Ukazh Journal of Arabic Studies
E-ISSN : 2797-1910
Open Acces
Ukazh: Journal of Arabic Studies, Vol. 6, No. 2, 2025 Page 303-316
DOI : 10.37274/ukazh.v6i2.219
http://ejournal.arraayah.ac.id

Analisis Validitas Dan Reliabilitas Instrumen Penilaian Tengah Semester Bahasa Arab Keterampilan Menyimak Muhammadiyah Boarding School Ar Fachrudin Yogyakarta

Alfi Rizqi Rahmawati[1], Moh. Mahbub[2], Siti Desi Hidayati[3]
[1,2,3]Universitas Islam Negeri Raden Mas Said, Surakarta, Indonesia
E-mail: alfirizqi1398@gmail.com1, muhmahbub03@gmail.com2, sitidesihidayati15@gmail.com3

Submission: 21-04-2025	*Revised: 22-05-2025*	*Accepted: 19-06-2025*	*Published: 28-06-2025*

Abstract

A good instrument test will show good data result. Measuring the result of this test uses the Validity and Reliability Test. Validity is measuring the extent to which measurements are carried out accurately, while reliabilty is measuring the extent to which measurements can be trusted because of their consistency. It can be said that Validity and reliability are interrelated. The aims of this research are: 1) to determine the validity of the mid-semester assessment questions for Arabic listening skills at MBS AR Fachrudin Yogyakarta; 2) to determine the reliability of the mid-semester assessment of arabic listening skills in the form of multiple choice question at MBS AR Fachrudin Yogyakarta. This research uses descriptive quantitative method with data collection techniques through mid-tests on the odd semester examination from a research sample of 30 students in class 10 L and 10 M for the 2023/2024 academic year. Validity and reliability analysis was carried out using SPSS 25 version. Based on the result of this study from 10 multiple choice questions items in the mid-semester assessment of Arabic listening skills at MBS AR Fachrudin in the validity test, it was found that 4 questions were classified as valid and 6 questions were invalid. Meanwhile, in the reliability test, the 4 valid questions were stated to have very low reliability.
Keywords: *Validity, Reliability, Mid-Semester Assessment, Arabic Language.*

Abstrak

Instrumen tes yang baik akan menunjukkan hasil data yang baik, pengukuran hasil tes ini menggunakan Uji Validitas dan Reliabilitas. Validitas adalah mengukur sejauh mana ketepatan pengukuran dilakukan, sedangkan reliabilitas ialah mengukur seberapa jauh sebuah konsistensi dalam pengukuran dapat dipercaya. Dapat dikatakan validitas dan reliabilitas saling berkaitan. Tujuan dari penelitian ini adalah: 1) untuk mengetahui Validitas pada butir soal penilaian tengah semester mata pelajaran Bahasa Arab keterampilan menyimak di MBS AR Fachrudin Yogyakarta; 2) untuk mengetahui reliabilitas penilaian tengah semester mata pelajaran Bahasa Arab keterampilan menyimak dalam bentuk soal pilihan ganda di MBS AR Fachrudin Yogyakarta. Peneliti dalam penelitian ini menggunakan metode penelitian kuantitatif deskriptif dengan teknik pengumpulan data melalui tes pada Ujian Tengah Semester Gasal dari sampel penelitian 30 orang siswa kelas 10 L dan 10 M tahun ajaran 2023/2024. Analisis validitas dan reliabilitas dilakukan menggunakan SPSS versi 25. Hasil penelitian dari 10 butir soal pilihan ganda pada penilaian tengah semester bahasa Arab keterampilan menyimak di MBS AR Fachrudin dalam uji validitas ditemukan 4 soal tergolong valid dan 6 soal tidak valid. Sedangkan pada uji reliabilitas dari 4 soal yang valid dinyatakan memiliki reliabilitas yang sangat rendah.
Kata Kunci: Validitas, Reliabilitas, Penilaian Tengah Semester, MBS Ar Fachrudin, Bahasa Arab.

© 2025 by the authors; This is an Open Access article distributed under the terms of the Creative Commons Attribution- ShareAlike 4.0 International License. (https://creativecommons.org/licenses/by-sa/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

A. PENDAHULUAN

Pendidikan didefinisikan dalam undang-undang republik indonesia tahun 2001 bahwa "pendidikan adalah suatu usaha yang secara sadar dan terencana guna mewujudkan suasana belajar dan proses pembelajaran supaya peserta didik secara aktif mengembangkan potensi dalam dirinya dan memiliki kekuatan spritual keagamaan, pengendalian diri, kepribadian, kecerdasan, akhlak mulia, serta keterampilan yang diperlukan untuk dirinya, masyarakat, bangsa dan negara" (Rulam, M. Pd, 2014). Dari definisi diatas dapat dipahami pendidikan itu suatu hubungan antara manusia dengan lingkungan sekitarnya bertujuan untuk mengembangkan segala potensi dari sisi jasmani dan juga ruhani yang dilakukan dengan kesadaran penuh secara bertahap dan terencana. Kemudian dari proses ini akan menghasilkan perubahan positif yang berkemajuan dari segi kognitif, afektif dan psikomotorik sehingga mampu menuntun seseorang untuk mencapai tujuan hidupnya. Maka pendidikan bisa dilihat sebagai proses dan tujuan, yaitu proses interaksi manusia dengan lingkungan sekitarnya dan hasil dari interaksi tersebut (Rulam, M. Pd, 2014).

Dalam dunia pendidikan ada dua entitas penting yang menjaga keberlangsungan pendidikan yaitu siswa dan guru. Guru sebagai subjek pembelajaran dan siswa menjadi objek yang diarahkan pembelajarannya, keduanya saling memengaruhi satu sama lain. Guru sebagai subjek pendidikan setidaknya menguasai beberapa hal: memahami prinsip dan wawasan pedagogis; memahami kondisi siswa; pengembangan kurikulum; perencanaan pembelajaran; pelaksanaan pembelajaran dengan dialogis dan pedagogis; penggunaan teknologi dalam pembelajaran; Evaluasi Hasil belajar (EHB); serta membantu siswa guna mencapai berbagai potensi mereka (Maghfirotunnisa et al., 2023).

Pada teori penyusunan dan perencanaan pembelajaran, terdapat tiga komponen pokok yang menggambarkan proses pembelajaran, yaitu: 1) perencanaan pembelajaran; 2) pelaksanaan pembelajaran; 3) penilaian (evaluasi). Ketiga komponen tersebut memiliki ikatan sebab akibat baik secara tidak langsung maupun langsung. Dengan adanya penilaian dapat diketahui sampai sejauh mana keberhasilan pembelajaran dan belajar dalam batas waktu tertentu. Walaupun penilaian hanya sebagian kecil dari rangkaian pembelajaran namun memiliki dampak yang besar untuk kualitas pembelajaran selanjutnya (Imawan et al., 2023). Penilaian menjadi tolok ukur untuk

mengetahui hasil dari proses pembelajaran apakah sesuai dengan tujuan dan kriteria yang telah ditetapkan. selain evaluasi, penilaian juga berkaitan dengan pengukuran. Pengukuran memberikan data dari aspek kuantitatif bagi penilaian sedangkan dari aspek kualitatif penilaian berupa interpretasi dan pertimbangan dari data kuantitatif tersebut. Penilaian dilakukan dengan menggunakan tes ataupun non-tes. Berdasarkan jenis penilaian tersebut, tes dibagi menjadi beberapa macam dua diantaranya yaitu Ujian Tengah Semester dan Ujian Akhir Semester yang kemudian kita kenal tes sumatif. Tes ini dilaksanakan ketika seluruh rangkaian kegiatan pembelajaran selesai (Handani, 2014).

Instrumen tes yang digunakan berpengaruh pada hasil olahan data, jika instrumen tes baik maka data yang dihasilkan juga baik dengan begitu guru dapat mengetahui kemampuan siswa dengan tepat. Arikunto (2016) berpendapat bahwa instrumen tes yang baik memiliki lima ketentuan: 1) validitas; 2) reliabilitas; 3) objektivitas; 4) praktikalitas dan 5) ekonomis (Dewi et al., 2020). Namun, sebagian besar peneliti hanya menggunakan validitas dan reliabilitas untuk mengukur baik tidaknya suatu instrumen tes. Validitas adalah mengukur sejauh mana ketepatan pengukuran dilakukan, sedangkan reliabilitas adalah mengukur seberapa jauh suatu konsistensi dari pengukuran dapat dipercaya. Dapat dikatakan validitas dan reliabilitas saling berkaitan (Yusup, 2018).

Validitas

Validitas atau keshahihan ialah suatu tes yang mengukur satu fungsi yang seharusnya diukur tanpa terdistraksi oleh fungsi lain yang hampir sama (Ambiya et al., 2022). Kata validitas berasal dari kata dalam bahasa Inggris validity yang berarti ketepatan, kecermatan dan dalam bahasa Arab berarti صحيح (shahih). Di dalam Bahasa Indonesia sendiri validitas sering disebut dengan valid atau sahih.

Scarvia B. Anderson dan kawan-kawan menulis sebuah buku berjudul Encyclopedia of Education Evaluation yang menjelaskan tentang validitas secara umum dibagi menjadi dua; validitas logis dan validitas empiris. Validitas Logis terdiri dari tiga macam; 1) validitas isi, 2) validitas muka dan 3) validitas konstruk. Sedangkan validitas empiris terdiri dari dua macam yaitu; validitas ramal dan validitas banding.

Validitas logis mempunyai sebutan lain yaitu validitas internal. Validitas logis adalah alat penilaian yang menilai suatu instrumen berdasarkan pada standar logika dan

hasil yang beralasan (Maghfirotunnisa et al., 2023). (1) Validitas isi, diperoleh setelah tes dilaksanakan dan dilihat sejauh mana tes ini mencerminkan kemampuan siswa pada tujuan pembelajaran yang ingin dicapai. Validitas ini sering disebut juga dengan validitas kurikulum karena tes dikatakan valid apabila telah merepresentasikan seluruh materi yang ada dalam kurikulum (Syaifudin, 2020). (2) Validitas Muka/Tampilan Validitas model ini ditinjau dari apa yang terlihat. Veliditas ini berhubungan dengan bagaimana seorang awam memberi tanggapan tentang alat ukur tersebut. Hal ini membuat validitas muka tidak terlalu ilmiah karna berdasar pada kebiasaan, contohnya format dalam menyusun pilihan jawaban pada soal pilihan ganda (Ida & Musyarofah, 2021). (3) Validitas Konstruksi, Validitas jenis ini memerlukan standar kompetensi, kompetensi dasar, indikator dan tujuan pembelajaran untuk mengukur soal-soal apakah telah memenuhi aspek dan definisi kerangka berpikir (Syaifudin, 2020). Jika pertanyaan dan pernyataan tiap item soal sesuai, oleh karenanya soal-soal tersebut dinyatakan valid secara validitas konstruk.

Reliabilitas

Sedangkan reliabilitas adalah kata serapan dari bahasa Inggris reliability/reliable, yang berarti dapat dipercaya sedangkan dalam bahasa Arab disebut ثبت (Tsabit) yang berarti sesuatu yang tetap, pasti dan tidak berubah. Jika alat ukur memberikan hasil yang konstan, ajeg maka alat ukur ini dapat dikatakan reliabel (Fikriyah & Muhammadiyah, 2021).

Jenis-Jenis Reliabilitas; (1) Test-Retest (Uji-Uji Ulang Reliabilitas) digunakan untuk pengujian reliabilitas ketika menentukan sejauh mana konsistensi hasil tes setelah diujikan beberapa kali dalam kurun waktu berbeda namun pada subjek yang sama dan soal tes yang sama. Instrumen soal disebut reliabel apabila koefisien korelasi positif dan signifikan. (2) Equivalent (Reliabilitas Ekuivalen) pada realibilitas ini dilakukan uji percobaan satu kali dengan dua buah instrumen yang berbeda tetapi sepadan (ekuivalen) dan diujikan pada subjek yang sama. Hasil dari uji instrumen ini apabila koefisien korelasi positif dan signifikan maka dikatakan reliabel. (3) Internal Consystency (Keandalan Internal) jenis reliabilitas ini disebut dengan konsistensi internal karena didasari oleh konsistensi setiap item tes benchmark. Pada pengujian ini dapat selesaikan dengan

teknik belah dua (split half) dari Spearman Brown, KR 20, KR 21 atau dengan teknik Alfa Cronbach.

Tes/Ujian

Tes adalah salah satu hasil kerja siswa yang berbeda-beda dari setiap masing-masing siswa secara terstruktur untuk dianalisis dan menghasilkan nilai numerik dari proses pengujian. Didalamnya tes memuat serangkaian tugas berupa pertanyaan-pertanyaan yang membutuhkan jawaban dari siswa untuk melihat kemampuan dan kecakapan masing-masing (Febriana, 2019).

Pada umumnya tes dibagi menjadi dua kategori: tes objektif dan tes subjektif. (1) Tes Objektif. Tes ini dinilai secara objektif tanpa ada unsur subjektivitas. Bentuk dari tes objektif lebih beragam yaitu: tes multiple choice (pilihan ganda), tes true-false (benar salah), matching (menjodohkan) (Yusrizal, 2016). (2) Tes Subjektif adalah butir-butir soal dalam tes yang banyak dipengaruhi dengan unsur-unsur subjektif. Pertanyaan-pertanyaannya biasanya diawali dengan terbuka seperti: mengapa, apa, bagaimana, siapa dan sebagainya. Bentuk dari tes subjektif berupa essay atau Uraian (Setiawan, 2018).

Bahasa Arab

Bahasa Arab memiliki kaidah yang lengkap dan keunggulan dibandingkan Bahasa asing lainnya, kedua hal ini membuat Bahasa Arab unik. Diantara ciri keunikannya adalah mampu menjadikan konsep yang rumit menjadi sederhana, yang buruk menjadi bagus, yang buram menjadi jelas dan perubahan-perubahan lain yang serupa. Dalam tes kebahasaan dapat diterapkan beberapa jenis tes untuk mengukur kemampuan berbahasa siswa diantaranya adalah tes kemampuan menyimak, berbicara, membaca dan menulis.

Keterampilan Menyimak (Maharah Istima')

Keterampilan menyimak adalah proses reseptif informasi/berita lewat perantara indera pendengar berupa lambang bunyi yang diucapkan lisan. Dalam kegiatan ini perlu adanya kemampuan mengetahui sistem bunyi Bahasa yang digunakan (Imawan et al., 2023). Pada tes keterampilan ini siswa diharapkan dapat megenali suara-suara dan

perbedaannya satu sama lain, memahami unsur bunyi, dan akrab dengan bunyi-bunyi Bahasa asing khususnya Bahasa Arab yang para siswa dengarkan (Ambiya et al., 2022).

Pada penelitian terdahulu tidak sedikit yang memilih fokus utamanya tentang evaluasi pembelajaran, lebih rinci lagi mengenai Validitas dan reliabilitas soal. Seperti penelitian yang diteliti oleh Zuhri, dkk (2024) mengenai validitas, reliabilitas, dan tingkat kesukaran soal Bahasa Arab tingkat SMP berbasis Artificial Intelligence (AI) melalui Platform Question Well, penelitian ini bertujuan untuk menguji validitas, reliabilitas dan tingkat kesukaran Soal yang dibuat menggunakan kecerdasan buatan (AI) dengan platform Question Well. Metode penelitian yang dipilih adalah deskriptif kuantitatif dan menghasilkan kesimpulan bahwa pembuatan soal dengan kecerdasan buatan bukan instrumen akhir melainkan sebagai sarana kolaborasi dengan guru untuk memanfaatkan peran AI dalam dunia pendidikan (Zuhri et al., 2024). Selain itu, penelitian tentang penilaian ideal dan evaluasi efektif dalam pembelajaran Bahasa Arab yang dikaji oleh Mahmudah dan Susilo, bertujuan untuk melihat prinsip-prinsip penilaian yang adil dan komprehensif melalui validitas, reliabilitas, keadilan, transparansi dan keterkaitan dengan tujuan pembelajaran dalam evaluasi akhir pembelajaran seperti penilaian formatif, sumatif dan autentik (Mahmudah & Susilo, 2024).

Adapun penelitian yang berkaitan dengan keterampilan menyimak dalam evaluasi pembelajaran ialah penelitian yang dilakukan oleh Suci Kirani, dkk (2024) tentang analisis instrumen evaluasi keterampilan menyimak Bahasa Arab pada platform Alef yang bertujuan untuk menganalisis instrumen evaluasi keterampilan menyimak Bahasa Arab yang ada pada platform Alef. Dengan metode penelitian deskriptif kualitatif dan pendekatan analisis konten, menyatakan bahwa penggunaan aplikasi dalam pembelajaran bahasa direkomendasikan untuk mengembangkan instrumen evaluasi sehingga lebih komprehensif dengan mengoptimalkan teknologi. Penelitian ini masih sangat relevan jika dilihat berdasarkan pada penelitian terdahulu (Kirani Adhistie t al., 2024).

Berdasarkan uraian diatas, berbeda dengan penelitian-penelitian yang ada sebelumnya maka artikel ini akan berfokus dan pada pengujian validitas dan reliabilitas dalam butir soal yang digunakan sebagai alat ukur kemampuan siswa pada penilaian tengah semester di MBS AR Fachrudin Yogyakarta.

B. METODOLOGI PENELITIAN

Peneliti menggunakan penelitian deskriptif kuantitatif. Menurut Yusuf, penelitian deskriptif kuantitatif merupakan usaha yang secara sadar juga sistematis guna memberikan jawaban dari sebuah masalah atau mendapatkan informasi yang spesifik sebuah fenomena dengan tahapan-tahapan penelitian pendekatan kuantitatif. Menurut Isaac dan Michael dalam Yusuf, penelitian deskriptif memiliki tujuan yaitu "to describe systematically the facts and characteristics of a given population or area of interest" (Yusuf, 2014).

Adapun populasi yang diujikan dalam penelitian ini adalah siswa-siswi kelas 10 MBS AR Fachrudin Yogyakarta tahun ajaran 2023/2024 sedangkan sampel penelitian ini adalah dua kelas yaitu siswa-siswi kelas 10(L) dan 10(M) dengan jumlah 30 siswa. Pengumpulan data pada penelitian ini dilakukan dengan tes dari hasil Penilaian Tengah Semester. Setelah mendapatkan keseluruhan sumber data atau responden, langkah selanjutnya adalah analisis data kuantitatif. Untuk menjawab rumusan masalah, tindakan-tindakannya yaitu mengelompokkan data menurut variabel dan jenis responden, tabulasi data, melihat data dan melakukan perhitungan (Noor, 2017). Butir-butir instrumen tes yang diuji dari segi validitas dan reliabilitas dievaluasi secara kuantitatif adalah penilaian akhir semester pengembangan Bahasa Arab menggunakan SPSS versi 29.1

C. HASIL DAN PEMBAHASAN

A. Validitas

Dalam menentukan validitas dan reliabilitas yang diinginkan maka peneliti menggunakan bantuan SPSS versi 25. Proses pengujian yang pertama dilakukan yaitu untuk mengetahui validitas butir soal berbentuk pilihan ganda kemudian dijabarkan hasilnya dengan ketentuan sebagai berikut:

1. Apabila r-hitung > r-tabel maka item yang diuji valid.
2. Apabila r-hitung < r-tabel maka item yang diuji tidak valid.

Setelah melakukan analisis melalui SPSS versi 25 menggunakan rumus pearson correlation dengan memasukkan sampel jawaban siswa sebanyak 30 orang dengan butir soal sebanyak 10 buah, maka didapatkan data sebagai berikut:

Correlations

Tabel 1. Hasil Uji menggunakan SPSS

		Soal 1	Soal 2	Soal 3	Soal 4	Soal 5	Soal 6	Soal 7	Soal 8	Soal 9	Soal 10	SoalTotal
Soal 1	Pearson Correlation	1	-.034	-.174	.[a]	.[a]	.308	-.050	-.073	.244	.[a]	.289
	Sig. (2-tailed)		.856	.359	.	.	.098	.795	.702	.194	.	.121
	N	30	30	30	30	30	30	30	30	30	30	30
Soal 2	Pearson Correlation	-.034	1	.199	.[a]	.[a]	.308	-.050	-.073	-.141	.[a]	.289
	Sig. (2-tailed)	.856		.293	.	.	.098	.795	.702	.456	.	.121
	N	30	30	30	30	30	30	30	30	30	30	30
Soal 3	Pearson Correlation	-.174	.199	1	.[a]	.[a]	.040	-.250	.026	-.157	.[a]	.374*
	Sig. (2-tailed)	.359	.293		.	.	.833	.183	.891	.407	.	.042
	N	30	30	30	30	30	30	30	30	30	30	30
Soal 4	Pearson Correlation	.[a]	.[a]	.[a]	.[a]	.[a]	.[a]	.[a]	.[a]	.[a]	.[a]	.[a]
	Sig. (2-tailed)	.	.	.		.	.	.	.	.	.	.
	N	30	30	30	30	30	30	30	30	30	30	30
Soal 5	Pearson Correlation	.[a]	.[a]	.[a]	.[a]	.[a]	.[a]	.[a]	.[a]	.[a]	.[a]	.[a]
	Sig. (2-tailed)	.	.	.	.		.	.	.	.	.	.
	N	30	30	30	30	30	30	30	30	30	30	30

Soal 6	Pearson Correlation	.308	.308	.040	.[a]	.[a]	1	.141	-.015	.167	.[a]	.652**
	Sig. (2-tailed)	.098	.098	.833	.	.		.457	.938	.378	.	<.001
	N	30	30	30	30	30	30	30	30	30	30	30
Soal 7	Pearson Correlation	-.050	-.050	-.250	.[a]	.[a]	.141	1	.288	.351	.[a]	.416*
	Sig. (2-tailed)	.795	.795	.183	.	.	.457		.122	.057	.	.022
	N	30	30	30	30	30	30	30	30	30	30	30
Soal 8	Pearson Correlation	-.073	-.073	.026	.[a]	.[a]	-.015	.288	1	-.095	.[a]	.331
	Sig. (2-tailed)	.702	.702	.891	.	.	.938	.122		.618	.	.074
	N	30	30	30	30	30	30	30	30	30	30	30
Soal 9	Pearson Correlation	.244	-.141	-.157	.[a]	.[a]	.167	.351	-.095	1	.[a]	.526**
	Sig. (2-tailed)	.194	.456	.407	.	.	.378	.057	.618		.	.003
	N	30	30	30	30	30	30	30	30	30	30	30
Soal 10	Pearson Correlation	.[a]	.[a]	.[a]	.[a]	.[a]	.[a]	.[a]	.[a]	.[a]	.[a]	.[a]
	Sig. (2-tailed)	.	.	.	.	.	.	.	.	.		.
	N	30	30	30	30	30	30	30	30	30	30	30
Soal Total	Pearson Correlation	.289	.289	.374*	.[a]	.[a]	.652**	.416*	.331	.526**	.[a]	1
	Sig. (2-tailed)	.121	.121	.042	.	.	<.001	.022	.074	.003	.	
	N	30	30	30	30	30	30	30	30	30	30	30

*. Correlation is significant at the 0.05 level (2-tailed).
**. Correlation is significant at the 0.01 level (2-tailed).
a. Cannot be computed because at least one of the variables is constant.

Adapun melalui tabel diatas bisa diketahui nilai total keseluruhan nomor soal sebagai bentuk analisis validitas butir soal yang dilakukan SPSS versi 25. Sesuai dengan acuan pengukuran untuk interpretasi hasil uji validitas butir soal, maka peneliti menguraikan pada tabel dibawah ini:

Tabel 1. Kesimpulan Hasil Uji SPSS versi 25

BUTIR SOAL MULTIPLE CHOICE	**HASIL UJI**		**INTERPRETASI**
	r-hitung	**r-tabel (N=30)**	
1	0,289	0,361	Tidak Valid
2	0,289	0,361	Tidak Valid
3	0,374	0,361	Valid
4	0,0	0,361	Tidak Valid
5	0,0	0,361	Tidak Valid
6	0,652	0,361	Valid
7	0,416	0,361	Valid
8	0,331	0,361	Tidak Valid
9	0,526	0,361	Valid
10	0,0	0,361	Tidak Valid

Berdasarkan pada tabel diatas, hasil interpretasi pada setiap butir soal tes pilihan ganda dengan taraf signifikan 0,05 maka dapat diklasifikasikan seperti berikut ini:

a. Butir soal tes yang termasuk valid sebanyak 4 butir soal secara keseluruhan.
b. Butir soal tes yang termasuk tidak valid sebanyak 6 butir soal secara keseluruhan.

Reliabilitas

Setelah melakukan pengujian pertama pada validitas, selanjutnya peneliti melakukan pengujian reliabilitas guna mengetahui reliabel atau tidak pada seluruh butir soal tes Penilaian Tengah Semester yang telah teruji valid. Pada uji reliabilitas interpretasi sebagai patokannya ditentukan sebagai berikut:

Tabel 13. Interpretasi Reliabilitas

Koefisien Korelasi	Kriteria Reliabilitas
$0{,}81 < r \leq 1{,}00$	Sangat Tinggi
$0{,}61 < r \leq 0{,}80$	Tinggi
$0{,}41 < r \leq 0{,}60$	Cukup
$0{,}21 < r \leq 0{,}40$	Rendah
$0{,}00 < r \leq 0{,}21$	Sanggat Rendah

Sumber: Google foto

Pada applikasi SPSS versi 25, untuk mengetahui reliabilitas butir soal tes dapat dilihat didalam tabel berjudul Case Processing Summary dan Reliability Statistic sebagai berikut:

Case Processing Summary

		N	%
Cases	Valid	30	100.0
	Excluded[a]	0	.0
	Total	30	100.0

a. Listwise deletion based on all variables in the procedure.

Reliability Statistics

Cronbach's Alpha	N of Items
.110	4

Dari tabel pertama Case Processing Summary menyatakan bahwa sampel yang berjumlah 30 orang adalah valid secara keseluruhan dengan persentase 100% dan tidak ada data sampel yang masuk dalam klasifikasi Excluded. Selanjutnya, dari tabel kedua Reliability Statistics dapat disimpulkan dengan interpretasi butir soal tes pilihan ganda yang telah teruji valid sebanyak 4 item berdasarkan Cronbach's Alpha bahwa hasil penilaian tengah semester Bahasa Arab adalah tidak reliabel atau memiliki tingkat reliabilitas yang sangat rendah karena nilai koefisien korelasinya hanya 0,110. Maka dapat dinyatakan butir soal penilaian tengah semester pada MBS AR Fachrudin Yogyakarta mempunyai reliabilitas yang sangat rendah.

D. SIMPULAN

Setelah melakukan analisis pada butir-butir soal tes Penilaian Tengah Semester (PTS) mata pelajaran Bahasa Arab kelas 10 L dan 10 M pada keterampilan menyimak (maharah istima'), maka peneliti dapat menyimpulkan bahwa analisis yang ditinjau dari segi validitas butir soal mata pelajaran Bahasa Arab keterampilan menyimak pada penilaian tengah semester secara kuantitatif menggunakan SPSS versi 25, hasil uji validitas dari 10 butir soal pilihan ganda adalah sebanyak 4 item soal memiliki validitas dengan taraf signifikan diatas 0,05 dan 6 item soal tidak valid dengan taraf signifikan dibawah 0,05.

Analisis yang ditinjau dari segi reliabilitas butir soal mata pelajaran Bahasa Arab keterampilan menyimak pada penilaian tengah semester secara kuantitatif menggunakan SPSS versi 25, hasil perolehan dari uji reliabilitas 4 butir soal yang terindeks valid memiliki nilai Cronbach's Alpha sebesar 0,110 sehingga dikatakan mempunyai reliabilitas yang sangat rendah.

Berdasarkan hasil tersebut maka saran untuk guru mata pelajaran Bahasa Arab khususnya pada keterampilan menyimak di MBS AR Fachrudin Yogyakarta yang menyusun butir soal tes penilaian tengah semester untuk memperhatikan dalam pembuatan soal dari segi kesukaran, keterbaruan dari soal-soal yang sudah ada sebelumnya sehingga dapat memperoleh butir-butir soal yang lebih valid dan lebih baik tingkat reliabilitasnya.

DAFTAR PUSTAKA

Ambiya, I. Z., Mulyawan, S., & Saefuloh, H. (2022). Analisis Soal Ujian Mata Pelajaran Bahasa Arab Kelas 12 di Madrasah Aliyyah. *El-Ibtikar*, *11*(1), 70–87.

Dewi, N. P., Rahmi, Y. L., Alberida, H., & Darussyamsu, R. (2020). Validitas dan Reliabilitas Instrumen Penilaian Kemampuan Berpikir Tingkat Tinggi ten-tang Materi Hereditas untuk Peserta Didik SMA/MA. *Jurnal Eksakta Pendidikan (Jep)*, *4*(2), 138. https://doi.org/10.24036/jep/vol4-iss2/512

Febriana, R. (2019). *Evaluasi Pembelajaran* (B. S. Fatmawati, Ed.; ed. 1). PT. Bumi Aksara.

Fikriyah, N., & Muhammadiyah, U. (2021). *Analisis Butir Soal Ulangan Tengah Semester Mata Pelajaran Bahasa Arab Kelas VII Semester Ganjil SMP Muhammadiyah 1 Yogyakarta Tahun Ajaran 2019 / 2020*. *3*(2), 128–140. https://doi.org/10.18196/mht.v3i2.10501

Handani, H. A. T. (2014). VALIDITAS DAN RELIABILITAS SOAL TENGAH SEMESTER GENAP KAITANNYA DENGAN KETERCAPAIAN TUJUAN PEMBELAJARAN BAHASA INDONESIA KELAS VIII A SMP NEGERI 2 BANYUDONO TAHUN PELAJARAN 2013/2014. *Publikasi Ums.Ac.Id*, *3*(2), 1–46. http://journal.stainkudus.ac.id/index.php/equilibrium/article/view/1268/1127

Ida, F. F., & Musyarofah, A. (2021). Validitas dan Reliabilitas dalam Analisis Butir Soal. *Al-Mu'Arrib: Journal of Arabic Education*, *1*(1), 34–44. https://doi.org/10.32923/al-muarrib.v1i1.2100

Imawan, Y., Baroroh, R. U., & Hikmatillah, Q. K. (2023). Inovasi Bentuk Penilaian Tes Kompetensi Reseptif Berbahasa Arab dalam Buku Al-Arabiyah Linnasyi'in Jilid 1. *Tsaqofiya*, *5*(1), 19–36. https://doi.org/10.21154/tsaqofiya.v5i1.177

Kirani Adhistia, S., & Islam Negeri Sutan Syarif Kasim Riau, U. (2024). Analisis Instrumen Evaluasi Keterampilan Menyimak Bahasa Arab pada Platform Alef. In *MUMTAZA : Journal of Arabic Teaching, Linguistic And Literature* (Vol. 4).

Maghfirotunnisa, A., Qomari, N., & Asbarin. (2023). pelajaran validitas isi dan reliabilitas instrumen penilaian harian mata pelajaran bahasa arab MI Manarul Islam Malang. *Shaut Al-'Arabiyah*, *11*(1), 19–28. https://doi.org/10.24252/saa.v11i1.34368

Mahmudah, & Susilo, Agus. (2024). Penilaian Ideal dan Evaluasi Efektif dalam Pembelajaran Bahasa Arab. *The 2nd Internasional Conference on Islamic Civilization and Humanities* Faculty of Adab and Humanity UIN Sunan Ampel Surabaya 28-29 May 2024. 787-800.

Noor, J. (2017). *Metodologi Penelitian: Skripsi, Tesis, Disertasi, dan Karya Ilmiah* (ed. 7). Kencana.

Rulam, M. Pd, Dr. Drs. A. (2014). *PENGANTAR PENDIDIKAN: Asas & Filsafat Pendidikan* (K. Rose, Ed.; cetakan 1). AR-RUZZ MEDIA.

Setiawan, D. F. (2018). *Prosedur Evaluasi dalam Pembelajaran* (ed. 1). Deepublish.

Syaifudin. (2020). Validitas dan Reliabilitas Instrumen Penilaian Pada Mata Pelajaran Bahasa Arab. *Cross-Border: Jurnal Kajian Perbatasan Antarnegara, Diplomasi Dan Hubungan Internasional*, *3*(2), 106–118.

Yusrizal, Prof. Dr. (2016). *Tanya Jawab Seputar Pengukuran, Penilaian dan Evaluasi Pendidikan* (Dr. Ramli, Ed.; ed. 1). Syiah Kuala University Press.

Yusuf, A. M. (2014). *Metode Penelitian: Kuantitatif, Kualitatif, dan Penelitian Gabungan* (ed. 1). Kencana.

Yusup, F. (2018). UJI VALIDITAS DAN RELIABILITAS INSTRUMEN PENELITIAN KUANTITATIF. *Jurnal Tarbiyah: Jurnal Ilmiah Kependidikan*, *7*(1), 17–23. https://doi.org/https://doi.org/10.18592/tarbiyah.v7i1.2100

Zuhri, N. Z., Syihabuddin, S., & Tatang, T. (2024). Analisis Validitas, Reliabilitas, dan Tingkat Kesukaran Soal Bahasa Arab Tingkat SMP Berbasis Artificial Intelligence (AI) melalui Platform QuestionWell. *Jurnal Pendidikan Dan Pembelajaran Indonesia (JPPI)*, *4*(2), 693–704. https://doi.org/10.53299/jppi.v4i2.576