

Analisis Instrumen Tes Menggunakan *Rasch Model* dan *Software SPSS 22.0*

Elsa Febrina Tarigan^{1✉}, Suriati Nilmarito², Khairani Islamiyah², Ayi Darmana², dan Retno Dwi Suyanti²

¹Pascasarjana Pendidikan Kimia, Universitas Negeri Medan, Medan, Indonesia

²Departemen Pendidikan Kimia, Universitas Negeri Medan, Medan Indonesia

Info Artikel

Diterima Maret 2022

Disetujui Mei 2022

Dipublikasikan Juli 2022

Keywords:

uji validitas
reliabilitas
instrumen tes
Rasch model
SPSS 22.0

Abstrak

Penelitian ini bertujuan untuk membandingkan perbedaan validitas dan reliabilitas instrumen tes menggunakan *Rasch model* dan *software SPSS 22.0*. Penelitian ini merupakan penelitian kualitatif dengan metode deskriptif, yang diujicobakan kepada 40 mahasiswa Universitas Negeri Medan yang berjumlah 40 butir soal dengan 5 opsi (a, b, c, d, dan e). Selanjutnya data yang diperoleh dianalisis melalui pendekatan teori tes klasik menggunakan bantuan program model *Rasch* dengan *software Winsteps* dan *SPSS* versi 22.0. Hasil penelitian menggunakan *SPSS 22.0* dan *Rasch model* memiliki keakuratan validitas yang baik yang menunjukkan taraf perbedaan hasil yang kecil pada hasil validitas butir soal. Berdasarkan analisis *SPSS 22.0* diperoleh 20 item yang valid dan 20 item yang tidak valid. Sedangkan menggunakan *Rasch model* dengan bantuan *software Winstep* diperoleh 17 item yang valid dan 23 item tidak valid. Berdasarkan keakuratan data, analisis validasi menggunakan *Rasch model* dengan bantuan *software Winstep* lebih akurat dibandingkan dengan menggunakan program *SPSS 22.0*. Faktor yang membedakan hasil reliabilitas *SPSS 22.0* dan *Rasch model* dengan bantuan dari hasil uji validasi. Sehingga, uji reliabilitas menggunakan *SPSS 22.0* tingkat reliabilitasnya 0,828 (kategori tinggi) dan *Rasch model* yang diperoleh 0,69 (kategori cukup).

Abstract

This study aims to compare the differences in the validity and reliability of test instruments using the Rasch Model and SPSS 22.0 software. This research is qualitative research with the descriptive method, tested on 40 students of Medan State University, totaling 40 questions with five options (a, b, c, d, and e). Furthermore, the data obtained were analyzed through the classical test theory approach using the Rasch model program with Winsteps software and SPSS version 22.0. The study results using SPSS 22.0 and the Rasch Model have good validity accuracy, which shows a trim level of difference in the results of the validity of the items. Based on the analysis of SPSS 22.0 obtained 20 valid items and 20 invalid items. In contrast, using the Rasch Model with the help of Winstep software, obtained 17 valid items and 23 invalid items. Based on the accuracy of the data, validation analysis using the Rasch Model with the help of Winstep software is more accurate than using the SPSS 22.0 program. Factors that distinguish the reliability results of SPSS 22.0 and the Rasch Model with the help of the validation test results. Thus, the reliability test using SPSS 22.0 has a reliability level of 0.828 (high category), and the Rasch Model obtained is 0.69 (enough category).

© 2022 Universitas Negeri Semarang

PENDAHULUAN

Evaluasi dalam pembelajaran adalah suatu proses atau kegiatan untuk mengukur dan menilai beberapa kemampuan siswa dalam pembelajaran seperti pengetahuan, sikap dan keterampilan guna membuat keputusan tentang status kemampuan siswa tersebut. Juga harus dibekali kemampuan mengevaluasi proses pembelajaran setelah kegiatan proses pembelajaran (Ratnawulan dan Rusdiana, 2014).

Suatu instrumen dikatakan valid ketika mengukur apa yang seharusnya diukur, atau dengan kata lain ketika suatu instrumen secara akurat mengukur setiap variabel yang ditentukan, itu dianggap sebagai instrumen yang valid untuk itu variabel tertentu. Ada empat jenis validitas; validitas wajah, validitas kriteria, validitas isi atau validitas konstruk (Mujis, 2011). Validitas wajah adalah melihat konsep apakah tes tersebut terlihat valid atau tidak pada tampilannya permukaan. Validitas kriteria adalah konsep yang akan ditunjukkan dalam penelitian yang sebenarnya untuk membangun pengetahuan yang baik tentang teori yang berkaitan dengan konsep sedangkan validitas isi melihat isi item apakah benar-benar mengukur konsep yang diukur dalam penelitian. Terakhir adalah validitas konstruk, yang mengukur sejauh mana instrumen secara akurat mengukur konstruk teoritis yang dirancang untuk diukur (Jackson, 2003).

Instrumen tes yang diberikan harus memiliki tingkat validitas dan reliabel agar pembuat tes mendapatkan hubungan tingkat kemampuan kesulitan instrument soal pada peserta didik. Sehingga dapat terukur kemampuan pembelajaran yang sebenarnya dalam proses evaluasi yang dilakukan. Tujuan utama uji reliabilitas instrumen penelitian ialah untuk mengukur konsistensi alat ukur yang digunakan peneliti kuantitatif. Dalam konteks ini, peneliti hendak mengetahui apakah terdapat ketepatan hasil pengukuran pada sampel yang sama dalam waktu yang berbeda. Dengan kata lain, sebuah instrumen penelitian, misalnya kuesioner dinyatakan reliabel jika instrumen tersebut dapat menyediakan hasil skor yang konsisten pada setiap pengukuran (Budiastuti dan Agustinus, 2018).

Untuk mengetahui valid tidaknya suatu instrumen penelitian maka diperlukan analisis statistik. Tes memiliki validitas yang tinggi jika hasilnya sesuai dengan kriteria, dalam arti memiliki kesejajaran antara tes dan kriteria. Terdapat banyak aplikasi yang dapat digunakan untuk melakukan uji validitas, salah satunya adalah *Winsteps*. *Winsteps* adalah perangkat lunak berbasis *Windows* yang membantu komputasi *Rasch model*, terutama di bidang evaluasi pendidikan, survei sikap, dan analisis skala.

Jenis instrumen yang umum atau sering digunakan untuk mengukur kemampuan peserta didik di sekolah adalah instrumen tes. Soal yang digunakan dalam kegiatan evaluasi haruslah memiliki kualitas yang baik, karena soal yang memiliki kualitas baik akan dapat memberikan informasi setepat-tepatnya tentang siswa mana yang sudah atau belum menguasai materi yang telah diajarkan oleh guru.

Instrumen tes dikatakan memiliki kualitas yang baik apabila memiliki validitas dan reliabilitas yang tinggi. Semakin tinggi nilai validitas dan reliabilitas suatu instrumen, maka akan semakin jitu data yang didapat dari suatu penelitian (Hayati dan Lailatussadiyah, 2016) Pernyataan tersebut didukung oleh Wahyuningsih (2015) yang menjelaskan validitas dan reliabilitas merupakan faktor penting dalam penentuan tes yang baik atau tidak.

Penelitian ini membahas mengenai perbandingan hasil validitas instrumen tes menggunakan *Rasch model* dengan *software Winsteps* dengan program SPSS versi 22.0. Permodelan *Rasch (Rasch model)* yang dicetuskan oleh *Dr. Georg Rasch* seorang ahli matematika dari Denmark. Permodelan *Rasch* memberikan pendekatan yang berbeda dalam penggunaan skor atau data mentah ujian pada konteks penilaian pendidikan. Tujuan utamanya adalah menghasilkan suatu skala pengukuran dengan interval yang sama yang nantinya dapat memberikan informasi secara akurat tentang peserta tes maupun kualitas soal yang dikerjakan. Dengan kata lain, analisis pada model *Rasch* akan menghasilkan informasi mengenai karakteristik butir dan siswa yang telah dibentuk menjadi metrik yang sama (Sumintono dan Widhiarso, 2015). Sedangkan SPSS adalah program aplikasi yang memiliki analisis data statistik yang cukup tinggi. SPSS memiliki sistem manajemen data pada lingkungan grafis dengan menggunakan menu-menu deskriptif dan kotak-kotak dialog yang sederhana, sehingga mudah mengoperasikan dan memahaminya. SPSS ini adalah salah satu program aplikasi yang paling banyak diminati dan digunakan oleh para analis dan peneliti untuk mengolah data-data statistik (Machali, 2015).

Meskipun suatu instrumen telah terstandar, tetapi hal itu tidak langsung membuat instrumen tersebut dapat digunakan dimana saja, kapan saja, kepada subjek siapa saja. Instrumen perlu diuji coba kembali setiap kali akan digunakan (Tavakol dan Dennick, 2011).

METODE

Penelitian ini menggunakan data sekunder dari soal UN mata pelajaran kimia Tahun 2012 yang diujicobakan kepada mahasiswa semester I kelas reguler Pendidikan kimia Universitas Negeri Medan. Teknik pengambilan data dilakukan secara purposive sampling dengan mengambil satu kelas sebanyak 40 mahasiswa dengan gender yang berbeda. Data sekunder diperoleh dengan dokumen tertulis dengan lima opsi (a, b, c, d, dan e) sebanyak 40 butir soal. Penelitian ini merupakan penelitian kualitatif dengan metode

deskriptif. Penelitian ini dimaksudkan untuk mencari informasi dan data yang dapat digunakan untuk mendeskripsikan kualitas instrumen tes secara empiris berdasarkan unsur validitas dan reliabel reliabel yang diolah dengan bantuan *Rasch model* dengan *software Winsteps* dan program SPSS versi 22.0.

Tabel 1. Hasil uji validasi SPSS 22,0 dan Rasch model

SPSS 22,0				Rasch model				
Item	Sig.5%	Korelasi	Keterangan	Item	MNSQ	ZSTD	P.MCorr	Keterangan
1	0,121	0,249	Tidak valid	1	1,13	0,77	0,28	Tidak valid
2	0,006	0,424	Valid	2	0,91	-0,44	0,46	Valid
3	0,042	0,323	Valid	3	0,94	-0,18	0,42	Valid
4	0,002	0,480	Valid	4	0,79	-1,04	0,54	Valid
5	0,145	0,235	Tidak valid	5	1,15	0,66	0,32	Tidak valid
6	0,032	0,340	Valid	6	1,05	0,34	0,41	Valid
7	0,000	0,599	Valid	7	0,76	-1,16	0,57	Valid
8	0,001	0,496	Valid	8	0,97	0,00	0,44	Valid
9	0,912	0,018	Tidak valid	9	1,40	1,85	0,12	Tidak valid
10	0,000	0,601	Valid	10	0,76	-1,37	0,59	Valid
11	0,000	0,597	Valid	11	0,73	-1,57	0,63	Valid
12	0,518	0,105	Tidak valid	12	1,24	0,58	0,12	Tidak valid
13	0,023	0,358	Valid	13	1,22	1,14	0,26	Tidak valid
14	0,032	0,340	Valid	14	1,12	0,67	0,34	Tidak valid
15	0,072	0,287	Tidak valid	15	1,11	0,66	0,32	Tidak valid
16	0,008	0,411	Valid	16	0,93	-0,32	0,46	Valid
17	0,103	0,262	Tidak valid	17	1,16	0,81	0,30	Tidak valid
18	0,000	0,535	Valid	18	0,72	-1,65	0,62	Valid
19	0,001	0,497	Valid	19	1,00	0,07	0,49	Valid
20	0,043	0,321	Valid	20	1,17	0,50	0,33	Tidak valid
21	0,012	0,392	Valid	21	0,82	-1,08	0,49	Valid
22	0,000	0,651	Valid	22	0,78	-1,31	0,55	Valid
23	0,001	0,488	Valid	23	0,92	-0,60	0,43	Valid
24	0,044	0,320	Valid	24	0,77	-0,66	0,44	Valid
25	0,013	0,389	Valid	25	0,95	-0,30	0,34	Tidak valid
26	0,101	0,263	Tidak valid	26	1,06	0,27	0,16	Tidak valid
27	0,973	-0,006	Tidak valid	27	1,25	1,75	0,06	Tidak valid
28	0,237	0,191	Tidak valid	28	1,04	0,32	0,29	Tidak valid
29	0,091	0,271	Tidak valid	29	0,77	-1,36	0,55	Valid
30	0,288	0,172	Tidak valid	30	1,00	0,15	0,22	Tidak valid
31	0,002	0,469	Valid	31	0,72	-1,92	0,62	Valid
32	0,340	-0,155	Tidak valid	32	1,20	0,52	0,06	Tidak valid
33	0,245	0,188	Tidak valid	33	1,03	0,21	0,23	Tidak valid
34	0,684	-0,066	Tidak valid	34	2,36	1,99	-0,24	Tidak valid
35	0,080	0,280	Tidak valid	35	0,98	-0,06	0,39	Tidak valid
36	0,892	0,022	Tidak valid	36	1,15	0,54	0,12	Tidak valid
37	0,245	0,188	Tidak valid	37	1,06	0,44	0,29	Tidak valid
38	0,051	0,310	Tidak valid	38	0,98	-0,06	0,30	Tidak valid
39	0,766	0,048	Tidak valid	39	1,38	1,40	0,06	Tidak valid
40	0,545	-0,099	Tidak valid	40	1,37	0,72	-0,01	Tidak valid

HASIL DAN PEMBAHASAN

Uji validitas

Uji validitas merupakan uji yang berfungsi untuk melihat apakah suatu alat ukur tersebut valid atau tidak valid. Uji validitas instrument tes berupa soal objektif menggunakan program SPSS 22.0 dan *Rasch Model* dengan bantuan *software Winstep*. Uji validitas dengan SPSS 22.0 dengan tingkat signifikansi 0,05, soal dikatakan valid jika $r_{hitung} < r_{tabel}$. Berdasarkan yang dikemukakan Sumintono dan Widhiarso (2015) uji validitas dengan *Rasch Model* dengan bantuan *software Winstep* untuk mengetahui soal valid atau tidak dengan memperhatikan kriteria berikut ini:

Nilai *outfit* MNSQ yang diterima : $0,5 < MNSQ < 1,5$

Nilai *outfit* ZSTD yang diterima : $-2,0 < ZSTD < +2,0$

Nilai *Pt Measure Corr* yang diterima : $0,4 < Pt Measure Corr < 0,85$

Setelah dilakukan uji validitas pada butir soal diperoleh hasil seperti yang tertuang pada Tabel 1.

Berdasarkan Tabel 1, diperoleh perbandingan bahwa hasil uji validitas dengan SPSS 22.0 diperoleh 20 item yang valid dan 20 item yang tidak valid. Sedangkan menggunakan *Rasch model* dengan bantuan *software Winstep* diperoleh 17 item yang valid dan 23 item tidak valid.

Hasil uji validitas dengan SPSS 22.0 dengan *Rasch model* dengan bantuan *software Winstep* pada item soal yang sama. Pada item soal 13, 14, 20, dan 25 berdasarkan analisis dengan program SPSS 22.0 dihasilkan keempat item tersebut valid, sedangkan analisis *Rasch model* dengan bantuan *software Winstep* keempat item tersebut diperoleh tidak valid. Pada item soal 29 berdasarkan analisis dengan program SPSS 22.0 dihasilkan item tersebut tidak valid, sedangkan analisis *Rasch model* dengan bantuan *software Winstep* item tersebut diperoleh valid.

Soal yang valid lebih banyak diperoleh dengan menggunakan SPSS 22.0 daripada dengan *Rasch model* disebabkan menggunakan SPSS untuk menentukan soal valid jika hanya ditentukan tingkat signifikannya ($r_{hitung} < r_{tabel}$). Sedangkan, analisis item soal dengan *Rasch model* untuk menentukan soal valid harus memenuhi tiga kriteria yaitu nilai *Outfit* MNSQ, nilai *Outfit* ZSTD, dan nilai *Point Measure Correlation*.

Oleh karena itu, berdasarkan penggunaan aplikasi dalam kegiatan analisis untuk data dengan jumlah yang lebih banyak (misalnya data lebih dari 25 atau data terdapat 40 item) maka analisis yang lebih mudah digunakan adalah dengan menggunakan SPSS dibandingkan dengan *Rasch model* dengan bantuan *software Winstep* data maksimum yang dianalisis sebanyak 25 item dan tidak lebih dari itu. Berdasarkan keakuratan data analisis validasi menggunakan *Rasch model* dengan bantuan *software Winstep* lebih akurat dibandingkan dengan menggunakan program SPSS 22.0.

Uji reliabilitas

Reliabilitas merupakan indeks yang menunjukkan sejauh mana suatu alat pengukur dapat dipercaya atau diandalkan. Sehingga uji reliabilitas dapat digunakan untuk mengetahui konsistensi alat ukur, apakah alat ukur tetap konsisten jika menghasilkan hasil yang sama meskipun dilakukan pengukuran berkali-kali (Widi, 2011). Tinggi rendahnya reliabilitas secara empirik ditunjukkan oleh suatu angka yang disebut koefisien reliabilitas. Semakin tinggi koefisien korelasi antara hasil ukur dua tes yang paralel, maka konsistensi antara keduanya semakin naik dan alat ukur disebut alat ukur yang reliabel (Purnamasari, 2016).

Walaupun terdapat beberapa metode uji reliabilitas, namun pada artikel ini menggunakan uji reliabilitas instrument tes dengan program SPSS 22.0 dengan tingkat signifikan 5% dan *Rasch model* dengan bantuan *software Winstep*. Dari hasil kedua metode tersebut maka diperoleh hasil reliabilitas seperti yang ditunjukkan pada tabel 2.

Uji reliabilitas menggunakan program SPSS 22.0 dengan *Cronbach's Alpha* (KR-20) soal dikatakan reliabel jika $r_{hitung} > 0,05$ (Tabel 2). Artinya analisis reliabilitas soal dengan SPSS 22.0 yang diperoleh 0,828 adalah reliabel dengan kategori tinggi. Sedangkan, uji reliabilitas soal dengan *Rasch model* dengan bantuan *software Winstep* diperoleh 0,69 yang artinya soal reliabel dengan kategori cukup.

Salah satu faktor yang membedakan hasil reliabilitas antara analisis menggunakan SPSS 22.0 dan *Rasch model* dengan bantuan *software Winstep* adalah dari hasil uji validasi menggunakan SPSS 22.0 ada 20 item yang valid dengan lebih banyak 3 item daripada analisis menggunakan *Rasch model* dengan jumlah 17 item yang valid. Sehingga, jika dilakukan uji reliabilitas menggunakan SPSS 22.0 tingkat reliabilitas yang diperoleh lebih besar juga yaitu 0,828 (kategori tinggi) dibandingkan dengan analisis menggunakan *Rasch model* yang diperoleh 0,69 (kategori cukup).

Tabel 2. Perbedaan uji reliabilitas dengan SPSS 22,0 dan *Rasch model*

SPSS 22.0		<i>Rasch model</i>	
Cronbach's alpha	Kategori	Item reliability	Kategori
0,828	Tinggi	0,69	Cukup

SIMPULAN

Perbandingan uji validitas menggunakan SPSS 22.0 diperoleh 20 item yang valid dan 20 item yang tidak valid. Sedangkan menggunakan *Rasch model* dengan bantuan *software Winstep* diperoleh 17 item yang valid dan 23 item tidak valid. Penggunaan SPSS untuk menentukan soal valid ditentukan tingkat signifikannya ($t_{hitung} < t_{tabel}$). Sedangkan, *Rasch model* harus memenuhi tiga kriteria yaitu nilai *Outfit* MNSQ, nilai *Outfit* ZSTD, dan nilai *Point Measure Correlation*. Berdasarkan keakuratan data analisis validasi menggunakan *Rasch model* dengan bantuan *software Winstep* lebih akurat dibandingkan dengan menggunakan program SPSS 22.0. Faktor yang membedakan hasil reliabilitas SPSS 22.0 dan *Rasch model* dengan bantuan *software Winstep* adalah dari hasil uji validasi. Sehingga, uji reliabilitas menggunakan SPSS 22.0 tingkat reliabilitasnya 0,828 (kategori tinggi) dan *Rasch model* yang diperoleh 0,69 (kategori cukup).

DAFTAR PUSTAKA

- Budiastusi, D., dan Agustinus, B. 2018. *Validitas dan Reliabilitas Penelitian*. Jakarta: Mitra Wacana Media.
- Hayati, S., dan Lailatussaadah, L. 2016. Validitas Dan Reliabilitas Instrumen Pengetahuan Pembelajaran Aktif, Kreatif Dan Menyenangkan (Pakem) Menggunakan Model Rasch. *Jurnal Ilmiah Didaktika*, 16(2): 169–179
- Jackson, S. L. 2003. *Research Methods and Statistics, A Critical Thinking Approach*. USA: Thomson Wadsworth.
- Machali, I. 2015. *Statistik Itu Mudah: Menggunakan SPSS Sebagai Alat bantu Statistik*. Yogyakarta : Ladang Kata
- Muijs, D. 2011. *Doing Quantitative Research in Education with SPSS*. London: SAGE Publications Ltd.
- Purnamasari, L. 2016. *Hubungan Antara Makna Kerja dengan Keterlibatan Kerja pada Karyawan di Bandung*. Bandung: Universitas Pendidikan Indonesia
- Ratnawulan, E., dan Rusdiana. 2014. *Evaluasi Pembelajaran*. Bandung: Pustaka Setia.
- Sumintono, B., dan Widhiarso, W. 2015. *Aplikasi Pemodelan Rasch pada Assessment Pendidikan*. Cimahi: Trim Komunikata.
- Susdelina, S., Perdana, S. A., dan Febrian, F. Analisis Perbandingan Teori Tes Klasik dan Rasch Model dalam Evaluasi Instrumen Pengukuran Pemahaman Konsep. Analisis Perbandingan Teori Tes Klasik dan Rasch Model dalam Evaluasi Instrumen Pengukuran Pemahaman Konsep.
- Tavakol, M. dan Dennick, R. 2011. Making Sense Of Cronbach's Alpha. *International Journal of Medical Education*. 2: 53-55
- Tri Wahyuningsih, E. 2015. *Analisis Butir Soal Tes Objektif Buatan Guru Ulangan Semester Ganjil Mata Pelajaran Ekonomi Kelas X di SMA Negeri 1 Melati Tahun Ajaran 2013/2014* (PhD Thesis). Fakultas Ekonomi.
- Widi, R. 2011. Uji Validitas dan Reliabilitas dalam Penelitian Epidemiologi Kedokteran Gigi. *J.K.G Unej*, 8(1):27-34