

**MENGENAL TEKNIK RELIABILITAS PENGUKURAN PSIKOLOGI  
MELALUI SIKAP TERHADAP PELAJARAN KIMIA**  
*FAMILIARIZE YOURSELF WITH PSYCHOLOGICAL MEASUREMENT REABILITY  
TECHNIQUES ATTITUDES TOWARD CHEMISTRY*

<sup>1</sup>Jelpa Periantalo, <sup>2</sup>Beni Rahim, <sup>3</sup>Marlitha Andika Rahman,  
<sup>4</sup>Yun Nina Ekawati <sup>5</sup>Fadzlul <sup>6</sup>Nick Jhowson

Jurusan Psikologi  
Fakultas Kedokteran dan Ilmu Kesehatan  
Universitas Jambi, Indonesia  
jelp.8487@unja.ac.id, beny.rahimpoeding@gmail.com, lita.rahman99@gmail.com,  
yun\_nina.e@unja.ac.id fadzlul\_fkik@unja.ac.id nickpanjaitan44@gmail.com

**ABSTRACT**

**Introduction** *The purpose of this article to give understanding about reliability of psychological measurement with the emperical data. The data used attitude toward chemistery scale from Charles Osgood theory.*

**Method** *It used semantic differensial format and continuum of score of item 1-7 with three items. It explained internal consistency, parallel test, test retest and composite score. There were N=278 subejcts from the Faculty of Medicine and Health Sciences The University of Jambi.*

**Result** *The internal consistency method used four technique with reliability coeffiecent  $r=0.717$  to  $0.824$ . the parallel test reliability showed equivalency coeffiecient  $r_{xx'}=0.851$  and the test retest reliability with  $r_{x1x2} = 0.754$ . Score composite reliability attitude toward Mathematics and Sciences showed  $r=0.914$ .*

**Conclusions** *The reliability analysis showed the good result, the data could be used in psychometric lecturing, validate other test, and the scale for the practical usage.*

**Keywords:** *reliability, parallel test, test retest, internal consistency, true score*

**Abstrak**

**Pendahuluan** Artikel bertujuan memberikan pemahaman teknik reliabilitas pengukuran psikologi dengan data empirik. Data menggunakan skala sikap terhadap pelajaran Kimia dari teori sikap Charles Osgood.

**Metode** Aitem menggunakan format penskalaan sematik differensial 1-7 dan berjumlah 3 aitem. Reliabilitas dijelaskan berupa konsistensi internal, tes paralel, tes ulang dan skor komposit. Subjek penelitian sebanyak N=278 dari mahasiswa Fakultas Kedokteran dan Ilmu Kesehatan Universitas Jambi.

**Hasil** Empat teknik reliabilitas konsistensi internal menghasilkan koefisien  $r=0.717$  s.d.  $0.824$ . Uji reliabilitas tes paralel menghasilkan koefisien ekivalensi sebesar  $r_{xx'}=0.851$ . Uji reliabilitas tes ulang menghasilkan koefisien stabilitas  $r_{x1x2}=0.754$ . Uji reliabilitas skor komposit dalam skala Sikap terhadap pelajaran Matematika dan Sains  $r=0,914$ .

**Kesimpulan** Hasil uji reliabilitas menunjukkan hasil memuaskan sehingga dapat digunakan sebagai data perkuliahan Psikometri, uji validiasi alat ukur lain, maupun kegunaan praktis.

**Kata kunci :** reliabilitas, tes paralel, tes ulang, konsistensi internal, skor murni

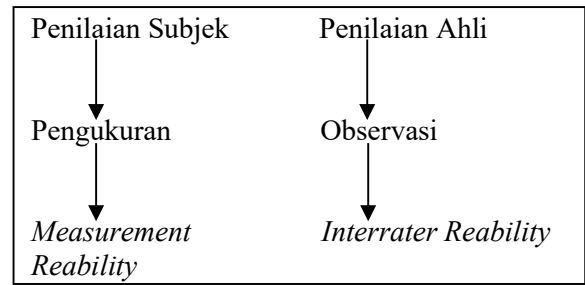
**Pendahuluan**

Dalam pengukuran psikologi, mengenal dua istilah utama validitas dan reliabilitas, yang merupakan syarat utama dalam alat ukur yang baik (Azwar, 2014; Kurpius & Stafford, 2006; Periantalo, 2017; Stangor, 2010; Supratiknya 2014). Secara tradisional, validitas diartikan sejauhmana alat ukur mampu mengungkap atribut yang menjadi sasaran ukurnya. Validitas dibagi menjadi tiga bagian yaitu isi, konstruk dan kriteria. Secara modern validitas berkaitan dengan interpretasi skor (AERA, APA, & NCME, 2014; Akhtar, 2020; Azwar, 2013). Artinya sejauhmana mampu membuktikan dukungan terhadap interpretasi skor baik melalui proses kualitatif maupun hasil kuantitatif.

Bukti validitas bersumber dari isi alat ukur, respons proses jawaban, struktur internal, hubungan dengan variabel lain dan konsekuensi tes. Pengukuran digunakan sebagai bagian penelitian ilmiah, mendiagnosis keadaan individu maupun memprediksi performansi masa depan. Apabila instrumen pengukuran digunakan memiliki berbagai informasi validitas yang baik, kekuatan kebenaran dari yang tujuan ukur adalah kuat (Azwar, 2017; Arifin, 2017; Periantalo, 2017).

Ada banyak penelitian ilmiah maupun pembahasan tentang validitas, berbagai hal tentang reliabilitas tidak sebanyak validitas. Reliabilitas dikaitkan dengan keterpercayaan terhadap skor hasil pengukuran (Matondang, 2009; Urbina, 2014; Widodo, 2008; Widhiarso, 2009). Sejahumana mampu mencerminkan variasi skor murni dan dengan variasi kesalahan (*error*). Reliabilitas yang baik adalah dimana variasi skor murni subjek adalah tinggi, sementara kesalahan pengukuran adalah minim.

Secara umum reliabilitas bergerak dari koefisien 0 – 1 dan bertanda positif; semakin mendekati satu, semakin bagus koefisien reliabilitas dari alat ukur tersebut. Ada toleransi terhadap skor reliabilitas yang baik disesuaikan dengan jenis pengukuran, tujuan, jumlah alat yang digunakan dan subjek. Tujuan diagnosis menghendaki sebaik 0.8; tujuan prognosis sebaiknya 0.9; tujuan penelitian bisa ditoleransi menjadi 0.7 (Azwar, 2016; Periantalo, 2017; Sumintoro & Widhiarso, 2014; Yusup, 2018).



**Gambar 1 Reabilitas Berdasarkan Respon**

Ada dua yang perlu dipahami berkaitan dengan proses respon reliabilitas (Azwar, 2013; Khumaedi, 2012; Kyriazod & Stalikas, 2018; Periantalo, 2016; Yusuf, 2010; Widi, 2010;). *Pertama*, aitem dinilai langsung oleh subjek yang menjadi tujuan ukur sehingga dia lah yang paling memahami isinya. Dia memberikan respon terhadap stimulus yang dihadapkan padanya sehingga pengukuran psikologi lah yang tepat. Kita dapat memberlakukannya pada pengisian tes potensi akademik, skala Komitmen organisasi, skala depresi remaja. *Kedua*, berkaitan dengan respon yang dinilai oleh ahli yang memahami konsep atau tujuan ukurnya, merekalah yang menilai apa yang dilakukan oleh subjek. Hasil dari penilaian para ahli yang akan diujikan reliabilitas yang dikenal dengan reliabilitas antar rater. Teknik Statistik yang dapat digunakan adalah koefisien Kappa untuk data dikotomi (1 dan 0) dan *interclass correlation coefficient* (ICC) untuk data politomi (0-4, 1-5, 1-7). Reliabilitas ini dapat diterapkan pada penilaian soal essay, lomba mengarang karya ilmiah, lomba menyanyi, lomba senam, lomba melukis.

**Tabel 2 Metode Reabilitas untuk Pengukuran Psikologi**

Metode	Teknik Statistik	Simbol	Keterangan
Test ulang	Korelasi	$r_{xx'}$	Satu alat tes, dua kali pengukuran
Test paralel	Korelasi	$r_{x_1x_2}$	Dua alat tes, satu kali pengukuran
Konsistensi internal	Alpha dkk	$\alpha$	Konsistensi antar aitem/ bagian
Komposit	Mosier	$\alpha$	Beberapa komponen/faktor

Metode reliabilitas yang pertama adalah tes ulang, dimana satu alat ukur diukur melalui lebih dari satu kurun waktu. Reliabilitas melihat stabilitas pengukuran antar waktu, yang dikenal

dengan koefisien stabilitas dengan simbol  $rx1x2$ . Metode ini cocok untuk variabel yang relatif menetap pada diri manusia. Metode tes paralel, dimana ada dua alat ukur yang setara baik secara konsep maupun statistika diuji dalam satu kurun waktu. Metode ini melihat kesetaraan pengukuran di antara dua alat ukur, skor yang dihasilkan disebut dengan koefisien ekivalensi dengan  $rxx'$ .

Kedua model ini menggunakan teknik statistika korelasi dengan skor total alat ukur. Metode konsistensi internal menggambarkan bagaimana konsistensi antar skor atau komponen di dalam pengukuran tersebut. Skor yang digunakan adalah skor dari setiap aitem dari alat ukur, bukan skor total alat ukur. Metode banyak digunakan karena efisien serta terdapat banyak rumus. Reliabilitas skor komposit diperuntukkan untuk sebuah instrumen yang pengukuran terdiri dari beberapa sub skala di dalamnya ataupun alat ukur multi dimensi dari analisis faktor (Azwar, 2016; Bolarinwa, 2015; Devellis, 2017; Periantalo dkk, 2014).

Artikel bertujuan untuk memberikan pemahaman tentang beberapa teknik reliabilitas dalam bentuk pengukuran psikologi melalui data riil berbasis riset. Teknik reliabilitas yang disajikan berupa konsistensi internal, tes paralel, tes ulang dan skor komposit. Data yang digunakan merupakan skala sikap terhadap pelajaran Kimia, yang menggunakan format aitem semantik differensial respon 1 – 7 untuk setiap aitem. Data dari alat ukur ini dapat dijadikan bahan pembelajaran mata kuliah psikometri di kelas. Dia juga bisa menjadi bahan uji validitas kriteria untuk alat ukur yang lain. Secara praktis, Alat ukur ini dapat digunakan dalam membantu siswa untuk menentukan pengembangan minat bakat maupun penjurusan kuliah.

## Metode

Variabel penelitian ini adalah sikap terhadap pelajaran Kimia. Sikap berkaitan dengan pandangan positif atau negatif terhadap suatu objek tertentu. Dasar teori dari sikap diambil dari teori Charles Osgood yang menyebutkan bahwa sikap terdiri dari tiga komponen, yaitu: evaluasi, potensi dan aktivitas (Azwar, 2014; Chráskaa & Chráskováa, 2016; Gluth dkk, 2010; Periantalo, 2016). Objek sikap dari penelitian ini adalah Kimia, salah satu pelajaran yang menjadi bahan kajian di peminatan Matematika dan Sains di SMA.

Instrumen peneltiian menggunakan skala sikap terhadap pelajaran Kimia, yang awalnya berjumlah 17 aitem. Skala tersebut dilakukan uji diskriminasi aitem dan analisis faktor dengan hasil yang memuaskan (Periantalo dkk, 2014). Skala final menggunakan 3 aitem dari setiap komponen sikap dari Charles Osgood. Format yang digunakan dalam sikap dari teori Osgood adalah semantif differensial, dimana dua kata sifat yang bertolak belakang disajikan dengan respon jenjang 1-7.

Skala Utama	Rumit	1 2 3 4 5 6 7	Sederhana
	Tidak Penting	1 2 3 4 5 6 7	Penting
	Jijik	1 2 3 4 5 6 7	Takjub
	Skala Paralel		
Berat	1 2 3 4 5 6 7	Ringan	
Tidak Penting	1 2 3 4 5 6 7	Penting	
Melesukan	1 2 3 4 5 6 7	Menggairahkan	

Terdapat dua bentuk skala final yaitu skala utama dan paralel. Skala utama adalah skala umum yang digunakan khususnya dalam penelitian, paralel adalah sebagai pengantinya. Skala paralel juga dibutuhkan sebagai bagian dari uji reliabilitas tes paralel. Skala utama digunakan bagian dari reliabilitas konstensi internal, tes ulang dan skor komposit. Subjek diminta untuk memberikan tanggapan dari kontinum dua kata sifat bertolak belakang tersebut. skor minimum yang didapat subjek adalah 3, skor maksimum adalah 21.

Teknik pengambilan sampel menggunakan teknik purposif, dimana subjek yang menjadi tujuan ukur dilibatkan dalam penelitian ini. Data diambil pada mahasiswa baru Fakultas Kedokteran dan Ilmu Kesehatan – Universitas Jambi terdiri dari Kedokteran, Keperawatan, Kesehatan Masyarakat dan Psikologi berjumlah N=278.

## Hasil

Terdapat berbagai teknik analisis data Statistika dalam memahami reliabilitas melalui skala sikap terhadap pelajaran Kimia. Pada pendekatan reliabilitas konsistensi internal menggunakan Cronbach's  $\alpha$ , Mc Donald's  $\omega$ , Guttman's  $\lambda_6$  dan greater lower bound. Pada uji reliabilitas skor komposit menggunakan teknik statistika formula Mosier. Pada reliabilitas tes ulang dan tes paralel

menggunakan teknik korelasi ( $r$ ), dimana Pearson untuk data normal serta Spearman Rho untuk data tidak normal. Dalam uji keparalelan menggunakan varians ( $S^2$ ), mean ( $M$ ), dan korelasi ( $r$ ). Analisis data dibantu oleh *software* JASP 0.10.10 (*Jeffrey's Amazing Statistics Program*) yang gratis, mudah dan berbasis penulisan APA.

**Tabel 3 Hasil Uji Reabilitas Konsistensi Internal Skala Sikap Terhadap Kimia**

Teknik	Koefisien
MC Donald	$\omega = 0.788$
Cronbach	$\alpha = 0.773$
Guttman	$\lambda_6 = 0.717$
Greater Lower Bound	$r = 0.824$

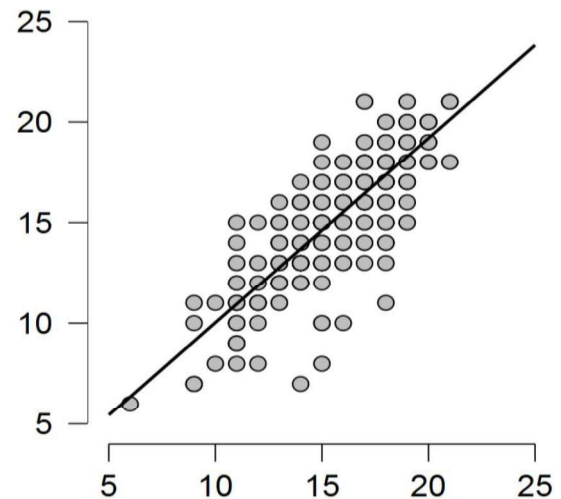
Dari keempat jenis uji reliabilitas konsistensi internal, terdapat kesetaraan di antara empat hasil koefisien reliabilitas tersebut bergerak dari 0.717 – 0.824. Reliabilitas dengan teknik *greater lower bound* memiliki koefisien reliabilitas paling tinggi yaitu  $r = 0.824$ . Hal tersebut menandakan bahwa sebanyak 82.4% variasi skor berasal dari variasi skor murni subjek, sebanyak 17.6% variasi skor berasal dari variasi kesalahan pengukuran. Keempat teknik tersebut tersedia di perangkat lunak JASP.

**Tabel 4 Hasil Uji Reabilitas Test Paralel Skala Sikap Terhadap Pelajaran Kimia**

	M	S <sup>2</sup>	rx <sub>y</sub>	rx <sub>x'</sub>
Skala Utama (X)	15.292	10.311	0.534	0.851
Skala Paralel (x')	15.743	8.263	0.571	

Untuk melihat apakah kedua alat ukur tersebut adalah paralel perlu diuji keparalelan melalui mean, varians, dan korelasi dengan variabel lain. Skala utama diberi simbol  $x$ , skala paralel/ kembaran diberi simbol  $x'$ , teknik statistika yang dilakukan berupa korelasi dengan simbol  $r$ . Hasil uji asumsi tersebut menunjukkan bahwa ketiga uji statistika tersebut memiliki kesetaraan dalam hasil sehingga reliabilitas alat ukur tersebut paralel. Skala utama dan skala paralel sama-sama memiliki nilai rata-rata sekitar 15; uji varians kedua skala cukup memuaskan karena menghasilkan skor varians untuk skala pertama  $S^2x = 10.311$  dan skala paralel  $S^2x' = 8.263$ . Variabel lain yang digunakan sebagai uji

keparalelan adalah skala sikap terhadap pelajaran Biologi dengan simbol  $y$ . Uji korelasi tersebut menghasilkan koefisien korelasi yang setara untuk kedua bentuk skala  $rx_{y'} = 0.534$  dan  $rx'y = 0.571$ . Koefisien ekivalensi yang dihasilkan



untuk reliabilitas sebesar  $rx_{x'} = 0.851$ .

**Gambar 1. Diagram Pencar Korelasi Dua Skala Sikap terhadap Pelajaran**

Hasil dari uji statistik dapat disajikan juga dalam bentuk gambar, selain angka. Dari gambar terlihat jelas menunjukkan korelasi positif sangat tinggi untuk koefisien reliabilitas tes ulang  $rx_{x'} = 0.851$ . Hasil bergerak dari kiri bawah menuju kanan atas secara teratur. Koefisien korelasi 0.800 – 0.999 berkategori positif sangat tinggi atau kuat. Hal tersebut menunjukkan bahwa apabila subjek memiliki skor yang tinggi pada skala utama, dia dipastikan mendapatkan skor yang tinggi pula pada skala paralel. Subjek yang mendapatkan skor yang rendah pada skala utama, dia dipastikan mendapatkan skor yang rendah pula pada skala paralel. Apabila seorang mendapatkan skor 17 pada skala utama, dia dipastikan mendapatkan skor sekitar 16, 17 atau 18 pada skala kedua. Apabila subjek mendapatkan skor 8 pada skala utama, dia dipastikan mendapatkan skor sekitar 6, 7 atau 8.

**Tabel 5 Hasil Reabilitas Tes Ulang Skala Sikap Terhadap Kimia**

	rx <sub>1</sub>
rx <sub>2</sub>	0.724

Pengukuran pertama diberi simbol  $x_1$  dan pengukuran kedua simbol  $x_2$ . Teknik Statistik yang digunakan berupa uji korelasi dengan simbol  $r$ . Uji reliabilitas skala sikap terhadap pelajaran Kimia dengan teknik tes ulang

menghasilkan koefisien stabilitas  $r_{x1x2} = 0.724$ . Koefisien korelasi tersebut berkategori positif tinggi. Hal tersebut menunjukkan bahwa subjek yang mendapatkan skor yang tinggi pada pengukuran pertama, dia pun mendapatkan skor yang tinggi pada pengukuran kedua. Subjek yang mendapatkan skor yang rendah pada pengukuran pertama, dia pun mendapatkan skor yang rendah sampai pengukuran kedua.

Skala sikap terhadap pelajaran Kimia dapat disatukan menjadi skala sikap terhadap pelajaran Matematika dan Sains. Dia terdiri dari sub skala sikap terhadap pelajaran Matematika, Fisika, Kimia, dan Biologi. Keempat skala tersebut merupakan alat ukur yang berbeda, tetapi bisa menjadi satu kesatuan. Setiap alat ukur memiliki tiga (3) aitem yang mewakili setiap komponen sehingga terdapat 12 aitem total. Dalam pengujian reliabilitas, skala sikap tersebut akan memiliki reliabilitas skor komposit dari rumus Mosier. Kita memerlukan informasi reliabilitas konsistensi internal dari setiap sub skala, korelasi antara sub skala, standar deviasi dan bobot relatif dari setiap sub skala.

**Tabel 7 Korelasi antar Komponen Skala Sikap Matematika dan Sains**

	Matematika	Fisika	Kimia	Biologi
Matematika	<b>(0.846)</b>	0.406	0.462	0.372
Fisika		<b>(0.751)</b>	0.493	0.408
Kimia			<b>(0.824)</b>	0.532
Biologi				<b>(0.794)</b>
r Komposit		0.914		

Tabel menjelaskan tentang reliabilitas konsistensi internal dari setiap skala dengan teknik *greater lower bound*. Skala Matematika memiliki koefisien reliabilitas  $r=0.846$ , skala fisika dengan koefisien  $r=0.751$ ; skala Kimia dengan koefisien  $r=0.824$  dan skala Biologi dengan koefisien  $r=0.794$ . Kita pun mendapati korelasi di antara empat komponen skala tersebut seperti korelasi skala Matematika dengan Fisika sebesar  $r_{xy} = 0.406$ ; korelasi skala Matematika Kimia sebesar  $r_{xz}=0.462$ . Bobot relatif dari setiap komponen tersebut adalah setara, dalam artian 25% untuk sub skala. Standar deviasi untuk sub skala Matematika  $S_x$ ; 3,205, Fisika  $S_y$  ; 3,478, Kimia  $S_z$  ; 3,287, Biologi  $S_a = 2,316$ . Setelah dimasukkan ke rumus, adapun koefisien reliabilitas skor komposit sebesar  $r=0,914$

### Pembahasan

Reliabilitas konsistensi internal yang dihasilkan melalui skala sikap terhadap pelajaran

Kimia berkategori bagus (Arifin, 2017; Azwar, 2016; Bolarinwa, 2015; DeVellis, 2015; Periantalo, 2015). Metode konsistensi internal dipengaruhi oleh faktor jumlah aitem. Semakin banyak jumlah aitem, semakin tinggi koefisien reliabilitas. Skala sikap terhadap pelajaran Kimia hanya berjumlah 3 aitem, apabila ditambah menjadi 10 aitem, koefisien reliabilitas di atas 0.90. Apa yang menyebabkan reliabilitas konsistensi internal yang baik ini, walaupun dengan sedikit aitem? Semua aitem memiliki daya beda  $rit > 0.600$  (Periantalo dkk, 2014). Semakin baik daya beda aitem, reliabilitas yang dihasilkan semakin tinggi.

Pada metode paralel reliabilitas yang dihasilkan juga bagus, kedua alat ukur tersebut mirip (Azwar, 2013; Howel, 2018; Khumaedi, 2012; Kyriazos & Stalikas, 2018; Periantalo, 2016). Secara teori memang berasal dari teori yang sama dan mumpuni, jumlah aitem pun serta kemiripan bunyi butir soal. Mean dan korelasi yang dihasilkan adalah setara, varians juga relatif sama. Secara teori dan angka statistik kedua sudah memiliki kemiripan yang kuat (paralel). Koefisien ekuivalensi yang dihasilkan sebesar  $r_{xx'} = 0.851$  yang berkategori positif sangat tinggi. Kita mempercayai bahwa terdapat konsistensi yang sangat tinggi dari kemurnian variasi skor murni dalam suatu hasil reliabilitas. Hal tersebut juga menandakan sumber error yang minim.

Pada reliabilitas tes ulang koefisien reliabilitas yang dihasilkan tetap berkategori baik. Koefisien stabilitas yang dihasilkan sebesar  $r_{x1x2} = 0.724$  yang berkategori positif tinggi dalam uji korelasi (Azwar, 2016; Periantalo, 2017; Yusup, 2018; Widi, 2011; Widhiarso, 2019). Hal tersebut menunjukkan ada konsistensi hasil pengukuran dalam kurun waktu yang berbeda. Variasi skor murni dari hasil pengukuran tersebut menunjukkan kestabilan hasil pengukuran antar waktu. Jumlah subjek dalam uji tersebut hanya 80 orang, reliabilitas tes ulang ikut dipengaruhi oleh subjek (Periantalo, 2015). Semakin banyak subjek, semakin tinggi koefisien reliabilitas yang dihasilkan. Apabila subjek dalam uji tes ulang ditambah menjadi 200 orang, dia mendapatkan koefisien reliabilitas  $> 0.800$

Reliabilitas skor komposit dari empat skala sikap terhadap pelajaran Matematika dan Sains menunjukkan hasil yang sangat memuaskan. Reliabilitas tersebut sebesar  $r=0,914$ . Reliabilitas yang sangat baik disebabkan karena masing-masing komponen memiliki reliabilitas konsistensi internal yang berkategori baik.

Begitu dengan korelasi diantara sub skala memiliki korelasi yang relatif sama di antara mereka. Hal tersebut menunjukkan adanya hubungan yang searah dalam skala tersebut (Azwar, 2014; Howel, 2018; Matondang, 2009; Periantalo, 2015; Urbina, 2014).

### Kesimpulan

Terdapat empat teknik dalam uji reliabilitas skala sikap sikap terhadap pelajaran Kimia yang menghasilkan data yang memuaskan. Uji reliabilitas konsistensi internal menghasilkan koefisien antara 0.717 sampai 0.824. Pada skala tes ulang paralel menghasilkan koefisien ekuivalensi sebesar  $r_{xx'} = 0.851$ . koefisien stabilitas yang dihasilkan pun dari uji reliabilitas tes ulang sebesar  $r_{1x2} = 0.728$ . Pada uji reliabilitas skor komposit menghasilkan koefisien sebesar  $r = 0.914$ . Data uji reliabilitas tersebut adalah bagus sehingga dapat digunakan oleh akademisi dalam penjelasan reliabilitas di mata kuliah Psikometri. Instrumen pun dapat digunakan untuk mendiagnosis sikap individu terhadap pelajaran tersebut. Instrumen juga dapat digunakan dalam validasi skala yang lain.

### Daftar Pustaka

AERA, APA, & NCME. (2014). *Standards for Educational and Psychological Testing*. Washington, DC: American Educational Research Association.

Akhtar. H. (2020, Agustus 15). Validitas dalam penyusunan alat ukur. <https://www.semestapsikometrika.com/2017/07/validitas.html>

Arifin, Z. (2017). Kriteria instrumen dalam suatu penelitian. *Jurnal THEOREMS*, 2 (1), 28-36.

Azwar, S. (2013). *Reliabilitas dan validitas*. Pustaka Pelajar: Yogyakarta.

Azwar, S. (2014). *Sikap manusia: Teori dan pengukurannya*. Pustaka Pelajar: Yogyakarta.

Azwar, S. (2014). *Penyusunan skala psikologi*. Ed. Ke-2. Pustaka Pelajar: Yogyakarta.

Azwar, S. (2014). *Konstruksi tes kognitif*. Pustaka Pelajar: Yogyakarta.

Azwar, S. (2016). *Dasar-dasar psikometri*. Ed. Ke-2. Pustaka Pelajar: Yogyakarta.

Azwar, S. (2017). *Metode penelitian kuantitatif psikologi*. Ed. Ke-2. Pustaka Pelajar: Yogyakarta.

Bolarinwa, O. A. (2015) Principles and methods of validity and reliability testing of questionnaires used in social and health

science researches. *Nigerian Postgraduate Medical Journal*, 22, 195-201.

Chráskaa, C. & Chráskováa, M. (2016). Semantic differential and its risks in the measurement of students' attitudes. *Procedia - Social and Behavioral Sciences*, 217, 820 – 829.

Devellis, R. F. (2017). *Scale development: Theory and application*. 2nd Ed. Thousand Oaks: Sage.

Gluth, S., Ebner, N. E., & Schmiedek, F. (2010). Attitudes toward younger and older adults: The German aging semantic differential. *International Journal of Behavioral Development*, 34 (2), 147–158.

Howel, C.D. (2018). *Statistical method for Psychology*. 7<sup>th</sup> Ed. Bellmon: Wadworth.

Khumaedi, M. (2012). Reliabilitas instrumen penelitian pendidikan. *Jurnal Penelitian Teknik Mesin*, 12 (1), 25-30.

Kurpius, S. E. R., Stafford, M.E. (2006). *Testing and measurement: A user-friendly guide*. California: Thousand Oaks.

Kyriazos, T. A., Stalikas, A. (2018). Applied Psychometrics: The steps of scale development and standardization Process. *Psychology*, 9, 2531-2560.

Matondang, Z. (2009). Validitas dan reliabilitas: suatu instrumen penelitian. *Jurnal Tabularasa*, 6 (1), 87-97.

Periantalo, J., Fadzlul, F., & Saputra, N. E. (2014). Konstruksi skala sikap terhadap pelajaran Matematika dan Sains. *Edu-Sains: Jurnal Pendidikan Matematika dan Ilmu Pengetahuan Alam*, 3 (2), 36-45.

Periantalo, J. (2015). *Penyusunan skala psikologi: asyik, mudah dan bermanfaat*. Pustaka Pelajar: Yogyakarta.

Periantalo, J. (2015). *Validitas alat ukur psikologi: Aplikasi praktis*. Yogyakarta: Pustaka Pelajar.

Periantalo, J. (2016). *Penelitian kuantitatif untuk psikologi*. Yogyakarta: Pustaka Pelajar.

Periantalo, J. (2017). *Statistika dasar untuk psikologi*. Yogyakarta: Pustaka Pelajar.

Sumintoro, B., & Widhiarso, W. (2015). *Aplikasi Rasch model untuk penelitian ilmu-ilmu sosial*. Bandung: Trim Komunikata.

Sumintoro, B., & Widhiarso, W. (2015). *Aplikasi pemodelan Rasch pada asesmen pendidikan*. Bandung: Trim Komunikata.

Supratiknya, A. (2014). *Pengukuran psikologis*. Yogyakarta: Penerbit Universitas Sanata Darma.

- Stangor, C. (2010). *Introduction to Psychology*. Boston: Flat World Knowledge.
- Urbina, S. (2014). *Essential of psychological testing*. New York: Mc Graw Hill.
- Widodo, P. B. (2006). Reliabilitas dan validitas konstruk skala konsep diri untuk mahasiswa Indonesia. *Jurnal Psikologi Universitas Diponegoro*, 3 (1), 1-9.
- Widhiarso, W. (2009). Estimasi reliabilitas pengukuran dalam pendekatan model persamaan struktural. *Buletin Psikologi*, 17 (1), 33-38.
- Widhiarso, W. (2020, September 21). Psychological Construct. <https://psikometrika.net/psymetrics/>
- Widi, R. (2011). Uji validitas dan reliabilitas dalam penelitian epidemiologi Kedokteran Gigi. *Jurnal Stomatognatic*, 8 (1), 27-34.
- Yusup, F. (2018). Uji validitas dan reliabilitas instrumen penelitian kuantitatif. *Jurnal Tarbiyah: Jurnal Ilmian Pendidikan*, 7 (1), 17-23.