

BAB I

PENDAHULUAN

A. Latar Belakang Masalah

Asesmen memegang peranan penting dalam proses pembelajaran karena menjadi dasar untuk memperoleh informasi mengenai ketercapaian kompetensi siswa serta mendukung pengambilan keputusan pembelajaran. Informasi tersebut hanya akan bermakna apabila diperoleh melalui instrumen yang mampu mengukur kompetensi secara tepat, konsisten, dan sesuai dengan tujuan yang hendak dicapai. Oleh sebab itu, kualitas instrumen menjadi salah satu faktor yang menentukan validitas dan reliabilitas data hasil asesmen. Haladyna dan Rodriguez (2013) menegaskan bahwa pengembangan instrumen perlu dilakukan secara sistematis dengan berlandaskan konstruk yang jelas, spesifikasi tes yang terdefinisi, serta prosedur validasi yang memadai agar hasil pengukuran dapat dipertanggungjawabkan.

Kebutuhan akan instrumen penilaian yang berkualitas semakin menguat seiring implementasi Kurikulum Merdeka yang menekankan pembelajaran bermakna dan penguasaan kompetensi. Salah satu arah kebijakan yang diperkuat melalui Permendikdasmen Nomor 13 Tahun 2025 adalah penerapan pendekatan *deep learning*, yang mendorong siswa untuk memahami konsep secara utuh, menghubungkan pengetahuan yang telah dimiliki dengan informasi baru, serta menerapkannya dalam berbagai konteks pembelajaran (Kementerian Pendidikan Dasar dan Menengah Republik Indonesia, 2025). Sejalan dengan kebijakan tersebut, Hattie dan Donoghue (2016) menjelaskan bahwa *deep learning* terjadi ketika siswa secara aktif membangun hubungan antara pengetahuan awal dan konsep baru sehingga terbentuk pemahaman yang lebih mendalam daripada sekadar mengingat informasi. Dengan demikian, asesmen tidak lagi cukup berfokus pada ketepatan jawaban akhir, tetapi juga perlu mampu menggambarkan kualitas pemahaman konseptual yang dimiliki siswa.

Urgensi pengembangan asesmen yang mampu mengukur pemahaman konseptual juga tercermin dari hasil *Programme for International Student*

Assessment (PISA) 2022. Hasil tersebut menunjukkan bahwa skor literasi sains siswa Indonesia mencapai 383, masih berada di bawah rata-rata *Organisation for Economic Co-operation and Development (OECD)*, yaitu 485 (OECD, 2023). Penilaian PISA tidak hanya mengukur penguasaan pengetahuan faktual, tetapi juga kemampuan menerapkan pengetahuan, menganalisis fenomena ilmiah, serta menggunakan bukti ilmiah untuk menyelesaikan permasalahan. Kemampuan tersebut sejalan dengan proses kognitif mengaplikasikan dan menganalisis dalam Taksonomi Bloom Revisi. Rendahnya capaian tersebut menunjukkan bahwa kemampuan siswa dalam menerapkan dan menganalisis konsep sains masih perlu ditingkatkan. Kondisi ini mengindikasikan perlunya instrumen asesmen yang mampu memberikan informasi yang lebih komprehensif tentang tingkat penguasaan konsep siswa.

Dalam Kurikulum Merdeka, Capaian Pembelajaran dijabarkan melalui Alur Tujuan Pembelajaran (ATP), Tujuan Pembelajaran (TP), dan Indikator Ketercapaian Tujuan Pembelajaran (IKTP). IKTP merupakan indikator operasional yang diturunkan dari TP dalam ATP, yang menggambarkan kompetensi yang diharapkan dikuasai siswa setelah mengikuti proses pembelajaran. Oleh karena itu, pengembangan instrumen penilaian perlu memperhatikan keterkaitan antara indikator soal dan IKTP agar asesmen yang dilakukan benar-benar mengukur kompetensi yang menjadi target pembelajaran. Penyusunan instrumen yang diturunkan secara sistematis dari IKTP diharapkan dapat meningkatkan kesesuaian antara tujuan pembelajaran dan informasi asesmen dalam menggambarkan pencapaian kompetensi siswa.

Salah satu kompetensi penting yang perlu diukur dalam pembelajaran kimia adalah pemahaman konseptual. Anderson dan Krathwohl (2001) menjelaskan bahwa pemahaman konseptual tidak hanya mencakup kemampuan mengenali suatu konsep, tetapi juga memahami hubungan antarkonsep, menerapkannya dalam berbagai situasi, serta menganalisis keterkaitan antarkonsep tersebut. Dalam penelitian ini, pemahaman konseptual diposisikan sebagai konstruk utama yang diukur melalui kemampuan memahami (C2), mengaplikasikan (C3), dan menganalisis (C4) sesuai dengan karakteristik materi. Pada pembelajaran kimia, pemahaman tersebut juga

menuntut kemampuan untuk menghubungkan representasi makroskopik, submikroskopik, dan simbolik secara terpadu (Johnstone, 1991; Treagust et al., 2003).

Salah satu topik yang sesuai untuk mengukur pemahaman konseptual adalah materi asam basa. Topik ini menjadi landasan bagi berbagai materi kimia lanjutan karena memuat konsep-konsep yang saling berkaitan, mulai dari teori asam-basa, kekuatan asam dan basa, derajat keasaman (pH), indikator asam basa, konstanta ionisasi (K_a/K_b), hingga derajat ionisasi (α). Penelitian ini membatasi ruang lingkup pada konsep-konsep tersebut, sedangkan larutan penyangga, hidrolisis garam, dan titrasi asam-basa tidak termasuk dalam cakupan penelitian karena diajarkan sebagai unit pembelajaran tersendiri pada jenjang SMA.

Pemilihan materi asam basa juga didasarkan pada karakteristik konseptualnya yang kompleks. Untuk memahami topik ini, siswa perlu mengaitkan berbagai konsep yang saling bergantung serta menghubungkan fenomena yang diamati dengan representasi submikroskopik dan simbolik. Kompleksitas tersebut menjadikan materi asam basa sesuai untuk mengukur kemampuan memahami, mengaplikasikan, dan menganalisis konsep secara terpadu.

Hasil diskusi awal dengan guru kimia serta pengamatan peneliti selama kegiatan akademik di SMA Negeri 48 Jakarta menunjukkan bahwa sebagian siswa masih mengalami kesulitan memahami keterkaitan antarkonsep, terutama ketika pembelajaran melibatkan representasi simbolik dan penyelesaian perhitungan. Guru juga menyampaikan bahwa asesmen yang digunakan masih didominasi oleh soal pilihan ganda konvensional sehingga informasi yang diperoleh lebih banyak menggambarkan ketepatan jawaban akhir daripada tingkat pemahaman konseptual siswa. Kondisi tersebut menunjukkan perlunya instrumen penilaian yang mampu memberikan informasi yang lebih rinci mengenai penguasaan konsep, khususnya pada materi yang memiliki karakteristik konseptual yang kompleks, seperti materi asam basa.

Indikasi awal di lapangan tersebut sejalan dengan berbagai hasil penelitian yang menunjukkan bahwa materi asam basa merupakan salah satu topik kimia yang paling sering menimbulkan miskonsepsi pada siswa. Kesulitan tersebut tidak hanya berkaitan dengan penggunaan konsep dan perhitungan, tetapi juga dengan kemampuan untuk menghubungkan konsep melalui representasi makroskopik, submikroskopik, dan simbolik.

Berbagai penelitian menunjukkan bahwa materi asam basa masih menjadi salah satu topik yang banyak menimbulkan miskonsepsi pada siswa. Penelitian Romine, Todd, dan Clark (2016) menunjukkan bahwa pemahaman siswa terhadap konsep asam basa melibatkan berbagai dimensi konseptual yang saling berkaitan dan berkembang secara progresif. Mubarakah, Mulyani, dan Indriyanti (2018) menemukan bahwa siswa Indonesia dan Thailand masih mengalami miskonsepsi pada konsep teori asam basa, kekuatan asam dan basa, serta pH, dengan persentase rata-rata miskonsepsi sebesar 30,76% pada siswa Indonesia. Djarwo dan Kafiar (2023) melaporkan bahwa miskonsepsi ditemukan pada berbagai subkonsep asam-basa dengan rentang 9%–45%, terutama pada konsep pH, kekuatan asam-basa, dan konstanta ionisasi, dengan 29% mahasiswa pendidikan kimia masih mengalami miskonsepsi. Temuan yang lebih mutakhir yang dilaporkan oleh Chimayah, Windayani, Kurniasih, Subarkah, dan Sukmawardani (2025) menunjukkan bahwa siswa masih mengalami miskonsepsi pada teori asam basa, indikator asam basa, dan perhitungan pH larutan, dengan persentase miskonsepsi tertinggi mencapai 30,96%. Hasil penelitian tersebut menunjukkan bahwa materi asam basa memerlukan instrumen yang mampu mengukur pemahaman konseptual secara lebih mendalam.

Berbagai instrumen telah dikembangkan untuk mengidentifikasi miskonsepsi, seperti *two-tier test* dan *four-tier test*. Kedua instrumen tersebut mampu mengungkap alasan di balik jawaban siswa sehingga memberikan informasi yang lebih rinci mengenai bentuk miskonsepsi yang dimiliki. Namun, penggunaannya memerlukan format respons yang lebih kompleks serta analisis hasil yang relatif lebih rumit dibandingkan dengan soal pilihan ganda. Di sisi lain, asesmen yang digunakan di sekolah masih didominasi oleh

soal pilihan ganda konvensional karena praktis, mudah dinilai, dan objektif. Meskipun demikian, sistem penskoran dikotomis pada soal pilihan ganda konvensional hanya membedakan jawaban benar dan salah sehingga belum mampu menggambarkan tingkat penguasaan konsep secara bertahap. Selain itu, peluang menebak jawaban yang benar menyebabkan skor yang diperoleh siswa tidak selalu mencerminkan kemampuan sebenarnya (Haladyna, 2004).

Salah satu alternatif yang berpotensi mengatasi keterbatasan tersebut adalah *Multiple Choice Complex* (MCC). Berbeda dengan pilihan ganda konvensional yang hanya memiliki satu jawaban benar, MCC memungkinkan lebih dari satu jawaban benar dalam satu butir soal (Haladyna & Rodriguez, 2013). Karakteristik tersebut memungkinkan siswa menunjukkan tingkat penguasaan konsep yang lebih beragam karena setiap respons dapat diberi skor secara parsial (*partial credit scoring*). Dengan demikian, hasil asesmen tidak hanya menunjukkan benar atau salah, tetapi juga memberikan informasi tentang tingkat penguasaan konsep yang dimiliki siswa.

Data yang dihasilkan melalui penskoran parsial bersifat politomis sehingga sesuai dianalisis menggunakan *Rasch Partial Credit Model* (PCM). Model ini dirancang untuk menganalisis respons dengan lebih dari dua kategori skor serta memberikan informasi mengenai karakteristik butir dan kemampuan siswa (Masters, 1982). Penggunaan PCM juga memungkinkan setiap kategori skor dievaluasi sehingga kualitas instrumen tidak hanya ditinjau dari validitas isi, tetapi juga dari karakteristik empiris setiap butir soal.

Pengembangan instrumen MCC tidak hanya bergantung pada format soal, tetapi juga pada kualitas konstruksi setiap butir. Haladyna (1993) mengkritik format *Type K* karena memungkinkan siswa memanfaatkan strategi eliminasi kombinasi jawaban sehingga skor yang diperoleh belum tentu mencerminkan pemahaman konsep. Penelitian ini menggunakan format "pilih semua jawaban yang benar", sehingga setiap opsi dievaluasi secara mandiri. Pendekatan tersebut sejalan dengan temuan Couch et al. (2018) yang menunjukkan bahwa format multirespons mampu memberikan informasi yang lebih rinci apabila setiap opsi disusun berdasarkan miskonsepsi yang telah teridentifikasi. Oleh karena itu, distraktor pada instrumen MCC dalam

penelitian ini dikembangkan berdasarkan miskonsepsi yang telah dilaporkan dalam berbagai penelitian, sehingga setiap alternatif jawaban merepresentasikan bentuk kesalahan konseptual yang nyata.

Berdasarkan kajian literatur, masih terdapat beberapa kesenjangan penelitian yang perlu diatasi. Pertama, soal pilihan ganda konvensional umumnya menggunakan penskoran dikotomis sehingga belum mampu menggambarkan tingkat penguasaan konsep secara parsial. Kedua, instrumen diagnostik seperti *two-tier test* dan *four-tier test* mampu mengidentifikasi miskonsepsi, tetapi memerlukan format respons dan analisis yang relatif lebih kompleks. Ketiga, penelitian yang mengembangkan instrumen *Multiple Choice Complex* pada materi asam basa dengan mengintegrasikan IKTP Kurikulum Merdeka, distraktor berbasis miskonsepsi, penskoran parsial, dan analisis Rasch Partial Credit Model (PCM) masih terbatas. Padahal, integrasi keempat aspek tersebut berpotensi menghasilkan instrumen yang tidak hanya memenuhi kriteria psikometrik, tetapi juga mampu memberikan informasi yang lebih komprehensif tentang tingkat pemahaman konseptual siswa.

Berdasarkan uraian tersebut, penelitian ini bertujuan untuk mengembangkan instrumen *Multiple Choice Complex* (MCC) yang valid dan reliabel untuk mengukur pemahaman konseptual siswa terhadap materi asam basa. Instrumen disusun berdasarkan IKTP Kurikulum Merdeka dengan mengacu pada kemampuan memahami (C2), mengaplikasikan (C3), dan menganalisis (C4), serta mempertimbangkan representasi makroskopik, submikroskopik, dan simbolik dalam bidang kimia. Kualitas instrumen dievaluasi melalui validitas isi menggunakan indeks Aiken's V dan analisis empiris menggunakan *Rasch Partial Credit Model* (PCM) dengan perangkat lunak jMetrik. Penelitian ini dibatasi pada tahap uji coba terbatas untuk memperoleh karakteristik empiris butir soal hasil pengembangan.

B. Fokus Penelitian

Berdasarkan latar belakang masalah yang telah diuraikan, penelitian ini difokuskan pada pengembangan instrumen asesmen berbentuk *Multiple Choice Complex* (MCC) untuk mengukur pemahaman konseptual siswa SMA pada materi asam basa. Fokus tersebut meliputi:

1. Menyusun *blueprint* dan butir soal MCC yang selaras dengan Indikator Ketercapaian Tujuan Pembelajaran (IKTP) Kurikulum Merdeka, kemampuan kognitif C2–C4, serta tiga level representasi kimia (makroskopik, submikroskopik, dan simbolik).
2. Menguji validitas isi instrumen melalui penilaian pakar (*expert judgment*) dengan menggunakan analisis Aiken's V.
3. Menguji kualitas empiris instrumen melalui uji coba terbatas dan menganalisis kualitas instrumen menggunakan pendekatan Rasch Partial Credit Model (PCM), yang meliputi tingkat kesulitan butir, kesesuaian butir terhadap model, reliabilitas instrumen, dan indeks separasi responden.
4. Menghasilkan produk final berupa 20 butir soal *Multiple Choice Complex* (MCC) yang telah memenuhi kriteria kualitas instrumen berdasarkan hasil validasi isi dan analisis empiris, beserta *blueprint* dan panduan penskoran.

Penelitian ini bukan merupakan penelitian eksperimen, bukan penelitian untuk mengetahui tingkat pemahaman atau profil miskonsepsi siswa, bukan penelitian implementasi kurikulum, dan bukan penelitian pengembangan media atau modul pembelajaran.

C. Perumusan Masalah

Berdasarkan fokus penelitian di atas, rumusan masalah dalam penelitian ini dirumuskan sebagai berikut:

1. Bagaimanakah prosedur pengembangan instrumen *Multiple Choice Complex* (MCC) untuk mengukur pemahaman konseptual siswa pada materi asam basa?
2. Bagaimanakah tingkat validitas isi instrumen *Multiple Choice Complex* (MCC) untuk mengukur pemahaman konseptual siswa pada materi asam basa berdasarkan penilaian pakar (*expert judgment*) dengan analisis Aiken's V?
3. Bagaimanakah kualitas empiris instrumen *Multiple Choice Complex* (MCC) untuk mengukur pemahaman konseptual siswa pada materi asam basa berdasarkan analisis Rasch *Partial Credit Model* (PCM) pada uji coba terbatas, yang meliputi:

- b) tingkat kesulitan butir (*item difficulty*),
- c) kesesuaian butir dengan model (*outfit mean square/MNSQ*),
- d) reliabilitas person (*person reliability*),
- e) reliabilitas item (*item reliability*), dan
- f) indeks separasi person (*person separation index*)?

D. Manfaat Penelitian

Hasil penelitian ini diharapkan dapat memberikan manfaat, baik secara teoretis maupun praktis, sebagai berikut:

1. Manfaat Teoritis

Penelitian ini diharapkan dapat memberikan kontribusi bagi pengembangan ilmu pengetahuan, khususnya dalam bidang asesmen pendidikan kimia, dengan cara:

- a) Memperkaya kajian teoretis mengenai pengembangan instrumen *Multiple Choice Complex* (MCC) yang valid dan reliabel.
- b) Memberikan sumbangan pemikiran dalam penerapan Rasch *Partial Credit Model* (PCM) untuk menganalisis kualitas butir soal berbasis data *polytomous* pada instrumen MCC.
- c) Mengintegrasikan teori tiga level representasi kimia (makroskopik, submikroskopik, dan simbolik) ke dalam konstruksi instrumen penilaian pemahaman konseptual.
- d) Menjadi referensi bagi peneliti selanjutnya yang akan mengembangkan instrumen penilaian serupa pada topik kimia lainnya.

2. Manfaat Praktis

a) Bagi Guru Kimia

- 1) Menyediakan instrumen penilaian alternatif berupa MCC yang telah teruji validitas dan reliabilitasnya untuk mengukur pemahaman konseptual siswa pada materi asam basa.
- 2) Memberikan contoh *blueprint* dan panduan penskoran yang dapat diadaptasi untuk kebutuhan asesmen di kelas.

b) Bagi Siswa

- 1) Memberikan pengalaman mengerjakan soal berbasis *Multiple Choice Complex* (MCC) yang menuntut kemampuan memahami, mengaplikasikan, dan menganalisis konsep kimia.
- 2) Mendukung proses asesmen pemahaman konseptual siswa pada materi asam basa melalui instrumen yang telah melalui proses validasi dan analisis kualitas empiris.

c) Bagi Peneliti Lain

- 1) Menjadi dasar atau acuan untuk penelitian lanjutan, seperti uji coba lapangan dengan sampel yang lebih besar dan heterogen, atau pengembangan instrumen MCC pada materi kimia lainnya.
- 2) Memberikan kontribusi pada pengembangan instrumen asesmen berbasis Teori Respons Butir (IRT) di bidang pendidikan kimia.

d) Bagi Program Studi Pendidikan Kimia dan Sekolah

- 1) Menambah khazanah hasil penelitian dalam bidang pengembangan instrumen penilaian kimia, khususnya instrumen *Multiple Choice Complex* (MCC) untuk mengukur pemahaman konseptual siswa.
- 2) Menjadi referensi dalam pengembangan asesmen kimia berbasis pemahaman konseptual di sekolah.
- 3) Sebagai bahan pengayaan pada mata kuliah Asesmen Pembelajaran Kimia dan Metodologi Penelitian di Program Studi Pendidikan Kimia.