



## Analisis Validitas Dan Reliabilitas Dalam Mendukung Ketuntasan Belajar Model STEM

Fatayah Fatayah\*, Ika Farida Yuliana, Layla Mufidah

Program Studi Pendidikan Kimia, Universitas Billfath Lamongan, Indonesia

\*Email: [fatayah.billfath@gmail.com](mailto:fatayah.billfath@gmail.com)

---

### Informasi Artikel    Abstrak

---

#### Kata kunci:

Validitas,  
Reliabilitas,  
Daya Beda,  
Taraf Kesukaran,  
STEM.

Diterima:

09-02-2022

Disetujui:

16-02-2022

Dipublikasikan:

26-02-2022

Penelitian dengan instrumen yang tidak valid dan tidak reliabel akan sangat mempengaruhi hasil penelitian. Minimnya pemaparan secara lengkap hasil analisis data serta belum ditemukan analisis validitas dan reliabilitas pada ulangan harian maka peneliti melakukan penelitian yang bertujuan untuk mendeskripsikan hasil validitas dan reliabilitas butir soal ulangan harian materi koloid dalam mendukung ketuntasan belajar model STEM. Jenis penelitian deskriptif kuantitatif. Peneliti ingin mendeskripsikan kevalidan butir soal kimia pada instrumen tes ulangan harian materi koloid dari segi: validitas, reliabilitas, tingkat kesukaran, dan daya pembeda. Hasil uji validitas isi diperoleh bahwa instrumen valid. Instrumen soal sebanyak 18 butir pilihan ganda dan 6 butir uraian diperoleh hasil semua valid dan reliabel. Soal pilihan ganda hanya nomor soal 6 dan 14 yang memiliki daya beda baik sedangkan soal 1,2, 3, 10, 12, 13, 15,16, 17, dan 18 memiliki daya beda yang cukup baik. Soal uraian dari nomor 1 sampai 6 memiliki daya beda yang baik. Berdasarkan kriteria taraf kesukaran diperoleh hasil soal pilihan ganda 6, 7, 8, 10, 13, dan 17 adalah soal yang tingkat kesukarannya sedang sedangkan soal nomor 1, 2, 3, 4, 5, 12, 14, 15, 16, dan 18. Instrumen yang dapat digunakan minimal kriterianya valid dan reliabel.

---

#### Abstract

---

Research with invalid and unreliable instruments will greatly affect the results of the study. The lack of complete exposure of the results of data analysis and the validity and reliability analysis of the daily tests have not been found, the researchers conducted a study that aims to describe the results of the validity and reliability of colloidal material daily test items in supporting the completeness of learning the STEM model. This type of quantitative descriptive research. The researcher wants to describe the validity of the chemistry questions on the colloidal material daily test instrument in terms of: validity, reliability, level of difficulty, and discriminating power. The results of the content validity test were obtained that the instrument was valid. The instrument questions were 18 multiple choice items and 6 description items, all valid and reliable results were obtained. Multiple-choice questions only have questions 6 and 14 which have good distinguishing power, while questions 1, 2, 3, 10, 12, 13, 15, 16, 17, and 18 have good distinguishing power. Explanation questions from numbers 1 to 6 have good discriminating power. Based on the criteria for the level of difficulty, the results of multiple choice questions 6, 7, 8, 10, 13, and 17 were questions with moderate level of difficulty, while questions numbered 1, 2, 3, 4, 5, 12, 14, 15, 16, and 18. Instruments that can be used at least have valid and reliable criteria.

---

## **PENDAHULUAN**

Pendidikan mempunyai peranan yang penting dalam pengembangan sumber daya manusia dan watak bangsa untuk kemajuan masyarakat dan bangsa. Harkat dan martabat bangsa ditentukan dari kualitas pendidikannya. Kualitas pendidikan dapat dilihat dari dua hal, yaitu proses pendidikan dan hasil pendidikan. Proses pendidikan yang berkualitas yaitu apabila seluruh komponen pendidikan terlibat dalam proses pendidikan itu sendiri (Irvan, 2017). Sedangkan hasil pendidikan yang berkualitas yaitu mengacu pada prestasi yang dicapai. Prestasi yang dicapai atau hasil pendidikan antara lain dapat berupa hasil tes kemampuan akademis (misalnya ulangan harian) (Irvan & Jauhari, 2018). Prestasi hasil pendidikan ini dapat dilihat berdasarkan keberhasilan belajar peserta didik (Irawan & Iasha, 2021).

Ketercapaian tujuan pembelajaran kognitif dapat diukur dengan menggunakan instrumen tes. Tes merupakan alat evaluasi atau penilaian yang berisi pertanyaan-pertanyaan yang diberikan kepada siswa untuk mendapatkan jawaban dari siswa dalam bentuk tulisan maupun lisan. Fungsi dari pemberian tes tersebut salah satunya dapat memberikan informasi terkait tingkat ketuntasan belajar siswa. Tes yang baik harus memenuhi syarat validitas, reliabilitas, objektivitas, praktikalitas, dan ekonomis (Arikunto, 2021). Tes sendiri dikatakan valid jika dapat mengukur apa yang hendak diukur. Sedangkan tes dikatakan reliabel jika memiliki keajegan. Suatu instrument evaluasi, dikatakan mempunyai nilai reliabilitas tinggi, apabila tes yang dibuat mempunyai hasil yang konsisten dalam mengukur yang hendak diukur (Arikunto, 2021). Selain validitas dan reliabilitas, yang tidak kalah penting juga melihat tingkat kesukaran dan daya beda. Tingkat kesukaran adalah peluang untuk menjawab benar suatu soal pada tingkat kemampuan tertentu yang biasanya dinyatakan dalam bentuk indeks. Indeks tingkat kesukaran ini pada umumnya dinyatakan dalam bentuk proporsi yang besarnya berkisar 0,00-1,00 (Sukidin & Basrowi, 2012). Sedangkan daya pembeda soal adalah kemampuan suatu butir soal dapat membedakan antara peserta didik yang telah menguasai materi yang ditanyakan dan peserta didik yang tidak/kurang/belum menguasai materi yang ditanyakan. Indeks daya pembeda berkisar antara -1,00 sampai +1,00 (Sukidin & Basrowi, 2012).

Mengukur tes (Butir soal) merupakan hal yang sangat penting dalam penelitian yang bertujuan untuk mengukur ketuntasan belajar. Penelitian dengan instrumen yang tidak valid dan tidak reliabel akan sangat mempengaruhi hasil penelitian sehingga sebelum memakai instrumen tersebut dalam suatu penelitian diperlukan uji validitas dan reliabilitas dengan tujuan agar dapat diketahui valid dan tidaknya suatu instrumen butir soal dan juga dapat diketahui reliabilitas dari butir soal serta agar hasil penelitian tidak terkontaminasi dengan faktor lain. Beberapa hasil penelitian terdahulu banyak yang mengungkap pentingnya kevalidan suatu instrumen seperti pada penelitian (Musayaroh, 2021) yang mengatakan bahwa instrumen tes literasi kimia berbasis HOTS yang valid mampu memberikan gambaran tingkatan literasi kimia berdasarkan levelnya dan memberikan gambaran kemampuan literasi kimia dalam menjawab soal khususnya soal HOTS. Hal senada juga diungkapkan (Utomo, 2019) yang berpendapat bahwa soal yang dijadikan alat evaluasi harus memiliki ciri-ciri soal yang baik salah satunya yaitu terkait kevalidan soal. Penelitian (Nuswowati et al., 2010) juga menceritakan betapa pentingnya validitas dan reliabilitas butir soal sebagai alat evaluasi. Hal senada juga diungkapkan (Riyani et al., 2017). Penelitian (Zein et al., 2013) dan (Sari et al., 2019) juga memaparkan pentingnya instrumen tes yang valid. Beberapa hasil dari uji literatur ditemukan bahwa pemaparan hasil validitas dan reliabilitas kurang lengkap (hanya dipaparkan hasil secara deskriptif sedangkan hasil perhitungan mentah tidak dicantumkan sebagai lampiran) seperti penelitian (Solichin, 2017) dan juga rata-rata analisis validitas hanya pada soal ujian akhir semester (UAS) seperti penelitian (Magdalena et al., 2021) dan juga penelitian (Ida, 2021). Peneliti belum menemukan penelitian yang menguji validitas dan reliabilitas instrumen tes untuk ulangan harian.

Berdasarkan latar belakang di atas yang mengungkap pentingnya validitas dan reliabilitas suatu tes maka meneliti dan masih minimnya pemaparan secara lengkap hasil analisis data serta belum ditemukan analisis validitas dan reliabilitas pada ulangan harian maka peneliti melakukan penelitian yang bertujuan untuk

mendiskripsikan hasil validitas dan reliabilitas butir soal ulangan harian materi koloid dalam mendukung ketuntasan belajar model STEM.

## METODE

Jenis penelitian yang digunakan dalam penelitian ini adalah penelitian diskriptif kuantitatif. Dalam penelitian ini peneliti ingin mendiskripsikan kevalidan butir soal kimia pada instrumen tes ulangan harian materi koloid dari segi: validitas, reliabilitas, tingkat kesukaran, dan daya pembeda. Subjek dari penelitian ini adalah tes ulangan harian materi koloid pada siswa MA yang berjumlah sebanyak 69 orang. Objek penelitian ini adalah kevalidan butir soal kimia pada tes ulangan harian materi koloid dilihat dari segi: validitas, reliabilitas, tingkat kesukaran dan daya pembeda. Data yang diperoleh dari penelitian ini yaitu data hasil tes materi koloid pada instrumen tes berupa lembar jawaban soal.

Pengumpulan data dilakukan dengan dua cara yaitu memberikan instrumen tes kepada validator ahli pada bidang kimia dan berpengalaman dalam mengajar untuk mengetahui validitas isi dari instrumen tersebut. Validator tes soal dalam penelitian ini terdiri dari 2 (dua) orang dosen dan 1 (satu) orang guru. Aspek yang menjadi perhatian adalah keterwakilan konsep soal yang diukur dan kesesuaiannya dengan indikator yang telah ditetapkan. Setiap validator diminta memberikan skor 2 untuk setiap butir soal yang susunan kalimatnya komunikatif dan mengandung konsep yang akan diukur. Skor 1 untuk setiap butir soal yang susunan kalimatnya komunikatif tetapi tidak mengandung konsep yang akan diukur. Skor 0 untuk setiap butir soal yang susunan kalimatnya tidak komunikatif dan tidak mengandung konsep yang akan diukur. Cara yang kedua adalah dengan memberi peserta didik tes materi koloid kemudian hasil tes tersebut dikumpulkan kemudian dianalisis dari segi validitas isi, validitas butir soal, reliabilitas, daya beda dan taraf kesukaran soal. Analisis data dilakukan dalam beberapa tahap yaitu:

### Validitas instrumen tes oleh ahli (validitas isi)

Validitas isi ini diberikan kepada tiga validator. Untuk mencari konsistensi antara penilai dilakukan dengan cara menghitung rata-rata persentase pemberian skor oleh tim penilai menggunakan persamaan (1).

$$P = \frac{\text{jumlah skor penilai}}{\text{jumlah skor maksimum}} \times 100\% \quad (1)$$

Persentase validitas soal yang didapatkan kemudian dikonversikan melalui Tabel 1.

**Tabel 1.** Kriteria Hasil Validasi Soal Tes oleh Ahli

Nilai P	Kriteria
81% - 100%	Sangat tinggi
61% - 80%	Tinggi
41% - 60%	Cukup
21% - 40%	Rendah
0% - 20%	Sangat rendah

(Kurniawati, 2011)

Berdasarkan kriteria pada Tabel 1 tersebut, instrumen dapat dikatakan valid jika hasil validasi masuk dalam kriteria “tinggi” sampai “sangat tinggi”.

### Validitas butir soal

Sebuah butir soal dapat dikatakan valid apabila mempunyai dukungan yang besar terhadap skor. Dengan kata lain butir soal memiliki validitas yang tinggi jika skor pada butir soal mempunyai kesejajaran dengan skor total. Validitas butir-butir soal ini dianalisis dengan persamaan korelasi product moment pearson pada persamaan (2).

$$r_{xy} = \frac{N \sum XY - (\sum X)(\sum Y)}{\sqrt{N\{\sum X^2 - (\sum X)^2\}\{N \sum Y^2 - (\sum Y)^2\}}} \tag{2}$$

Perhitungan dengan menggunakan bantuan program SPSS 16 for windows. Pengujian menggunakan uji dua sisi dengan taraf signifikansi 0,05 dan kriteria pengujian adalah sebagai berikut. Jika nilai signifikansi (sig.) < 0,05 dan bernilai positif maka instrumen atau butir-butir pertanyaan berkorelasi signifikan terhadap skor total (dinyatakan valid). Jika nilai signifikansi (sig.) > 0,05 maka instrumen atau butir-butir pertanyaan tidak berkorelasi signifikan terhadap skor total (dinyatakan tidak valid).

**Reliabilitas tes**

Reliabilitas berhubungan dengan keandalan dan kejelasan hasil. Butir soal akan memiliki tingkat keandalan yang tinggi, apabila tes memberikan hasil yang korelatif tetap (ajeg). Reliabilitas butir soal dihitung dengan menggunakan program SPSS 16.0 dengan cara analisis Alpha Cronbach dan menggunakan persamaan (3) dan (4).

$$r_{11} = \left(\frac{n}{n-1}\right) \left(\frac{S^2 - \sum pq}{S^2}\right) \tag{3}$$

$$S^2 = \frac{\sum x^2 - \frac{(\sum x)^2}{N}}{N} \tag{4}$$

Instrumen dikatakan reliabel apabila Cronbach's Alpha > 0,600.

**Daya beda**

Daya beda butir soal merupakan kemampuan suatu soal untuk membedakan antara siswa yang berkemampuan tinggi dengan siswa yang berkemampuan rendah. Persamaan daya beda ditunjukkan pada persamaan (5).

$$D = \frac{BA}{JA} - \frac{BB}{JB} = PA - PB \tag{5}$$

Keterangan:

- D = daya beda
- JA = jumlah testee kelompok atas
- JB = jumlah testee kelompok bawah
- BA = banyaknya testee kelompok atas yang menjawab benar
- BB = banyaknya testee kelompok bawah yang menjawab benar
- PA = proporsi kelompok atas yang menjawab benar
- PB = proporsi kelompok bawah yang menjawab benar

Kriteria penentuan daya beda butir soal disajikan pada tabel 2.

Harga D	Kriteria Daya Beda Butir Soal
0,00 ≤ D ≤ 0,20	Jelek
0,21 ≤ D ≤ 0,40	Cukup
0,41 ≤ D ≤ 0,70	Baik
0,71 ≤ D ≤ 1,00	Baik sekali

(Kurniawati, 2011)

Berdasarkan kriteria pada Tabel 2 tersebut, instrumen dapat dikatakan valid jika tingkat daya beda butir soal masuk dalam kriteria “cukup” sampai “baik sekali”

### Tingkat Kesukaran Butir Soal

Soal yang baik adalah soal yang ditingkat kesukarannya tidak terlalu sulit dan tidak terlalu mudah. Tingkat kesukaran (P) butir soal ditentukan berdasarkan perbedaan antara siswa yang menjawab benar dengan jumlah peserta tes. Rumus yang digunakan adalah:

$$P = \frac{B}{J_s} \quad (6)$$

Keterangan:

P = tingkat kesukaran

B = jumlah peserta didik yang menjawab benar

J<sub>s</sub> = jumlah peserta didik yang mengikuti tes

Kriteria penentuan tingkat kesukaran butir soal dapat dilihat pada tabel berikut.

**Tabel 3.** Kriteria Penentuan Tingkat Kesukaran Butir Soal

Harga P	Kriteria Tingkat Kesukaran
0,00 – 0,30	Soal sukar
0,31 – 0,70	Soal sedang
0,71 – 1,00	Soal mudah

(Kurniawati, 2011)

Berdasarkan kriteria pada Tabel 3 tersebut, instrumen dapat dikatakan valid jika tingkat kesukaran butir soal masuk dalam kriteria “sedang”

## HASIL DAN PEMBAHASAN

Hasil penelitian berikut ini menjelaskan tentang gambaran umum dari data yang telah diperoleh berdasarkan penelitian. Data-data yang telah diperoleh adalah data validitas isi, validitas butir soal, reliabilitas, taraf kesukaran, dan daya beda. Hasil dari penelitian ini dapat dipaparkan sebagai berikut:

### Validitas instrumen tes oleh ahli (validitas isi)

Validitas isi dinilai oleh 3 orang pakar bidang keahlian pendidikan kimia dan diperoleh hasil seperti yang ditunjukkan pada tabel 4.

**Tabel 4.** Hasil Validasi Ahli Instrumen Tes

No Soal	Skor			Skor Total	Persentase (%)	Kriteria Kevalidan
	Validator 1	Validator 2	Validato 3			
1	3	3	3	9	100	Sangat tinggi
2	1	3	3	7	77,7	Tinggi
3	3	3	3	9	100	Sangat tinggi
4	1	1	3	5	55,5	Sedang
5	1	1	3	5	55,5	Sedang
6	3	3	3	9	100	Sangat tinggi
7	1	3	2	6	66,6	Tinggi
8	1	3	3	7	77,7	Tinggi
9	3	3	3	9	100	Sangat tinggi
10	3	3	3	9	100	Sangat tinggi
11	3	3	2	8	88,8	Sangat tinggi
12	3	3	3	9	100	Sangat tinggi
13	3	3	3	9	100	Sangat tinggi

14	3	3	3	9	100	Sangat tinggi
15	3	3	3	9	100	Sangat tinggi
16	3	3	3	9	100	Sangat tinggi
17	3	3	2	8	88,8	Sangat tinggi
18	3	3	2	8	88,8	Sangat tinggi
19	2	3	3	8	88,8	Sangat tinggi
20	3	3	3	9	100	Sangat tinggi
21	3	3	3	9	100	Sangat tinggi
22	3	3	2	8	88,8	Sangat tinggi
23	3	3	3	9	100	Sangat tinggi
24	2	3	3	8	88,8	Sangat tinggi
25	2	3	2	7	77,7	Tinggi

Dari Tabel 4 di atas, hasil validasi ahli dapat diringkas seperti yang ditunjukkan pada tabel 5.

**Tabel 5.** Persentasi Jumlah Skor Validator

No	Validator	Persentase Jumlah Skor (%)
1	Validator 1	82,66
2	Validator 2	94,66
3	Validator 3	92,00
<b>Rata-rata Persentasi</b>		<b>89,77</b>
<b>Kriteria</b>		<b>Sangat Tinggi</b>

Hasil skor validator pada Tabel 5 dapat disimpulkan bahwa instrumen valid dan dapat digunakan untuk penelitian, karena berada pada kriteria “sangat tinggi”.

**Validitas Butir Soal Tes**

Untuk mengetahui kevalidan butir soal dilakukan dengan cara memberikan instrumen soal kepada siswa yang pernah dibelajarkan materi koloid. Soal diujikan di MA Darur Istiqomah dengan jumlah siswa sebanyak 69 siswa. Bentuk soal ada dua macam yaitu soal pilihan ganda sebanyak 18 butir sedangkan uraian sebanyak 6 butir. Uji validitas dilakukan dengan bantuan program SPSS 21. Adapun hasil validitas butir soal uraian dapat dilihat pada Tabel 6.

**Tabel 6.** Hasil Validitas Soal Uraian

		x1	x2	x3	x4	x5	x6	Total
<b>x1</b>	Pearson Correlation	1	,251*	,330**	,369**	,282*	,321**	,525**
	Sig. (2-tailed)		0.037	0.006	0.002	0.019	0.007	0.000
	N	69	69	69	69	69	69	69
<b>x2</b>	Pearson Correlation	,251*	1	,601**	0.190	,402**	,498**	,581**
	Sig. (2-tailed)	0.037		0.000	0.118	0.001	0.000	0.000
	N	69	69	69	69	69	69	69
<b>x3</b>	Pearson Correlation	,330**	,601**	1	,564**	,693**	,760**	,759**
	Sig. (2-tailed)	0.006	0.000		0.000	0.000	0.000	0.000
	N	69	69	69	69	69	69	69
<b>x4</b>	Pearson Correlation	,369**	0.190	,564**	1	,601**	,661**	,621**
	Sig. (2-tailed)	0.002	0.118	0.000		0.000	0.000	0.000
	N	69	69	69	69	69	69	69
<b>x5</b>	Pearson Correlation	,282*	,402**	,693**	,601**	1	,650**	,672**
	Sig. (2-tailed)	0.019	0.001	0.000	0.000		0.000	0.000
	N	69	69	69	69	69	69	69
<b>x6</b>	Pearson Correlation	,321**	,498**	,760**	,661**	,650**	1	,776**
	Sig. (2-tailed)	0.007	0.000	0.000	0.000	0.000		0.000
	N	69	69	69	69	69	69	69

<b>Total</b>	Pearson Correlation	,525**	,581**	,759**	,621**	,672**	,776**	1
	Sig. (2-tailed)	0.000	0.000	0.000	0.000	0.000	0.000	
	N	69	69	69	69	69	69	69

**\*. Correlation is significant at the 0.05 level (2-tailed).**

**\*\*.** Correlation is significant at the 0.01 level (2-tailed).

Membaca hasil output SPSS ada dua cara yaitu (1) dengan melihat harga r Pearson Correlation pada kolom total dibandingkan dengan nilai significant. Jika r Pearson Correlation  $\geq$  nilai sig. maka valid tetapi jika harga r Pearson Correlation  $<$  nilai sig. maka tidak valid. Cara (2) yaitu dengan melihat harga r Pearson Correlation pada kolom total yang ada tandanya bintang satu atau dua berarti valid tetapi jika tidak ada tanda bintang berarti tidak valid.

Dari Tabel 6 dapat diketahui bahwa ke enam soal uraian diperoleh harga r Pearson Correlation  $>$  nilai sig. sehingga dapat disimpulkan bahwa keenam butir soal uraian adalah valid. Hasil validitas soal pilihan ganda dipaparkan pada Tabel 7.

**Tabel 7.** Validitas Soal pilihan ganda

		Total
<b>x1</b>	Pearson Correlation	,701**
	Sig. (2-tailed)	0.000
	N	69
<b>x2</b>	Pearson Correlation	,616**
	Sig. (2-tailed)	0.000
	N	69
<b>x3</b>	Pearson Correlation	,671**
	Sig. (2-tailed)	0.000
	N	69
<b>x4</b>	Pearson Correlation	,339**
	Sig. (2-tailed)	0.004
	N	69
<b>x5</b>	Pearson Correlation	,386**
	Sig. (2-tailed)	0.001
	N	69
<b>x6</b>	Pearson Correlation	,549**
	Sig. (2-tailed)	0.000
	N	69
<b>x7</b>	Pearson Correlation	,270*
	Sig. (2-tailed)	0.025
	N	69
<b>x8</b>	Pearson Correlation	,346**
	Sig. (2-tailed)	0.004
	N	69
<b>x9</b>	Pearson Correlation	,288*
	Sig. (2-tailed)	0.016
	N	69
<b>x10</b>	Pearson Correlation	,503**
	Sig. (2-tailed)	0.000
	N	69
<b>x11</b>	Pearson Correlation	,304*
	Sig. (2-tailed)	0.011

	N	69
<b>x12</b>	Pearson Correlation	,707**
	Sig. (2-tailed)	0.000
	N	69
<b>x13</b>	Pearson Correlation	,438**
	Sig. (2-tailed)	0.000
	N	69
<b>x14</b>	Pearson Correlation	,603**
	Sig. (2-tailed)	0.000
	N	69
<b>x15</b>	Pearson Correlation	,660**
	Sig. (2-tailed)	0.000
	N	69
<b>x16</b>	Pearson Correlation	,598**
	Sig. (2-tailed)	0.000
	N	69
<b>x17</b>	Pearson Correlation	,589**
	Sig. (2-tailed)	0.000
	N	69
<b>x18</b>	Pearson Correlation	,281*
	Sig. (2-tailed)	0.020
	N	69
<b>total</b>	Pearson Correlation	1
	Sig. (2-tailed)	
	N	69

Membaca hasil output SPSS ada dua cara yaitu (1) dengan melihat harga r Pearson Correlation pada kolom total dibandingkan dengan nilai significant. Jika r Pearson Correlation  $\geq$  nilai sig. maka valid tetapi jika harga r Pearson Correlation  $<$  nilai sig. maka tidak valid. Cara (2) yaitu dengan melihat harga r Pearson Correlation pada kolom total yang ada tandanya bintang satu atau dua berarti valid tetapi jika tidak ada tanda bintang berarti tidak valid.

Dari Tabel 7 dapat diketahui bahwa ke delapanbelas soal pilihan ganda diperoleh harga r Pearson Correlation  $>$  nilai sig. sehingga dapat disimpulkan bahwa ke delapanbelas butir soal pilihan ganda adalah valid. Output validitas soal pilihan ganda hanya disajikan pada kolom total karena kalau disajikan semua item soal maka akan terlalu panjang.

### Reliabilitas

Reliabilitas digunakan untuk mengetahui keajegan suatu instrumen tes. Untuk menghitung reliabilitas digunakan bantuan SPSS 21. Hasil dari reliabilitas butir soal uraian dan pilihan ganda dapat dilihat pada Tabel 8 dan 9.

Tabel 8. Hasil Uji Reliabilitas Soal Uraian

Reliability Statistics	
Cronbach's Alpha	N of Items
0,844	6

Cara mengetahui instrumen tes tersebut reliabel atau tidak dengan cara melihat harga Cronbach's Alpha. Jika harga Cronbach's Alpha  $\geq$  ambang batas Cronbach's Alpha (0,60) dan bernilai positif maka instrumen tes tersebut reliabel tetapi jika harga Cronbach's Alpha  $<$  ambang batas Cronbach's Alpha maka

instrumen tes tidak reliabel. Berdasarkan Tabel 8 diperoleh harga Cronbach's Alpha sebesar  $0,844 > 0,60$  maka instrumen soal uraian berarti reliabel. Reliabilitas soal pilihan ganda dapat dilihat pada Tabel 9.

**Tabel 9.** Hasil Uji Reliabilitas Butir Soal Pilihan Ganda  
Reliability Statistics

Cronbach's Alpha	N of Items
0,806	18

Butir soal pilihan ganda dikatakan reliabel jika nilai Cronbach's Alpha  $\geq$  ambang batas Cronbach's Alpha (0,60) dan bernilai positif tetapi jika nilai Cronbach's Alpha  $<$  ambang batas Cronbach's Alpha maka tidak reliabel. Dari Tabel 9 diperoleh harga Cronbach's Alpha sebesar 0,806 artinya lebih besar dari ambang batas Cronbach's Alpha sehingga dapat disimpulkan bahwa butir soal pilihan ganda reliabel.

### Daya Beda

Selain melihat validitas dan reliabel, sebuah instrumen tes yang baik harus juga memiliki daya beda yang baik dan baik sekali. Hasil daya beda soal pilihan ganda dan uraian dapat dilihat pada Tabel 10.

**Tabel 10.** Hasil Uji Daya Beda Soal Pilihan Ganda dan Uraian

No. Soal	BA	BB	JA	JB	PA	PB	PA-PB	Kriteria
1	34	23	35	34	0,97	0,68	0,29	Cukup
2	35	25	35	34	1	0,74	0,26	Cukup
3	34	22	35	34	0,97	0,65	0,32	Cukup
4	35	30	35	34	1	0,88	0,12	Jelek
5	35	31	35	34	1	0,91	0,09	Jelek
6	30	14	35	34	0,86	0,41	0,45	Baik
7	24	20	35	34	0,69	0,59	0,1	Jelek
8	23	21	35	34	0,66	0,62	0,04	Jelek
9	10	5	35	34	0,29	0,15	0,14	Jelek
10	26	17	35	34	0,74	0,5	0,24	Cukup
11	32	29	35	34	0,91	0,85	0,06	Jelek
12	33	21	35	34	0,94	0,62	0,33	Cukup
13	25	11	35	34	0,71	0,32	0,39	Cukup
14	33	16	35	34	0,94	0,47	0,47	Baik
15	35	24	35	34	1	0,71	0,29	Cukup
16	34	22	35	34	0,97	0,65	0,32	Cukup
17	30	17	35	34	0,86	0,5	0,36	Cukup
18	30	19	35	34	0,86	0,56	0,3	Cukup
1	95,5	77	35	34	2,729	2,265	0,464	Baik
2	81,5	63,5	35	34	2,329	1,868	0,461	Baik
3	105	71,5	35	34	3	2,103	0,897	Baik
4	95	64	35	34	2,714	1,882	0,832	Baik
5	88,5	62	35	34	2,529	1,824	0,705	Baik
6	90,5	57,5	35	34	2,586	1,691	0,895	Baik

Soal pilihan ganda ada pada nomor 1 sampai 18 yang atas sedangkan nomor 1 sampai 6 yang bawa merupakan soal uraian. Dari Tabel 10 dapat diperoleh bahwa soal pilihan ganda hanya nomor soal 6 dan 14 yang memiliki daya beda baik sedangkan soal 1,2, 3, 10, 12, 13, 15,16, 17, dan 18 memiliki daya beda yang cukup baik dan nomor soal 4, 5, dan 11 memiliki daya beda yang jelek. Soal uraian dari nomor 1 sampai 6 memiliki daya beda yang baik. Daya beda yang dapat digunakan sebagai alat uji dalam penelitian sebaiknya yang memiliki daya beda cukup baik, baik dan baik sekali.

### Tingkat kesukaran

Tingkat kesukaran juga merupakan hal yang dipersyaratkan dalam sebuah syarat instrumen tes. Instrumen tes dikatakan baik jika memiliki tingkat kesukaran sedang. Adapun hasil dari uji taraf kesukaran dapat dipaparkan pada Tabel 11.

**Tabel 11.** Hasil Uji Taraf Kesukaran Soal Pilihan Ganda dan Uraian

No. Soal	B	Js	P	Kriteria
1	57	69	0,83	Soal Mudah
2	60	69	0,87	Soal Mudah
3	56	69	0,81	Soal Mudah
4	65	69	0,94	Soal Mudah
5	66	69	0,96	Soal Mudah
6	44	69	0,64	Soal Sedang
7	44	69	0,64	Soal Sedang
8	44	69	0,64	Soal Sedang
9	15	69	0,22	Soal Sulit
10	43	69	0,62	Soal Sedang
11	61	69	0,88	Soal Mudah
12	54	69	0,78	Soal Mudah
13	36	69	0,52	Soal Sedang
14	49	69	0,71	Soal Mudah
15	59	69	0,86	Soal Mudah
16	56	69	0,81	Soal Mudah
17	47	69	0,68	Soal Sedang
18	49	69	0,71	Soal Mudah
1	173	69	0,14	Soal Sulit
2	145	69	0,12	Soal Sulit
3	177	69	0,14	Soal Sulit
4	159	69	0,13	Soal Sulit
5	151	69	0,12	Soal Sulit
6	148	69	0,12	Soal Sulit

Soal nomor 1 sampai 18 yang atas merupakan soal pilihan ganda sedangkan soal nomor 1 sampai 6 yang bawah merupakan soal uraian. Dari Tabel 11 diperoleh bahwa soal pilihan ganda 6, 7, 8, 10, 13, dan 17 adalah soal yang tingkat kesukarannya sedang sedangkan soal nomor 1, 2, 3, 4, 5, 12, 14, 15, 16, dan 18 merupakan soal dengan kriteria mudah dan soal nomor 9 merupakan soal yang sulit. Untuk soal uraian nomor 1 sampai 6 tergolong taraf kesukarannya dengan kriteria soal sulit. Soal yang digunakan sebagai instrumen tes sebaiknya soal yang memiliki tingkat kesukaran yang sedang.

Secara keseluruhan hasil validitas butir soal, reliabilitas, daya beda, dan tingkat kesukaran dapat dilihat pada Tabel 12.

**Tabel 12.** Hasil Uji Validitas Butir Soal, Reliabilitas, Daya Beda, dan Taraf Kesukaran

Nomor Soal	Sig.	Nilai $r_{hitung}$	Validitas	Cronbach's Alpha	Reliabilitas	Tingkat Kesukaran	Daya Beda
<i>Pilihan Ganda</i>							
1	0,000	0,701	Valid	0,782	Reliabel	Soal Mudah	Cukup
2	0,000	0,616	Valid	0,789	Reliabel	Soal Mudah	Cukup
3	0,000	0,671	Valid	0,784	Reliabel	Soal Mudah	Cukup
4	0,004	0,339	Valid	0,803	Reliabel	Soal Mudah	Jelek
5	0,001	0,386	Valid	0,802	Reliabel	Soal Mudah	Jelek
6	0,000	0,549	Valid	0,793	Reliabel	Soal Sedang	Baik
7	0,025	0,270	Valid	0,815	Reliabel	Soal Sedang	Jelek
8	0,004	0,346	Valid	0,809	Reliabel	Soal Sedang	Jelek
9	0,0016	0,288	Valid	0,810	Reliabel	Soal Sulit	Jelek
10	0,000	0,503	Valid	0,797	Reliabel	Soal Sedang	Cukup

11	0,011	0,304	Valid	0,806	Reliabel	Soal Mudah	Jelek
12	0,000	0,707	Valid	0,781	Reliabel	Soal Mudah	Cukup
13	0,000	0,438	Valid	0,803	Reliabel	Soal Sedang	Cukup
14	0,000	0,603	Valid	0,789	Reliabel	Soal Mudah	Baik
15	0,000	0,660	Valid	0,786	Reliabel	Soal Mudah	Cukup
16	0,000	0,598	Valid	0,789	Reliabel	Soal Mudah	Cukup
17	0,000	0,589	Valid	0,790	Reliabel	Soal Sedang	Cukup
18	0,020	0,281	Valid	0,812	Reliabel	Soal Mudah	Cukup
<i>Essai</i>							
1	0,000	0,525	Valid	0,864	Reliabel	Soal Sulit	Baik
2	0,000	0,581	Valid	0,843	Reliabel	Soal Sulit	Baik
3	0,000	0,759	Valid	0,787	Reliabel	Soal Sulit	Baik
4	0,000	0,621	Valid	0,817	Reliabel	Soal Sulit	Baik
5	0,000	0,672	Valid	0,803	Reliabel	Soal Sulit	Baik
6	0,000	0,776	Valid	0,784	Reliabel	Soal Sulit	Baik

**Keterangan:**Tarf Ketidakpercayaan ( $\alpha$ ) = 0,05Nilai batas *Cronbach's Alpha* = 0,600

Dari hasil keseluruhan (validitas, reliabilitas, daya beda, taraf kesukaran) diperoleh bahwa soal yang memenuhi ke empat kriteria tersebut adalah soal nomor 10,13, dan 17 pada soal pilihan ganda sedangkan soal uraian tidak ada yang memenuhi ke empat kriteria tersebut. Akan tetapi jika soal tersebut masih valid dan reliabel maka soal tersebut masih layak digunakan sebagai instrumen penelitian cuma perlu diperhatikan jangan terlalu memberi banyak porsi soal yang taraf kesukarannya sulit dan juga yang terlalu mudah dan juga jangan memakai terlalu banyak soal dengan daya beda yang jelek.

**KESIMPULAN**

Berdasarkan hasil penelitian dan untuk menjawab rumusan masalah dapat disimpulkan bahwa: hasil uji validitas isi oleh tiga ahli bidang pendidikan kimia diperoleh bahwa instrumen valid dan dapat digunakan untuk penelitian, karena berada pada kriteria sangat tinggi. Instrumen soal sebanyak 18 butir pilihan ganda dan 6 butir uraian diperoleh hasil semua valid dan reliabel. Soal pilihan ganda hanya nomor soal 6 dan 14 yang memiliki daya beda baik sedangkan soal 1,2, 3, 10, 12, 13, 15,16, 17, dan 18 memiliki daya beda yang cukup baik dan nomor soal 4, 5, dan 11 memiliki daya beda yang jelek. Soal uraian dari nomor 1 sampai 6 memiliki daya beda yang baik. Berdasarkan kriteria taraf kesukaran diperoleh hasil soal pilihan ganda 6, 7, 8, 10 13, dan 17 adalah soal yang tingkat kesukarannya sedang sedangkan soal nomor 1, 2, 3, 4, 5, 12, 14, 15, 16, dan 18 merupakan soal dengan kriteria mudah dan soal nomor 9 merupakan soal yang sulit. Untuk soal uraian nomor 1 sampai 6 tergolong taraf kesukarannya dengan kriteria soal sulit. Instrumen yang dapat digunakan minimal kriterianya valid dan reliabel.

**UCAPAN TERIMAKASIH**

Ucapan terimakasih tidak terhingga buat validator yang bersedia menvalidasi instrumen tes. Tim penulis juga mengucapkan terimakasih kepada sekolah tempat melaksanakan penelitian yaitu MA Darul Istiqomah yang telah memberikan kesempatan melakukan penelitian dan mendukung dengan semua sarana dan prasarananya selama proses pengambilan data penelitian.

**DAFTAR PUSTAKA**

- Arikunto, S. (2021). *Dasar-Dasar Evaluasi Pendidikan Edisi 3*. Bumi Aksara.
- Ida, F. F. (2021). Validitas dan Reliabilitas dalam Analisis Butir Soal. *AL-MU'ARRIB: JOURNAL OF ARABIC EDUCATION*, 1(1), 34–44.

- Irawan, S., & Iasha, V. (2021). Core Learning Model and Mathematical Disposition, Against Mathematics Problem Solving Ability of Elementary School Students. *Buana Pendidikan: Jurnal Fakultas Keguruan Dan Ilmu Pendidikan*, 17(2), 122–129.
- Irvan, M. (2017). Pengetahuan mahasiswa pg-paud unipa surabaya tentang pendidikan inklusif. *Buana Pendidikan: Jurnal Fakultas Keguruan Dan Ilmu Pendidikan*, 13(24), 155–164.
- Irvan, M., & Jauhari, M. N. (2018). Implementasi Pendidikan Inklusif Sebagai Perubahan Paradigma Pendidikan Di Indonesia. *Buana Pendidikan: Jurnal Fakultas Keguruan Dan Ilmu Pendidikan*, 14(26), 175–187.
- Kurniawati, D. (2011). Pengaruh penggunaan model STAD-PETA konsep terhadap prestasi belajar siswa kelas X SMK Putra Indonesia Malang pada materi pokok materi dan perubahannya. Universitas Negeri Malang.
- Magdalena, I., Fauziah, S. N., Faziah, S. N., & Nopus, F. S. (2021). Analisis Validitas, Reliabilitas, Tingkat Kesulitan dan Daya Beda Butir Soal Ujian Akhir Semester Tema 7 Kelas III SDN Karet 1 Sepatan. *BINTANG*, 3(2), 198–214.
- Musayaroh, T. (2021). HOTS-Based Chemical Literacy Test Instrument Validation in Terms of Content Validity by Experts. *UNESA Journal of Chemical Education*, 10(3), 243–251.
- Nuswowati, M., Binadja, A., & Ifada, K. E. N. (2010). Pengaruh validitas dan reliabilitas butir soal ulangan akhir semester bidang studi kimia terhadap pencapaian kompetensi. *Jurnal Inovasi Pendidikan Kimia*, 4(1).
- Riyani, R., Maizora, S., & Hanifah, H. (2017). Uji Validitas Pengembangan Tes Untuk Mengukur Kemampuan Pemahaman Relasional Pada Materi Persamaan Kuadrat Siswa Kelas Viii Smp. *Jurnal Penelitian Pembelajaran Matematika Sekolah (JP2MS)*, 1(1), 60–65.
- Sari, W. M., Riswanto, R., & Partono, P. (2019). Validitas mobile pocket book berbasis android menggunakan adobe flash pada materi suhu dan kalor. *Berkala Ilmiah Pendidikan Fisika*, 7(1), 35–42.
- Solichin, M. (2017). Analisis daya beda soal, taraf kesukaran, validitas butir tes, interpretasi hasil tes dan validitas ramalan dalam evaluasi pendidikan. *Dirasat: Jurnal Manajemen Dan Pendidikan Islam*, 2(2), 192–213.
- Sukidin, B., & Basrowi, S. (2012). Manajemen Pendidikan Penelitian. Penerbit Insan Cendekia.
- Utomo, B. (2019). Analisis validitas isi butir soal sebagai salah satu upaya peningkatan kualitas pembelajaran di madrasah berbasis nilai-nilai islam. *Jurnal Pendidikan Matematika (Kudus)*, 1(2).
- Zein, A., Fadillah, M., & Novianti, R. (2013). Hubungan Antara Validitas Butir, Reliabilitas, Tingkat Kesukaran Dan Daya Pembeda Soal Ujian Semester Genap Bidang Studi Biologi Kelas Xi Sma/Ma Negeri Di Kota Padang Tahun Pelajaran 2010/2011.