Kesimpulannya, analisis derajat dan sentralitas perlu dirancang sebagai proses argumentatif. Data tidak berbicara sendiri; maknanya bergantung pada definisi konstruk, kualitas prosedur, dan konteks. Oleh sebab itu, klaim akhir sebaiknya menggunakan bahasa yang sesuai dengan kekuatan bukti, misalnya “menunjukkan pola”, “mengindikasikan”, atau “mendukung penjelasan”.

7.17.11 Deteksi Proposisi Miskonsepsi

Keputusan tentang deteksi proposisi miskonsepsi harus dibuat sebelum data utama dikumpulkan, bukan ditambahkan setelah hasil diketahui. Fokus ini berada dalam representasi dan analisis hubungan antarkonsep untuk mengungkap organisasi pengetahuan, proposisi keliru, konsep pusat, dan perubahan struktur konseptual. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Nakhleh (1992) serta AERA, APA, dan NCME (2014) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Tahap kerja dimulai dengan penegasan model jaringan. Setelah itu, peneliti menyusun aturan untuk daftar konsep dan menentukan bagaimana kata penghubung akan diperiksa. Bukti sebaiknya berasal dari lebih dari satu bentuk, misalnya peta sebelum-sesudah dan proposisi. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Pada topik kesetimbangan-laju, peneliti juga perlu memperhatikan bahasa. Istilah yang sama dapat diberi makna berbeda oleh peserta didik. Klarifikasi melalui pertanyaan tindak lanjut, gambar, atau contoh tandingan membantu membedakan masalah terminologi dari kesalahan konseptual.

Keputusan yang harus didokumentasikan:

- Definisikan model jaringan dalam istilah operasional.
- Tentukan bukti utama dari peta konsep siswa dan bukti pendukung dari urutan konstruksi.
- Periksa kejelasan instruksi sebelum membuat kategori atau model.
- Catat potensi jumlah tautan sebagai pemahaman beserta strategi mitigasinya.

- Sajikan indeks struktural bersama batas interpretasi dan ketidakpastian.

Kualitas penelitian meningkat ketika peneliti menunjukkan rantai bukti dari peta sebelum-sesudah menuju perubahan struktur, menyatakan asumsi, dan memeriksa alternatif penjelasan. Sebaliknya, menganggap sentralitas sebagai sebab dapat menghasilkan klaim yang tampak tegas tetapi tidak stabil. Rekomendasi praktis harus proporsional dengan bukti.

Pada akhirnya, keputusan tentang deteksi proposisi miskonsepsi menentukan apakah penelitian menghasilkan daftar jawaban salah atau pemahaman yang dapat ditindaklanjuti mengenai cara peserta didik berpikir. Sasaran utamanya ialah bukti yang cukup untuk merancang diagnosis, pembelajaran, dan penelitian lanjutan secara bertanggung jawab.

7.17.12 Perbandingan Peta Sebelum-Sesudah

Perbandingan peta sebelum-sesudah perlu dipahami sebagai keputusan metodologis yang mempunyai konsekuensi pada desain, analisis, dan pelaporan. Fokus ini berada dalam representasi dan analisis hubungan antarkonsep untuk mengungkap organisasi pengetahuan, proposisi keliru, konsep pusat, dan perubahan struktur konseptual. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Novak dan Gowin (1984) serta Novak dan Cañas (2008) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Peneliti sebaiknya menyiapkan matriks yang menghubungkan pertanyaan penelitian dengan jenis peta, peta konsep siswa, teknik analisis, dan peta proposisi. Matriks ini mencegah pengumpulan data yang banyak tetapi tidak relevan. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Ilustrasi atom-ion-molekul juga menunjukkan bahwa kualitas instrumen memengaruhi temuan. Distraktor yang tidak berasal dari pemikiran peserta didik dapat menghasilkan pilihan acak, sedangkan pertanyaan yang terlalu menuntut matematika dapat mengukur keterampilan hitung alih-alih konsepsi. Peneliti perlu menilai kesesuaian bukti dengan konstruk yang dimaksud.

Pertanyaan audit metodologis:

- Definisikan jenis peta dalam istilah operasional.
- Tentukan bukti utama dari proposisi dan bukti pendukung dari peta ahli.
- Periksa validitas kriteria skor sebelum membuat kategori atau model.
- Catat potensi mengabaikan label hubungan beserta strategi mitigasinya.
- Sajikan jaringan konsep bersama batas interpretasi dan ketidakpastian.

Peneliti sebaiknya memisahkan keputusan apriori dari keputusan eksploratori. Pemisahan tersebut membantu mengendalikan menilai estetika dan mencegah membandingkan peta tidak ekuivalen. Selanjutnya, kejelasan instruksi dan reliabilitas pengodean dilaporkan bersama ketidakpastian, sehingga pembaca dapat menilai kekuatan temuan tanpa bergantung pada kesimpulan penulis semata.

Dengan demikian, perbandingan peta sebelum-sesudah berfungsi menghubungkan keputusan desain dengan interpretasi. Produk yang diharapkan bukan hanya peta proposisi, tetapi juga penjelasan mengenai bagaimana produk tersebut diperoleh dan dalam kondisi apa ia dapat digunakan. Penelitian miskonsepsi yang baik selalu menyisakan ruang bagi penjelasan alternatif dan verifikasi lanjutan.

7.17.13 Pemetaan Jaringan Kelas

Ketelitian penelitian pada bagian pemetaan jaringan kelas sangat menentukan apakah pola jawaban dapat ditafsirkan sebagai bukti konseptual. Fokus ini berada dalam representasi dan analisis hubungan antarkonsep untuk mengungkap organisasi pengetahuan, proposisi keliru, konsep pusat, dan perubahan struktur konseptual. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Novak dan Cañas (2008) serta AERA, APA, dan NCME (2014) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Dokumentasi metodologis hendaknya memuat siapa yang melakukan proposisi, kapan data diperoleh, bagaimana kata penghubung dikendalikan, dan atas dasar apa indeks struktural disusun. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Contoh pada kajian redoks-muatan-elektron memperlihatkan pentingnya pemisahan antara ketidaktahuan, kesalahan prosedural, dan miskonsepsi. Seorang peserta dapat gagal karena lupa rumus, sedangkan peserta lain memilih jawaban yang sama karena menggunakan model mental yang tidak ilmiah. Perbedaan ini baru terlihat ketika proposisi dibaca bersama peta ahli dan bukti tentang alasan.

Bukti minimum yang perlu tersedia:

- Definisikan daftar konsep dalam istilah operasional.
- Tentukan bukti utama dari urutan konstruksi dan bukti pendukung dari wawancara penjas.
- Periksa reliabilitas pengodean sebelum membuat kategori atau model.

- Catat potensi membandingkan peta tidak ekuivalen beserta strategi mitigasinya.
- Sajikan simpul miskonsepsi bersama batas interpretasi dan ketidakpastian.

Standar metodologis menuntut koherensi antara konstruk, prosedur, analisis, dan klaim. Dalam konteks ini, validitas kriteria skor tidak dapat digantikan oleh indeks tunggal. Demikian pula, kesetaraan kesempatan harus dinilai dengan bukti yang sesuai. Novak dan Cañas (2008) dan AERA, APA, dan NCME (2014) memberikan dasar untuk pelaporan yang lengkap dan dapat ditelaah.

Inti metodologis subbagian ini adalah keterlacakan. Pembaca harus dapat mengikuti perjalanan dari pertanyaan, data, pengodean, analisis, sampai indeks struktural. Ketika jejak itu jelas, temuan dapat digunakan untuk memperbaiki pembelajaran tanpa mengubah diagnosis konseptual menjadi label bagi peserta didik.

7.17.14 Visualisasi dan Interpretasi

Peneliti sering memusatkan perhatian pada teknik, padahal visualisasi dan interpretasi justru mengatur batas inferensi yang dapat dibuat. Fokus ini berada dalam representasi dan analisis hubungan antarkonsep untuk mengungkap organisasi pengetahuan, proposisi keliru, konsep pusat, dan perubahan struktur konseptual. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Ruiz-Primo dan Shavelson (1996) serta Novak dan Cañas (2008) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Rancangan yang dapat diaudit mencantumkan alasan pemilihan arah hubungan, prosedur penanganan skema skor, dan kriteria keputusan untuk jenis peta. Tanpa penjelasan itu, pembaca sulit menilai apakah urutan konstruksi benar-benar mendukung klaim. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Kasus gas-tekanan-suhu dapat digunakan untuk memperlihatkan hubungan antara pola individual dan pola kelas. Sebagian miskonsepsi mungkin hanya dimiliki beberapa peserta tetapi sangat kuat, sedangkan pola lain tersebar luas karena bersumber dari analogi atau buku teks yang sama. Analisis perlu menjaga kedua tingkat tersebut.

Kriteria pemeriksaan sebelum analisis:

- Definisikan arah hubungan dalam istilah operasional.
- Tentukan bukti utama dari peta ahli dan bukti pendukung dari peta sebelum-sesudah.
- Periksa kesetaraan kesempatan sebelum membuat kategori atau model.
- Catat potensi jaringan terlalu padat beserta strategi mitigasinya.

- Sajikan perbandingan dengan peta ahli bersama batas interpretasi dan ketidakpastian. Pemeriksaan kualitas dapat dilakukan melalui audit terhadap reliabilitas pengodean, analisis kasus yang bertentangan, dan evaluasi interpretasi struktur. Apabila ditemukan mengabaikan label hubungan, peneliti perlu menurunkan kekuatan klaim atau menambah bukti. Transparansi seperti ini meningkatkan kegunaan temuan bagi penelitian lanjutan dan praktik kelas.

Kesimpulannya, visualisasi dan interpretasi perlu dirancang sebagai proses argumentatif. Data tidak berbicara sendiri; maknanya bergantung pada definisi konstruk, kualitas prosedur, dan konteks. Oleh sebab itu, klaim akhir sebaiknya menggunakan bahasa yang sesuai dengan kekuatan bukti, misalnya “menunjukkan pola”, “mengindikasikan”, atau “mendukung penjelasan”.

7.17.15 Sintesis Pemetaan Jaringan Konsep

Kajian yang kuat memperlakukan sintesis pemetaan jaringan konsep sebagai bagian dari rantai argumentasi, bukan sebagai unsur pelengkap. Fokus ini berada dalam representasi dan analisis hubungan antarkonsep untuk mengungkap organisasi pengetahuan, proposisi keliru, konsep pusat, dan perubahan struktur konseptual. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Talanquer (2011) serta AERA, APA, dan NCME (2014) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Pada tahap perencanaan, kata penghubung dan model jaringan harus dibedakan secara eksplisit. Pada tahap pelaksanaan, konsistensi peta ahli perlu diperiksa; pada tahap interpretasi, kesetaraan kesempatan menjadi dasar pembatasan klaim. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Dalam kelas yang mempelajari larutan-konsentrasi-kelarutan, data agregat mungkin menunjukkan mayoritas jawaban salah. Akan tetapi, wawancara dapat memperlihatkan beberapa mekanisme berbeda: bahasa sehari-hari, generalisasi aturan, salah membaca simbol, atau model partikel yang tidak tepat. Karena itu, satu persentase tidak boleh menggantikan analisis struktur alasan.

Langkah operasional yang disarankan:

- Definisikan kata penghubung dalam istilah operasional.
- Tentukan bukti utama dari wawancara penjelas dan bukti pendukung dari peta konsep siswa.

- Periksa interpretasi struktur sebelum membuat kategori atau model.
- Catat potensi menganggap sentralitas sebagai sebab beserta strategi mitigasinya.
- Sajikan perubahan struktur bersama batas interpretasi dan ketidakpastian.

Kriteria keberhasilan tidak cukup berupa banyaknya data. Yang lebih penting adalah apakah kesetaraan kesempatan dapat ditelusuri, apakah triangulasi makna hubungan diperiksa, dan apakah validitas kriteria skor dijelaskan. Peneliti harus menghindari membandingkan peta tidak ekuivalen; jika risiko itu tidak dapat dihilangkan, dampaknya perlu dianalisis dan dicantumkan sebagai batas penelitian.

Pada akhirnya, keputusan tentang sintesis pemetaan jaringan konsep menentukan apakah penelitian menghasilkan daftar jawaban salah atau pemahaman yang dapat ditindaklanjuti mengenai cara peserta didik berpikir. Sasaran utamanya ialah bukti yang cukup untuk merancang diagnosis, pembelajaran, dan penelitian lanjutan secara bertanggung jawab.

7.18 Etika Penelitian

Subbab ini membahas perlindungan hak, martabat, privasi, pembelajaran, dan kesejahteraan peserta ketika peneliti mengumpulkan data tentang kesalahan konseptual. Fokus utamanya adalah menjaga keselarasan antara pertanyaan penelitian, bentuk bukti, teknik analisis, dan kekuatan klaim. Dalam konteks miskonsepsi kimia, keputusan metodologis harus mampu membedakan pola konseptual dari respons yang timbul karena keterbatasan bahasa, kesalahan hitung, atau kondisi administrasi. Pembahasan berikut dipecah menjadi lima belas bagian agar setiap keputusan dapat ditelaah secara operasional.

7.18.1 Prinsip Dasar Etika Penelitian Pendidikan

Subbagian ini menempatkan prinsip dasar etika penelitian pendidikan sebagai simpul penting dalam etika penelitian. Fokus ini berada dalam perlindungan hak, martabat, privasi, pembelajaran, dan kesejahteraan peserta ketika peneliti mengumpulkan data tentang kesalahan konseptual. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Levitt et al. (2018) serta Belmont Report (1979) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Secara operasional, peneliti perlu menetapkan hak menarik diri, menjelaskan persetujuan etik, dan mendokumentasikan minimisasi risiko. Keputusan tersebut kemudian diterjemahkan ke dalam pilihan data demografis, formulir persetujuan, serta rekaman audio-video. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Sebagai ilustrasi, pada topik penelitian pada siswa minor, peserta didik dapat memberikan jawaban benar pada tingkat simbolik tetapi alasan submikroskopiknya keliru. Peneliti tidak cukup menghitung benar-salah; ia perlu membandingkan data demografis dengan formulir persetujuan, memeriksa tingkat keyakinan, dan menguji konsistensi pada konteks lain. Hasil analisis mungkin menunjukkan bahwa kesalahan hanya muncul pada satu representasi, atau justru merupakan hubungan konseptual yang stabil.

Keputusan yang harus didokumentasikan:

- Definisikan hak menarik diri dalam istilah operasional.
- Tentukan bukti utama dari log digital dan bukti pendukung dari identitas terkode.
- Periksa akuntabilitas sebelum membuat kategori atau model.
- Catat potensi stigmatisasi siswa beserta strategi mitigasinya.
- Sajikan protokol etik bersama batas interpretasi dan ketidakpastian.

Mutu inferensi terutama ditentukan oleh minimisasi data, respect for persons, dan justice. Ancaman utama ialah penyimpanan tidak aman serta koersi oleh guru. Karena itu, laporan perlu menyajikan bukti yang mendukung sekaligus bukti yang membatasi interpretasi. Prinsip ini konsisten dengan penekanan Levitt et al. (2018) dan Belmont Report (1979) bahwa validitas melekat pada interpretasi dan penggunaan hasil, bukan sekadar pada nama instrumen atau desain.

Dengan demikian, prinsip dasar etika penelitian pendidikan berfungsi menghubungkan keputusan desain dengan interpretasi. Produk yang diharapkan bukan hanya prosedur insiden, tetapi juga penjelasan mengenai bagaimana produk tersebut diperoleh dan dalam kondisi apa ia dapat digunakan. Penelitian miskonsepsi yang baik selalu menyisakan ruang bagi penjelasan alternatif dan verifikasi lanjutan.

7.18.2 Persetujuan Etik Institusional

Pembahasan mengenai persetujuan etik institusional diperlukan agar etika penelitian tidak berhenti pada prosedur administratif. Fokus ini berada dalam perlindungan hak, martabat, privasi, pembelajaran, dan kesejahteraan peserta ketika peneliti mengumpulkan data tentang kesalahan konseptual. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Mertler (2020) serta Levitt et al. (2018) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Penerapan praktis dapat disusun dalam urutan: mendefinisikan rencana penggunaan data, mengoperasionalkan consent dan assent, mengumpulkan log digital, memeriksa akuntabilitas, lalu melaporkan strategi umpan balik. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data

hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Misalnya, pada guru sebagai peneliti, perubahan skor sesudah pembelajaran belum otomatis menandakan perubahan konsep. Peserta didik mungkin mempelajari pola jawaban tes. Bukti yang lebih meyakinkan mencakup transfer ke situasi baru, penjelasan kausal, representasi partikel, dan retensi setelah jeda waktu.

Pertanyaan audit metodologis:

- Definisikan rencana penggunaan data dalam istilah operasional.
- Tentukan bukti utama dari formulir persetujuan dan bukti pendukung dari rekaman audio-video.
- Periksa respect for persons sebelum membuat kategori atau model.
- Catat potensi koersi oleh guru beserta strategi mitigasinya.
- Sajikan lembar informasi bersama batas interpretasi dan ketidakpastian.

Temuan yang baik tidak selalu berarti hasil yang besar atau sesuai harapan. Temuan bernilai ketika akuntabilitas terjaga, beneficence dijelaskan, dan ketidakpastian disajikan. Kerangka Mertler (2020) serta Levitt et al. (2018) menolong peneliti membedakan kekuatan bukti dari daya tarik narasi.

Inti metodologis subbagian ini adalah keterlacakan. Pembaca harus dapat mengikuti perjalanan dari pertanyaan, data, pengodean, analisis, sampai strategi umpan balik. Ketika jejak itu jelas, temuan dapat digunakan untuk memperbaiki pembelajaran tanpa mengubah diagnosis konseptual menjadi label bagi peserta didik.

7.18.3 Informed Consent dan Assent

Dalam metodologi penelitian miskonsepsi kimia, informed consent dan assent menentukan mutu hubungan antara data dan kesimpulan. Fokus ini berada dalam perlindungan hak, martabat, privasi, pembelajaran, dan kesejahteraan peserta ketika peneliti mengumpulkan data tentang kesalahan konseptual. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. BERA (2024) serta Belmont Report (1979) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Prosedur minimal mencakup definisi persetujuan etik, alasan memilih formulir persetujuan, aturan mengubah respons menjadi kategori, serta mekanisme pemeriksaan respect for persons. Semua penyimpangan dari rencana awal perlu dicatat. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya

berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Pada penelitian mengenai rekaman video praktikum, peneliti dapat menyusun beberapa tugas ekuivalen yang berubah konteks, representasi, atau tuntutan penalarannya. Jika pola respons tetap berulang dan disertai keyakinan tinggi, terdapat dasar yang lebih kuat untuk menyebut indikasi miskonsepsi. Jika pola berubah-ubah, interpretasi yang lebih hati-hati adalah pengetahuan parsial atau ketergantungan konteks.

Bukti minimum yang perlu tersedia:

- Definisikan persetujuan etik dalam istilah operasional.
- Tentukan bukti utama dari identitas terkode dan bukti pendukung dari hasil tes.
- Periksa beneficence sebelum membuat kategori atau model.
- Catat potensi nilai akademik terpengaruh beserta strategi mitigasinya.
- Sajikan rencana anonimisasi bersama batas interpretasi dan ketidakpastian.

Pemeriksaan mutu hendaknya dilakukan selama penelitian, bukan hanya pada akhir. Catatan tentang respect for persons, keputusan terkait justice, dan kejadian yang berpotensi menimbulkan stigmatisasi siswa perlu dipelihara dalam *audit trail*. Hal ini memungkinkan koreksi dini dan pelaporan yang jujur.

Kesimpulannya, informed consent dan assent perlu dirancang sebagai proses argumentatif. Data tidak berbicara sendiri; maknanya bergantung pada definisi konstruk, kualitas prosedur, dan konteks. Oleh sebab itu, klaim akhir sebaiknya menggunakan bahasa yang sesuai dengan kekuatan bukti, misalnya “menunjukkan pola”, “mengindikasikan”, atau “mendukung penjelasan”.

7.18.4 Penelitian dengan Peserta Didik Minor

Keputusan tentang penelitian dengan peserta didik minor harus dibuat sebelum data utama dikumpulkan, bukan ditambahkan setelah hasil diketahui. Fokus ini berada dalam perlindungan hak, martabat, privasi, pembelajaran, dan kesejahteraan peserta ketika peneliti mengumpulkan data tentang kesalahan konseptual. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Belmont Report (1979) serta Levitt et al. (2018) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Tahap kerja dimulai dengan penegasan consent dan assent. Setelah itu, peneliti menyusun aturan untuk kerahasiaan dan menentukan bagaimana rencana penggunaan data akan diperiksa. Bukti sebaiknya berasal dari lebih dari satu bentuk, misalnya identitas terkode dan hasil tes. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama

penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Pada topik tes diagnostik bernilai nol, peneliti juga perlu memperhatikan bahasa. Istilah yang sama dapat diberi makna berbeda oleh peserta didik. Klarifikasi melalui pertanyaan tindak lanjut, gambar, atau contoh tandingan membantu membedakan masalah terminologi dari kesalahan konseptual.

Kriteria pemeriksaan sebelum analisis:

- Definisikan consent dan assent dalam istilah operasional.
- Tentukan bukti utama dari rekaman audio-video dan bukti pendukung dari data demografis.
- Periksa justice sebelum membuat kategori atau model.
- Catat potensi identitas dapat dikenali beserta strategi mitigasinya.
- Sajikan jadwal retensi data bersama batas interpretasi dan ketidakpastian.

Kualitas penelitian meningkat ketika peneliti menunjukkan rantai bukti dari identitas terkode menuju lembar informasi, menyatakan asumsi, dan memeriksa alternatif penjelasan. Sebaliknya, koersi oleh guru dapat menghasilkan klaim yang tampak tegas tetapi tidak stabil. Rekomendasi praktis harus proporsional dengan bukti.

Pada akhirnya, keputusan tentang penelitian dengan peserta didik minor menentukan apakah penelitian menghasilkan daftar jawaban salah atau pemahaman yang dapat ditindaklanjuti mengenai cara peserta didik berpikir. Sasaran utamanya ialah bukti yang cukup untuk merancang diagnosis, pembelajaran, dan penelitian lanjutan secara bertanggung jawab.

7.18.5 Relasi Kuasa Guru-Peneliti

Relasi kuasa guru-peneliti perlu dipahami sebagai keputusan metodologis yang mempunyai konsekuensi pada desain, analisis, dan pelaporan. Fokus ini berada dalam perlindungan hak, martabat, privasi, pembelajaran, dan kesejahteraan peserta ketika peneliti mengumpulkan data tentang kesalahan konseptual. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. AERA, APA, dan NCME (2014) serta Belmont Report (1979) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Peneliti sebaiknya menyiapkan matriks yang menghubungkan pertanyaan penelitian dengan minimisasi risiko, rekaman audio-video, teknik analisis, dan rencana anonimisasi. Matriks ini mencegah pengumpulan data yang banyak tetapi tidak relevan. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya

berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Ilustrasi wawancara setelah kelas juga menunjukkan bahwa kualitas instrumen memengaruhi temuan. Distraktor yang tidak berasal dari pemikiran peserta didik dapat menghasilkan pilihan acak, sedangkan pertanyaan yang terlalu menuntut matematika dapat mengukur keterampilan hitung alih-alih konsepsi. Peneliti perlu menilai kesesuaian bukti dengan konstruk yang dimaksud.

Langkah operasional yang disarankan:

- Definisikan minimisasi risiko dalam istilah operasional.
- Tentukan bukti utama dari hasil tes dan bukti pendukung dari log digital.
- Periksa transparansi sebelum membuat kategori atau model.
- Catat potensi penyimpanan tidak aman beserta strategi mitigasinya.
- Sajikan prosedur insiden bersama batas interpretasi dan ketidakpastian.

Peneliti sebaiknya memisahkan keputusan apriori dari keputusan eksploratori. Pemisahan tersebut membantu mengendalikan nilai akademik terpengaruh dan mencegah publikasi kutipan sensitif. Selanjutnya, justice dan minimisasi data dilaporkan bersama ketidakpastian, sehingga pembaca dapat menilai kekuatan temuan tanpa bergantung pada kesimpulan penulis semata.

Dengan demikian, relasi kuasa guru-peneliti berfungsi menghubungkan keputusan desain dengan interpretasi. Produk yang diharapkan bukan hanya rencana anonimisasi, tetapi juga penjelasan mengenai bagaimana produk tersebut diperoleh dan dalam kondisi apa ia dapat digunakan. Penelitian miskonsepsi yang baik selalu menyisakan ruang bagi penjelasan alternatif dan verifikasi lanjutan.

7.18.6 Risiko Stigmatisasi Miskonsepsi

Ketelitian penelitian pada bagian risiko stigmatisasi miskonsepsi sangat menentukan apakah pola jawaban dapat ditafsirkan sebagai bukti konseptual. Fokus ini berada dalam perlindungan hak, martabat, privasi, pembelajaran, dan kesejahteraan peserta ketika peneliti mengumpulkan data tentang kesalahan konseptual. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Republik Indonesia (2022) serta Levitt et al. (2018) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Dokumentasi metodologis hendaknya memuat siapa yang melakukan hasil tes, kapan data diperoleh, bagaimana rencana penggunaan data dikendalikan, dan atas dasar apa jadwal retensi data disusun. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan

selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Contoh pada kajian data platform daring memperlihatkan pentingnya pemisahan antara ketidaktahuan, kesalahan prosedural, dan miskonsepsi. Seorang peserta dapat gagal karena lupa rumus, sedangkan peserta lain memilih jawaban yang sama karena menggunakan model mental yang tidak ilmiah. Perbedaan ini baru terlihat ketika hasil tes dibaca bersama log digital dan bukti tentang alasan.

Keputusan yang harus didokumentasikan:

- Definisikan kerahasiaan dalam istilah operasional.
- Tentukan bukti utama dari data demografis dan bukti pendukung dari formulir persetujuan.
- Periksa minimisasi data sebelum membuat kategori atau model.
- Catat potensi publikasi kutipan sensitif beserta strategi mitigasinya.
- Sajikan strategi umpan balik bersama batas interpretasi dan ketidakpastian.

Standar metodologis menuntut koherensi antara konstruk, prosedur, analisis, dan klaim. Dalam konteks ini, transparansi tidak dapat digantikan oleh indeks tunggal. Demikian pula, akuntabilitas harus dinilai dengan bukti yang sesuai. Republik Indonesia (2022) dan Levitt et al. (2018) memberikan dasar untuk pelaporan yang lengkap dan dapat ditelaah.

Inti metodologis subbagian ini adalah keterlacakan. Pembaca harus dapat mengikuti perjalanan dari pertanyaan, data, pengodean, analisis, sampai jadwal retensi data. Ketika jejak itu jelas, temuan dapat digunakan untuk memperbaiki pembelajaran tanpa mengubah diagnosis konseptual menjadi label bagi peserta didik.

7.18.7 Privasi dan Kerahasiaan

Peneliti sering memusatkan perhatian pada teknik, padahal privasi dan kerahasiaan justru mengatur batas inferensi yang dapat dibuat. Fokus ini berada dalam perlindungan hak, martabat, privasi, pembelajaran, dan kesejahteraan peserta ketika peneliti mengumpulkan data tentang kesalahan konseptual. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Levitt et al. (2018) serta Belmont Report (1979) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Rancangan yang dapat diaudit mencantumkan alasan pemilihan hak menarik diri, prosedur penanganan persetujuan etik, dan kriteria keputusan untuk minimisasi risiko. Tanpa penjelasan itu, pembaca sulit menilai apakah data demografis benar-benar mendukung klaim. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean,

penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Kasus publikasi gambar peta konsep dapat digunakan untuk memperlihatkan hubungan antara pola individual dan pola kelas. Sebagian miskonsepsi mungkin hanya dimiliki beberapa peserta tetapi sangat kuat, sedangkan pola lain tersebar luas karena bersumber dari analogi atau buku teks yang sama. Analisis perlu menjaga kedua tingkat tersebut.

Pertanyaan audit metodologis:

- Definisikan hak menarik diri dalam istilah operasional.
- Tentukan bukti utama dari log digital dan bukti pendukung dari identitas terkode.
- Periksa akuntabilitas sebelum membuat kategori atau model.
- Catat potensi stigmatisasi siswa beserta strategi mitigasinya.
- Sajikan protokol etik bersama batas interpretasi dan ketidakpastian.

Pemeriksaan kualitas dapat dilakukan melalui audit terhadap minimisasi data, analisis kasus yang bertentangan, dan evaluasi respect for persons. Apabila ditemukan penyimpanan tidak aman, peneliti perlu menurunkan kekuatan klaim atau menambah bukti. Transparansi seperti ini meningkatkan kegunaan temuan bagi penelitian lanjutan dan praktik kelas.

Kesimpulannya, privasi dan kerahasiaan perlu dirancang sebagai proses argumentatif. Data tidak berbicara sendiri; maknanya bergantung pada definisi konstruk, kualitas prosedur, dan konteks. Oleh sebab itu, klaim akhir sebaiknya menggunakan bahasa yang sesuai dengan kekuatan bukti, misalnya “menunjukkan pola”, “mengindikasikan”, atau “mendukung penjelasan”.

7.18.8 Pelindungan Data Pribadi

Kajian yang kuat memperlakukan pelindungan data pribadi sebagai bagian dari rantai argumentasi, bukan sebagai unsur pelengkap. Fokus ini berada dalam perlindungan hak, martabat, privasi, pembelajaran, dan kesejahteraan peserta ketika peneliti mengumpulkan data tentang kesalahan konseptual. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Mertler (2020) serta Levitt et al. (2018) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Pada tahap perencanaan, rencana penggunaan data dan consent dan assent harus dibedakan secara eksplisit. Pada tahap pelaksanaan, konsistensi log digital perlu diperiksa; pada tahap interpretasi, akuntabilitas menjadi dasar pembatasan klaim. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur

data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Dalam kelas yang mempelajari berbagi dataset teranonim, data agregat mungkin menunjukkan mayoritas jawaban salah. Akan tetapi, wawancara dapat memperlihatkan beberapa mekanisme berbeda: bahasa sehari-hari, generalisasi aturan, salah membaca simbol, atau model partikel yang tidak tepat. Karena itu, satu persentase tidak boleh menggantikan analisis struktur alasan.

Bukti minimum yang perlu tersedia:

- Definisikan rencana penggunaan data dalam istilah operasional.
- Tentukan bukti utama dari formulir persetujuan dan bukti pendukung dari rekaman audio-video.
- Periksa respect for persons sebelum membuat kategori atau model.
- Catat potensi koersi oleh guru beserta strategi mitigasinya.
- Sajikan lembar informasi bersama batas interpretasi dan ketidakpastian.

Kriteria keberhasilan tidak cukup berupa banyaknya data. Yang lebih penting adalah apakah akuntabilitas dapat ditelusuri, apakah beneficence diperiksa, dan apakah transparansi dijelaskan. Peneliti harus menghindari publikasi kutipan sensitif; jika risiko itu tidak dapat dihilangkan, dampaknya perlu dianalisis dan dicantumkan sebagai batas penelitian.

Pada akhirnya, keputusan tentang perlindungan data pribadi menentukan apakah penelitian menghasilkan daftar jawaban salah atau pemahaman yang dapat ditindaklanjuti mengenai cara peserta didik berpikir. Sasaran utamanya ialah bukti yang cukup untuk merancang diagnosis, pembelajaran, dan penelitian lanjutan secara bertanggung jawab.

7.18.9 Rekaman Audio, Video, dan Gambar

Subbagian ini menempatkan rekaman audio, video, dan gambar sebagai simpul penting dalam etika penelitian. Fokus ini berada dalam perlindungan hak, martabat, privasi, pembelajaran, dan kesejahteraan peserta ketika peneliti mengumpulkan data tentang kesalahan konseptual. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. BERA (2024) serta Belmont Report (1979) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Secara operasional, peneliti perlu menetapkan persetujuan etik, menjelaskan minimisasi risiko, dan mendokumentasikan hak menarik diri. Keputusan tersebut kemudian diterjemahkan ke dalam pilihan formulir persetujuan, rekaman audio-video, serta data demografis. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan

pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Sebagai ilustrasi, pada topik penelitian pada siswa minor, peserta didik dapat memberikan jawaban benar pada tingkat simbolik tetapi alasan submikroskopiknya keliru. Peneliti tidak cukup menghitung benar-salah; ia perlu membandingkan formulir persetujuan dengan rekaman audio-video, memeriksa tingkat keyakinan, dan menguji konsistensi pada konteks lain. Hasil analisis mungkin menunjukkan bahwa kesalahan hanya muncul pada satu representasi, atau justru merupakan hubungan konseptual yang stabil.

Kriteria pemeriksaan sebelum analisis:

- Definisikan persetujuan etik dalam istilah operasional.
- Tentukan bukti utama dari identitas terkode dan bukti pendukung dari hasil tes.
- Periksa beneficence sebelum membuat kategori atau model.
- Catat potensi nilai akademik terpengaruh beserta strategi mitigasinya.
- Sajikan rencana anonimisasi bersama batas interpretasi dan ketidakpastian.

Mutu inferensi terutama ditentukan oleh respect for persons, justice, dan minimisasi data. Ancaman utama ialah stigmatisasi siswa serta identitas dapat dikenali. Karena itu, laporan perlu menyajikan bukti yang mendukung sekaligus bukti yang membatasi interpretasi. Prinsip ini konsisten dengan penekanan BERA (2024) dan Belmont Report (1979) bahwa validitas melekat pada interpretasi dan penggunaan hasil, bukan sekadar pada nama instrumen atau desain.

Dengan demikian, rekaman audio, video, dan gambar berfungsi menghubungkan keputusan desain dengan interpretasi. Produk yang diharapkan bukan hanya protokol etik, tetapi juga penjelasan mengenai bagaimana produk tersebut diperoleh dan dalam kondisi apa ia dapat digunakan. Penelitian miskonsepsi yang baik selalu menyisakan ruang bagi penjelasan alternatif dan verifikasi lanjutan.

7.18.10 Hak Menolak dan Menarik Diri

Pembahasan mengenai hak menolak dan menarik diri diperlukan agar etika penelitian tidak berhenti pada prosedur administratif. Fokus ini berada dalam perlindungan hak, martabat, privasi, pembelajaran, dan kesejahteraan peserta ketika peneliti mengumpulkan data tentang kesalahan konseptual. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Belmont Report (1979) serta Levitt et al. (2018) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Penerapan praktis dapat disusun dalam urutan: mendefinisikan consent dan assent, mengoperasionalkan kerahasiaan, mengumpulkan identitas terkode, memeriksa beneficence, lalu melaporkan lembar informasi. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Misalnya, pada guru sebagai peneliti, perubahan skor sesudah pembelajaran belum otomatis menandakan perubahan konsep. Peserta didik mungkin mempelajari pola jawaban tes. Bukti yang lebih meyakinkan mencakup transfer ke situasi baru, penjelasan kausal, representasi partikel, dan retensi setelah jeda waktu.

Langkah operasional yang disarankan:

- Definisikan consent dan assent dalam istilah operasional.
- Tentukan bukti utama dari rekaman audio-video dan bukti pendukung dari data demografis.
- Periksa justice sebelum membuat kategori atau model.
- Catat potensi identitas dapat dikenali beserta strategi mitigasinya.
- Sajikan jadwal retensi data bersama batas interpretasi dan ketidakpastian.

Temuan yang baik tidak selalu berarti hasil yang besar atau sesuai harapan. Temuan bernilai ketika beneficence terjaga, transparansi dijelaskan, dan ketidakpastian disajikan. Kerangka Belmont Report (1979) serta Levitt et al. (2018) menolong peneliti membedakan kekuatan bukti dari daya tarik narasi.

Inti metodologis subbagian ini adalah keterlacakan. Pembaca harus dapat mengikuti perjalanan dari pertanyaan, data, pengodean, analisis, sampai lembar informasi. Ketika jejak itu jelas, temuan dapat digunakan untuk memperbaiki pembelajaran tanpa mengubah diagnosis konseptual menjadi label bagi peserta didik.

7.18.11 Insentif dan Kompensasi

Dalam metodologi penelitian miskonsepsi kimia, insentif dan kompensasi menentukan mutu hubungan antara data dan kesimpulan. Fokus ini berada dalam perlindungan hak, martabat, privasi, pembelajaran, dan kesejahteraan peserta ketika peneliti mengumpulkan data tentang kesalahan konseptual. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. AERA, APA, dan NCME (2014) serta Belmont Report (1979) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Prosedur minimal mencakup definisi minimisasi risiko, alasan memilih rekaman audio-video, aturan mengubah respons menjadi kategori, serta mekanisme pemeriksaan justice. Semua penyimpangan dari rencana awal perlu dicatat. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Pada penelitian mengenai rekaman video praktikum, peneliti dapat menyusun beberapa tugas ekuivalen yang berubah konteks, representasi, atau tuntutan penalarannya. Jika pola respons tetap berulang dan disertai keyakinan tinggi, terdapat dasar yang lebih kuat untuk menyebut indikasi miskonsepsi. Jika pola berubah-ubah, interpretasi yang lebih hati-hati adalah pengetahuan parsial atau ketergantungan konteks.

Keputusan yang harus didokumentasikan:

- Definisikan minimisasi risiko dalam istilah operasional.
- Tentukan bukti utama dari hasil tes dan bukti pendukung dari log digital.
- Periksa transparansi sebelum membuat kategori atau model.
- Catat potensi penyimpanan tidak aman beserta strategi mitigasinya.
- Sajikan prosedur insiden bersama batas interpretasi dan ketidakpastian.

Pemeriksaan mutu hendaknya dilakukan selama penelitian, bukan hanya pada akhir. Catatan tentang justice, keputusan terkait minimisasi data, dan kejadian yang berpotensi menimbulkan nilai akademik terpengaruh perlu dipelihara dalam *audit trail*. Hal ini memungkinkan koreksi dini dan pelaporan yang jujur.

Kesimpulannya, insentif dan kompensasi perlu dirancang sebagai proses argumentatif. Data tidak berbicara sendiri; maknanya bergantung pada definisi konstruk, kualitas prosedur, dan konteks. Oleh sebab itu, klaim akhir sebaiknya menggunakan bahasa yang sesuai dengan kekuatan bukti, misalnya “menunjukkan pola”, “mengindikasikan”, atau “mendukung penjelasan”.

7.18.12 Umpan Balik Hasil Diagnosis

Keputusan tentang umpan balik hasil diagnosis harus dibuat sebelum data utama dikumpulkan, bukan ditambahkan setelah hasil diketahui. Fokus ini berada dalam perlindungan hak, martabat, privasi, pembelajaran, dan kesejahteraan peserta ketika peneliti mengumpulkan data tentang kesalahan konseptual. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Republik Indonesia (2022) serta Levitt et al. (2018) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Tahap kerja dimulai dengan penegasan kerahasiaan. Setelah itu, peneliti menyusun aturan untuk rencana penggunaan data dan menentukan bagaimana consent dan assent akan diperiksa. Bukti sebaiknya berasal dari lebih dari satu bentuk, misalnya hasil tes dan log digital. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Pada topik tes diagnostik bernilai nol, peneliti juga perlu memperhatikan bahasa. Istilah yang sama dapat diberi makna berbeda oleh peserta didik. Klarifikasi melalui pertanyaan tindak lanjut, gambar, atau contoh tandingan membantu membedakan masalah terminologi dari kesalahan konseptual.

Pertanyaan audit metodologis:

- Definisikan kerahasiaan dalam istilah operasional.
- Tentukan bukti utama dari data demografis dan bukti pendukung dari formulir persetujuan.
- Periksa minimisasi data sebelum membuat kategori atau model.
- Catat potensi publikasi kutipan sensitif beserta strategi mitigasinya.
- Sajikan strategi umpan balik bersama batas interpretasi dan ketidakpastian.

Kualitas penelitian meningkat ketika peneliti menunjukkan rantai bukti dari hasil tes menuju jadwal retensi data, menyatakan asumsi, dan memeriksa alternatif penjelasan. Sebaliknya, identitas dapat dikenali dapat menghasilkan klaim yang tampak tegas tetapi tidak stabil. Rekomendasi praktis harus proporsional dengan bukti.

Pada akhirnya, keputusan tentang umpan balik hasil diagnosis menentukan apakah penelitian menghasilkan daftar jawaban salah atau pemahaman yang dapat ditindaklanjuti mengenai cara peserta didik berpikir. Sasaran utamanya ialah bukti yang cukup untuk merancang diagnosis, pembelajaran, dan penelitian lanjutan secara bertanggung jawab.

7.18.13 Etika Berbagi Data

Etika berbagi data perlu dipahami sebagai keputusan metodologis yang mempunyai konsekuensi pada desain, analisis, dan pelaporan. Fokus ini berada dalam perlindungan hak, martabat, privasi, pembelajaran, dan kesejahteraan peserta ketika peneliti mengumpulkan data tentang kesalahan konseptual. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Levitt et al. (2018) serta Belmont Report (1979) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Peneliti sebaiknya menyiapkan matriks yang menghubungkan pertanyaan penelitian dengan hak menarik diri, data demografis, teknik analisis, dan prosedur insiden. Matriks ini mencegah pengumpulan data yang banyak tetapi tidak relevan. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Ilustrasi wawancara setelah kelas juga menunjukkan bahwa kualitas instrumen memengaruhi temuan. Distraktor yang tidak berasal dari pemikiran peserta didik dapat menghasilkan pilihan acak, sedangkan pertanyaan yang terlalu menuntut matematika dapat mengukur keterampilan hitung alih-alih konsepsi. Peneliti perlu menilai kesesuaian bukti dengan konstruk yang dimaksud.

Bukti minimum yang perlu tersedia:

- Definisikan hak menarik diri dalam istilah operasional.
- Tentukan bukti utama dari log digital dan bukti pendukung dari identitas terkode.
- Periksa akuntabilitas sebelum membuat kategori atau model.
- Catat potensi stigmatisasi siswa beserta strategi mitigasinya.
- Sajikan protokol etik bersama batas interpretasi dan ketidakpastian.

Peneliti sebaiknya memisahkan keputusan apriori dari keputusan eksploratori. Pemisahan tersebut membantu mengendalikan penyimpanan tidak aman dan mencegah koersi oleh guru. Selanjutnya, minimisasi data dan respect for persons dilaporkan bersama ketidakpastian, sehingga pembaca dapat menilai kekuatan temuan tanpa bergantung pada kesimpulan penulis semata.

Dengan demikian, etika berbagi data berfungsi menghubungkan keputusan desain dengan interpretasi. Produk yang diharapkan bukan hanya prosedur insiden, tetapi juga penjelasan mengenai bagaimana produk tersebut diperoleh dan dalam kondisi apa ia dapat digunakan. Penelitian miskonsepsi yang baik selalu menyisakan ruang bagi penjelasan alternatif dan verifikasi lanjutan.

7.18.14 Etika Publikasi dan Kepengarangan

Ketelitian penelitian pada bagian etika publikasi dan kepengarangan sangat menentukan apakah pola jawaban dapat ditafsirkan sebagai bukti konseptual. Fokus ini berada dalam perlindungan hak, martabat, privasi, pembelajaran, dan kesejahteraan peserta ketika peneliti mengumpulkan data tentang kesalahan konseptual. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Mertler

(2020) serta Levitt et al. (2018) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Dokumentasi metodologis hendaknya memuat siapa yang melakukan log digital, kapan data diperoleh, bagaimana consent dan assent dikendalikan, dan atas dasar apa strategi umpan balik disusun. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Contoh pada kajian data platform daring memperlihatkan pentingnya pemisahan antara ketidaktahuan, kesalahan prosedural, dan miskonsepsi. Seorang peserta dapat gagal karena lupa rumus, sedangkan peserta lain memilih jawaban yang sama karena menggunakan model mental yang tidak ilmiah. Perbedaan ini baru terlihat ketika log digital dibaca bersama identitas terkode dan bukti tentang alasan.

Kriteria pemeriksaan sebelum analisis:

- Definisikan rencana penggunaan data dalam istilah operasional.
- Tentukan bukti utama dari formulir persetujuan dan bukti pendukung dari rekaman audio-video.
- Periksa respect for persons sebelum membuat kategori atau model.
- Catat potensi koersi oleh guru beserta strategi mitigasinya.
- Sajikan lembar informasi bersama batas interpretasi dan ketidakpastian.

Standar metodologis menuntut koherensi antara konstruk, prosedur, analisis, dan klaim. Dalam konteks ini, akuntabilitas tidak dapat digantikan oleh indeks tunggal. Demikian pula, beneficence harus dinilai dengan bukti yang sesuai. Mertler (2020) dan Levitt et al. (2018) memberikan dasar untuk pelaporan yang lengkap dan dapat ditelaah.

Inti metodologis subbagian ini adalah keterlacakan. Pembaca harus dapat mengikuti perjalanan dari pertanyaan, data, pengodean, analisis, sampai strategi umpan balik. Ketika jejak itu jelas, temuan dapat digunakan untuk memperbaiki pembelajaran tanpa mengubah diagnosis konseptual menjadi label bagi peserta didik.

7.18.15 Sintesis Protokol Etika Penelitian Miskonsepsi

Peneliti sering memusatkan perhatian pada teknik, padahal sintesis protokol etika penelitian miskonsepsi justru mengatur batas inferensi yang dapat dibuat. Fokus ini berada dalam perlindungan hak, martabat, privasi, pembelajaran, dan kesejahteraan peserta ketika peneliti mengumpulkan data tentang kesalahan konseptual. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. BERA (2024)

serta Belmont Report (1979) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Rancangan yang dapat diaudit mencantumkan alasan pemilihan persetujuan etik, prosedur penanganan minimisasi risiko, dan kriteria keputusan untuk hak menarik diri. Tanpa penjelasan itu, pembaca sulit menilai apakah formulir persetujuan benar-benar mendukung klaim. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Kasus publikasi gambar peta konsep dapat digunakan untuk memperlihatkan hubungan antara pola individual dan pola kelas. Sebagian miskonsepsi mungkin hanya dimiliki beberapa peserta tetapi sangat kuat, sedangkan pola lain tersebar luas karena bersumber dari analogi atau buku teks yang sama. Analisis perlu menjaga kedua tingkat tersebut.

Langkah operasional yang disarankan:

- Definisikan persetujuan etik dalam istilah operasional.
- Tentukan bukti utama dari identitas terkode dan bukti pendukung dari hasil tes.
- Periksa beneficence sebelum membuat kategori atau model.
- Catat potensi nilai akademik terpengaruh beserta strategi mitigasinya.
- Sajikan rencana anonimisasi bersama batas interpretasi dan ketidakpastian.

Pemeriksaan kualitas dapat dilakukan melalui audit terhadap respect for persons, analisis kasus yang bertentangan, dan evaluasi justice. Apabila ditemukan stigmatisasi siswa, peneliti perlu menurunkan kekuatan klaim atau menambah bukti. Transparansi seperti ini meningkatkan kegunaan temuan bagi penelitian lanjutan dan praktik kelas.

Kesimpulannya, sintesis protokol etika penelitian miskonsepsi perlu dirancang sebagai proses argumentatif. Data tidak berbicara sendiri; maknanya bergantung pada definisi konstruk, kualitas prosedur, dan konteks. Oleh sebab itu, klaim akhir sebaiknya menggunakan bahasa yang sesuai dengan kekuatan bukti, misalnya “menunjukkan pola”, “mengindikasikan”, atau “mendukung penjelasan”.

7.19 Penyajian Temuan Miskonsepsi

Subbab ini membahas komunikasi temuan secara presisi, transparan, tidak menstigma, dan sesuai dengan kekuatan bukti agar dapat digunakan oleh peneliti serta guru. Fokus utamanya adalah menjaga keselarasan antara pertanyaan penelitian, bentuk bukti, teknik analisis, dan kekuatan klaim. Dalam konteks miskonsepsi kimia, keputusan metodologis harus mampu membedakan pola konseptual dari respons yang timbul karena keterbatasan bahasa, kesalahan hitung, atau kondisi administrasi. Pembahasan berikut dipecah menjadi lima belas bagian agar setiap keputusan dapat ditelaah secara operasional.

7.19.1 Prinsip Penyajian yang Bertanggung Jawab

Kajian yang kuat memperlakukan prinsip penyajian yang bertanggung jawab sebagai bagian dari rantai argumentasi, bukan sebagai unsur pelengkap. Fokus ini berada dalam komunikasi temuan secara presisi, transparan, tidak menstigma, dan sesuai dengan kekuatan bukti agar dapat digunakan oleh peneliti serta guru. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Levitt et al. (2018) serta Cumming (2014) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Pada tahap perencanaan, unit pelaporan dan tingkat detail harus dibedakan secara eksplisit. Pada tahap pelaksanaan, konsistensi kutipan perlu diperiksa; pada tahap interpretasi, kelengkapan menjadi dasar pembatasan klaim. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Dalam kelas yang mempelajari infografik untuk guru, data agregat mungkin menunjukkan mayoritas jawaban salah. Akan tetapi, wawancara dapat memperlihatkan beberapa mekanisme berbeda: bahasa sehari-hari, generalisasi aturan, salah membaca simbol, atau model partikel yang tidak tepat. Karena itu, satu persentase tidak boleh menggantikan analisis struktur alasan.

Keputusan yang harus didokumentasikan:

- Definisikan unit pelaporan dalam istilah operasional.
- Tentukan bukti utama dari pola respons dan bukti pendukung dari ukuran efek.
- Periksa keterbacaan sebelum membuat kategori atau model.
- Catat potensi grafik terpotong beserta strategi mitigasinya.
- Sajikan matriks lintas representasi bersama batas interpretasi dan ketidakpastian.

Kriteria keberhasilan tidak cukup berupa banyaknya data. Yang lebih penting adalah apakah kelengkapan dapat ditelusuri, apakah proporsionalitas klaim diperiksa, dan apakah aksesibilitas visual dijelaskan. Peneliti harus menghindari persentase tanpa denominator; jika risiko itu tidak dapat dihilangkan, dampaknya perlu dianalisis dan dicantumkan sebagai batas penelitian.

Pada akhirnya, keputusan tentang prinsip penyajian yang bertanggung jawab menentukan apakah penelitian menghasilkan daftar jawaban salah atau pemahaman yang dapat ditindaklanjuti mengenai cara peserta didik berpikir. Sasaran utamanya ialah bukti yang cukup

untuk merancang diagnosis, pembelajaran, dan penelitian lanjutan secara bertanggung jawab.

7.19.2 Penyajian Tabel Prevalensi

Subbagian ini menempatkan penyajian tabel prevalensi sebagai simpul penting dalam penyajian temuan miskonsepsi. Fokus ini berada dalam komunikasi temuan secara presisi, transparan, tidak menstigma, dan sesuai dengan kekuatan bukti agar dapat digunakan oleh peneliti serta guru. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. O'Brien et al. (2014) serta Levitt et al. (2018) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Secara operasional, peneliti perlu menetapkan bentuk visual, menjelaskan bahasa ketidakpastian, dan mendokumentasikan audiens. Keputusan tersebut kemudian diterjemahkan ke dalam pilihan pola respons, ukuran efek, serta prevalensi. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

$$\hat{p} = x/n; SE(\hat{p}) = \sqrt{[\hat{p}(1 - \hat{p})/n]}$$

Sebagai ilustrasi, pada topik tabel miskonsepsi asam-basa, peserta didik dapat memberikan jawaban benar pada tingkat simbolik tetapi alasan submikroskopiknya keliru. Peneliti tidak cukup menghitung benar-salah; ia perlu membandingkan pola respons dengan ukuran efek, memeriksa tingkat keyakinan, dan menguji konsistensi pada konteks lain. Hasil analisis mungkin menunjukkan bahwa kesalahan hanya muncul pada satu representasi, atau justru merupakan hubungan konseptual yang stabil.

Pertanyaan audit metodologis:

- Definisikan bentuk visual dalam istilah operasional.
- Tentukan bukti utama dari peta konsep dan bukti pendukung dari joint display.
- Periksa proporsionalitas klaim sebelum membuat kategori atau model.
- Catat potensi label deficit beserta strategi mitigasinya.
- Sajikan joint display bersama batas interpretasi dan ketidakpastian.

Mutu inferensi terutama ditentukan oleh keterbacaan, konsistensi angka-narasi, dan akurasi. Ancaman utama ialah kutipan dipilih sepihak serta diskusi melebihi data. Karena itu, laporan perlu menyajikan bukti yang mendukung sekaligus bukti yang membatasi interpretasi. Prinsip ini konsisten dengan penekanan O'Brien et al. (2014) dan Levitt et al. (2018) bahwa validitas

melekat pada interpretasi dan penggunaan hasil, bukan sekadar pada nama instrumen atau desain.

Dengan demikian, penyajian tabel prevalensi berfungsi menghubungkan keputusan desain dengan interpretasi. Produk yang diharapkan bukan hanya kutipan analitik, tetapi juga penjelasan mengenai bagaimana produk tersebut diperoleh dan dalam kondisi apa ia dapat digunakan. Penelitian miskonsepsi yang baik selalu menyisakan ruang bagi penjelasan alternatif dan verifikasi lanjutan.

7.19.3 Grafik dan Visualisasi

Pembahasan mengenai grafik dan visualisasi diperlukan agar penyajian temuan miskonsepsi tidak berhenti pada prosedur administratif. Fokus ini berada dalam komunikasi temuan secara presisi, transparan, tidak menstigma, dan sesuai dengan kekuatan bukti agar dapat digunakan oleh peneliti serta guru. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Page et al. (2021) serta Cumming (2014) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Penerapan praktis dapat disusun dalam urutan: mendefinisikan tingkat detail, mengoperasionalkan anonimisasi contoh, mengumpulkan peta konsep, memeriksa proporsionalitas klaim, lalu melaporkan matriks lintas representasi. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Misalnya, pada heatmap butir redoks, perubahan skor sesudah pembelajaran belum otomatis menandakan perubahan konsep. Peserta didik mungkin mempelajari pola jawaban tes. Bukti yang lebih meyakinkan mencakup transfer ke situasi baru, penjelasan kausal, representasi partikel, dan retensi setelah jeda waktu.

Bukti minimum yang perlu tersedia:

- Definisikan tingkat detail dalam istilah operasional.
- Tentukan bukti utama dari ukuran efek dan bukti pendukung dari prevalensi.
- Periksa konsistensi angka-narasi sebelum membuat kategori atau model.
- Catat potensi diskusi melebihi data beserta strategi mitigasinya.
- Sajikan ringkasan implikasi bersama batas interpretasi dan ketidakpastian.

Temuan yang baik tidak selalu berarti hasil yang besar atau sesuai harapan. Temuan bernilai ketika proporsionalitas klaim terjaga, aksesibilitas visual dijelaskan, dan ketidakpastian

disajikan. Kerangka Page et al. (2021) serta Cumming (2014) menolong peneliti membedakan kekuatan bukti dari daya tarik narasi.

Inti metodologis subbagian ini adalah keterlacakan. Pembaca harus dapat mengikuti perjalanan dari pertanyaan, data, pengodean, analisis, sampai matriks lintas representasi. Ketika jejak itu jelas, temuan dapat digunakan untuk memperbaiki pembelajaran tanpa mengubah diagnosis konseptual menjadi label bagi peserta didik.

7.19.4 Penyajian Kategori Konsepsi

Dalam metodologi penelitian miskonsepsi kimia, penyajian kategori konsepsi menentukan mutu hubungan antara data dan kesimpulan. Fokus ini berada dalam komunikasi temuan secara presisi, transparan, tidak menstigma, dan sesuai dengan kekuatan bukti agar dapat digunakan oleh peneliti serta guru. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Cumming (2014) serta Levitt et al. (2018) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Prosedur minimal mencakup definisi bahasa ketidakpastian, alasan memilih ukuran efek, aturan mengubah respons menjadi kategori, serta mekanisme pemeriksaan konsistensi angka-narasi. Semua penyimpangan dari rencana awal perlu dicatat. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Pada penelitian mengenai alur perubahan konsepsi, peneliti dapat menyusun beberapa tugas ekuivalen yang berubah konteks, representasi, atau tuntutan penalarannya. Jika pola respons tetap berulang dan disertai keyakinan tinggi, terdapat dasar yang lebih kuat untuk menyebut indikasi miskonsepsi. Jika pola berubah-ubah, interpretasi yang lebih hati-hati adalah pengetahuan parsial atau ketergantungan konteks.

Kriteria pemeriksaan sebelum analisis:

- Definisikan bahasa ketidakpastian dalam istilah operasional.
- Tentukan bukti utama dari joint display dan bukti pendukung dari kutipan.
- Periksa aksesibilitas visual sebelum membuat kategori atau model.
- Catat potensi judul sensasional beserta strategi mitigasinya.
- Sajikan tabel kategori bersama batas interpretasi dan ketidakpastian.

Pemeriksaan mutu hendaknya dilakukan selama penelitian, bukan hanya pada akhir. Catatan tentang konsistensi angka-narasi, keputusan terkait akurasi, dan kejadian yang berpotensi

menimbulkan label deficit perlu dipelihara dalam *audit trail*. Hal ini memungkinkan koreksi dini dan pelaporan yang jujur.

Kesimpulannya, penyajian kategori konsepsi perlu dirancang sebagai proses argumentatif. Data tidak berbicara sendiri; maknanya bergantung pada definisi konstruk, kualitas prosedur, dan konteks. Oleh sebab itu, klaim akhir sebaiknya menggunakan bahasa yang sesuai dengan kekuatan bukti, misalnya “menunjukkan pola”, “mengindikasikan”, atau “mendukung penjelasan”.

7.19.5 Penyajian Data Keyakinan

Keputusan tentang penyajian data keyakinan harus dibuat sebelum data utama dikumpulkan, bukan ditambahkan setelah hasil diketahui. Fokus ini berada dalam komunikasi temuan secara presisi, transparan, tidak menstigma, dan sesuai dengan kekuatan bukti agar dapat digunakan oleh peneliti serta guru. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. BERA (2024) serta Cumming (2014) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Tahap kerja dimulai dengan penegasan anonimisasi contoh. Setelah itu, peneliti menyusun aturan untuk unit pelaporan dan menentukan bagaimana tingkat detail akan diperiksa. Bukti sebaiknya berasal dari lebih dari satu bentuk, misalnya *joint display* dan kutipan. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Pada topik peta jaringan ikatan, peneliti juga perlu memperhatikan bahasa. Istilah yang sama dapat diberi makna berbeda oleh peserta didik. Klarifikasi melalui pertanyaan tindak lanjut, gambar, atau contoh tandingan membantu membedakan masalah terminologi dari kesalahan konseptual.

Langkah operasional yang disarankan:

- Definisikan anonimisasi contoh dalam istilah operasional.
- Tentukan bukti utama dari prevalensi dan bukti pendukung dari pola respons.
- Periksa akurasi sebelum membuat kategori atau model.
- Catat potensi persentase tanpa denominator beserta strategi mitigasinya.
- Sajikan grafik prevalensi bersama batas interpretasi dan ketidakpastian.

Kualitas penelitian meningkat ketika peneliti menunjukkan rantai bukti dari *joint display* menuju ringkasan implikasi, menyatakan asumsi, dan memeriksa alternatif penjelasan.

Sebaliknya, diskusi melebihi data dapat menghasilkan klaim yang tampak tegas tetapi tidak stabil. Rekomendasi praktis harus proporsional dengan bukti.

Pada akhirnya, keputusan tentang penyajian data keyakinan menentukan apakah penelitian menghasilkan daftar jawaban salah atau pemahaman yang dapat ditindaklanjuti mengenai cara peserta didik berpikir. Sasaran utamanya ialah bukti yang cukup untuk merancang diagnosis, pembelajaran, dan penelitian lanjutan secara bertanggung jawab.

7.19.6 Penyajian Pola Respons

Penyajian pola respons perlu dipahami sebagai keputusan metodologis yang mempunyai konsekuensi pada desain, analisis, dan pelaporan. Fokus ini berada dalam komunikasi temuan secara presisi, transparan, tidak menstigma, dan sesuai dengan kekuatan bukti agar dapat digunakan oleh peneliti serta guru. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Appelbaum et al. (2018) serta Levitt et al. (2018) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Peneliti sebaiknya menyiapkan matriks yang menghubungkan pertanyaan penelitian dengan audiens, prevalensi, teknik analisis, dan tabel kategori. Matriks ini mencegah pengumpulan data yang banyak tetapi tidak relevan. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Ilustrasi kutipan wawancara kesetimbangan juga menunjukkan bahwa kualitas instrumen memengaruhi temuan. Distraktor yang tidak berasal dari pemikiran peserta didik dapat menghasilkan pilihan acak, sedangkan pertanyaan yang terlalu menuntut matematika dapat mengukur keterampilan hitung alih-alih konsepsi. Peneliti perlu menilai kesesuaian bukti dengan konstruk yang dimaksud.

Keputusan yang harus didokumentasikan:

- Definisikan audiens dalam istilah operasional.
- Tentukan bukti utama dari kutipan dan bukti pendukung dari peta konsep.
- Periksa kelengkapan sebelum membuat kategori atau model.
- Catat potensi kutipan dipilih sepihak beserta strategi mitigasinya.
- Sajikan kutipan analitik bersama batas interpretasi dan ketidakpastian.

Peneliti sebaiknya memisahkan keputusan apriori dari keputusan eksploratori. Pemisahan tersebut membantu mengendalikan judul sensasional dan mencegah grafik terpotong.

Selanjutnya, akurasi dan keterbacaan dilaporkan bersama ketidakpastian, sehingga pembaca dapat menilai kekuatan temuan tanpa bergantung pada kesimpulan penulis semata.

Dengan demikian, penyajian pola respons berfungsi menghubungkan keputusan desain dengan interpretasi. Produk yang diharapkan bukan hanya tabel kategori, tetapi juga penjelasan mengenai bagaimana produk tersebut diperoleh dan dalam kondisi apa ia dapat digunakan. Penelitian miskonsepsi yang baik selalu menyisakan ruang bagi penjelasan alternatif dan verifikasi lanjutan.

7.19.7 Penyajian Kutipan Kualitatif

Ketelitian penelitian pada bagian penyajian kutipan kualitatif sangat menentukan apakah pola jawaban dapat ditafsirkan sebagai bukti konseptual. Fokus ini berada dalam komunikasi temuan secara presisi, transparan, tidak menstigma, dan sesuai dengan kekuatan bukti agar dapat digunakan oleh peneliti serta guru. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Levitt et al. (2018) serta Cumming (2014) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Dokumentasi metodologis hendaknya memuat siapa yang melakukan kutipan, kapan data diperoleh, bagaimana tingkat detail dikendalikan, dan atas dasar apa grafik prevalensi disusun. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Contoh pada kajian forest plot intervensi memperlihatkan pentingnya pemisahan antara ketidaktahuan, kesalahan prosedural, dan miskonsepsi. Seorang peserta dapat gagal karena lupa rumus, sedangkan peserta lain memilih jawaban yang sama karena menggunakan model mental yang tidak ilmiah. Perbedaan ini baru terlihat ketika kutipan dibaca bersama peta konsep dan bukti tentang alasan.

Pertanyaan audit metodologis:

- Definisikan unit pelaporan dalam istilah operasional.
- Tentukan bukti utama dari pola respons dan bukti pendukung dari ukuran efek.
- Periksa keterbacaan sebelum membuat kategori atau model.
- Catat potensi grafik terpotong beserta strategi mitigasinya.
- Sajikan matriks lintas representasi bersama batas interpretasi dan ketidakpastian.

Standar metodologis menuntut koherensi antara konstruk, prosedur, analisis, dan klaim. Dalam konteks ini, kelengkapan tidak dapat digantikan oleh indeks tunggal. Demikian pula,

proporsionalitas klaim harus dinilai dengan bukti yang sesuai. Levitt et al. (2018) dan Cumming (2014) memberikan dasar untuk pelaporan yang lengkap dan dapat ditelaah.

Inti metodologis subbagian ini adalah keterlacakan. Pembaca harus dapat mengikuti perjalanan dari pertanyaan, data, pengodean, analisis, sampai grafik prevalensi. Ketika jejak itu jelas, temuan dapat digunakan untuk memperbaiki pembelajaran tanpa mengubah diagnosis konseptual menjadi label bagi peserta didik.

7.19.8 Penyajian Gambar dan Peta Konsep

Peneliti sering memusatkan perhatian pada teknik, padahal penyajian gambar dan peta konsep justru mengatur batas inferensi yang dapat dibuat. Fokus ini berada dalam komunikasi temuan secara presisi, transparan, tidak menstigma, dan sesuai dengan kekuatan bukti agar dapat digunakan oleh peneliti serta guru. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. O'Brien et al. (2014) serta Levitt et al. (2018) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Rancangan yang dapat diaudit mencantumkan alasan pemilihan bentuk visual, prosedur penanganan bahasa ketidakpastian, dan kriteria keputusan untuk audiens. Tanpa penjelasan itu, pembaca sulit menilai apakah pola respons benar-benar mendukung klaim. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Kasus profil kelas dapat digunakan untuk memperlihatkan hubungan antara pola individual dan pola kelas. Sebagian miskonsepsi mungkin hanya dimiliki beberapa peserta tetapi sangat kuat, sedangkan pola lain tersebar luas karena bersumber dari analogi atau buku teks yang sama. Analisis perlu menjaga kedua tingkat tersebut.

Bukti minimum yang perlu tersedia:

- Definisikan bentuk visual dalam istilah operasional.
- Tentukan bukti utama dari peta konsep dan bukti pendukung dari joint display.
- Periksa proporsionalitas klaim sebelum membuat kategori atau model.
- Catat potensi label deficit beserta strategi mitigasinya.
- Sajikan joint display bersama batas interpretasi dan ketidakpastian.

Pemeriksaan kualitas dapat dilakukan melalui audit terhadap keterbacaan, analisis kasus yang bertentangan, dan evaluasi konsistensi angka-narasi. Apabila ditemukan kutipan dipilih sepihak, peneliti perlu menurunkan kekuatan klaim atau menambah bukti. Transparansi seperti ini meningkatkan kegunaan temuan bagi penelitian lanjutan dan praktik kelas.

Kesimpulannya, penyajian gambar dan peta konsep perlu dirancang sebagai proses argumentatif. Data tidak berbicara sendiri; maknanya bergantung pada definisi konstruk, kualitas prosedur, dan konteks. Oleh sebab itu, klaim akhir sebaiknya menggunakan bahasa yang sesuai dengan kekuatan bukti, misalnya “menunjukkan pola”, “mengindikasikan”, atau “mendukung penjelasan”.

7.19.9 Joint Display Mixed Methods

Kajian yang kuat memperlakukan *joint display mixed methods* sebagai bagian dari rantai argumentasi, bukan sebagai unsur pelengkap. Fokus ini berada dalam komunikasi temuan secara presisi, transparan, tidak menstigma, dan sesuai dengan kekuatan bukti agar dapat digunakan oleh peneliti serta guru. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Page et al. (2021) serta Cumming (2014) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Pada tahap perencanaan, tingkat detail dan anonimisasi contoh harus dibedakan secara eksplisit. Pada tahap pelaksanaan, konsistensi peta konsep perlu diperiksa; pada tahap interpretasi, proporsionalitas klaim menjadi dasar pembatasan klaim. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Dalam kelas yang mempelajari infografik untuk guru, data agregat mungkin menunjukkan mayoritas jawaban salah. Akan tetapi, wawancara dapat memperlihatkan beberapa mekanisme berbeda: bahasa sehari-hari, generalisasi aturan, salah membaca simbol, atau model partikel yang tidak tepat. Karena itu, satu persentase tidak boleh menggantikan analisis struktur alasan.

Kriteria pemeriksaan sebelum analisis:

- Definisikan tingkat detail dalam istilah operasional.
- Tentukan bukti utama dari ukuran efek dan bukti pendukung dari prevalensi.
- Periksa konsistensi angka-narasi sebelum membuat kategori atau model.
- Catat potensi diskusi melebihi data beserta strategi mitigasinya.
- Sajikan ringkasan implikasi bersama batas interpretasi dan ketidakpastian.

Kriteria keberhasilan tidak cukup berupa banyaknya data. Yang lebih penting adalah apakah proporsionalitas klaim dapat ditelusuri, apakah aksesibilitas visual diperiksa, dan apakah kelengkapan dijelaskan. Peneliti harus menghindari grafik terpotong; jika risiko itu tidak dapat dihilangkan, dampaknya perlu dianalisis dan dicantumkan sebagai batas penelitian.

Pada akhirnya, keputusan tentang joint display mixed methods menentukan apakah penelitian menghasilkan daftar jawaban salah atau pemahaman yang dapat ditindaklanjuti mengenai cara peserta didik berpikir. Sasaran utamanya ialah bukti yang cukup untuk merancang diagnosis, pembelajaran, dan penelitian lanjutan secara bertanggung jawab.

7.19.10 Penyajian Hasil Eksperimen

Subbagian ini menempatkan penyajian hasil eksperimen sebagai simpul penting dalam penyajian temuan miskonsepsi. Fokus ini berada dalam komunikasi temuan secara presisi, transparan, tidak menstigma, dan sesuai dengan kekuatan bukti agar dapat digunakan oleh peneliti serta guru. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Cumming (2014) serta Levitt et al. (2018) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Secara operasional, peneliti perlu menetapkan bahasa ketidakpastian, menjelaskan audiens, dan mendokumentasikan bentuk visual. Keputusan tersebut kemudian diterjemahkan ke dalam pilihan ukuran efek, prevalensi, serta pola respons. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Sebagai ilustrasi, pada topik tabel miskonsepsi asam-basa, peserta didik dapat memberikan jawaban benar pada tingkat simbolik tetapi alasan submikroskopiknya keliru. Peneliti tidak cukup menghitung benar-salah; ia perlu membandingkan ukuran efek dengan prevalensi, memeriksa tingkat keyakinan, dan menguji konsistensi pada konteks lain. Hasil analisis mungkin menunjukkan bahwa kesalahan hanya muncul pada satu representasi, atau justru merupakan hubungan konseptual yang stabil.

Langkah operasional yang disarankan:

- Definisikan bahasa ketidakpastian dalam istilah operasional.
- Tentukan bukti utama dari joint display dan bukti pendukung dari kutipan.
- Periksa aksesibilitas visual sebelum membuat kategori atau model.
- Catat potensi judul sensasional beserta strategi mitigasinya.
- Sajikan tabel kategori bersama batas interpretasi dan ketidakpastian.

Mutu inferensi terutama ditentukan oleh konsistensi angka-narasi, akurasi, dan keterbacaan. Ancaman utama ialah label deficit serta persentase tanpa denominator. Karena itu, laporan perlu menyajikan bukti yang mendukung sekaligus bukti yang membatasi interpretasi. Prinsip ini konsisten dengan penekanan Cumming (2014) dan Levitt et al. (2018) bahwa validitas

melekat pada interpretasi dan penggunaan hasil, bukan sekadar pada nama instrumen atau desain.

Dengan demikian, penyajian hasil eksperimen berfungsi menghubungkan keputusan desain dengan interpretasi. Produk yang diharapkan bukan hanya joint display, tetapi juga penjelasan mengenai bagaimana produk tersebut diperoleh dan dalam kondisi apa ia dapat digunakan. Penelitian miskonsepsi yang baik selalu menyisakan ruang bagi penjelasan alternatif dan verifikasi lanjutan.

7.19.11 Penyajian SLR dan Meta-Analisis

Pembahasan mengenai penyajian slr dan meta-analisis diperlukan agar penyajian temuan miskonsepsi tidak berhenti pada prosedur administratif. Fokus ini berada dalam komunikasi temuan secara presisi, transparan, tidak menstigma, dan sesuai dengan kekuatan bukti agar dapat digunakan oleh peneliti serta guru. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. BERA (2024) serta Cumming (2014) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Penerapan praktis dapat disusun dalam urutan: mendefinisikan anonimisasi contoh, mengoperasionalkan unit pelaporan, mengumpulkan joint display, memeriksa aksesibilitas visual, lalu melaporkan ringkasan implikasi. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

$$Q = \sum w_i(\theta_i - \bar{\theta})^2; I^2 = \max[0, (Q - df)/Q] \times 100\%$$

Misalnya, pada heatmap butir redoks, perubahan skor sesudah pembelajaran belum otomatis menandakan perubahan konsep. Peserta didik mungkin mempelajari pola jawaban tes. Bukti yang lebih meyakinkan mencakup transfer ke situasi baru, penjelasan kausal, representasi partikel, dan retensi setelah jeda waktu.

Keputusan yang harus didokumentasikan:

- Definisikan anonimisasi contoh dalam istilah operasional.
- Tentukan bukti utama dari prevalensi dan bukti pendukung dari pola respons.
- Periksa akurasi sebelum membuat kategori atau model.
- Catat potensi persentase tanpa denominator beserta strategi mitigasinya.
- Sajikan grafik prevalensi bersama batas interpretasi dan ketidakpastian.

Temuan yang baik tidak selalu berarti hasil yang besar atau sesuai harapan. Temuan bernilai ketika aksesibilitas visual terjaga, kelengkapan dijelaskan, dan ketidakpastian disajikan. Kerangka BERA (2024) serta Cumming (2014) menolong peneliti membedakan kekuatan bukti dari daya tarik narasi.

Inti metodologis subbagian ini adalah keterlacakan. Pembaca harus dapat mengikuti perjalanan dari pertanyaan, data, pengodean, analisis, sampai ringkasan implikasi. Ketika jejak itu jelas, temuan dapat digunakan untuk memperbaiki pembelajaran tanpa mengubah diagnosis konseptual menjadi label bagi peserta didik.

7.19.12 Bahasa Ketidakpastian

Dalam metodologi penelitian miskonsepsi kimia, bahasa ketidakpastian menentukan mutu hubungan antara data dan kesimpulan. Fokus ini berada dalam komunikasi temuan secara presisi, transparan, tidak menstigma, dan sesuai dengan kekuatan bukti agar dapat digunakan oleh peneliti serta guru. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Appelbaum et al. (2018) serta Levitt et al. (2018) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Prosedur minimal mencakup definisi audiens, alasan memilih prevalensi, aturan mengubah respons menjadi kategori, serta mekanisme pemeriksaan akurasi. Semua penyimpangan dari rencana awal perlu dicatat. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Pada penelitian mengenai alur perubahan konsepsi, peneliti dapat menyusun beberapa tugas ekuivalen yang berubah konteks, representasi, atau tuntutan penalarannya. Jika pola respons tetap berulang dan disertai keyakinan tinggi, terdapat dasar yang lebih kuat untuk menyebut indikasi miskonsepsi. Jika pola berubah-ubah, interpretasi yang lebih hati-hati adalah pengetahuan parsial atau ketergantungan konteks.

Pertanyaan audit metodologis:

- Definisikan audiens dalam istilah operasional.
- Tentukan bukti utama dari kutipan dan bukti pendukung dari peta konsep.
- Periksa kelengkapan sebelum membuat kategori atau model.
- Catat potensi kutipan dipilih sepihak beserta strategi mitigasinya.
- Sajikan kutipan analitik bersama batas interpretasi dan ketidakpastian.

Pemeriksaan mutu hendaknya dilakukan selama penelitian, bukan hanya pada akhir. Catatan tentang akurasi, keputusan terkait keterbacaan, dan kejadian yang berpotensi menimbulkan judul sensasional perlu dipelihara dalam *audit trail*. Hal ini memungkinkan koreksi dini dan pelaporan yang jujur.

Kesimpulannya, bahasa ketidakpastian perlu dirancang sebagai proses argumentatif. Data tidak berbicara sendiri; maknanya bergantung pada definisi konstruk, kualitas prosedur, dan konteks. Oleh sebab itu, klaim akhir sebaiknya menggunakan bahasa yang sesuai dengan kekuatan bukti, misalnya “menunjukkan pola”, “mengindikasikan”, atau “mendukung penjelasan”.

7.19.13 Bahasa Nonstigmatisasi

Keputusan tentang bahasa nonstigmatisasi harus dibuat sebelum data utama dikumpulkan, bukan ditambahkan setelah hasil diketahui. Fokus ini berada dalam komunikasi temuan secara presisi, transparan, tidak menstigma, dan sesuai dengan kekuatan bukti agar dapat digunakan oleh peneliti serta guru. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Levitt et al. (2018) serta Cumming (2014) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Tahap kerja dimulai dengan penegasan unit pelaporan. Setelah itu, peneliti menyusun aturan untuk tingkat detail dan menentukan bagaimana anonimisasi contoh akan diperiksa. Bukti sebaiknya berasal dari lebih dari satu bentuk, misalnya kutipan dan peta konsep. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Pada topik peta jaringan ikatan, peneliti juga perlu memperhatikan bahasa. Istilah yang sama dapat diberi makna berbeda oleh peserta didik. Klarifikasi melalui pertanyaan tindak lanjut, gambar, atau contoh tandingan membantu membedakan masalah terminologi dari kesalahan konseptual.

Bukti minimum yang perlu tersedia:

- Definisikan unit pelaporan dalam istilah operasional.
- Tentukan bukti utama dari pola respons dan bukti pendukung dari ukuran efek.
- Periksa keterbacaan sebelum membuat kategori atau model.
- Catat potensi grafik terpotong beserta strategi mitigasinya.
- Sajikan matriks lintas representasi bersama batas interpretasi dan ketidakpastian.

Kualitas penelitian meningkat ketika peneliti menunjukkan rantai bukti dari kutipan menuju grafik prevalensi, menyatakan asumsi, dan memeriksa alternatif penjelasan. Sebaliknya, persentase tanpa denominator dapat menghasilkan klaim yang tampak tegas tetapi tidak stabil. Rekomendasi praktis harus proporsional dengan bukti.

Pada akhirnya, keputusan tentang bahasa nonstigmatisasi menentukan apakah penelitian menghasilkan daftar jawaban salah atau pemahaman yang dapat ditindaklanjuti mengenai cara peserta didik berpikir. Sasaran utamanya ialah bukti yang cukup untuk merancang diagnosis, pembelajaran, dan penelitian lanjutan secara bertanggung jawab.

7.19.14 Pelaporan Sesuai Pedoman

Pelaporan sesuai pedoman perlu dipahami sebagai keputusan metodologis yang mempunyai konsekuensi pada desain, analisis, dan pelaporan. Fokus ini berada dalam komunikasi temuan secara presisi, transparan, tidak menstigma, dan sesuai dengan kekuatan bukti agar dapat digunakan oleh peneliti serta guru. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. O'Brien et al. (2014) serta Levitt et al. (2018) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Peneliti sebaiknya menyiapkan matriks yang menghubungkan pertanyaan penelitian dengan bentuk visual, pola respons, teknik analisis, dan kutipan analitik. Matriks ini mencegah pengumpulan data yang banyak tetapi tidak relevan. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Ilustrasi kutipan wawancara kesetimbangan juga menunjukkan bahwa kualitas instrumen memengaruhi temuan. Distraktor yang tidak berasal dari pemikiran peserta didik dapat menghasilkan pilihan acak, sedangkan pertanyaan yang terlalu menuntut matematika dapat mengukur keterampilan hitung alih-alih konsepsi. Peneliti perlu menilai kesesuaian bukti dengan konstruk yang dimaksud.

Kriteria pemeriksaan sebelum analisis:

- Definisikan bentuk visual dalam istilah operasional.
- Tentukan bukti utama dari peta konsep dan bukti pendukung dari joint display.
- Periksa proporsionalitas klaim sebelum membuat kategori atau model.
- Catat potensi label deficit beserta strategi mitigasinya.
- Sajikan joint display bersama batas interpretasi dan ketidakpastian.

Peneliti sebaiknya memisahkan keputusan apriori dari keputusan eksploratori. Pemisahan tersebut membantu mengendalikan kutipan dipilih sepihak dan mencegah diskusi melebihi data. Selanjutnya, keterbacaan dan konsistensi angka-narasi dilaporkan bersama ketidakpastian, sehingga pembaca dapat menilai kekuatan temuan tanpa bergantung pada kesimpulan penulis semata.

Dengan demikian, pelaporan sesuai pedoman berfungsi menghubungkan keputusan desain dengan interpretasi. Produk yang diharapkan bukan hanya kutipan analitik, tetapi juga penjelasan mengenai bagaimana produk tersebut diperoleh dan dalam kondisi apa ia dapat digunakan. Penelitian miskonsepsi yang baik selalu menyisakan ruang bagi penjelasan alternatif dan verifikasi lanjutan.

7.19.15 Sintesis Struktur Pelaporan Temuan

Ketelitian penelitian pada bagian sintesis struktur pelaporan temuan sangat menentukan apakah pola jawaban dapat ditafsirkan sebagai bukti konseptual. Fokus ini berada dalam komunikasi temuan secara presisi, transparan, tidak menstigma, dan sesuai dengan kekuatan bukti agar dapat digunakan oleh peneliti serta guru. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Page et al. (2021) serta Cumming (2014) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Dokumentasi metodologis hendaknya memuat siapa yang melakukan peta konsep, kapan data diperoleh, bagaimana anonimisasi contoh dikendalikan, dan atas dasar apa matriks lintas representasi disusun. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Contoh pada kajian forest plot intervensi memperlihatkan pentingnya pemisahan antara ketidaktahuan, kesalahan prosedural, dan miskonsepsi. Seorang peserta dapat gagal karena lupa rumus, sedangkan peserta lain memilih jawaban yang sama karena menggunakan model mental yang tidak ilmiah. Perbedaan ini baru terlihat ketika peta konsep dibaca bersama joint display dan bukti tentang alasan.

Langkah operasional yang disarankan:

- Definisikan tingkat detail dalam istilah operasional.
- Tentukan bukti utama dari ukuran efek dan bukti pendukung dari prevalensi.
- Periksa konsistensi angka-narasi sebelum membuat kategori atau model.
- Catat potensi diskusi melebihi data beserta strategi mitigasinya.

- Sajikan ringkasan implikasi bersama batas interpretasi dan ketidakpastian.

Standar metodologis menuntut koherensi antara konstruk, prosedur, analisis, dan klaim. Dalam konteks ini, proporsionalitas klaim tidak dapat digantikan oleh indeks tunggal. Demikian pula, aksesibilitas visual harus dinilai dengan bukti yang sesuai. Page et al. (2021) dan Cumming (2014) memberikan dasar untuk pelaporan yang lengkap dan dapat ditelaah.

Inti metodologis subbagian ini adalah keterlacakan. Pembaca harus dapat mengikuti perjalanan dari pertanyaan, data, pengodean, analisis, sampai matriks lintas representasi. Ketika jejak itu jelas, temuan dapat digunakan untuk memperbaiki pembelajaran tanpa mengubah diagnosis konseptual menjadi label bagi peserta didik.

7.20 Keterbatasan Penelitian Miskonsepsi

Subbab ini membahas identifikasi sistematis terhadap batas desain, instrumen, sampel, analisis, konteks, dan interpretasi agar pembaca memahami apa yang dapat dan tidak dapat disimpulkan. Fokus utamanya adalah menjaga keselarasan antara pertanyaan penelitian, bentuk bukti, teknik analisis, dan kekuatan klaim. Dalam konteks miskonsepsi kimia, keputusan metodologis harus mampu membedakan pola konseptual dari respons yang timbul karena keterbatasan bahasa, kesalahan hitung, atau kondisi administrasi. Pembahasan berikut dipecah menjadi lima belas bagian agar setiap keputusan dapat ditelaah secara operasional.

7.20.1 Perbedaan Keterbatasan dan Delimitasi

Peneliti sering memusatkan perhatian pada teknik, padahal perbedaan keterbatasan dan delimitasi justru mengatur batas inferensi yang dapat dibuat. Fokus ini berada dalam identifikasi sistematis terhadap batas desain, instrumen, sampel, analisis, konteks, dan interpretasi agar pembaca memahami apa yang dapat dan tidak dapat disimpulkan. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Page et al. (2021) serta Appelbaum et al. (2018) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Rancangan yang dapat diaudit mencantumkan alasan pemilihan alternatif penjelasan, prosedur penanganan sumber keterbatasan, dan kriteria keputusan untuk analisis sensitivitas. Tanpa penjelasan itu, pembaca sulit menilai apakah karakteristik sampel benar-benar mendukung klaim. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Kasus data keyakinan *self-report* dapat digunakan untuk memperlihatkan hubungan antara pola individual dan pola kelas. Sebagian miskonsepsi mungkin hanya dimiliki beberapa peserta tetapi sangat kuat, sedangkan pola lain tersebar luas karena bersumber dari analogi atau buku teks yang sama. Analisis perlu menjaga kedua tingkat tersebut.

Keputusan yang harus didokumentasikan:

- Definisikan alternatif penjelasan dalam istilah operasional.
- Tentukan bukti utama dari kasus bertentangan dan bukti pendukung dari profil missing data.
- Periksa rencana mitigasi sebelum membuat kategori atau model.
- Catat potensi daftar klise beserta strategi mitigasinya.
- Sajikan matriks keterbatasan bersama batas interpretasi dan ketidakpastian.

Pemeriksaan kualitas dapat dilakukan melalui audit terhadap proporsi, analisis kasus yang bertentangan, dan evaluasi kejujuran pelaporan. Apabila ditemukan mengabaikan bias publikasi, peneliti perlu menurunkan kekuatan klaim atau menambah bukti. Transparansi seperti ini meningkatkan kegunaan temuan bagi penelitian lanjutan dan praktik kelas.

Kesimpulannya, perbedaan keterbatasan dan delimitasi perlu dirancang sebagai proses argumentatif. Data tidak berbicara sendiri; maknanya bergantung pada definisi konstruk, kualitas prosedur, dan konteks. Oleh sebab itu, klaim akhir sebaiknya menggunakan bahasa yang sesuai dengan kekuatan bukti, misalnya “menunjukkan pola”, “mengindikasikan”, atau “mendukung penjelasan”.

7.20.2 Keterbatasan Konstruk

Kajian yang kuat memperlakukan keterbatasan konstruk sebagai bagian dari rantai argumentasi, bukan sebagai unsur pelengkap. Fokus ini berada dalam identifikasi sistematis terhadap batas desain, instrumen, sampel, analisis, konteks, dan interpretasi agar pembaca memahami apa yang dapat dan tidak dapat disimpulkan. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Messick (1995) serta Page et al. (2021) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Pada tahap perencanaan, agenda perbaikan dan tingkat dampak harus dibedakan secara eksplisit. Pada tahap pelaksanaan, konsistensi kasus bertentangan perlu diperiksa; pada tahap interpretasi, rencana mitigasi menjadi dasar pembatasan klaim. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Dalam kelas yang mempelajari hasil tidak konsisten lintas representasi, data agregat mungkin menunjukkan mayoritas jawaban salah. Akan tetapi, wawancara dapat memperlihatkan beberapa mekanisme berbeda: bahasa sehari-hari, generalisasi aturan, salah membaca simbol, atau model partikel yang tidak tepat. Karena itu, satu persentase tidak boleh menggantikan analisis struktur alasan.

Pertanyaan audit metodologis:

- Definisikan agenda perbaikan dalam istilah operasional.
- Tentukan bukti utama dari catatan deviasi dan bukti pendukung dari fidelitas.
- Periksa kejujuran pelaporan sebelum membuat kategori atau model.
- Catat potensi menyembunyikan kelemahan utama beserta strategi mitigasinya.
- Sajikan analisis sensitivitas bersama batas interpretasi dan ketidakpastian.

Kriteria keberhasilan tidak cukup berupa banyaknya data. Yang lebih penting adalah apakah rencana mitigasi dapat ditelusuri, apakah spesifisitas diperiksa, dan apakah analisis dampak dijelaskan. Peneliti harus menghindari tidak membedakan limitasi dan delimitasi; jika risiko itu tidak dapat dihilangkan, dampaknya perlu dianalisis dan dicantumkan sebagai batas penelitian.

Pada akhirnya, keputusan tentang keterbatasan konstruk menentukan apakah penelitian menghasilkan daftar jawaban salah atau pemahaman yang dapat ditindaklanjuti mengenai cara peserta didik berpikir. Sasaran utamanya ialah bukti yang cukup untuk merancang diagnosis, pembelajaran, dan penelitian lanjutan secara bertanggung jawab.

7.20.3 Keterbatasan Instrumen

Subbagian ini menempatkan keterbatasan instrumen sebagai simpul penting dalam keterbatasan penelitian miskonsepsi. Fokus ini berada dalam identifikasi sistematis terhadap batas desain, instrumen, sampel, analisis, konteks, dan interpretasi agar pembaca memahami apa yang dapat dan tidak dapat disimpulkan. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. AERA, APA, dan NCME (2014) serta Appelbaum et al. (2018) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Secara operasional, peneliti perlu menetapkan sumber keterbatasan, menjelaskan analisis sensitivitas, dan mendokumentasikan alternatif penjelasan. Keputusan tersebut kemudian diterjemahkan ke dalam pilihan catatan deviasi, fidelitas, serta karakteristik sampel. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Sebagai ilustrasi, pada topik tes hanya satu topik, peserta didik dapat memberikan jawaban benar pada tingkat simbolik tetapi alasan submikroskopiknya keliru. Peneliti tidak cukup menghitung benar-salah; ia perlu membandingkan catatan deviasi dengan fidelitas, memeriksa tingkat keyakinan, dan menguji konsistensi pada konteks lain. Hasil analisis mungkin menunjukkan bahwa kesalahan hanya muncul pada satu representasi, atau justru merupakan hubungan konseptual yang stabil.

Bukti minimum yang perlu tersedia:

- Definisikan sumber keterbatasan dalam istilah operasional.
- Tentukan bukti utama dari profil missing data dan bukti pendukung dari reliabilitas.
- Periksa spesifisitas sebelum membuat kategori atau model.
- Catat potensi menyatakan semua keterbatasan ringan beserta strategi mitigasinya.
- Sajikan batas inferensi bersama batas interpretasi dan ketidakpastian.

Mutu inferensi terutama ditentukan oleh kejujuran pelaporan, hubungan dengan klaim, dan proporsi. Ancaman utama ialah daftar klise serta menggunakan keterbatasan untuk membenarkan hasil. Karena itu, laporan perlu menyajikan bukti yang mendukung sekaligus bukti yang membatasi interpretasi. Prinsip ini konsisten dengan penekanan AERA, APA, dan NCME (2014) dan Appelbaum et al. (2018) bahwa validitas melekat pada interpretasi dan penggunaan hasil, bukan sekadar pada nama instrumen atau desain.

Dengan demikian, keterbatasan instrumen berfungsi menghubungkan keputusan desain dengan interpretasi. Produk yang diharapkan bukan hanya matriks keterbatasan, tetapi juga penjelasan mengenai bagaimana produk tersebut diperoleh dan dalam kondisi apa ia dapat digunakan. Penelitian miskonsepsi yang baik selalu menyisakan ruang bagi penjelasan alternatif dan verifikasi lanjutan.

7.20.4 Keterbatasan Sampel

Pembahasan mengenai keterbatasan sampel diperlukan agar keterbatasan penelitian miskonsepsi tidak berhenti pada prosedur administratif. Fokus ini berada dalam identifikasi sistematis terhadap batas desain, instrumen, sampel, analisis, konteks, dan interpretasi agar pembaca memahami apa yang dapat dan tidak dapat disimpulkan. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Appelbaum et al. (2018) serta Page et al. (2021) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Penerapan praktis dapat disusun dalam urutan: mendefinisikan tingkat dampak, mengoperasionalkan batas generalisasi, mengumpulkan profil missing data, memeriksa spesifisitas, lalu melaporkan analisis sensitivitas. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data

hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Misalnya, pada sampel satu sekolah, perubahan skor sesudah pembelajaran belum otomatis menandakan perubahan konsep. Peserta didik mungkin mempelajari pola jawaban tes. Bukti yang lebih meyakinkan mencakup transfer ke situasi baru, penjelasan kausal, representasi partikel, dan retensi setelah jeda waktu.

Kriteria pemeriksaan sebelum analisis:

- Definisikan tingkat dampak dalam istilah operasional.
- Tentukan bukti utama dari fidelitas dan bukti pendukung dari karakteristik sampel.
- Periksa hubungan dengan klaim sebelum membuat kategori atau model.
- Catat potensi menggunakan keterbatasan untuk membenarkan hasil beserta strategi mitigasinya.
- Sajikan prioritas replikasi bersama batas interpretasi dan ketidakpastian.

Temuan yang baik tidak selalu berarti hasil yang besar atau sesuai harapan. Temuan bernilai ketika spesifisitas terjaga, analisis dampak dijelaskan, dan ketidakpastian disajikan. Kerangka Appelbaum et al. (2018) serta Page et al. (2021) menolong peneliti membedakan kekuatan bukti dari daya tarik narasi.

Inti metodologis subbagian ini adalah keterlacakan. Pembaca harus dapat mengikuti perjalanan dari pertanyaan, data, pengodean, analisis, sampai analisis sensitivitas. Ketika jejak itu jelas, temuan dapat digunakan untuk memperbaiki pembelajaran tanpa mengubah diagnosis konseptual menjadi label bagi peserta didik.

7.20.5 Keterbatasan Desain Deskriptif

Dalam metodologi penelitian miskonsepsi kimia, keterbatasan desain deskriptif menentukan mutu hubungan antara data dan kesimpulan. Fokus ini berada dalam identifikasi sistematis terhadap batas desain, instrumen, sampel, analisis, konteks, dan interpretasi agar pembaca memahami apa yang dapat dan tidak dapat disimpulkan. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Shadish et al. (2002) serta Appelbaum et al. (2018) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Prosedur minimal mencakup definisi analisis sensitivitas, alasan memilih fidelitas, aturan mengubah respons menjadi kategori, serta mekanisme pemeriksaan hubungan dengan klaim. Semua penyimpangan dari rencana awal perlu dicatat. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna

untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Pada penelitian mengenai wawancara setelah tes, peneliti dapat menyusun beberapa tugas ekuivalen yang berubah konteks, representasi, atau tuntutan penalarannya. Jika pola respons tetap berulang dan disertai keyakinan tinggi, terdapat dasar yang lebih kuat untuk menyebut indikasi miskonsepsi. Jika pola berubah-ubah, interpretasi yang lebih hati-hati adalah pengetahuan parsial atau ketergantungan konteks.

Langkah operasional yang disarankan:

- Definisikan analisis sensitivitas dalam istilah operasional.
- Tentukan bukti utama dari reliabilitas dan bukti pendukung dari kasus bertentangan.
- Periksa analisis dampak sebelum membuat kategori atau model.
- Catat potensi mengabaikan bias publikasi beserta strategi mitigasinya.
- Sajikan rekomendasi desain bersama batas interpretasi dan ketidakpastian.

Pemeriksaan mutu hendaknya dilakukan selama penelitian, bukan hanya pada akhir. Catatan tentang hubungan dengan klaim, keputusan terkait proporsi, dan kejadian yang berpotensi menimbulkan menyatakan semua keterbatasan ringan perlu dipelihara dalam *audit trail*. Hal ini memungkinkan koreksi dini dan pelaporan yang jujur.

Kesimpulannya, keterbatasan desain deskriptif perlu dirancang sebagai proses argumentatif. Data tidak berbicara sendiri; maknanya bergantung pada definisi konstruk, kualitas prosedur, dan konteks. Oleh sebab itu, klaim akhir sebaiknya menggunakan bahasa yang sesuai dengan kekuatan bukti, misalnya “menunjukkan pola”, “mengindikasikan”, atau “mendukung penjelasan”.

7.20.6 Keterbatasan Eksperimen dan Kuasi-Eksperimen

Keputusan tentang keterbatasan eksperimen dan kuasi-eksperimen harus dibuat sebelum data utama dikumpulkan, bukan ditambahkan setelah hasil diketahui. Fokus ini berada dalam identifikasi sistematis terhadap batas desain, instrumen, sampel, analisis, konteks, dan interpretasi agar pembaca memahami apa yang dapat dan tidak dapat disimpulkan. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Levitt et al. (2018) serta Page et al. (2021) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Tahap kerja dimulai dengan penegasan batas generalisasi. Setelah itu, peneliti menyusun aturan untuk agenda perbaikan dan menentukan bagaimana tingkat dampak akan diperiksa. Bukti sebaiknya berasal dari lebih dari satu bentuk, misalnya reliabilitas dan kasus bertentangan. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama

penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Pada topik intervensi dua minggu, peneliti juga perlu memperhatikan bahasa. Istilah yang sama dapat diberi makna berbeda oleh peserta didik. Klarifikasi melalui pertanyaan tindak lanjut, gambar, atau contoh tandingan membantu membedakan masalah terminologi dari kesalahan konseptual.

Keputusan yang harus didokumentasikan:

- Definisikan batas generalisasi dalam istilah operasional.
- Tentukan bukti utama dari karakteristik sampel dan bukti pendukung dari catatan deviasi.
- Periksa proporsi sebelum membuat kategori atau model.
- Catat potensi tidak membedakan limitasi dan delimitasi beserta strategi mitigasinya.
- Sajikan agenda penelitian bersama batas interpretasi dan ketidakpastian.

Kualitas penelitian meningkat ketika peneliti menunjukkan rantai bukti dari reliabilitas menuju prioritas replikasi, menyatakan asumsi, dan memeriksa alternatif penjelasan. Sebaliknya, menggunakan keterbatasan untuk membenarkan hasil dapat menghasilkan klaim yang tampak tegas tetapi tidak stabil. Rekomendasi praktis harus proporsional dengan bukti.

Pada akhirnya, keputusan tentang keterbatasan eksperimen dan kuasi-eksperimen menentukan apakah penelitian menghasilkan daftar jawaban salah atau pemahaman yang dapat ditindaklanjuti mengenai cara peserta didik berpikir. Sasaran utamanya ialah bukti yang cukup untuk merancang diagnosis, pembelajaran, dan penelitian lanjutan secara bertanggung jawab.

7.20.7 Keterbatasan Data Kualitatif

Keterbatasan data kualitatif perlu dipahami sebagai keputusan metodologis yang mempunyai konsekuensi pada desain, analisis, dan pelaporan. Fokus ini berada dalam identifikasi sistematis terhadap batas desain, instrumen, sampel, analisis, konteks, dan interpretasi agar pembaca memahami apa yang dapat dan tidak dapat disimpulkan. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Page et al. (2021) serta Appelbaum et al. (2018) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Peneliti sebaiknya menyiapkan matriks yang menghubungkan pertanyaan penelitian dengan alternatif penjelasan, karakteristik sampel, teknik analisis, dan rekomendasi desain. Matriks ini mencegah pengumpulan data yang banyak tetapi tidak relevan. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya

berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Ilustrasi attrition tinggi juga menunjukkan bahwa kualitas instrumen memengaruhi temuan. Distraktor yang tidak berasal dari pemikiran peserta didik dapat menghasilkan pilihan acak, sedangkan pertanyaan yang terlalu menuntut matematika dapat mengukur keterampilan hitung alih-alih konsepsi. Peneliti perlu menilai kesesuaian bukti dengan konstruk yang dimaksud.

Pertanyaan audit metodologis:

- Definisikan alternatif penjelasan dalam istilah operasional.
- Tentukan bukti utama dari kasus bertentangan dan bukti pendukung dari profil missing data.
- Periksa rencana mitigasi sebelum membuat kategori atau model.
- Catat potensi daftar klise beserta strategi mitigasinya.
- Sajikan matriks keterbatasan bersama batas interpretasi dan ketidakpastian.

Peneliti sebaiknya memisahkan keputusan apriori dari keputusan eksploratori. Pemisahan tersebut membantu mengendalikan mengabaikan bias publikasi dan mencegah menyembunyikan kelemahan utama. Selanjutnya, proporsi dan kejujuran pelaporan dilaporkan bersama ketidakpastian, sehingga pembaca dapat menilai kekuatan temuan tanpa bergantung pada kesimpulan penulis semata.

Dengan demikian, keterbatasan data kualitatif berfungsi menghubungkan keputusan desain dengan interpretasi. Produk yang diharapkan bukan hanya rekomendasi desain, tetapi juga penjelasan mengenai bagaimana produk tersebut diperoleh dan dalam kondisi apa ia dapat digunakan. Penelitian miskonsepsi yang baik selalu menyisakan ruang bagi penjelasan alternatif dan verifikasi lanjutan.

7.20.8 Keterbatasan Mixed Methods

Ketelitian penelitian pada bagian keterbatasan *mixed methods* sangat menentukan apakah pola jawaban dapat ditafsirkan sebagai bukti konseptual. Fokus ini berada dalam identifikasi sistematis terhadap batas desain, instrumen, sampel, analisis, konteks, dan interpretasi agar pembaca memahami apa yang dapat dan tidak dapat disimpulkan. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Messick (1995) serta Page et al. (2021) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Dokumentasi metodologis hendaknya memuat siapa yang melakukan kasus bertentangan, kapan data diperoleh, bagaimana tingkat dampak dikendalikan, dan atas dasar apa agenda penelitian disusun. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan

pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Contoh pada kajian instrumen belum invarians memperlihatkan pentingnya pemisahan antara ketidaktahuan, kesalahan prosedural, dan miskonsepsi. Seorang peserta dapat gagal karena lupa rumus, sedangkan peserta lain memilih jawaban yang sama karena menggunakan model mental yang tidak ilmiah. Perbedaan ini baru terlihat ketika kasus bertentangan dibaca bersama profil missing data dan bukti tentang alasan.

Bukti minimum yang perlu tersedia:

- Definiskan agenda perbaikan dalam istilah operasional.
- Tentukan bukti utama dari catatan deviasi dan bukti pendukung dari fidelitas.
- Periksa kejujuran pelaporan sebelum membuat kategori atau model.
- Catat potensi menyembunyikan kelemahan utama beserta strategi mitigasinya.
- Sajikan analisis sensitivitas bersama batas interpretasi dan ketidakpastian.

Standar metodologis menuntut koherensi antara konstruk, prosedur, analisis, dan klaim. Dalam konteks ini, rencana mitigasi tidak dapat digantikan oleh indeks tunggal. Demikian pula, spesifisitas harus dinilai dengan bukti yang sesuai. Messick (1995) dan Page et al. (2021) memberikan dasar untuk pelaporan yang lengkap dan dapat ditelaah.

Inti metodologis subbagian ini adalah keterlacakan. Pembaca harus dapat mengikuti perjalanan dari pertanyaan, data, pengodean, analisis, sampai agenda penelitian. Ketika jejak itu jelas, temuan dapat digunakan untuk memperbaiki pembelajaran tanpa mengubah diagnosis konseptual menjadi label bagi peserta didik.

7.20.9 Keterbatasan SLR

Peneliti sering memusatkan perhatian pada teknik, padahal keterbatasan slr justru mengatur batas inferensi yang dapat dibuat. Fokus ini berada dalam identifikasi sistematis terhadap batas desain, instrumen, sampel, analisis, konteks, dan interpretasi agar pembaca memahami apa yang dapat dan tidak dapat disimpulkan. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. AERA, APA, dan NCME (2014) serta Appelbaum et al. (2018) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Rancangan yang dapat diaudit mencantumkan alasan pemilihan sumber keterbatasan, prosedur penanganan analisis sensitivitas, dan kriteria keputusan untuk alternatif penjelasan. Tanpa penjelasan itu, pembaca sulit menilai apakah catatan deviasi benar-benar mendukung klaim. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean,

penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Kasus data keyakinan *self-report* dapat digunakan untuk memperlihatkan hubungan antara pola individual dan pola kelas. Sebagian miskonsepsi mungkin hanya dimiliki beberapa peserta tetapi sangat kuat, sedangkan pola lain tersebar luas karena bersumber dari analogi atau buku teks yang sama. Analisis perlu menjaga kedua tingkat tersebut.

Kriteria pemeriksaan sebelum analisis:

- Definisikan sumber keterbatasan dalam istilah operasional.
- Tentukan bukti utama dari profil missing data dan bukti pendukung dari reliabilitas.
- Periksa spesifisitas sebelum membuat kategori atau model.
- Catat potensi menyatakan semua keterbatasan ringan beserta strategi mitigasinya.
- Sajikan batas inferensi bersama batas interpretasi dan ketidakpastian.

Pemeriksaan kualitas dapat dilakukan melalui audit terhadap kejujuran pelaporan, analisis kasus yang bertentangan, dan evaluasi hubungan dengan klaim. Apabila ditemukan daftar klise, peneliti perlu menurunkan kekuatan klaim atau menambah bukti. Transparansi seperti ini meningkatkan kegunaan temuan bagi penelitian lanjutan dan praktik kelas.

Kesimpulannya, keterbatasan slr perlu dirancang sebagai proses argumentatif. Data tidak berbicara sendiri; maknanya bergantung pada definisi konstruk, kualitas prosedur, dan konteks. Oleh sebab itu, klaim akhir sebaiknya menggunakan bahasa yang sesuai dengan kekuatan bukti, misalnya “menunjukkan pola”, “mengindikasikan”, atau “mendukung penjelasan”.

7.20.10 Keterbatasan Analisis Statistik

Kajian yang kuat memperlakukan keterbatasan analisis statistik sebagai bagian dari rantai argumentasi, bukan sebagai unsur pelengkap. Fokus ini berada dalam identifikasi sistematis terhadap batas desain, instrumen, sampel, analisis, konteks, dan interpretasi agar pembaca memahami apa yang dapat dan tidak dapat disimpulkan. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Appelbaum et al. (2018) serta Page et al. (2021) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Pada tahap perencanaan, tingkat dampak dan batas generalisasi harus dibedakan secara eksplisit. Pada tahap pelaksanaan, konsistensi profil missing data perlu diperiksa; pada tahap interpretasi, spesifisitas menjadi dasar pembatasan klaim. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur

data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Dalam kelas yang mempelajari hasil tidak konsisten lintas representasi, data agregat mungkin menunjukkan mayoritas jawaban salah. Akan tetapi, wawancara dapat memperlihatkan beberapa mekanisme berbeda: bahasa sehari-hari, generalisasi aturan, salah membaca simbol, atau model partikel yang tidak tepat. Karena itu, satu persentase tidak boleh menggantikan analisis struktur alasan.

Langkah operasional yang disarankan:

- Definisikan tingkat dampak dalam istilah operasional.
- Tentukan bukti utama dari fidelitas dan bukti pendukung dari karakteristik sampel.
- Periksa hubungan dengan klaim sebelum membuat kategori atau model.
- Catat potensi menggunakan keterbatasan untuk membenarkan hasil beserta strategi mitigasinya.
- Sajikan prioritas replikasi bersama batas interpretasi dan ketidakpastian.

Kriteria keberhasilan tidak cukup berupa banyaknya data. Yang lebih penting adalah apakah spesifisitas dapat ditelusuri, apakah analisis dampak diperiksa, dan apakah rencana mitigasi dijelaskan. Peneliti harus menghindari menyembunyikan kelemahan utama; jika risiko itu tidak dapat dihilangkan, dampaknya perlu dianalisis dan dicantumkan sebagai batas penelitian.

Pada akhirnya, keputusan tentang keterbatasan analisis statistik menentukan apakah penelitian menghasilkan daftar jawaban salah atau pemahaman yang dapat ditindaklanjuti mengenai cara peserta didik berpikir. Sasaran utamanya ialah bukti yang cukup untuk merancang diagnosis, pembelajaran, dan penelitian lanjutan secara bertanggung jawab.

7.20.11 Keterbatasan Interpretasi Keyakinan

Subbagian ini menempatkan keterbatasan interpretasi keyakinan sebagai simpul penting dalam keterbatasan penelitian miskonsepsi. Fokus ini berada dalam identifikasi sistematis terhadap batas desain, instrumen, sampel, analisis, konteks, dan interpretasi agar pembaca memahami apa yang dapat dan tidak dapat disimpulkan. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Shadish et al. (2002) serta Appelbaum et al. (2018) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Secara operasional, peneliti perlu menetapkan analisis sensitivitas, menjelaskan alternatif penjelasan, dan mendokumentasikan sumber keterbatasan. Keputusan tersebut kemudian diterjemahkan ke dalam pilihan fidelitas, karakteristik sampel, serta catatan deviasi. Setiap

keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Sebagai ilustrasi, pada topik tes hanya satu topik, peserta didik dapat memberikan jawaban benar pada tingkat simbolik tetapi alasan submikroskopiknya keliru. Peneliti tidak cukup menghitung benar-salah; ia perlu membandingkan fidelitas dengan karakteristik sampel, memeriksa tingkat keyakinan, dan menguji konsistensi pada konteks lain. Hasil analisis mungkin menunjukkan bahwa kesalahan hanya muncul pada satu representasi, atau justru merupakan hubungan konseptual yang stabil.

Keputusan yang harus didokumentasikan:

- Definisikan analisis sensitivitas dalam istilah operasional.
- Tentukan bukti utama dari reliabilitas dan bukti pendukung dari kasus bertentangan.
- Periksa analisis dampak sebelum membuat kategori atau model.
- Catat potensi mengabaikan bias publikasi beserta strategi mitigasinya.
- Sajikan rekomendasi desain bersama batas interpretasi dan ketidakpastian.

Mutu inferensi terutama ditentukan oleh hubungan dengan klaim, proporsi, dan kejujuran pelaporan. Ancaman utama ialah menyatakan semua keterbatasan ringan serta tidak membedakan limitasi dan delimitasi. Karena itu, laporan perlu menyajikan bukti yang mendukung sekaligus bukti yang membatasi interpretasi. Prinsip ini konsisten dengan penekanan Shadish et al. (2002) dan Appelbaum et al. (2018) bahwa validitas melekat pada interpretasi dan penggunaan hasil, bukan sekadar pada nama instrumen atau desain.

Dengan demikian, keterbatasan interpretasi keyakinan berfungsi menghubungkan keputusan desain dengan interpretasi. Produk yang diharapkan bukan hanya batas inferensi, tetapi juga penjelasan mengenai bagaimana produk tersebut diperoleh dan dalam kondisi apa ia dapat digunakan. Penelitian miskonsepsi yang baik selalu menyisakan ruang bagi penjelasan alternatif dan verifikasi lanjutan.

7.20.12 Keterbatasan Lintas Budaya dan Bahasa

Pembahasan mengenai keterbatasan lintas budaya dan bahasa diperlukan agar keterbatasan penelitian miskonsepsi tidak berhenti pada prosedur administratif. Fokus ini berada dalam identifikasi sistematis terhadap batas desain, instrumen, sampel, analisis, konteks, dan interpretasi agar pembaca memahami apa yang dapat dan tidak dapat disimpulkan. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Levitt et al. (2018) serta Page et al. (2021) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Penerapan praktis dapat disusun dalam urutan: mendefinisikan batas generalisasi, mengoperasionalkan agenda perbaikan, mengumpulkan reliabilitas, memeriksa analisis dampak, lalu melaporkan prioritas replikasi. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Misalnya, pada sampel satu sekolah, perubahan skor sesudah pembelajaran belum otomatis menandakan perubahan konsep. Peserta didik mungkin mempelajari pola jawaban tes. Bukti yang lebih meyakinkan mencakup transfer ke situasi baru, penjelasan kausal, representasi partikel, dan retensi setelah jeda waktu.

Pertanyaan audit metodologis:

- Definisikan batas generalisasi dalam istilah operasional.
- Tentukan bukti utama dari karakteristik sampel dan bukti pendukung dari catatan deviasi.
- Periksa proporsi sebelum membuat kategori atau model.
- Catat potensi tidak membedakan limitasi dan delimitasi beserta strategi mitigasinya.
- Sajikan agenda penelitian bersama batas interpretasi dan ketidakpastian.

Temuan yang baik tidak selalu berarti hasil yang besar atau sesuai harapan. Temuan bernilai ketika analisis dampak terjaga, rencana mitigasi dijelaskan, dan ketidakpastian disajikan. Kerangka Levitt et al. (2018) serta Page et al. (2021) menolong peneliti membedakan kekuatan bukti dari daya tarik narasi.

Inti metodologis subbagian ini adalah keterlacakan. Pembaca harus dapat mengikuti perjalanan dari pertanyaan, data, pengodean, analisis, sampai prioritas replikasi. Ketika jejak itu jelas, temuan dapat digunakan untuk memperbaiki pembelajaran tanpa mengubah diagnosis konseptual menjadi label bagi peserta didik.

7.20.13 Keterbatasan Ekologis dan Implementasi

Dalam metodologi penelitian miskonsepsi kimia, keterbatasan ekologis dan implementasi menentukan mutu hubungan antara data dan kesimpulan. Fokus ini berada dalam identifikasi sistematis terhadap batas desain, instrumen, sampel, analisis, konteks, dan interpretasi agar pembaca memahami apa yang dapat dan tidak dapat disimpulkan. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Page et al. (2021) serta Appelbaum et al. (2018) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Prosedur minimal mencakup definisi alternatif penjelasan, alasan memilih karakteristik sampel, aturan mengubah respons menjadi kategori, serta mekanisme pemeriksaan proporsi.

Semua penyimpangan dari rencana awal perlu dicatat. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Pada penelitian mengenai wawancara setelah tes, peneliti dapat menyusun beberapa tugas ekuivalen yang berubah konteks, representasi, atau tuntutan penalarannya. Jika pola respons tetap berulang dan disertai keyakinan tinggi, terdapat dasar yang lebih kuat untuk menyebut indikasi miskonsepsi. Jika pola berubah-ubah, interpretasi yang lebih hati-hati adalah pengetahuan parsial atau ketergantungan konteks.

Bukti minimum yang perlu tersedia:

- Definiskan alternatif penjelasan dalam istilah operasional.
- Tentukan bukti utama dari kasus bertentangan dan bukti pendukung dari profil missing data.
- Periksa rencana mitigasi sebelum membuat kategori atau model.
- Catat potensi daftar klise beserta strategi mitigasinya.
- Sajikan matriks keterbatasan bersama batas interpretasi dan ketidakpastian.

Pemeriksaan mutu hendaknya dilakukan selama penelitian, bukan hanya pada akhir. Catatan tentang proporsi, keputusan terkait kejujuran pelaporan, dan kejadian yang berpotensi menimbulkan mengabaikan bias publikasi perlu dipelihara dalam *audit trail*. Hal ini memungkinkan koreksi dini dan pelaporan yang jujur.

Kesimpulannya, keterbatasan ekologis dan implementasi perlu dirancang sebagai proses argumentatif. Data tidak berbicara sendiri; maknanya bergantung pada definisi konstruk, kualitas prosedur, dan konteks. Oleh sebab itu, klaim akhir sebaiknya menggunakan bahasa yang sesuai dengan kekuatan bukti, misalnya “menunjukkan pola”, “mengindikasikan”, atau “mendukung penjelasan”.

7.20.14 Strategi Mitigasi dan Analisis Sensitivitas

Keputusan tentang strategi mitigasi dan analisis sensitivitas harus dibuat sebelum data utama dikumpulkan, bukan ditambahkan setelah hasil diketahui. Fokus ini berada dalam identifikasi sistematis terhadap batas desain, instrumen, sampel, analisis, konteks, dan interpretasi agar pembaca memahami apa yang dapat dan tidak dapat disimpulkan. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. Messick (1995) serta Page et al. (2021) sama-sama menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Tahap kerja dimulai dengan penegasan agenda perbaikan. Setelah itu, peneliti menyusun aturan untuk tingkat dampak dan menentukan bagaimana batas generalisasi akan diperiksa. Bukti sebaiknya berasal dari lebih dari satu bentuk, misalnya kasus bertentangan dan profil missing data. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Pada topik intervensi dua minggu, peneliti juga perlu memperhatikan bahasa. Istilah yang sama dapat diberi makna berbeda oleh peserta didik. Klarifikasi melalui pertanyaan tindak lanjut, gambar, atau contoh tandingan membantu membedakan masalah terminologi dari kesalahan konseptual.

Kriteria pemeriksaan sebelum analisis:

- Definisikan agenda perbaikan dalam istilah operasional.
- Tentukan bukti utama dari catatan deviasi dan bukti pendukung dari fidelitas.
- Periksa kejujuran pelaporan sebelum membuat kategori atau model.
- Catat potensi menyembunyikan kelemahan utama beserta strategi mitigasinya.
- Sajikan analisis sensitivitas bersama batas interpretasi dan ketidakpastian.

Kualitas penelitian meningkat ketika peneliti menunjukkan rantai bukti dari kasus bertentangan menuju agenda penelitian, menyatakan asumsi, dan memeriksa alternatif penjelasan. Sebaliknya, tidak membedakan limitasi dan delimitasi dapat menghasilkan klaim yang tampak tegas tetapi tidak stabil. Rekomendasi praktis harus proporsional dengan bukti.

Pada akhirnya, keputusan tentang strategi mitigasi dan analisis sensitivitas menentukan apakah penelitian menghasilkan daftar jawaban salah atau pemahaman yang dapat ditindaklanjuti mengenai cara peserta didik berpikir. Sasaran utamanya ialah bukti yang cukup untuk merancang diagnosis, pembelajaran, dan penelitian lanjutan secara bertanggung jawab.

7.20.15 Sintesis Pelaporan Keterbatasan

Sintesis pelaporan keterbatasan perlu dipahami sebagai keputusan metodologis yang mempunyai konsekuensi pada desain, analisis, dan pelaporan. Fokus ini berada dalam identifikasi sistematis terhadap batas desain, instrumen, sampel, analisis, konteks, dan interpretasi agar pembaca memahami apa yang dapat dan tidak dapat disimpulkan. Peneliti perlu menyatakan sejak awal apa yang dimaksud, indikator apa yang dapat diamati, dan bagaimana indikator tersebut dibedakan dari kesalahan hitung, ketidaktahuan, atau respons yang muncul karena salah membaca soal. Penegasan tersebut penting karena miskonsepsi merupakan inferensi mengenai struktur pemahaman, bukan objek yang dapat dilihat secara langsung. AERA, APA, dan NCME (2014) serta Appelbaum et al. (2018) sama-sama

menekankan perlunya hubungan yang dapat dipertanggungjawabkan antara konstruk, bukti, dan penggunaan hasil.

Peneliti sebaiknya menyiapkan matriks yang menghubungkan pertanyaan penelitian dengan sumber keterbatasan, catatan deviasi, teknik analisis, dan matriks keterbatasan. Matriks ini mencegah pengumpulan data yang banyak tetapi tidak relevan. Setiap keputusan dicatat dalam protokol, termasuk versi instrumen, aturan pengodean, penanganan respons ganda, prosedur data hilang, dan alasan perubahan selama penelitian. Dokumentasi ini tidak hanya berguna untuk replikasi, tetapi juga membantu peneliti membedakan analisis yang telah direncanakan dari eksplorasi yang muncul setelah melihat data.

Ilustrasi attrition tinggi juga menunjukkan bahwa kualitas instrumen memengaruhi temuan. Distraktor yang tidak berasal dari pemikiran peserta didik dapat menghasilkan pilihan acak, sedangkan pertanyaan yang terlalu menuntut matematika dapat mengukur keterampilan hitung alih-alih konsepsi. Peneliti perlu menilai kesesuaian bukti dengan konstruk yang dimaksud.

Langkah operasional yang disarankan:

- Definisikan sumber keterbatasan dalam istilah operasional.
- Tentukan bukti utama dari profil missing data dan bukti pendukung dari reliabilitas.
- Periksa spesifisitas sebelum membuat kategori atau model.
- Catat potensi menyatakan semua keterbatasan ringan beserta strategi mitigasinya.
- Sajikan batas inferensi bersama batas interpretasi dan ketidakpastian.

Peneliti sebaiknya memisahkan keputusan apriori dari keputusan eksploratori. Pemisahan tersebut membantu mengendalikan daftar klise dan mencegah menggunakan keterbatasan untuk membenarkan hasil. Selanjutnya, kejujuran pelaporan dan hubungan dengan klaim dilaporkan bersama ketidakpastian, sehingga pembaca dapat menilai kekuatan temuan tanpa bergantung pada kesimpulan penulis semata.

Dengan demikian, sintesis pelaporan keterbatasan berfungsi menghubungkan keputusan desain dengan interpretasi. Produk yang diharapkan bukan hanya matriks keterbatasan, tetapi juga penjelasan mengenai bagaimana produk tersebut diperoleh dan dalam kondisi apa ia dapat digunakan. Penelitian miskonsepsi yang baik selalu menyisakan ruang bagi penjelasan alternatif dan verifikasi lanjutan.

7.21 Sintesis Metodologi Penelitian Miskonsepsi Kimia

Metodologi penelitian miskonsepsi kimia harus dibangun sebagai satu sistem keputusan. Paradigma memberi dasar bagi cara peneliti memandang konsepsi; desain menentukan bentuk pertanyaan yang dapat dijawab; sampling menetapkan jangkauan bukti; instrumen dan teknik pengumpulan data menghubungkan konstruk dengan respons; analisis mengubah data menjadi pola; sedangkan pelaporan menunjukkan kekuatan serta batas klaim. Kelemahan pada satu komponen dapat merusak keseluruhan rantai inferensi.

Tidak ada satu metode yang selalu paling baik. Tes bertingkat efisien untuk pemetaan luas, tetapi wawancara lebih mampu menjelaskan mekanisme penalaran. Studi kasus memberikan kedalaman, sementara survei memperluas cakupan. Eksperimen menguji perubahan akibat intervensi, sedangkan *design-based research* menghubungkan pengembangan desain dan pembentukan teori lokal. Mixed methods diperlukan ketika pertanyaan menuntut keluasan dan kedalaman sekaligus, bukan sekadar karena peneliti ingin memakai banyak teknik.

Penelitian yang etis dan bermutu tinggi menggunakan bahasa inferensial secara hati-hati. Respons salah tunggal tidak cukup untuk menyatakan miskonsepsi. Bukti yang lebih kuat berasal dari keberulangan pola, alasan yang dapat diidentifikasi, konsistensi lintas konteks, keyakinan, dan ketahanan terhadap koreksi sederhana. Bahkan ketika indikator tersebut tersedia, peneliti tetap perlu menyatakan bahwa temuan merupakan inferensi yang terikat pada instrumen, konteks, dan populasi.

Bagi peneliti pemula, prosedur yang paling aman adalah menyusun matriks keselarasan sejak proposal. Setiap pertanyaan penelitian dipasangkan dengan konstruk, unit analisis, sumber data, prosedur pengumpulan, teknik analisis, kriteria mutu, bentuk hasil, dan batas klaim. Matriks tersebut kemudian digunakan sebagai alat audit sepanjang penelitian. Perubahan tetap dapat dilakukan, tetapi alasan dan dampaknya harus terdokumentasi.

Bagi guru dan peneliti kelas, metodologi yang ketat tidak harus berarti prosedur yang rumit. Ketelitian dapat diwujudkan melalui pertanyaan yang jelas, lebih dari satu sumber bukti, aturan kategori yang eksplisit, pencatatan konteks, evaluasi transfer, serta refleksi terhadap alternatif penjelasan. Prinsip utamanya adalah memperoleh bukti yang cukup untuk mengambil tindakan pembelajaran tanpa mengubah peserta didik menjadi objek pelabelan.

Akhirnya, penelitian miskonsepsi kimia perlu bergerak dari inventarisasi kesalahan menuju penjelasan dan perbaikan. Temuan seharusnya menunjukkan konsep mana yang rapuh, hubungan apa yang keliru, representasi mana yang gagal dihubungkan, kondisi apa yang mengaktifkan miskonsepsi, dan intervensi apa yang membantu rekonstruksi. Dengan metodologi yang transparan, penelitian dapat memberikan kontribusi teoretis sekaligus praktis bagi pendidikan kimia.

Daftar Pustaka Bab 7

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). Standards for educational and psychological testing. *American Educational Research Association*.

Anderson, T., & Shattuck, J. (2012). Design-based research: A decade of progress in education research? *Educational Researcher*, 41(1), 16–25. doi: 10.3102/0013189X11428813.

Appelbaum, M., Cooper, H., Kline, R. B., Mayo-Wilson, E., Nezu, A. M., & Rao, S. M. (2018). Journal article reporting standards for quantitative research in psychology: The APA Publications and Communications Board task force report. *American Psychologist*, 73(1), 3–25. doi: 10.1037/amp0000191.

- Barbera, J. (2013). A psychometric analysis of the Chemical Concepts Inventory. *Journal of Chemical Education*, 90(5), 546–553. doi: 10.1021/ed3004353.
- Boateng, G. O., Neilands, T. B., Frongillo, E. A., Melgar-Quiñonez, H. R., & Young, S. L. (2018). Best practices for developing and validating scales for health, social, and behavioral research: A primer. *Frontiers in Public Health*, 6, Article 149. doi: 10.3389/fpubh.2018.00149.
- Bock, R. D. (1972). Estimating item parameters and latent ability when responses are scored in two or more nominal categories. *Psychometrika*, 37, 29–51. doi: 10.1007/BF02291411.
- Brandriet, A. R., & Bretz, S. L. (2014). Measuring meta-ignorance through the lens of confidence: Examining students' redox misconceptions about oxidation numbers, charge, and electron transfer. *Chemistry Education Research and Practice*, 15(4), 729–746. doi: 10.1039/C4RP00129J.
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. doi: 10.1191/1478088706qp063oa.
- Braun, V., & Clarke, V. (2021). One size fits all? What counts as quality practice in (reflexive) thematic analysis? *Qualitative Research in Psychology*, 18(3), 328–352. doi: 10.1080/14780887.2020.1769238.
- British Educational Research Association. (2024). Ethical guidelines for educational research (5th ed.). *British Educational Research Association*.
- Campbell, D. T., & Stanley, J. C. (1963). Experimental and quasi-experimental designs for research. *Houghton Mifflin*.
- Carter, N., Bryant-Lukosius, D., DiCenso, A., Blythe, J., & Neville, A. J. (2014). The use of triangulation in qualitative research. *Oncology Nursing Forum*, 41(5), 545–547. doi: 10.1188/14.ONF.545-547.
- Cochran, W. G. (1977). Sampling techniques (3rd ed.). *Wiley*.
- Collins, L. M., & Lanza, S. T. (2010). Latent class and latent transition analysis: With applications in the social, behavioral, and health sciences. *Wiley*.
- Creswell, J. W., & Creswell, J. D. (2018). Research design: Qualitative, quantitative, and mixed methods approaches (5th ed.). *SAGE Publications*.
- Creswell, J. W., & Plano Clark, V. L. (2018). Designing and conducting mixed methods research (3rd ed.). *SAGE Publications*.
- Cumming, G. (2014). The new statistics: Why and how. *Psychological Science*, 25(1), 7–29. doi: 10.1177/0956797613504966.

- Des Jarlais, D. C., Lyles, C., & Crepaz, N. (2004). Improving the reporting quality of nonrandomized evaluations of behavioral and public health interventions: The TREND statement. *American Journal of Public Health, 94*(3), 361–366. doi: 10.2105/AJPH.94.3.361.
- DeVellis, R. F., & Thorpe, C. T. (2021). Scale development: Theory and applications (5th ed.). *SAGE Publications*.
- Dillman, D. A., Smyth, J. D., & Christian, L. M. (2014). Internet, phone, mail, and mixed-mode surveys: The tailored design method (4th ed.). *Wiley*.
- Embretson, S. E., & Reise, S. P. (2000). Item response theory for psychologists. *Lawrence Erlbaum Associates*.
- Eysenbach, G. (2004). Improving the quality of web surveys: The Checklist for Reporting Results of Internet E-Surveys (CHERRIES). *Journal of Medical Internet Research, 6*(3), e34. doi: 10.2196/jmir.6.3.e34.
- Farmer, T., Robinson, K., Elliott, S. J., & Eyles, J. (2006). Developing and implementing a triangulation protocol for qualitative health research. *Qualitative Health Research, 16*(3), 377–394. doi: 10.1177/1049732305285708.
- Fetters, M. D., Curry, L. A., & Creswell, J. W. (2013). Achieving integration in mixed methods designs—Principles and practices. *Health Services Research, 48*(6), 2134–2156. doi: 10.1111/1475-6773.12117.
- Fowler, F. J. (2014). Survey research methods (5th ed.). *SAGE Publications*.
- Greene, J. C., Caracelli, V. J., & Graham, W. F. (1989). Toward a conceptual framework for mixed-method evaluation designs. *Educational Evaluation and Policy Analysis, 11*(3), 255–274. doi: 10.3102/01623737011003255.
- Hasan, S., Bagayoko, D., & Kelley, E. L. (1999). Misconceptions and the certainty of response index (CRI). *Physics Education, 34*(5), 294–299. doi: 10.1088/0031-9120/34/5/304.
- Johnson, R. B., & Onwuegbuzie, A. J. (2004). Mixed methods research: A research paradigm whose time has come. *Educational Researcher, 33*(7), 14–26. doi: 10.3102/0013189X033007014.
- Kaltakci Gurel, D., Eryilmaz, A., & McDermott, L. C. (2015). A review and comparison of diagnostic instruments to identify students' misconceptions in science. *EURASIA Journal of Mathematics, Science and Technology Education, 11*(5), 989–1008. doi: 10.12973/eurasia.2015.1369a.
- Kane, M. T. (2013). Validating the interpretations and uses of test scores. *Journal of Educational Measurement, 50*(1), 1–73. doi: 10.1111/jedm.12000.
- Kemmis, S., McTaggart, R., & Nixon, R. (2014). The action research planner: Doing critical participatory action research. *Springer*. doi: 10.1007/978-981-4560-67-2.

- Lakens, D. (2022). Sample size justification. *Collabra: Psychology*, 8(1), Article 33267. doi: 10.1525/collabra.33267.
- Levitt, H. M., Bamberg, M., Creswell, J. W., Frost, D. M., Josselson, R., & Suárez-Orozco, C. (2018). Journal article reporting standards for qualitative primary, qualitative meta-analytic, and mixed methods research in psychology. *American Psychologist*, 73(1), 26–46. doi: 10.1037/amp0000151.
- Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. SAGE Publications.
- Little, R. J. A., & Rubin, D. B. (2019). *Statistical analysis with missing data* (3rd ed.). Wiley.
- Lohr, S. L. (2021). *Sampling: Design and analysis* (3rd ed.). Chapman and Hall/CRC.
- Malterud, K., Siersma, V. D., & Guassora, A. D. (2016). Sample size in qualitative interview studies: Guided by information power. *Qualitative Health Research*, 26(13), 1753–1760. doi: 10.1177/1049732315617444.
- McKenney, S., & Reeves, T. C. (2018). *Conducting educational design research* (2nd ed.). Routledge.
- Merriam, S. B., & Tisdell, E. J. (2016). *Qualitative research: A guide to design and implementation* (4th ed.). Jossey-Bass.
- Mertler, C. A. (2020). *Action research: Improving schools and empowering educators* (6th ed.). SAGE Publications.
- Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist*, 50(9), 741–749. doi: 10.1037/0003-066X.50.9.741.
- Miles, M. B., Huberman, A. M., & Saldaña, J. (2014). *Qualitative data analysis: A methods sourcebook* (3rd ed.). SAGE Publications.
- Nakhleh, M. B. (1992). Why some students don't learn chemistry: Chemical misconceptions. *Journal of Chemical Education*, 69(3), 191–196. doi: 10.1021/ed069p191.
- National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research. (1979). *The Belmont report: Ethical principles and guidelines for the protection of human subjects of research*. U.S. Department of Health and Human Services.
- Novak, J. D., & Cañas, A. J. (2008). *The theory underlying concept maps and how to construct and use them*. Institute for Human and Machine Cognition.
- Novak, J. D., & Gowin, D. B. (1984). *Learning how to learn*. Cambridge University Press.
- Nowell, L. S., Norris, J. M., White, D. E., & Moules, N. J. (2017). Thematic analysis: Striving to meet the trustworthiness criteria. *International Journal of Qualitative Methods*, 16, 1–13. doi: 10.1177/1609406917733847.

- Nyachwaya, J. M., Mohamed, A.-R., Roehrig, G. H., Wood, N. B., Kern, A. L., & Schneider, J. L. (2011). The development of an open-ended drawing tool: An alternative diagnostic tool for assessing students' understanding of the particulate nature of matter. *Chemistry Education Research and Practice*, 12(2), 121–132. doi: 10.1039/C1RP90017J.
- O'Brien, B. C., Harris, I. B., Beckman, T. J., Reed, D. A., & Cook, D. A. (2014). Standards for reporting qualitative research: A synthesis of recommendations. *Academic Medicine*, 89(9), 1245–1251. doi: 10.1097/ACM.0000000000000388.
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., et al. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, 372, n71. doi: 10.1136/bmj.n71.
- Page, M. J., Moher, D., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., et al. (2021). PRISMA 2020 explanation and elaboration: Updated guidance and exemplars for reporting systematic reviews. *BMJ*, 372, n160. doi: 10.1136/bmj.n160.
- Palinkas, L. A., Horwitz, S. M., Green, C. A., Wisdom, J. P., Duan, N., & Hoagwood, K. (2015). Purposeful sampling for qualitative data collection and analysis in mixed method implementation research. *Administration and Policy in Mental Health*, 42(5), 533–544. doi: 10.1007/s10488-013-0528-y.
- Patton, M. Q. (2015). *Qualitative research & evaluation methods* (4th ed.). SAGE Publications.
- Republik Indonesia. (2022). Undang-Undang Republik Indonesia Nomor 27 Tahun 2022 tentang Pelindungan Data Pribadi. *Sekretariat Negara Republik Indonesia*.
- Robinson, O. C. (2014). Sampling in interview-based qualitative research: A theoretical and practical guide. *Qualitative Research in Psychology*, 11(1), 25–41. doi: 10.1080/14780887.2013.801543.
- Ruiz-Primo, M. A., & Shavelson, R. J. (1996). Problems and issues in the use of concept maps in science assessment. *Journal of Research in Science Teaching*, 33(6), 569–600. doi: 10.1002/(SICI)1098-2736(199608)33:6<569::AID-TEA1>3.0.CO;2-M.
- Saldaña, J. (2021). *The coding manual for qualitative researchers* (4th ed.). SAGE Publications.
- Schulz, K. F., Altman, D. G., & Moher, D. (2010). CONSORT 2010 statement: Updated guidelines for reporting parallel group randomised trials. *BMJ*, 340, c332. doi: 10.1136/bmj.c332.
- Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Houghton Mifflin.
- Shea, B. J., Reeves, B. C., Wells, G., Thuku, M., Hamel, C., Moran, J., et al. (2017). AMSTAR 2: A critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. *BMJ*, 358, j4008. doi: 10.1136/bmj.j4008.

Taber, K. S. (2018). The use of Cronbach's alpha when developing and reporting research instruments in science education. *Research in Science Education*, 48(6), 1273–1296. doi: 10.1007/s11165-016-9602-2.

Talanquer, V. (2011). Macro, submicro, and symbolic: The many faces of the chemistry "triplet". *International Journal of Science Education*, 33(2), 179–195. doi: 10.1080/09500690903386435.

The Design-Based Research Collective. (2003). Design-based research: An emerging paradigm for educational inquiry. *Educational Researcher*, 32(1), 5–8. doi: 10.3102/0013189X032001005.

Tong, A., Sainsbury, P., & Craig, J. (2007). Consolidated criteria for reporting qualitative research (COREQ): A 32-item checklist for interviews and focus groups. *International Journal for Quality in Health Care*, 19(6), 349–357. doi: 10.1093/intqhc/mzm042.

Treagust, D. F. (1988). Development and use of diagnostic tests to evaluate students' misconceptions in science. *International Journal of Science Education*, 10(2), 159–169. doi: 10.1080/0950069880100204.

van Buuren, S. (2018). Flexible imputation of missing data (2nd ed.). *CRC Press*.

von Elm, E., Altman, D. G., Egger, M., Pocock, S. J., Gøtzsche, P. C., & Vandenbroucke, J. P. (2007). The strengthening the reporting of observational studies in epidemiology (STROBE) statement: Guidelines for reporting observational studies. *PLoS Medicine*, 4(10), e296. doi: 10.1371/journal.pmed.0040296.

Wang, F., & Hannafin, M. J. (2005). Design-based research and technology-enhanced learning environments. *Educational Technology Research and Development*, 53(4), 5–23. doi: 10.1007/BF02504682.

Wasserstein, R. L., & Lazar, N. A. (2016). The ASA statement on p-values: Context, process, and purpose. *The American Statistician*, 70(2), 129–133. doi: 10.1080/00031305.2016.1154108.

Whiting, P., Savović, J., Higgins, J. P. T., Caldwell, D. M., Reeves, B. C., Shea, B., et al. (2016). ROBIS: A new tool to assess risk of bias in systematic reviews. *Journal of Clinical Epidemiology*, 69, 225–234. doi: 10.1016/j.jclinepi.2015.08.016.

Wren, D., & Barbera, J. (2014). Psychometric analysis of the thermochemistry concept inventory. *Chemistry Education Research and Practice*, 15(3), 380–390. doi: 10.1039/C3RP00170A.

Yin, R. K. (2018). Case study research and applications: Design and methods (6th ed.). *SAGE Publications*.

GLOSARIUM

A

Analisis butir — Proses mengevaluasi kualitas setiap butir instrumen berdasarkan tingkat kesukaran, daya pembeda, pola distraktor, dan karakteristik respons. Dalam buku ini, analisis butir digunakan untuk memperbaiki instrumen diagnosis miskonsepsi. *Padanan asing: item analysis.*

Analisis isi kualitatif — Teknik menganalisis data teks, gambar, atau dokumen melalui pengodean dan pengelompokan isi. Digunakan untuk mengidentifikasi pola konsepsi dan alasan peserta didik. *Padanan asing: qualitative content analysis.*

Analisis pola respons — Kajian terhadap kombinasi jawaban, alasan, keyakinan, dan konsistensi respons peserta didik. Analisis ini membantu membedakan miskonsepsi dari kesalahan yang tidak menetap. *Padanan asing: response-pattern analysis.*

Analisis sensitivitas — Pemeriksaan apakah kesimpulan penelitian tetap stabil ketika asumsi, kriteria, atau prosedur analisis diubah. Dalam buku ini, analisis sensitivitas digunakan untuk menilai kekuatan inferensi penelitian.

Asesmen diagnostik — Proses sistematis untuk mengumpulkan dan menafsirkan bukti tentang pemahaman peserta didik. Dalam pembelajaran kimia, asesmen ini digunakan untuk menemukan konsepsi awal, pengetahuan parsial, dan miskonsepsi. *Padanan asing: diagnostic assessment.*

Asesmen formatif — Asesmen yang dilakukan selama pembelajaran untuk memberikan umpan balik dan memperbaiki proses belajar. Hasil diagnosis miskonsepsi dapat menjadi dasar asesmen formatif. *Padanan asing: formative assessment.*

Assent — Persetujuan sadar dari peserta didik yang belum memiliki kewenangan hukum penuh, terutama peserta didik minor. Dalam penelitian, *assent* melengkapi persetujuan orang tua atau wali.

Audit trail — Dokumentasi terperinci mengenai keputusan pengumpulan, pengodean, analisis, dan interpretasi data. Dalam penelitian miskonsepsi, *audit trail* memperkuat keterlacakan dan kredibilitas temuan.

B

Bahasa nonstigmatisasi — Penggunaan bahasa yang tidak memberi cap negatif kepada peserta didik. Dalam buku ini, miskonsepsi diperlakukan sebagai karakteristik respons, bukan identitas tetap seseorang.

Beban kognitif — Tuntutan mental yang harus diproses oleh memori kerja ketika peserta didik menyelesaikan tugas. Beban yang terlalu tinggi dapat menghasilkan kesalahan yang bukan miskonsepsi. *Padanan asing: cognitive load.*

Bias — Penyimpangan sistematis yang memengaruhi pengumpulan, analisis, atau interpretasi data. Buku ini membahas bias instrumen, budaya, seleksi, penilai, dan pelaporan.

Blueprint instrumen — Rancangan atau kisi-kisi yang menghubungkan domain konsep, indikator, bentuk butir, dan bukti yang hendak dikumpulkan. *Padanan asing: instrument blueprint atau test specification.*

Bukti diagnosis — Data berupa jawaban, alasan, keyakinan, gambar, percakapan, atau tindakan yang digunakan untuk menyimpulkan keadaan pemahaman peserta didik. Bukti harus ditafsirkan secara hati-hati.

Bukti validitas — Informasi yang mendukung ketepatan interpretasi dan penggunaan hasil asesmen. Buku ini membahas bukti berbasis isi, proses respons, struktur internal, hubungan dengan variabel lain, dan konsekuensi penggunaan.

C

Case study — Penelitian mendalam terhadap satu kasus atau beberapa kasus yang dibatasi secara jelas. Dalam penelitian miskonsepsi, kasus dapat berupa individu, kelas, sekolah, atau episode pembelajaran. *Padanan Indonesia: studi kasus.*

Certainty of Response Index (CRI) — Indeks yang menunjukkan tingkat keyakinan peserta didik terhadap jawabannya. CRI membantu membedakan ketidaktahuan, tebakan, dan indikasi miskonsepsi.

Cluster analysis — Teknik statistik untuk mengelompokkan peserta didik atau pola respons berdasarkan kemiripan karakteristik. Digunakan untuk menemukan profil konseptual yang tidak ditentukan sebelumnya. *Padanan Indonesia: analisis klaster.*

Codebook — Dokumen yang berisi nama kode, definisi, aturan penggunaan, dan contoh penerapan kode. Dalam analisis kualitatif, *codebook* menjaga konsistensi pengodean.

Coding — Proses memberi label pada bagian data yang mengandung makna tertentu. Dalam buku ini, pengodean digunakan untuk menganalisis alasan terbuka, wawancara, gambar, dan percakapan kelas. *Padanan Indonesia: pengodean.*

Concept inventory — Instrumen yang dirancang untuk mengukur pemahaman terhadap konsep-konsep inti dan konsepsi alternatif dalam suatu bidang. *Padanan Indonesia: inventori konsep.*

Construct-irrelevant variance — Variasi skor yang berasal dari faktor di luar konstruk yang hendak diukur, seperti bahasa yang terlalu sulit atau kemampuan membaca. Kondisi ini mengancam validitas diagnosis.

D

Data hilang — Data yang tidak tersedia karena peserta tidak menjawab, keluar dari penelitian, atau mengalami kendala teknis. Penanganannya harus dijelaskan secara transparan. *Padanan asing: missing data.*

Design-Based Research (DBR) — Penelitian iteratif yang mengembangkan dan menguji desain pembelajaran dalam konteks autentik. Dalam buku ini, DBR digunakan untuk mengembangkan diagnosis dan remediasi miskonsepsi.

Diagnosis miskonsepsi — Proses membuat inferensi tentang adanya pola pemahaman yang tidak sesuai dengan konsepsi ilmiah. Diagnosis memerlukan lebih dari satu jawaban salah.

Differential Item Functioning (DIF) — Kondisi ketika suatu butir bekerja berbeda pada kelompok yang memiliki tingkat kemampuan sama. DIF digunakan untuk memeriksa potensi bias dan keadilan instrumen.

Distraktor — Opsi jawaban salah dalam tes pilihan ganda. Dalam instrumen diagnostik, distraktor harus mewakili konsepsi alternatif yang benar-benar mungkin dimiliki peserta didik.

Drawing test — Instrumen yang meminta peserta didik menggambar partikel, struktur, proses, atau model kimia. Gambar digunakan untuk mengungkap representasi mental yang sulit terlihat melalui jawaban pilihan ganda. *Padanan Indonesia: tes menggambar.*

E

Eksperimen — Desain penelitian untuk menguji pengaruh suatu intervensi melalui manipulasi variabel dan penggunaan pembanding. Dalam penelitian miskonsepsi, eksperimen digunakan untuk menguji efektivitas remediasi.

Elisitasi — Proses memunculkan gagasan, pengetahuan awal, atau model mental peserta didik. Elisitasi dapat dilakukan melalui pertanyaan, fenomena pemantik, gambar, atau wawancara. *Padanan asing: elicitation.*

Epistemologi — Kajian tentang asal, bentuk, dan dasar keabsahan pengetahuan. Dalam buku ini, epistemologi berkaitan dengan cara memperoleh bukti tentang pemahaman peserta didik.

Etika penelitian — Prinsip yang mengatur perlindungan partisipan, persetujuan, kerahasiaan, keadilan, dan penggunaan data. Diagnosis tidak boleh digunakan untuk memperlakukan peserta didik.

Exit ticket — Pertanyaan atau tugas singkat yang dikerjakan pada akhir pembelajaran. Instrumen ini dapat digunakan untuk mendeteksi pemahaman dan miskonsepsi secara cepat.

Expert judgment — Penilaian pakar terhadap relevansi, kejelasan, dan keterwakilan isi instrumen. *Expert judgment* merupakan bagian penting dari validasi isi. *Padanan Indonesia: pertimbangan ahli.*

F

Fairness — Prinsip bahwa instrumen dan penggunaannya harus memberikan kesempatan yang adil kepada seluruh kelompok peserta didik. *Fairness* menuntut pemeriksaan bias bahasa, budaya, dan akses.

False negative — Kesalahan diagnosis ketika miskonsepsi yang sebenarnya ada tidak berhasil terdeteksi. Hal ini dapat terjadi karena butir kurang sensitif atau bukti terlalu terbatas.

False positive — Kesalahan diagnosis ketika respons peserta didik dikategorikan sebagai miskonsepsi, padahal disebabkan oleh ketidaktahuan, salah baca, atau kelalaian.

Fidelitas implementasi — Tingkat kesesuaian pelaksanaan intervensi dengan rancangan yang telah ditetapkan. Fidelitas diperlukan agar perubahan hasil dapat dikaitkan secara wajar dengan intervensi. *Padanan asing: implementation fidelity.*

Five-tier diagnostic test — Tes diagnostik lima tingkat yang memperluas tes empat tingkat dengan satu dimensi tambahan. Instrumen ini digunakan untuk memperoleh bukti diagnosis yang lebih terperinci.

Four-tier diagnostic test — Tes yang memisahkan jawaban, keyakinan terhadap jawaban, alasan, dan keyakinan terhadap alasan. Pemisahan tersebut membantu memperkuat klasifikasi pemahaman.

G

Gagasan awal — Pemahaman yang telah dimiliki peserta didik sebelum mempelajari suatu konsep secara formal. Gagasan awal dapat sesuai, belum lengkap, atau bertentangan dengan konsep ilmiah. *Padanan asing: prior conception.*

Generalisasi analitik — Penerapan temuan studi kasus pada teori atau proposisi, bukan langsung pada populasi statistik. Digunakan untuk menjelaskan keterterapan temuan di luar kasus yang diteliti.

Graf — Representasi jaringan yang terdiri atas simpul dan hubungan antarsimpul. Dalam pemetaan konsep, konsep menjadi simpul dan relasi menjadi sisi. *Padanan asing: graph.*

Grafik diagnostik — Penyajian visual pola jawaban, prevalensi miskonsepsi, keyakinan, atau perubahan konseptual. Grafik harus menampilkan ketidakpastian dan tidak menyesatkan pembaca.

H

Heterogenitas — Keragaman karakteristik peserta didik, instrumen, konteks, atau hasil penelitian. Dalam meta-analisis, heterogenitas menunjukkan tingkat perbedaan antarpelaksanaan penelitian.

Hipotesis — Pernyataan sementara yang dapat diuji melalui data. Dalam penelitian eksperimen, hipotesis mengarahkan pengujian pengaruh intervensi terhadap perubahan konsepsi.

Hubungan konsep — Keterkaitan bermakna antara dua atau lebih konsep. Diagnosis tidak hanya menilai penguasaan konsep tunggal, tetapi juga ketepatan relasi antarkonsep.

I

Indeks konsistensi — Ukuran yang menunjukkan kestabilan pola respons pada beberapa butir atau konteks. Konsistensi yang kuat dapat menjadi bukti tambahan adanya miskonsepsi.

Inferensi — Kesimpulan yang dibuat berdasarkan bukti yang tersedia. Dalam buku ini, miskonsepsi dipahami sebagai inferensi mengenai struktur pemahaman, bukan sesuatu yang dapat diamati langsung.

Informed consent — Persetujuan sukarela setelah partisipan memperoleh informasi tentang tujuan, prosedur, risiko, manfaat, dan haknya. *Padanan Indonesia: persetujuan setelah penjelasan.*

Instrumen diagnostik — Alat untuk mengumpulkan bukti tentang pemahaman, alasan, keyakinan, atau model mental peserta didik. Instrumen dapat berbentuk tes, wawancara, gambar, observasi, atau tugas kinerja.

Interpretivisme — Paradigma yang menekankan pemahaman makna berdasarkan pengalaman dan konteks partisipan. Digunakan ketika peneliti ingin memahami cara peserta didik memaknai konsep.

Item Response Theory (IRT) — Keluarga model statistik yang menghubungkan karakteristik butir dengan kemampuan atau atribut laten peserta didik. IRT digunakan dalam pengembangan dan evaluasi instrumen.

J

Jawaban–alasan–keyakinan — Tiga sumber bukti yang digunakan secara terpadu dalam diagnosis miskonsepsi. Kombinasinya lebih informatif daripada jawaban benar atau salah saja.

Jaringan konsep — Struktur yang menunjukkan konsep-konsep dan hubungan di antara konsep tersebut. Pemetaan jaringan membantu menemukan konsep pusat dan relasi yang mengandung miskonsepsi.

Jejak penyelesaian — Urutan langkah yang dilakukan peserta didik ketika menyelesaikan masalah. Jejak ini membantu menemukan lokasi kesalahan prosedural atau konseptual.

Joint display — Penyajian terpadu data kuantitatif dan kualitatif dalam tabel, matriks, atau diagram. Dalam *mixed methods*, *joint display* membantu menghasilkan meta-inferensi.

K

Kalibrasi keyakinan — Tingkat kesesuaian antara keyakinan peserta didik dan kebenaran jawabannya. Keyakinan tinggi pada jawaban salah dapat menjadi indikator miskonsepsi.

Kategori pemahaman — Pengelompokan respons menjadi memahami konsep, pengetahuan parsial, tidak tahu, menebak, atau mengalami miskonsepsi. Aturan kategorisasi harus dinyatakan secara transparan.

Kejelasan butir — Tingkat kemudahan stimulus dan pertanyaan dipahami sesuai maksud penyusun. Ketidakjelasan bahasa dapat menimbulkan kesalahan yang tidak berkaitan dengan konsep kimia.

Keyakinan — Penilaian peserta didik terhadap kepastian jawaban atau alasannya. Dalam tes bertingkat, keyakinan menjadi bukti tambahan, bukan satu-satunya dasar diagnosis. *Padanan asing: confidence*.

Konsepsi alternatif — Penjelasan atau pemahaman yang berbeda dari konsepsi ilmiah yang diterima. Istilah ini sering digunakan sebagai alternatif yang lebih netral daripada miskonsepsi. *Padanan asing: alternative conception*.

Konstruk — Atribut teoretis yang hendak diukur, seperti pemahaman konsep atau keyakinan. Konstruk harus didefinisikan secara operasional sebelum instrumen dikembangkan.

Kredibilitas — Tingkat keterpercayaan temuan penelitian kualitatif. Kredibilitas diperkuat melalui triangulasi, pemeriksaan kasus negatif, dan dokumentasi analisis.

Kuasi-eksperimen — Penelitian yang menguji pengaruh intervensi tanpa pengacakan penuh terhadap peserta atau kelompok. *Padanan asing: quasi-experiment*.

L

Latent Class Analysis (LCA) — Teknik statistik untuk menemukan kelompok laten berdasarkan pola respons. Dalam buku ini, LCA dapat digunakan untuk mengidentifikasi profil konseptual peserta didik.

Latent Transition Analysis (LTA) — Pengembangan LCA yang menganalisis perpindahan peserta didik dari satu kelas laten ke kelas lain dari waktu ke waktu.

Literatur abu-abu — Sumber ilmiah di luar publikasi komersial, seperti tesis, disertasi, laporan lembaga, dan prosiding. Literatur ini dapat digunakan dalam *systematic literature review*. *Padanan asing: grey literature*.

Log aktivitas digital — Rekaman tindakan peserta didik dalam sistem digital, seperti urutan klik, waktu respons, revisi jawaban, dan penggunaan sumber. Data ini dapat memperkaya diagnosis.

Longitudinal — Pendekatan yang mengamati perkembangan peserta didik dalam lebih dari satu waktu pengukuran. Digunakan untuk memeriksa kestabilan atau perubahan konsepsi.

M

Makroskopik — Tingkat representasi kimia yang berkaitan dengan gejala yang dapat diamati secara langsung, seperti perubahan warna, gas, atau endapan.

Meta-analisis — Teknik statistik untuk menggabungkan hasil beberapa penelitian yang sebanding. Digunakan untuk memperkirakan kecenderungan efek dan heterogenitas temuan.

Meta-ignorance — Keadaan ketika peserta didik tidak menyadari bahwa pengetahuannya keliru atau terbatas. Kondisi ini dapat terlihat dari jawaban salah dengan keyakinan tinggi.

Meta-inferensi — Kesimpulan terpadu yang diperoleh melalui integrasi temuan kuantitatif dan kualitatif dalam penelitian *mixed methods*.

Metakognisi — Kesadaran dan pengaturan seseorang terhadap proses berpikirnya sendiri. Tingkat keyakinan merupakan salah satu indikator metakognitif.

Miskonsepsi — Pola pemahaman yang tidak sesuai dengan konsepsi ilmiah, relatif stabil, dan digunakan secara konsisten dalam menjelaskan fenomena. Satu jawaban salah belum cukup membuktikan miskonsepsi.

Mixed methods — Pendekatan penelitian yang mengintegrasikan data kuantitatif dan kualitatif untuk memperoleh keluasan dan kedalaman pemahaman. *Padanan Indonesia: metode campuran*.

Model mental — Representasi internal yang digunakan seseorang untuk menjelaskan, memprediksi, atau menafsirkan suatu fenomena. Model mental dapat diungkap melalui gambar, penjelasan, dan wawancara.

Model Rasch — Model pengukuran yang menghubungkan tingkat kemampuan peserta dengan tingkat kesulitan butir pada satu skala logit.

N

Nonrespons — Keadaan ketika peserta yang dipilih tidak memberikan respons atau tidak menyelesaikan penelitian. Nonrespons dapat menimbulkan bias apabila memiliki pola tertentu.

Nonstigmatisasi — Prinsip menghindari pemberian label negatif kepada peserta didik berdasarkan hasil diagnosis. Hasil seharusnya digunakan untuk perbaikan pembelajaran.

Nominal Response Model — Model IRT untuk menganalisis respons dengan kategori yang tidak memiliki urutan. Model ini dapat digunakan untuk mengkaji pilihan distraktor.

O

Observasi kelas — Pengamatan sistematis terhadap percakapan, tindakan, penggunaan representasi, dan interaksi peserta didik selama pembelajaran.

Omega — Koefisien reliabilitas yang mengestimasi konsistensi skor berdasarkan model faktor. Omega dapat menjadi alternatif koefisien alfa. *Padanan asing: coefficient omega*.

Ontologi — Kajian mengenai hakikat realitas yang diteliti. Dalam penelitian miskonsepsi, ontologi berkaitan dengan cara peneliti memandang konsep, konsepsi, dan miskonsepsi.

Open-ended diagnostic test — Tes diagnostik yang meminta peserta didik memberikan jawaban atau alasan dengan kata-katanya sendiri. Instrumen ini memungkinkan munculnya konsepsi yang belum diperkirakan penyusun.

P

Paradigma penelitian — Kerangka dasar mengenai realitas, pengetahuan, dan cara penelitian dilakukan. Buku ini membahas postpositivisme, konstruktivisme, interpretivisme, pragmatisme, dan realisme kritis.

Pelaporan diagnostik — Penyajian hasil diagnosis secara transparan, proporsional, dan dapat ditindaklanjuti. Pelaporan harus membedakan bukti, interpretasi, ketidakpastian, dan rekomendasi.

Pemetaan miskonsepsi — Proses menggambarkan jenis, prevalensi, hubungan, dan distribusi miskonsepsi pada individu atau kelompok.

Pengetahuan parsial — Keadaan ketika sebagian unsur konsep telah dipahami, tetapi hubungan atau penerapannya belum lengkap. Pengetahuan parsial tidak boleh langsung dikategorikan sebagai miskonsepsi.

Peta ahli — Peta konsep yang disusun atau disepakati oleh pakar sebagai pembanding terhadap peta konsep peserta didik. *Padanan asing: expert map.*

Peta konsep — Representasi grafis yang menunjukkan konsep, hubungan, dan kata penghubung. Peta konsep digunakan untuk menilai struktur pengetahuan peserta didik. *Padanan asing: concept map.*

Postpositivisme — Paradigma yang mengakui bahwa realitas dapat diteliti melalui bukti, tetapi setiap pengukuran dan kesimpulan memiliki ketidakpastian.

Power statistik — Peluang suatu analisis mendeteksi efek yang benar-benar ada. *Power* dipengaruhi oleh ukuran sampel, ukuran efek, variabilitas, dan tingkat signifikansi.

Predict–Observe–Explain (POE) — Strategi diagnostik yang meminta peserta didik memprediksi, mengamati, dan menjelaskan suatu fenomena. Perbedaan antara prediksi dan pengamatan membantu mengungkap model mental.

Prevalensi miskonsepsi — Proporsi peserta didik atau respons yang menunjukkan kategori miskonsepsi pada populasi atau sampel tertentu.

PRISMA 2020 — Pedoman pelaporan *systematic review* yang menekankan transparansi pencarian, seleksi, penilaian, dan sintesis literatur.

Probing — Pertanyaan lanjutan untuk memperjelas alasan, asumsi, atau hubungan konsep dalam jawaban peserta didik. *Probing* banyak digunakan dalam wawancara diagnostik.

Q

Qualitative research — Penelitian yang berfokus pada makna, proses, pengalaman, dan konteks melalui data seperti wawancara, gambar, serta percakapan. Dalam buku ini digunakan untuk memahami mekanisme terbentuknya konsepsi. *Padanan Indonesia: penelitian kualitatif.*

Quantitative research — Penelitian yang menggunakan data numerik dan analisis statistik untuk memperkirakan prevalensi, hubungan, perubahan, atau kualitas instrumen. *Padanan Indonesia: penelitian kuantitatif.*

R

Rasch analysis — Analisis menggunakan model Rasch untuk menguji kesesuaian butir, responden, kategori, dan fungsi pengukuran. *Padanan Indonesia: analisis Rasch.*

Realisme kritis — Paradigma yang membedakan kejadian yang diamati dari mekanisme mendasar yang menyebabkannya. Dalam penelitian miskonsepsi, respons dipandang sebagai indikator mekanisme pemahaman.

Refleksivitas — Kesadaran peneliti terhadap pengaruh posisi, asumsi, pengalaman, dan keputusan dirinya terhadap penelitian.

Reliabilitas — Tingkat konsistensi dan ketepatan skor atau hasil pengodean untuk penggunaan tertentu. Reliabilitas tinggi tidak otomatis menjamin validitas.

Remediasi miskonsepsi — Tindakan pembelajaran untuk menantang dan merekonstruksi pemahaman yang tidak sesuai dengan konsep ilmiah. *Padanan asing: misconception remediation.*

Representasi kimia — Cara menyatakan konsep kimia melalui tingkat makroskopik, submikroskopik, simbolik, matematis, atau grafis.

Respons waktu nyata — Jawaban atau tindakan peserta didik yang dikumpulkan langsung selama pembelajaran melalui perangkat digital atau teknik asesmen cepat. *Padanan asing: real-time response.*

Retensi — Kemampuan mempertahankan pemahaman setelah jeda waktu. Tes retensi digunakan untuk membedakan perubahan konseptual yang stabil dari peningkatan sesaat.

S

Sampling — Proses memilih sebagian anggota populasi sebagai sumber data. Strategi sampling harus sesuai dengan tujuan dan desain penelitian.

Sentralitas — Ukuran posisi penting suatu konsep dalam jaringan. Konsep dengan sentralitas tinggi memiliki banyak hubungan atau peran strategis dalam struktur pengetahuan.

Simple random sampling — Teknik pemilihan sampel yang memberikan peluang sama kepada setiap anggota populasi. *Padanan Indonesia: sampling acak sederhana.*

Statistik deskriptif — Teknik meringkas data melalui frekuensi, persentase, rerata, median, atau ukuran penyebaran.

Stimulus — Informasi awal berupa pertanyaan, fenomena, gambar, data, atau narasi yang memicu respons peserta didik. Stimulus harus relevan dan tidak menimbulkan beban yang tidak diperlukan.

Stratified sampling — Teknik sampling yang membagi populasi ke dalam strata sebelum memilih sampel. *Padanan Indonesia: sampling berstrata.*

Studi deskriptif — Penelitian yang menggambarkan distribusi, karakteristik, atau pola miskonsepsi tanpa menguji hubungan sebab akibat.

Submikroskopik — Tingkat representasi kimia yang menjelaskan zat dan perubahan melalui partikel, atom, ion, atau molekul yang tidak terlihat langsung.

Survei — Desain penelitian untuk memperoleh gambaran dari sampel yang mewakili populasi. Dalam penelitian miskonsepsi, survei dapat memetakan prevalensi dan variasi konsepsi.

Systematic Literature Review (SLR) — Kajian literatur dengan prosedur pencarian, seleksi, penilaian, dan sintesis yang direncanakan secara sistematis.

T

Teaching to the test — Pembelajaran yang terlalu diarahkan pada pola butir tes sehingga peningkatan skor belum tentu mencerminkan perubahan pemahaman.

Tes diagnostik bertingkat — Tes yang menggabungkan jawaban, alasan, dan/atau keyakinan dalam beberapa tingkat. *Padanan asing: multi-tier diagnostic test.*

Test-retest — Prosedur memberikan instrumen yang sama pada dua waktu untuk memeriksa kestabilan skor.

Think-Aloud Protocol — Teknik meminta peserta didik mengucapkan proses berpikirnya ketika menyelesaikan tugas. Digunakan untuk menelusuri langkah penalaran secara langsung.

Three-tier diagnostic test — Tes tiga tingkat yang umumnya menggabungkan pilihan jawaban, alasan, dan tingkat keyakinan.

Transfer pemahaman — Kemampuan menerapkan konsep pada konteks, representasi, atau masalah baru. Transfer menjadi bukti bahwa perubahan konsep tidak terbatas pada hafalan soal.

Triangulasi — Penggunaan beberapa sumber, metode, peneliti, teori, atau waktu untuk memeriksa kesesuaian dan perbedaan bukti.

Two-tier diagnostic test — Tes dua tingkat yang menggabungkan pilihan jawaban dan alasan. Instrumen ini lebih informatif daripada pilihan ganda konvensional.

U

Uji coba instrumen — Pelaksanaan awal instrumen pada kelompok terbatas untuk menemukan masalah butir, prosedur, waktu, dan sistem penskoran. *Padanan asing: instrument piloting.*

Uji kognitif — Prosedur untuk mengetahui cara peserta didik memahami stimulus, pertanyaan, opsi, dan proses menjawab. *Padanan asing: cognitive interviewing atau cognitive testing.*

Ukuran efek — Besaran yang menunjukkan kekuatan perbedaan atau hubungan, bukan hanya keberadaan signifikansi statistik. *Padanan asing: effect size.*

Umpan balik diagnostik — Informasi spesifik berdasarkan hasil diagnosis yang membantu peserta didik dan guru menentukan langkah perbaikan.

Underrepresentation konstruk — Keadaan ketika instrumen hanya mengukur sebagian kecil konstruk yang seharusnya dinilai. Hal ini membatasi validitas kesimpulan.

Unit analisis — Satuan utama yang dianalisis dalam penelitian, seperti peserta didik, butir, kelas, sekolah, atau episode pembelajaran.

V

Validasi isi — Proses menilai kesesuaian butir dengan domain konsep dan tujuan instrumen melalui pertimbangan pakar serta bukti pendukung lainnya. *Padanan asing: content validation.*

Validitas — Derajat dukungan bukti terhadap ketepatan interpretasi dan penggunaan hasil asesmen. Validitas bukan sekadar sifat tetap suatu instrumen.

Validitas eksternal — Tingkat keterterapan hasil penelitian pada populasi, konteks, waktu, atau situasi lain.

Validitas internal — Tingkat keyakinan bahwa perubahan hasil disebabkan oleh intervensi, bukan oleh faktor pengganggu.

Variabel — Karakteristik yang nilainya dapat berbeda antarpeserta, kelompok, atau waktu. Dalam penelitian, variabel dapat berperan sebagai prediktor, hasil, kontrol, atau moderator.

Visualisasi pola respons — Penyajian grafis kombinasi jawaban, alasan, keyakinan, atau perubahan kategori untuk memudahkan interpretasi data.

W

Wawancara berbasis tugas — Wawancara yang dilakukan ketika peserta didik mengerjakan tugas tertentu. Pertanyaan diarahkan pada alasan, strategi, dan keputusan yang dibuat.

Wawancara klinis — Wawancara mendalam dan fleksibel untuk menelusuri struktur pemikiran peserta didik melalui pertanyaan dan *probing*.

Wawancara kognitif — Wawancara untuk menilai bagaimana responden memahami dan menjawab butir instrumen. Teknik ini digunakan dalam pengembangan dan revisi butir.

Waktu respons — Lama waktu yang digunakan peserta didik untuk memberikan jawaban. Data ini dapat membantu mendeteksi tebakan, keraguan, atau beban pemrosesan.

Within-case analysis — Analisis mendalam terhadap satu kasus sebelum dilakukan perbandingan dengan kasus lain. *Padanan Indonesia: analisis dalam kasus.*

X

x pada estimasi prevalensi — Lambang statistik untuk jumlah kasus atau respons yang memenuhi kategori tertentu dalam rumus ($\hat{p}=x/n$). Dalam buku ini, nilai x dapat menunjukkan jumlah respons yang dikategorikan sebagai miskonsepsi.

Y

Yakin–ragu — Kategori sederhana untuk mencatat tingkat kepastian peserta didik terhadap jawaban atau alasan. Kategori ini dapat digunakan ketika skala keyakinan yang lebih terperinci tidak diperlukan.

Yakin–tidak yakin — Pemisahan respons berdasarkan kepastian subjektif peserta didik. Jawaban salah yang disertai keyakinan tinggi memerlukan pemeriksaan lebih lanjut sebagai indikasi miskonsepsi.

Z

Zat — Materi yang mempunyai komposisi dan sifat tertentu. Dalam diagnosis kimia, pemahaman tentang zat sering diperiksa pada tingkat makroskopik, submikroskopik, dan simbolik.

Zat murni — Materi yang tersusun atas satu jenis unsur atau senyawa dengan komposisi tetap. Konsep ini perlu dibedakan dari campuran untuk mencegah kesalahan klasifikasi materi.

- KASMUI -
Dosen Kimia, Komputasi, IT, AI UNNES

<https://hisabmu.com/>
<https://kasmui.cloud/>

Ketua PCM
Gunungpati 2

Anggota Majelis Tabligh
PDM Kota Semarang
& PWM Jawa Tengah

Tim Pengembang
Software KHGT
MTT PP Muhammadiyah

Praktisi
Ilmu Falak

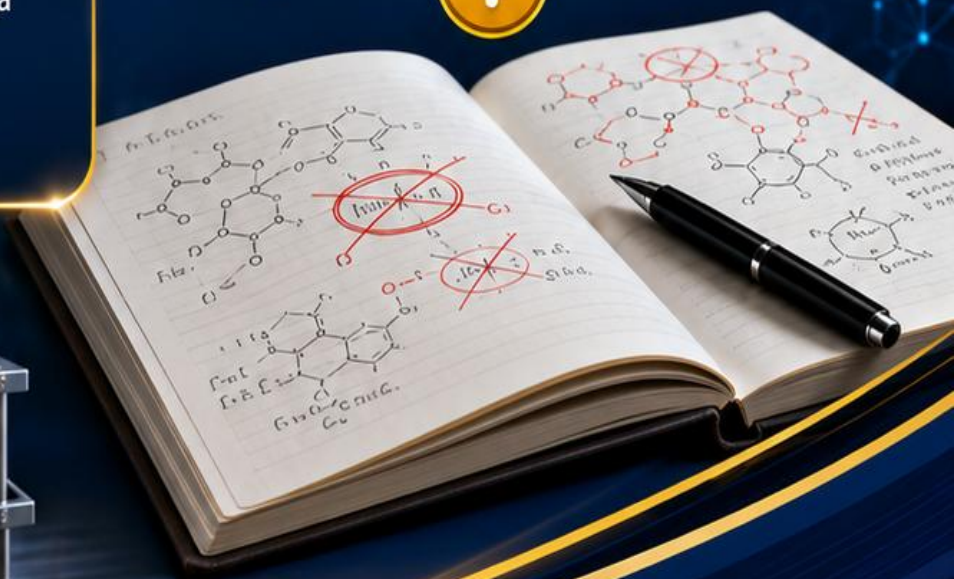
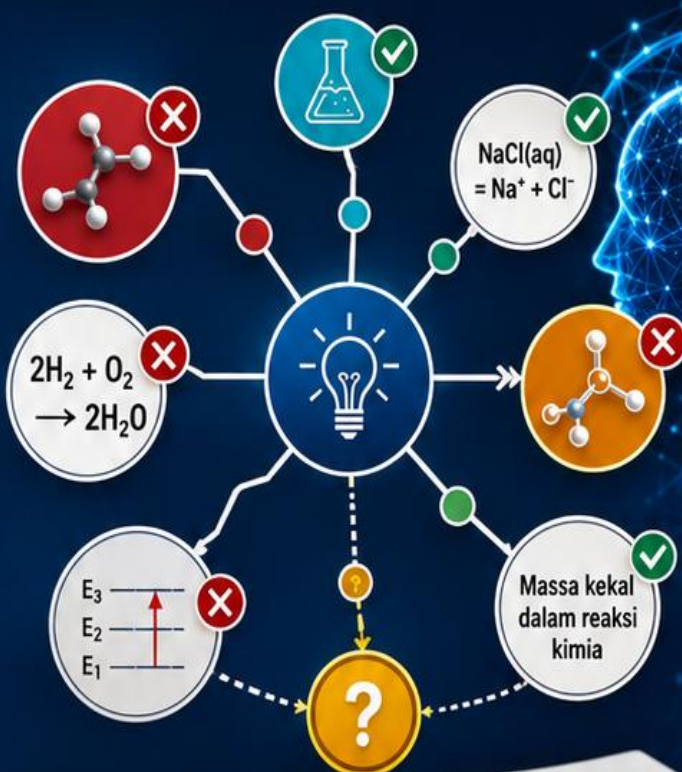
KASMUI

MISKONSEPSI KONSEP KIMIA SMA

BAGIAN 2: DIAGNOSIS DAN PEMETAAN MISKONSEPSI KIMIA

SINOPSIS

Buku ini membahas secara sistematis cara mendiagnosis dan memetakan miskonsepsi kimia pada jenjang SMA. Uraian mencakup prinsip diagnosis miskonsepsi, ragam instrumen asesmen diagnostik, seperti tes pilihan ganda, tes beralasan, two-tier hingga five-tier diagnostic test, CRI, wawancara, peta konsep, analisis butir, validasi ahli, serta pelaporan hasil diagnosis. Buku ini juga mengulas pendekatan penelitian untuk mengidentifikasi pola kesalahan konsep peserta didik sehingga guru, mahasiswa, dan peneliti dapat merancang pembelajaran yang lebih tepat sasaran.



KASMUI