

Uji Validitas Sebagai Penentuan Kualitas Butir Soal Pilihan Ganda "Pacu Kimia" FMIPAK Universitas Negeri Manado

Freetje Waworuntu^{a*}, Jeane M. Tuerah^a, Sonandar M. Tengker^a, Abdon Saiya^a, Vlagia Paat^a, Stefan Rumengan^a

^a Jurusan Kimia, Universitas Negeri Manado, 95618, Minahasa, Sulawesi Utara, Indonesia

INFO ARTIKEL

Diterima 1 Juli 2024
Disetujui 29 September
2024

Key word:

Validity Test,
Discrimination Index,
Level of Difficulty,
Multiple Choice Test,
Chemistry lesson

Kata kunci:

Uji Validitas,
Indeks Deskriminasi,
Taraf Kesulitan,
Tes Pilihan Ganda,
Pelajaran Kimia

ABSTRACT

The purpose of this research is to determine the quality of multiple-choice items in the "Chemistry Boost" test from FMIPAK Unima, based on responses from 80 high school students answering 40 items. The validity test used a combination of Discrimination Index (DI) and Difficulty Level (DL). DI has 6 categories: Very Good, Good Enough, Marginal, Poor Discrimination, Negative, while DL consists of 6 categories: Discarded, Very Easy, Easy, Maintained, Very Difficult, and Reviewed for deficiencies or taught. The results obtained are 23 valid items, or 57.5%. Based on the results, the cognitive processes of Anderson's revised Bloom Taxonomy were determined, with valid items in C1, C2, C3, C4, C5, and C6 being 2, 2, 12, 6, 0, respectively. Items categorized as Higher Order Thinking Skills (HOTS) are in C4 and C5, totaling 18 items. Valid chemistry topics include Acid-Base, Atomic Structure, Stoichiometry, and Reaction Rates. Topics less understood by participants include Compounds and Chemical Formulas, Solubility, Chemical Equilibrium, Ionization Energy, Redox, Electrochemistry, and Hydrocarbons.

ABSTRAK

Tujuan penelitian ini adalah untuk menentukan kualitas butir soal pilihan ganda "Pacu Kimia" FMIPAK Unima dari 80 siswa SMA dalam menjawab 40 butir soal. Uji validitas yang digunakan adalah perpaduan Indeks Deskriminasi (ID) dengan Taraf Kesulitan (TK). ID memiliki 6 kategori yakni Sangat Bagus, Cukup Bagus, Marjinal, Diskriminasi buruk, Negatif sedangkan TK terdiri dari 6 kategori yakni Dibuang, Sangat Mudah, Mudah, Dipertahankan, Sangat Sulit dan Ditinjau kekurangannya atau diajarkan. Hasil yang diperoleh adalah kualitas butir soal yang valid sebanyak 23 item atau 57,5 %. Berdasarkan hasil dilanjutkan menentukan proses kognitif dari Taksonomi Bloom revisi Anderson yakni butir soal yang valid dari C1, C2, C3, C4, C5 dan C6 adalah 2, 2, 12, 6, 0. Butir soal untuk kategori HOTS adalah C4, C5 sebanyak 18 item. Berdasarkan isi/materi pelajaran kimia yang valid adalah Asam basa, Struktur Atom, Stoikhiometri, dan Laju Reaksi, Kemudian yang kurang dipahami oleh peserta meliputi Senyawa dan Rumus Kimianya, Kelarutan, Keseimbangan Kimia, Energi Ionisasi, Redoks, Elektrokimia, dan Hidrokarbon.

*e-mail:

freetjewaworuntu@unima
.ac.id

*Telp:

082192373778

Pendahuluan

Kegiatan Pacu Kimia merupakan acara perlombaan kimia secara periodik tahunan dari Jurusan Kimia FMIPAK Unima. Acara ini menghadirkan siswa-siswa terbaik yang mempelajari kimia di tingkat SMA se Sulawesi Utara. Pada babak awalnya dilakukan seleksi kemampuan siswa menggunakan tes pilihan ganda sebagai seleksi penguasaan pelajaran kimia mereka di SMA. Analisis respon siswa pada tes pilihan ganda sebagai hasil belajar kimia ini dijadikan kajian penelitian instrumen tentang validitas berdasarkan tabel matriks spesifikasi dimensi proses kognitif dan isi pengetahuan kimia. Kemudian penerapan validitas isi pada soal pilihan ganda ini menggunakan cara Teori Test Klasik (*Classical Test Theory*, CTT).

Beberapa penjelasan tentang validitas adalah sebagai berikut:

- CTT terutama berfokus pada kinerja secara keseluruhan (misalnya, skor yang diamati, varians skor, dll.). Dalam istilah praktis, menerapkan CTT lebih mudah karena seseorang dapat dengan cepat memperoleh sejumlah statistik yang berguna (misalnya, skor rata-rata, indeks diskriminasi, point-biserial, dll.) [1].
- Korelasi point-biserial adalah nilai korelasi *product moment* Pearson jika salah satu variabel bersifat dikotomis dan variabel lainnya bersifat metrik. Nilai berkisar dari +1, hubungan positif sempurna; sampai nol, tidak ada hubungan sama sekali; ke -1, korelasi negatif sempurna [2].
- Pengertian uji validitas adalah sejauh mana ketepatan dan kecermatan suatu alat ukur dalam melakukan fungsi ukurnya. Suatu tes atau instrumen pengukur dapat dikatakan mempunyai validitas yang tinggi apabila alat tersebut menjalankan fungsi ukurnya sesuai dengan maksud dilakukannya pengukuran tersebut [3].
- Validitas adalah sejauhmana tes itu mengukur apa yang dimaksudkan untuk diukur sedangkan validitas isi adalah validitas yang fokus kepada elemen-elemen apa yang diukur. Validitas isi adalah penjelasan tentang suatu alat ukur secara substantif. Hasil akhir dari validitas isi adalah penilaian tentang kelayakan isi

tes [4].

- Kualitas Tes Pacu kimia dalam bentuk pilihan ganda merujuk pada validitas isi.

Kimia adalah studi tentang materi dan perubahan yang dialaminya [5]. Kimia adalah disiplin ilmu eksperimental. Kimia mencoba menjelaskan fenomena makroskopis dalam bentuk struktur molekul materi. Kimia menyelidiki dinamika proses dan reaksi. Kimia menyelidiki perubahan energi yang terjadi dalam reaksi kimia. Kimia bertujuan memahami dan menjelaskan kehidupan dikaitkan dengan struktur kimia dan proses dalam sistem kehidupan. Kimia menggunakan bahasa khusus [6].

Hakekat ilmu kimia adalah menguraikan peristiwa kimia yang telah ada dan dapat terjadi (sintesis) di alam semesta. Kimia berupaya memahami perilaku materi dengan mempelajari perilaku atom dan molekul. Artinya mempelajari peristiswa kimia yang berkaitan dengan perilaku susunan dan struktur materi serta perubahan energi yang terjadi disetiap atom dan molekul.

Kimia diungkap secara verbal maupun kuantitatif serta divisualisasi dalam bentuk teks dan simbol-simbol unsur kimia, rumus kimia dan rumus persamaan matematika serta simbol lainnya. Isi dan materi pelajaran kimia memiliki keunikan tersendiri. Termasuk keunikan dari pelajaran kimia adalah merepresentasikan dalam tingkat makroskopik, submikroskopik, dan representatif simbolik. Kimia penuh dengan simbol dan notasi dan membentuk susunan simbol yang bermakna. Beberapa pernyataan yang terkait adalah sebagai berikut:

- Istilah makroskopik, submikroskopik, dan simbolik mengikuti segitiga tingkat representasi dari Johnstone [7].
- Segitiga tersebut merupakan keterkaitan dari a) demonstrasi percobaan laboratorium, b) ilustrasi tingkat partikel, dan c) rumus kimia dan matematika.
- Sebutan triplet kimia karena gagasan pengetahuan kimia yang direpresentasikan dalam tiga ciri utama: makro, submikro, dan simbolik [8].
- Memahami hubungan antara ketiga tingkatan tersebut dan bagaimana keterkaitannya satu sama lain, diperlukan

agar siswa dapat memahami penjelasan lengkap tentang fenomena kimia [9].

Pencapaian belajar kimia seperti mata pelajaran lainnya mengacu pada Taksonomi Bloom Revisi yang memiliki 2 (dua) dimensi yakni a) dimensi proses kognisi dan b) dimensi pengetahuan.

Taksonomi Bloom Revisi pada dimensi kognitif (*Cognitive*) terdiri dari enam level yakni C1 = *remembering* (mengingat), C2 = *understanding* (memahami), C3 = *applying* (menerapkan), C4 = *analyzing* (menganalisis, mengurai), C5 = *evaluating* (menilai), dan C6 = *creating* (mencipta). Pada dimensi pengetahuan terdiri dari konseptual, faktual, prosedural, dan metakognitif. Dimensi pengetahuan hanya memuat jenis-jenis pengetahuan sedangkan dimensi kognitif memuat macam-macam proses kognitif [10].

Pengertian dan contoh dalam pelajaran kimia dari proses kognitif [11] adalah sebagai berikut:

1. Mengingat (C1) adalah kembali pengetahuan relevan yang bersumber dari memori waktu lama (*long term memory*), seperti: mengenal, menyebutkan. Sebagai contoh mengenal lambang unsur kimia, dan menyebutkan waktu penemuan kimia
2. Memahami (C2) adalah membangun pengertian pesan instruksional mencakup lisan, tulisan, dan komunikasi grafik (*interpreting, exemplifying, classifying, summarizing, inferring, comparing, explaining*). Sebagai contoh menguraikan konsep kimia dengan ungkapan sendiri, menyebutkan contoh senyawa organik, menggolongkan jenis karbohidrat, menduga struktur senyawa organik, membandingkan unsur pada tabel periodik, menjelaskan arah reaksi.
3. Mengaplikasikan (C3) adalah menggunakan prosedur (*executing, implementing*). Sebagai contoh mempraktekkan penyulingan, menggunakan hukum gas ideal.
4. Menganalisis (C4) adalah membagi materi menjadi bagian pokok dan menentukan cara bagian tersebut saling berhubungan satu sama lain dan berhubungan dengan keseluruhan struktur (*differentiating, organizing, attributing*). Sebagai contoh

mengidentifikasi unsur penting dari masalah, menganalisis laporan penelitian kimia, memberitahukan sikap penulis artikel kimia.

5. Mengevaluasi (C5) adalah menyatakan pendapat berdasarkan kriteria dan standar (*checking, critiquing*). Sebagai contoh memeriksa rasionalitas pemecahan masalah, dan mengkritisi perbedaan metode ilmiah.
6. Mengkreasi (C6) adalah menempatkan unsur bersama dengan bentuk koherensi atau keseluruhan fungsi, mengorganisir unsur ke dalam pola ataususunan baru (*generating, planning, producing*). Sebagai contoh menghasilkan hipotesis, merencanakan metode ilmiah, menulis karya ilmiah

Dimensi Pengetahuan yang merupakan konten materi pelajaran [12] adalah sebagai berikut:

1. Faktual adalah a) Pengetahuan tentang terminologi; dan b) Pengetahuan tentang bagian detail dan unsur-unsur.
2. Konseptual adalah a) Pengetahuan tentang klasifikasi dan kategori; b) Pengetahuan tentang prinsip dan generalisasi; dan c) Pengetahuan tentang teori, model, dan struktur.
3. Prosedural adalah a) Pengetahuan tentang keterampilan khusus dengan suatu bidang tertentu dan pengetahuan tentang algoritme; b) Pengetahuan tentang teknik dan metode yang berhubungan dengan suatu bidang tertentu; dan c) Pengetahuan tentang kriteria untuk menentukan kapan suatu prosedur tepat untuk digunakan
4. Metakognitif adalah a) Pengetahuan tentang strategi; b) Pengetahuan tentang konteks dan kondisi yang sesuai; dan c) Pengetahuan tentang diri sendiri.

Matriks penggabungan proses kognitif dan pengetahuan sebagai unit tabel spesifikasi butir soal pilihan ganda yang dapat dilihat pada Tabel 1. Tabel spesifikasi untuk menentukan indikator pencapaian belajar dan digunakan dalam penetapan distribusi kisi-kisi soal. Tabel ini sangat berguna untuk melihat penyebaran soal-soal yang ada.

Dengan demikian matriks dimensi proses kognitif dan pengetahuan dapat

digunakan sebagai tabel spesifikasi atau indikator untuk pengukuran kemampuan siswa yang dinyatakan sebagai hasil belajar. Dimensi proses kognitif yang dimulai dari hierarki yang sederhana sampai yang paling kompleks berdasarkan taksonomi Bloom revisi yakni C1 = mengingat, C2 = memahami, C3 = menerapkan, C4 = menganalisis, C5 = mengevaluasi, dan C6 = mencipta. Kemudian dimensi pengetahuan yang bersifat dari konkret hingga abstrak yakni a) faktual, b) konseptual, c) prosedural dan d) metakognitif.

Taksonomi Bloom dibagi menjadi a) keterampilan berpikir tingkat rendah (*Lower-Order Thinking Skills*) untuk C1, C2, dan C3; dan b) keterampilan berpikir tingkat tinggi (*Higher-Order Thinking Skills*, HOTS) untuk C4, C5, dan C6 [13]. Dengan demikian, dapat dikatakan bahwa C1 atau mengingat sebagai tingkat kemampuan terendah dan C6 atau mencipta sebagai tingkat kemampuan tinggi.

Tabel 1. Matriks Indikator Proses Kognitif dan Materi/Isi Suatu Mata Pelajaran

Proses Kognitif	Dimensi Pengetahuan			
	Fakta	Konsep	Prosedur	Meta-kognisi
C1 Mengingat				
C2 Memahami				
C3 Menerapkan				
C4 Menganalisis				
C5 Mengevaluasi				
C6 Mencipta				

Tabel 1 sebagai tabel spesifikasi tes yang berbentuk matriks merupakan langkah awal dalam penyusunan tes berdasarkan tujuan pembelajaran. Langkah pertama ini menunjukkan gambaran keseluruhan karakteristik yang dimiliki sebuah tes. Spesifikasi yang jelas akan mempermudah dalam mengidentifikasi butir soal. Dalam Kisi-kisi merupakan tabel matriks yang berisi spesifikasi soal-soal yang akan dibuat. Hasil dalam matrik merupakan kisi-kisi butir soal [14, 15, 16, 17, 18].

Kisi-kisi merupakan acuan penulisan butir soal tes pilihan ganda. Tes pilihan ganda dianggap sebagai salah satu bentuk penilaian pendidikan yang paling efektif dan bertahan lama yang tetap dipraktikkan hingga saat ini [19]. Bentuk tes pilihan ganda tergolong instrumen respon seleksi. Bentuk ini mudah diakses [20]. Soal tes bentuk pilihan ganda/*multiple choice test* terdiri dari stem, yaitu bentuk pertanyaan/bentuk pernyataan yang belum sempurna, dan opsi atau pilihan jawaban yang berbentuk perkataan, bilangan, ataupun kalimat. soal tes bentuk pilihan ganda/*multiple choice test* dapat digunakan menilai hasil belajar yang lebih kompleks dan berkaitan dengan aspek ingatan, pengertian, aplikasi, analisis, sintesis, dan evaluasi [21].

Metode Penelitian

Tujuan penelitian adalah untuk memperoleh gambaran validitas setiap butir soal. Metode penelitian ini berfokus pada deskriptif kuantitatif yakni: a) perhitungan indeks daya butir soal dengan teknik korelasi *point biserial*, dan b) perhitungan taraf kesulitan butir soal dengan menghitung proporsi benar. Adanya perhitungan ini akan menghasilkan butir soal yang dinyatakan valid kemudian sebagai masukan dalam pengembangan instrumen pendidikan kimia.

Lembar jawaban yang diolah dari responden sebanyak 88 siswa yang mengisi pilihan ganda sebanyak 40 butir soal. Mereka terdiri dari Kelas 11 dan 12. Mereka berasal dari beberapa SMA Negeri dan Swasta di Sulawesi Utara yakni 1) Manado Internasional School, 2) SMA Eber Haezar, 3) SMA Negeri 1 Tagulandang, 4) SMA Negeri 9 Manado, 5) SMA Negeri 1 Langowan, 6) SMA Kristen 2 Tomohon, 7) SMA Rex Mundi Manado, 8) SMA Kristen 1 Tomohon, 9) SDH Holand Village Manado, 10) SMA Khatolik Theodorus KTM, dan 11) SMA Negeri 1 Manado.

Berdasarkan Tabel 1 Matriks Proses Kognitif dan Pengetahuan maka dikembangkan kisi-kisi 40 butir soal pilihan ganda pelajaran kimia yang dapat dilihat pada Tabel XX.

Pengolahan data untuk indeks diskriminasi dengan menghitung koefisien korelasi *point-biserial* (r_{pb}) setiap butir soal.

Rumus korelasi point-biserial tersebut adalah:

$$r_{pb} = \frac{\bar{x}_n - \bar{x}_t}{s_t} \sqrt{p/q}$$

Keterangan:

r_{pb} = koefisien korelasi *point-biserial* setiap butir soal

$\bar{x}_n$ = rata-rata skor pada butir soal

$\bar{x}_t$ = rata-rata skor total

s_t = standar deviasi dari skor total

p = proporsi jawaban benar

q = (1 - p) atau proporsi jawaban salah

Korelasi antara variabel kontinu dan dikotomis menggunakan r_{pb} . Diawali dengan mendefinisikan variabel kontinu, x , sebagai skor siswa dalam ujian, dan variabel dikotomi, y , sebagai indikasi apakah siswa menjawab pertanyaan dengan salah (mendapat skor 0) atau benar (mendapat skor 1). Karena kita memilih 0 dan 1 sebagai nilai y , maka rata-rata y adalah proporsi siswa yang menjawab pertanyaan dengan benar, dilambangkan dengan p . Sebaliknya, proporsi siswa yang menjawab salah, dinotasikan q , adalah 1 - p [22].

Pengujian ini kompetibel dengan tes pilihan ganda karena mengkorelasikan jumlah (benar) yang berada pada skor total (x_n) dari setiap butir soal yang berada pada skor total (x_t) dengan jumlah skor total ($\bar{x}_t$).

Kemudian menghitung taraf kesulitan yakni proporsi (p) benar setiap butir soal dengan rumus p = jumlah benar butir soal (B) per jumlah responden (JR) dengan rumus sebagai berikut:

$$p = \frac{B}{JR}$$

Perhitungan koefisien *point biserial* dan proporsi benar setiap butir soal dibantu oleh Program aplikasi MS Excel.

Kriteria sebagai rujukan interpretasi data diperoleh diantaranya,

1. Rujukan Interpretasi statistik inferensial berdasarkan r_{tabel} pada $n = 88$ dan $\alpha = 0,05$, koefisien korelasi point-biserial (r_{pb}) = 0,20. Apabila $r_{hitung} \geq 0,20$ maka dinyatakan bahwa indeks diskriminasi memadai pada setiap butir soal yang dihitung.
2. Rujukan interpretasi statistik deskripsi untuk indeks diskriminasi berdasarkan Tabel 2 [23].
3. Rujukan interpretasi statistik deskripsi

untuk taraf kesulitan berdasarkan Tabel 3 [23].

Tabel 2. Interpretasi Koefisien Korelasi *Point Biserial* (r_{pb}) sebagai Indeks Diskriminasi (DB)

Nilai r_{pb}	Rekomendasi	Kode
0.40-1.00	Sangat Bagus	SB
0.30-0.39	Cukup Bagus	CB
0.20-0.29	Marjinal atau dapat diterima	M
0.00-0.19	Diskriminasi yang buruk	TB
Negatif	Siswa yang berprestasi rendah mendapat jawaban benar lebih banyak dibandingkan siswa yang berprestasi tinggi	N

Tabel 3. Interpretasi Taraf Kesulitan Berdasarkan Proporsi

Nilai p	Rekomendasi	Kode
0.90-1.00	Dibuang	X
0.75-0.89	Sangat Mudah	SM
0.71-0.74	Mudah	M
0.30-0.70	Dipertahankan	B
0.20-0.29	Sangat sulit	SS
0.00-0.19	Meninjau kekurangan atau mengajar kembali	T

Hasil dan Pembahasan

Data empirik menunjukkan sebagai berikut:

1. Responden 88 orang (88 lembar kertas jawaban). Jawaban benar maksimum pada nomor soal PG-01 = 68 sedangkan jawaban minimum pada nomor PG-15 dan PG-36 = 11.
2. Jumlah butir soal pilihan ganda sebanyak 40 item. Skor benar maksimum = 30 sedangkan skor benar minimum = 3. Rata-rata skor total = 12,16 dan Standar Deviasi = 5,42.
3. Hasil analisis indeks diskriminasi (ID) berdasarkan koefisien korelasi point-biserial dengan perhitungan dibantu Excel pada Tabel 4 yang memadai dari nomor 1 - 32 = 32 item dan yang tidak memadai = 8 item dari nomor 24 - 40.

Tabel 4. Keputusan “Valid” dan “Tidak Valid” Butir Soal Pilihan Ganda

No	Soal	ID	B	TK
Valid				
1	PG-17	0.5	SB	33
2	PG-13	0.5	SB	23
3	PG-30	0.49	SB	24
4	PG-27	0.44	SB	18
5	PG-18	0.43	SB	43
6	PG-25	0.42	SB	21
7	PG-03	0.41	SB	46
8	PG-23	0.4	SB	39
9	PG-19	0.4	SB	20
10	PG-40	0.38	CB	18
11	PG-08	0.37	CB	62
12	PG-34	0.35	CB	34
13	PG-12	0.35	CB	19
14	PG-02	0.33	CB	31
15	PG-26	0.33	CB	26
16	PG-09	0.32	CB	25
17	PG-28	0.3	CB	28
18	PG-05	0.29	M	33
19	PG-06	0.29	M	55
20	PG-38	0.28	M	25
21	PG-20	0.27	M	49
22	PG-01	0.21	M	68
23	PG-32	0.21	M	22
Tidak Valid				
24	PG-10	0.52	SB	16
25	PG-39	0.49	SB	16
26	PG-31	0.48	SB	16
27	PG-07	0.47	SB	16
28	PG-35	0.42	SB	15
29	PG-24	0.35	CB	16
30	PG-21	0.33	CB	15
31	PG-29	0.3	CB	17
32	PG-36	0.25	M	11
33	PG-04	0.18	TB	38
34	PG-37	0.16	TB	16
35	PG-11	0.14	TB	19
36	PG-16	0.14	TB	20
37	PG-33	0.08	TB	23
38	PG-22	0.03	TB	24
39	PG-15	0	TB	11
30	PG-14	-0.05	N	19

4. Berdasarkan rekomendasi indeks diskriminasi (ID) butir soal pada Tabel 4 dengan kategori:
- Sangat Bagus (SB) = 14 item (35%),

- Cukup Bagus (CB) = 11 item (27,5%),
- Marjinal (M) = 7 item (17,5%),
- Buruk (TB) = 7 item (17,5), dan
- Negatif (N) = 1 item (2,5%).

Indeks diskriminasi butir soal diterima 32 item (80%) dan yang dibuang 8 item (20%)

Tabel 5. Butir Soal “Valid” Berdasarkan Matriks Proses Kognitif dan Materi/Isi Pelajaran Kimia

No	Soal	Indikator
1	PG-06	Memahami Fakta Larutan Elektrolit
2	PG-05	Memahami Konsep Ikatan Kimia
3	PG-23	Menerapkan Konsep Asam Basa
4	PG-26	Menerapkan Konsep Asam Basa
5	PG-27	Menerapkan Konsep Asam Basa
6	PG-28	Menerapkan Konsep Asam Basa
7	PG-32	Menerapkan Konsep Larutan Penyangga
8	PG-30	Menerapkan Konsep pH Larutan
9	PG-34	Menerapkan Konsep Sifat Koligatif Larutan
10	PG-13	Menerapkan Konsep Stoikiometri
11	PG-03	Menerapkan Konsep Struktur Atom
12	PG-08	Menerapkan Prinsip Bilangan Oksidasi
13	PG-02	Menerapkan Prinsip Struktur Atom
14	PG-17	Menerapkan Prinsip Termokimia
15	PG-18	Menganalisis Prinsip Laju Reaksi
16	PG-19	Menganalisis Prinsip Laju Reaksi
17	PG-20	Menganalisis Prinsip Laju Reaksi
18	PG-25	Menganalisis Prinsip Larutan Asam Basa
19	PG-38	Menganalisis Prinsip Organik
20	PG-12	Menganalisis Prinsip Stoikiometri
21	PG-09	Mengevaluasi Prinsip Stoikiometri
22	PG-40	Mengingat Fakta Kimia Unsur
23	PG-01	Mengingat Konsep Struktur Atom

5. Berdasarkan rekomendasi taraf kesulitan (TK) pada Tabel 4 maka diperoleh butir soal dengan kategori, yakni:
- Dibuang (X) = 0 item (0%),
 - Sangat Mudah (SM) = 1 item (2.5%),

- Mudah (M) = 0 item (0%),
- Dipertahankan (B) = 13 item (32.5%),
- Sangat sulit (SS) = 15 item (37.5%), dan
- Ditinjau Kembali (T) = 11 item (27.5%).

Taraf kesulitan butir soal yang diterima 29 item (%) dan yang dibuang 11 item (%).

- Keputusan “valid” dan “tidak valid” butir soal pada Tabel 4 nilai indeks diskriminasi (ID) pada Tabel 2 dikaitkan dengan nilai taraf kesulitan (TK) pada Tabel 3. Dengan demikian butir soal valid 23 item (57.5%) dan tidak valid 17 item (42.5%).

Tabel 6. Butir Soal “Tidak Valid” berdasarkan Matriks Proses Kognitif dan Materi/Isi Pelajaran Kimia

No	Soal	Indikator
1	PG-07	Memahami Prinsip Senyawa dan Rumus Kimianya
2	PG-14	Menerapkan Konsep Hidrokarbon
3	PG-35	Menerapkan Konsep Larutan
4	PG-22	Menerapkan Prinsip Kesetimbangan Kimia
5	PG-11	Menerapkan Prinsip Stoikiometri
6	PG-16	Menerapkan Prinsip Termokimia
7	PG-37	Menganalisis Konsep Elektrokimia
8	PG-04	Menganalisis Prinsip Energi Ionisasi
9	PG-33	Menganalisis Prinsip Kelarutan
10	PG-31	Menganalisis Prinsip Larutan Penyangga
11	PG-29	Menganalisis Prinsip pH Larutan
12	PG-36	Mengevaluasi Prinsip Elektrokimia
13	PG-15	Mengevaluasi Prinsip Hidrokarbon
14	PG-21	Mengevaluasi Prinsip Kesetimbangan Kimia
15	PG-24	Mengevaluasi Prinsip Larutan Asam Basa
16	PG-10	Mengevaluasi Prinsip Redoks
17	PG-39	Mengevaluasi Prinsip Organik

- Merujuk pada Tabel 4 maka butir soal valid dan tidak valid berdasarkan proses kognisi

(Tabel 5) & Isi/Materi Pelajaran (Tabel 6).

Tabel 7. Butir Soal Valid dan Tidak Valid Berdasarkan Proses Kognisi

No	Proses Kognisi	Valid	Tidak Valid
1	Mengingat (C1)	2	-
2	Memahami (C2)	2	1
3	Menerapkan (C3)	12	5
4	Menganalisis (C4)	6	5
5	Mengevaluasi (C5)	1	6
6	Mencipta (C6)	-	-
		23	17

Tabel 8. Butir Soal Valid dan Tidak Valid Berdasarkan Isi/Materi Pelajaran

No	Isi/Materi Pelajaran	Valid	Tidak Valid
1	Struktur Atom	3	-
2	Bilangan oksidasi	1	-
3	Senyawa dan Rumus Kimianya	-	1
4	Stoikiometri	3	1
5	Ikatan Kimia	1	-
6	Laju Reaksi	3	-
7	Termokimia	1	1
8	Kelarutan	-	2
9	Kesetimbangan Kimia	-	2
10	Larutan Elektrolit	1	-
11	Asam Basa	5	1
12	pH Larutan	1	1
13	Larutan Penyangga	1	1
14	Sifat Koligatif Larutan	1	-
15	Energi Ionisasi	-	1
16	Redoks	-	1
17	Elektrokimia	-	2
18	Hidrokarbon	-	2
19	Kimia Organik	1	1
20	Kimia Unsur	1	-
		23	17

Pembahasan

Pembahasan yang dapat disampaikan adalah sebagai berikut:

- Indeks Deskriminasi (ID) yang tergolong memadai pada kategori Sangat Bagus, SB = 14 item (35%), Cukup Bagus, CB = 11 item

- (17,5%), dan Marjinal, $M = 17$ item (17,5%). Hal ini dapat dikatakan bahwa berdasarkan ID dengan $SB > CB > M$ dapat dikatakan butir soal dengan urutan seperti ini baik adanya. Butir soal yang tidak memadai pada kategori Buruk, $TB = 7$ item (17,5%) dan Negatif, $N = 1$ item (2,5%). Negatif berarti siswa yang berprestasi rendah mendapat jawaban benar lebih banyak dibandingkan siswa yang berprestasi tinggi.
2. Taraf Kesulitan (TK) tergolong memadai sangat muda, mudah, dipertahankan, dan sangat sulit sedangkan yang tidak memadai adalah dengan nilai 0,00-0,19 berarti isi/materi pelajarannya perlu ditinjau atau diajarkan kembali.
 3. Kualitas butir soal berdasarkan indeks deskriminasi (ID) yang memadai dikaitkan dengan taraf kesulitan (TK) maka butir soal yang valid 23 item (57,5%), jumlah butir soal terbuang 17 item (42,5%). Hal ini yang terbuang tergolong sangat besar. Artinya kualitas menggunakan ID yang dikoreksi dengan TK menunjukkan jumlah butir soal berkurang hampir 50%. Hal ini disebabkan peserta tes memiliki kemampuan yang sangat bervariasi dari Kelas XI dan Kelas XII sedangkan isi/materi pelajaran kimia dari Kelas X sampai kelas XII.
 4. Berdasarkan Taksonomi Bloom Revisi Anderson (Tabel 7) bahwa hasil butir soal proses kognitif pada level C1, C2, C3, C4, C5, dan C6 adalah 2, 2, 12, 6, 1, 0. Hasil ini menunjukkan relatif sesuai dengan teori penyusunan soal yang mana pada tingkat menengah jumlah butir soal lebih banyak dari mudah dan sulit. Soal kategori mencipta (C6) tidak valid dan kategori HOTS hanya pada C4 dan C5.
 5. Berdasarkan isi/materi pelajaran (Tabel 8) terlihat 20 isi/materi pelajaran kimia tetapi butir soal yang valid 14 isi/materi pelajaran. Butir soal yang valid 5 item pada Asam Basa, kemudian 3 item masing-masing untuk Struktur Atom, Stoikiometri, dan Laju Reaksi. Hal ini menunjukkan bahwa peserta lebih dominan mempelajari isi/materi pelajaran kimia tersebut.
 6. Isi/materi pelajaran Senyawa dan Rumus Kimianya, Kelarutan, Kestimbangan Kimia, Energi Ionisasi, Redoks, Elektrokimia, dan Hidrokarbon kurang dipahami oleh peserta.
 7. Contoh pembahasan Soal PG No 39 adalah sebagai berikut:
Produk reaksi antara $\text{HNO}_3 + \text{H}_2\text{SO}_4$ dengan asam benzoat menghasilkan
A. Asam o-amino benzoat
B. Asam m-amino benzoat
C. Asam p-amino benzoat
D. Asam o-nitro benzoat
E. Asam m-nitro benzoat
Pilihan yang benar adalah E.
Sesuai dengan perhitungan korelasi *point biserial* yang menunjukkan indeks deskriminasi (ID) = 0,48 dan Taraf Kesulitan (TK) = 0,18. Apabila hanya melihat ID, koefisiennya signifikan dengan kategori Sangat Bagus. Setelah disandingkan dengan TK dengan kategori melebihi Sangat Sulit atau perlu diajarkan kembali. Dengan demikian dinyatakan tidak valid.
Berkaitan dengan Matriks Indikator Proses Kognitif dan Materi/Isi Suatu Mata Pelajaran adalah Mengevaluasi (C5) Prinsip dalam pelajaran organik dan termasuk kategori HOTS.
Sesuai dengan stemnya maka peluang benar ada pada pilihan D dan E tetapi perlu memahami prinsip penentuan orto-meta-para substituen pada senyawa turunan benzena.

Kesimpulan

Kesimpulan penentuan kualitas butir soal pilihan ganda "Pacu Kimia" FMIPAK Unima adalah sebagai berikut:

1. Responden 80 siswa SMA untuk menjawab 40 butir soal.
2. Penentuan kualitas butir soal menggunakan teori tes klasik (*classical test theory*) yakni a) Indeks Deskriminasi (ID) menggunakan korelasi *point biserial* untuk menghitung jumlah skor total pada butir soal tertentu yang dikorelasikan dengan jumlah skor total keseluruhan butir soal dan b) Taraf Kesulitan (TK) dengan menggunakan proporsi benar pada setiap butir soal.
3. ID memiliki 6 kategori yakni Sangat Bagus,

Cukup Bagus, Marjinal, Diskriminasi buruk, Negatif sedangkan TK terdiri dari 6 kategori yakni Dibuang, Sangat Mudah, Mudah, Dipertahankan, Sangat Sulit dan Ditinjau kekurangannya atau diajarkan.

4. Berdasarkan ID yang dikaitkan dengan TK maka kualitas butir soal yang valid sebanyak 23 item atau 57,5 %.
5. Berdasarkan uji validitas tersebut ternyata untuk proses kognitif dari Taksonomi Bloom revisi Anderson butir soal yang valid dari C1, C2, C3, C4, C5 dan C6 adalah 2, 2, 12, 6, 0. Butir soal untuk kategori HOTS adalah C4, C5 sebanyak 18 item.
6. Dari 20 isi/materi pelajaran kimia yang valid sebanyak 14 yakni Asam basa, Struktur Atom, Stoikiometri, dan Laju Reaksi, Kemudian yang kurang dipahami oleh peserta meliputi Senyawa dan Rumus Kimianya, Kelarutan, Kesetimbangan Kimia, Energi Ionisasi, Redoks, Elektrokimia, dan Hidrokarbon.

Daftar Pustaka

1. Sorenson B, Hanson K. Using classical test theory and rasch modeling to improve general chemistry exams on a per instructor basis. *Journal of chemical education*. 2021 Apr 19;98(5):1529-38.
2. Kornbrot D. Point biserial correlation. *Wiley StatsRef: Statistics Reference Online*. 2014 Apr 14.
3. Amelia, M. A. (2016). Analisis Soal Tes Hasil Belajar High Order Thinking Skills (HOTS) Matematika Materi Pecahan untuk Kelas 5 Sekolah Dasar. *Jurnal Penelitian*, 20(2).
4. Ihsan H. Validitas Isi Alat Ukur Penelitian: Konsep Dan Panduan Penilaiannya. *Pedagogia*. 2015;13(3):173-9.
5. Goldsby, K., & Chang, R. (2015). *Chemistry*. McGraw-Hill Higher Education.
6. Rahayu S. Mengoptimalkan aspek literasi dalam pembelajaran kimia abad 21. *InProsiding Seminar Nasional Kimia UNY 2017 Oct 14 (Vol. 21, No. 4, pp. 183-188)*.
7. Wicaksono AG. Johnstone's levels of representation in science learning. *SPEKTRA: Jurnal Kajian Pembelajaran Sains*. 2022 May 1;8(1):19-26.
8. Talanquer, V. (2011). Macro, submicro, and symbolic: The many faces of the chemistry "triplet". *International Journal of Science Education*, 33(2), 179-195.
9. Kapici, H. Ö. (2023). From Symbolic Representation to Submicroscopic One: Preservice Science Teachers' Struggle with Chemical Representation Levels in Chemistry. *International Journal of Research in Education and Science*, 9(1), 134-147.
10. Ritonga, H. M., & Yusmaita, E. (2019) Analisis Kurikulum 2013 pada Kompetensi Inti 3 Kelas XII SMA Berdasarkan Dimensi Proses Kognitif dan Pengetahuan. *Edukimia*, 1 (3), 61-66.
11. Iskandar, D., & Senam, S. (2015). Studi kemampuan guru kimia sma lulusan UNY dalam mengembangkan soal UAS berbasis HOTS. *Jurnal Inovasi Pendidikan IPA*, 1(1), 65-72.
12. Khoirunnisah, K., & Yusmaita, E. (2019). Analysis of Core Competency 3 of Curriculum 2013 for 11th Grade of Senior High School Based on Taxonomy Bloom (Revised Version). *Edukimia*, 1(3), 67-72.
13. Juhandi, A. (2016). Analisis soal jenjang kognitif taksonomi bloom revisi pada buku sekolah elektronik (BSE) biologi SMA. *Jurnal Pengajaran MIPA*, 21(1), 61-66.
14. Purnomo, P., & Palupi, M. S. (2016). Pengembangan tes hasil belajar matematika materi menyelesaikan masalah yang berkaitan dengan waktu, jarak dan kecepatan untuk siswa kelas V. *Jurnal Penelitian*, 20(2).
15. Rohmah, F. N. (2017). Urgensi Evaluasi untuk Pendidikan. *Fikrotuna: Jurnal Pendidikan dan Manajemen Islam*, 5(1).
16. Wati, M., & Mahtari, S. (2017). Pengembangan Instrumen Kognitif Fisika Siswa SMP. *Jurnal Penelitian dan Pengkajian Ilmu Pendidikan: e-Saintika*, 1(1), 45-56.
17. Taufiq Satria, M., & Edi, I. (2018). Instrumen penilaian kemampuan berpikir kritis peserta didik SMA negeri mata pelajaran Biologi Kelas X. *BIOEDUKASI: Jurnal Pendidikan Biologi*, 11(2), 105-110.
18. Mulatsih, B. (2021). PENGEMBANGAN TES DIAGNOSTIK KESULITAN BELAJAR KIMIA MODEL TESTLET. *Volume XII/Nomor 01/April 2021*, 12(01), 1.

19. Breakall, J., Randles, C., & Tasker, R. (2019). Development and use of a multiple-choice item writing flaws evaluation instrument in the context of general chemistry. *Chemistry Education Research and Practice*, 20(2), 369-382.
20. Zara, C., Smith, C., Coombs, S. J., Herbert, G., & Isaacs, T. (2013). Key concepts in educational assessment. *Key Concepts in Educational Assessment*, 1-176.
21. Objektif Tes Pilihan Ganda (Multiple Choicetest) pada Mata Pelajaran Matematika Pendidikan Dasar pada Masa Pandemi Covid 19 (Doctoral dissertation, IAIN KUDUS).
22. Leblanc LeBlanc V, Cox M. Interpretation of the point-biserial correlation coefficient in the context of a school examination. *The Quantitative Methods for Psychology*. 2017 Jan 1;13(1):46-56.
22. Dichoso AA, Joy MR, Cabauatan M. Test item analyzer using point-biserial correlation and p-values. *International Journal Of Scientific & Technology Research*. 2020;9(4):2122-6.



© 2024 by the authors. Licensee Oxygenius Journal Of Chem. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).