



Edukatif: Jurnal Ilmu Pendidikan

Volume 6 Nomor 1 Februari 2024 Halaman 965 - 977

<https://edukatif.org/index.php/edukatif/index>

Analisis Validitas dan Reliabilitas Butir Soal Sumatif Akhir Semester Ganjil Mata Pelajaran PAI

Muhammad Isa Anshari^{1✉}, Rodiah Nasution², Muhammad Irsyad³, Alifia Zuhriatul Alifa⁴, Indah Aminatus Zuhriyah⁵

Universitas Islam Negeri Maulana Malik Ibrahim Malang, Indonesia^{1,2,3,4,5}

e-mail : anshariaan14@gmail.com¹, rodiahnasution01@gmail.com², irsyadmuhammm@gmail.com³, alifiazuhrialifa01@gmail.com⁴, zuhriyah@pgmi.uin-malang.ac.id⁵

Abstrak

Penilaian soal berbentuk Sumatif merupakan salah satu jalan dalam tahap evaluasi pada ranah pendidikan, Adapun soal tersebut dikatakan memenuhi syarat jika memenuhi unsur validitas dan reliabilitas. Tujuan dari penelitian ini adalah mengetahui validitas dan reliabilitas butir soal, dengan harapan guru dapat mengidentifikasi butir soal yang kurang valid dan reliabel, serta merancang perbaikan yang sesuai. Metode dalam penelitian ini yaitu dokumentasi dan wawancara dengan guru bidang studi PAI untuk mendapatkan data yang diperlukan dalam penelitian dan datanya berasal dari 50 butir soal sumatif kelas X di SMA Muhammadiyah Tanah Grogot. Dalam rangka menguji tingkat validitas dan reliabilitas menggunakan SPSS maka dapat diperoleh hasil bahwa validitas soal sumatif akhir semester untuk kelas X dapat diketahui 36 butir soal dinyatakan valid sedangkan 14 butir soal dinyatakan tidak valid. Sedangkan tingkat reliabilitas untuk kelas X memiliki nilai reliabilitas sebesar 0,855 sehingga instrumen soal dinyatakan reliabel. Berdasarkan hasil analisis validitas dan reliabilitas butir soal ujian sumatif akhir semester ganjil mata pelajaran PAI tahun ajaran 2022/2023 SMA Muhammadiyah Tanah Grogot, maka dapat disimpulkan berdasarkan uji validitas dan reliabilitas bahwa soal pilihan ganda ujian sumatif akhir semester (UAS) untuk kelas X dinyatakan valid dan reliabel, artinya instrumen soal ini baik digunakan sebagai alat evaluasi.

Kata Kunci: Validitas, Reliabilitas, Soal Sumatif.

Abstract

Summative assessment of questions is one way in the evaluation stage in the educational realm. The questions are said to meet the requirements if they meet the elements of validity and reliability. The aim of this research is to determine the validity and reliability of the question items, with the hope that teachers can identify question items that are less valid and reliable, and design appropriate improvements. The method in this research is documentation and interviews with teachers in the PAI field of study to obtain the data needed for the research and the data comes from 50 class X summative questions at SMA Muhammadiyah Tanah Grogot. In order to test the level of validity and reliability using SPSS, the results can be obtained that the validity of the final semester summative questions for class X can be seen that 36 questions were declared valid while 14 questions were declared invalid. Meanwhile, the reliability level for class X has a reliability value of 0.855 so that the question instrument is declared reliable. Based on the results of the validity and reliability analysis of the final odd semester summative exam questions for PAI subjects for the 2022/2023 academic year at SMA Muhammadiyah Tanah Grogot, it can be concluded based on the validity and reliability tests that the multiple choice questions for the final semester summative exam (UAS) for class X are declared valid and reliable, meaning that this question instrument is good to use as an evaluation tool.

Keywords: Validity, Reliability, Summative Questions.

Copyright (c) 2024 Muhammad Isa Anshari, Rodiah Nasution, Muhammad Irsyad, Alifia Zuhriatul Alifa, Indah Aminatus Zuhriyah

✉ Corresponding author :

Email : anshariaan14@gmail.com

DOI : <https://doi.org/10.31004/edukatif.v6i1.5931>

ISSN 2656-8063 (Media Cetak)

ISSN 2656-8071 (Media Online)

PENDAHULUAN

Salah satu unsur yang sangat vital dalam proses kependidikan adalah evaluasi pembelajaran. Unsur evaluasi dalam pendidikan muncul karena terdapat *goals* yang ingin diraih dalam proses pendidikan itu sendiri, jadi evaluasi hadir sebagai tolak ukur sampai mana langkah menuju *goals* tersebut telah diaruhi. Maka dari itu puncak kesuksesan dalam kegiatan kependidikan sangat diperlukan adanya unsur evaluasi (Realita, 2015). Tetapi pada praktek *real life* dalam kegiatan kependidikan tingkat keakuratan atau ketepatan dalam mengukur kurang maksimal, salah satunya dikarenakan kelemahan pada bagaian instrumen tes. Padahal dengan semakin tepat dan akurat dalam mengevaluasi tentunya akan berpengaruh banyak terhadap urusan yang diambil dalam kegiatan kependidikan tersebut. Salah satu jawaban dalam permasalahan tersebut adalah adanya *standardized test*. Gagasan *standardized test* diharapkan agar tes tersebut memenuhi unsur validitas dan reliabilitas. Di Indonesia gagasan tersebut diimplementasikan dalam ujian nasional dan berbagai yang dibuat oleh guru itu sendiri (Pahri, 2021).

Definisi dari suatu tes adalah berbagai soal yang mempunyai jawaban salah maupun benar yang digunakan sebagai tolak ukur pemahaman dan *skill* suatu individu (Indah et al., 2021). Tes adalah salah satu metode yang digunakan untuk mengukur dan menilai pencapaian siswa dalam proses belajar mereka, tes sendiri terdiri dari tugas atau latihan yang harus dilakukan siswa (Magdalena et al., 2021) Tes-tes yang ada dalam berbagai mapel tersebut sangatlah *urgent* kedudukan dan fungsinya, oleh karena itu diharapkan berbagai tes tersebut memenuhi standar yang telah disebutkan diatas. Adapun ukuran dan aspek yang digunakan yaitu taraf pemahaman peserta didik terhadap materi. Selain itu, pencapaian siswa dapat dinilai melalui tiga elemen inti dari hasil belajar, yakni aspek kognitif yang berkaitan dengan pengetahuan, aspek psikomotorik yang melibatkan keterampilan fisik, dan aspek afektif yang mencerminkan sikap, nilai, dan emosi siswa terhadap materi pelajaran.

Tes Sumatif adalah *assesment* pada kurun waktu ujung semester pada kalender pendidikan. Dan hasilnya akan dicantumkan dalam raport sekaligus penentuan dalam keputusan naik atau tidaknya siswa (Shodiyah & Septi, 2014). Kemudian, hasil tes sumatif siswa dibandingkan untuk menentukan tujuan pembelajaran yang telah dirancang dan ditentukan untuk mencapai kriteria minimal derajat Kesempurnaan atau biasa disebut (KKM). Kemudian, hasil tes sumatif siswa dibandingkan untuk menentukan tujuan pembelajaran yang telah ditetapkan untuk mencapai kriteria minimal derajat kesempurnaan, atau KKM. Selain itu, tes sumatif juga menjadi alat penting dalam menilai sejauh mana kemampuan siswa menguasai keterampilan yang diajarkan dalam suatu mata pelajaran. Tes sumatif ini bertujuan untuk memberikan gambaran komprehensif tentang pencapaian siswa dalam memahami dan menerapkan materi pelajaran tertentu (Alfani, 2022).

Pilihan ganda merupakan salah satu bentuk dan cara dalam tes dimana guru memberikan pertanyaan dengan beberapa pilihan jawaban. Dalam menghadapi tes seperti ini, siswa ditantang untuk mengidentifikasi jawaban yang tepat di antara pilihan-pilihan yang tersedia. Hal ini mendorong mereka untuk mempertimbangkan secara kritis dan memilih jawaban yang paling sesuai dengan pertanyaan atau permasalahan yang diberikan. Oleh sebab itu, tes haruslah memiliki kualitas yaitu tes yang mempunyai validitas dan reliabilitas dalam butir soalnya. Dengan melakukan analisis terhadap berbagai butir soal, dapat ditemukan kelemahan atau potensi yang perlu ditingkatkan. Apabila terdapat pertanyaan yang belum valid dan reliabel, dapat dilakukan penyesuaian agar mencapai tingkat kualitas yang lebih baik. Ini memungkinkan pengembangan dan penyempurnaan butir soal untuk memastikan pengukuran pencapaian siswa menjadi lebih valid dan reliabel (Hayati & Lailatussaadah, 2016). Untuk meningkatkan kualitas soal yang telah disiapkan, guru harus melakukan analisis elemen (*item analysis*). Tujuan lain adalah untuk menemukan pertanyaan yang baik, kurang baik, dan buruk. Analisis pertanyaan dapat menentukan tingkat keparahan masalah dan

"petunjuk" (Farida & Musyarofah, 2021). Maka menjadi suatu keharusan dalam memberlakukan analisis terlebih dahulu pada butir soal agar dapat diketahui kualitas soal tersebut

Menurut (Fadli, R., Hidayati, S., Cholifah, M., Siroj, R. A., & Afgani, 2023) dalam penelitian sebelumnya menemukan tentang analisis validitas dan reliabilitas instrumen bahwa validitas dan reliabilitas instrumen menentukan tingkat kualitas instrumen yang digunakan. Faktor-faktor seperti respons, kondisi atau keadaan lokasi penelitian, penggunaan alat yang tidak ideal, dan lain-lain sangat memengaruhi proses validasi dan reliabilitas penelitian ini. Oleh karena itu, kualitas dan kuantitas penelitian akan dipengaruhi oleh validitas dan kredibilitas hasil. Selain itu, hasil penelitian (Cahyaningrum et al., 2023) menunjukkan bahwa tes yang digunakan memenuhi syarat untuk reliabilitas, kekhasan, dan tingkat kesukaran software Anates. Selanjutnya, hasil pengukuran tersebut akan digunakan sebagai penilaian dan dasar pertanyaan berikutnya. Menurut penelitian (Hikmah & Muslimah, 2021) alat uji yang sah dan dapat diandalkan tidak hanya memberikan data yang valid dan dapat diandalkan tetapi juga menghasilkan kesimpulan yang mencerminkan keadaan sebenarnya.

Penelitian ini penting untuk dilakukan karena fakta dilapangan menunjukkan bahwa beberapa tenaga pendidik khususnya guru PAI di SMA Muhammadiyah Tanah Grogot belum mengimplementasikan analisis butir soal tes sumatif akhir semester ganjil tahun ajaran 22/23, dikarenakan beberapa alasan misalnya seperti alokasi waktu yang dimiliki oleh sebagian guru mapel terbatas akibat pemadatan materi juga tidak ditunjangnya oleh penguasaan guru terhadap teknologi. Sehingga peserta didik telah menyelesaikan asesmen sumatif akhir semester ganjil dengan instrumen soal yang belum melalui tahapan analisis butir soal itu sendiri. Padahal mengingat tingkat urgensitas yang telah dipaparkan diatas maka wajib untuk seorang pendidik agar dapat menganalisis secara maksimal tes tersebut, sebagai bentuk evaluasi terhadap keberhasilan instrumen tes yang diujikan terutama dalam penyusunan soal.

Penelitian ini bertujuan untuk mendorong guru-guru PAI untuk melakukan refleksi terhadap kualitas butir soal sumatif akhir semester ganjil. Fokus utama penelitian adalah pada analisis validitas dan reliabilitas butir soal, dengan harapan hasil penelitian dapat memberikan manfaat praktis bagi para guru. Melalui refleksi tersebut, diharapkan guru dapat mengidentifikasi butir soal yang kurang valid dan reliabel, serta merancang perbaikan yang sesuai. Dengan demikian, penelitian ini tidak hanya memberikan wawasan tentang kualitas instrumen tes PAI, tetapi juga memberikan panduan konkret bagi guru untuk meningkatkan efektivitas dan akurasi evaluasi terhadap peserta didik, mencapai tujuan pendidikan PAI dengan lebih baik.

Berdasarkan paparan diatas maka sekiranya menarik jika peneliti lebih dalam membahas tentang "Analisis Validitas dan Reliabilitas Butir Soal Sumatif Akhir Semester Ganjil Mata Pelajaran PAI". Melalui penelitian ini, diharapkan bisa memberikan manfaat kepada pendidik dan guru PAI di SMA Muhammadiyah Tanah Grogot guna merefleksikan dan melakukan perbaikan jika terdapat beberapa butir soal yang kurang valid dan reliabel agar tercapainya *goals* dalam suatu proses kependidikan khususnya PAI.

METODE

Dalam penelitian ini, data-data yang diperlukan diperoleh melalui dua metode, yaitu dokumentasi dan wawancara dengan para guru PAI. Pendekatan yang digunakan dalam penelitian ini adalah *ex post facto*, yang berarti pengumpulan data dilakukan setelah peristiwa-peristiwa terkait telah terjadi. Dengan demikian, penelitian ini mengeksplorasi hubungan antara variabel-variabel yang telah ada, tanpa intervensi peneliti pada variabel tersebut.

Penelitian dilaksanakan di SMA Muhammadiyah Tanah Grogot yang beralamat di jalan Pangeran Menteri kompleks Masjid Syuhada, Kota Tanah Grogot, Provinsi Kalimantan Timur. Penelitian dilaksanakan setelah penilaian akhir semester ganjil tahun ajaran 2022/2023.

Partisipan dalam penelitian ini terdiri dari 62 siswa yang terdaftar dalam kelas XA, XB, dan XC, serta guru-guru PAI yang berperan sebagai informan. Sampel penelitian ini adalah kumpulan soal ujian sumatif akhir semester ganjil mata pelajaran PAI, yang terdiri dari 50 soal pilihan ganda.

Langkah pelaksanaan mencakup pengumpulan data melalui dokumentasi soal ujian dan lembar jawaban siswa, serta mencatat identitas siswa dari kelas XA, XB, dan XC. Dengan demikian, langkah ini bertujuan untuk mendapatkan informasi yang diperlukan dari materi dan hasil ujian dan identitas siswa untuk dianalisis lebih lanjut terkait hasil belajar dalam penelitian. Tahap selanjutnya melibatkan analisis dengan menguji validitas dan reliabilitas setiap butir soal berdasarkan jawaban benar dan salah dari ujian sumatif akhir semester ganjil. Proses ini menggunakan pendekatan deskriptif kuantitatif dengan mengkonversi data tersebut menjadi angka-angka untuk mengevaluasi pencapaian akhir siswa dalam menjawab soal. Dengan demikian, penelitian ini bertujuan untuk memastikan keandalan dan kevalidan dari setiap item soal dalam mengukur pemahaman siswa terhadap materi pelajaran PAI.

Adapun tahapan pengelolaan dan penganalisisannya sebagai berikut:

Pertama, Menentukan nilai validitas butir item dengan rumus:

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

Keterangan :

r_{xy} : koefisien korelasi yang dicari

N: banyaknya peserta tes

X: nilai variabel X (skor item)

Y: nilai variabel Y (skor item)

Adapun dasar pengambilan keputusan pada uji validitas adalah:

Jika $r_{hitung} > r_{tabel}$ maka butir item valid.

Jika $r_{hitung} < r_{tabel}$ maka butir item tidak valid.

Kedua, menentukan nilai reliabilitas. Cara menghitung reliabilitas suatu tes adalah dengan menggunakan rumus:

$$r_{11} = \left(\frac{k}{k-1} \right) \left(1 - \frac{\sum S_i}{S_t} \right)$$

Keterangan:

r_{11} = Reliabilitas instrumen

k = Jumlah butir soal

$\sum S_i$ = Jumlah varians masing-masing item

S_t = Varians soal

Adapun dasar pengambilan keputusan pada uji reliabilitas adalah:

Jika Cronbach alpha $> 0,6$ maka instrumen tes dapat dikatakan reliabel

Jika Cronbach alpha $< 0,6$ maka instrumen tes dapat dikatakan tidak reliabel.

HASIL DAN PEMBAHASAN

Pengertian Uji Validitas

Validitas tes memiliki fungsi untuk mengukur suatu objek yang akan digunakan untuk mengukur pula. Dalam aspek validitas tersebut ada 3 unsur yang akan dilihat, yakni: isi, rekaan teoritis dan kriteria dari alat ukur tersebut (Khafidin, 2014).

Orang yang pertama kali dalam mengemukakan definisi dari validitas adalah Garret pada tahun 1937, dia mendefinisikan validitas sebagai keakuratan sebuah tes dalam mengukur suatu objek. Lalu, pada tahun 1971 Cronbach memiliki interpretasi bahwa validitas dibangun berdasarkan beberapa bukti terlebih dahulu agar tes semakin akurat. Untuk definisi validitas ini sendiri masih menjadi perdebatan, tetapi di tahun 1987 Mehrens menganggap validitas paling tepat didefinisikan sebagai takaran seberapa tepat sarannya suatu pengukuran dalam membuat sebuah tes. Lain halnya dengan Anastasi mengemukakan pendapat bahwa validitas tes berhubungan dengan objek yang akan diukur dan seberapa maksimal tes yang dipakai untuk mengukur objek tersebut (Muzaffir, 2016).

Di *real life* dalam dunia pendidikan sendiri banyak ditemui narasi-narasi: "Tes ini bagus karena sudah tervalidasi," narasi tersebut tidak tepat, seharusnya narasi itu "Tes ini sudah bagus karena sudah teruji validitasnya" atau "Tes ini bagus karena validitasnya tinggi." Penjelasan ini menegaskan bahwa tes dinilai baik karena telah terbukti valid, yang mengindikasikan bahwa instrumen tersebut dapat diandalkan untuk mengukur apa yang dimaksudkan". Maka seyogyanya jika membicarakan tentang evaluasi kata "valid" cocok jika disandingkan dengan kata "data".

Salah satu ciri dalam mengenali instrument yang valid adalah dengan melihat terlebih dahulu valid atau tidaknya suatu data. Oleh karena itu dapat dipastikan jika suatu data sudah memenuhi syarat valid maka akan menghasilkan *output* berupa instrument yang valid pula.

Sebenarnya diksi valid dapat diganti dengan kata sah. Tetapi karena kata sah tersebut belum secara maksimal dalam perannya untuk menggantikan kata valid maka kata valid tetap masyhur digunakan sampai sekarang.

Jenis-jenis Validitas

Secara garis besar jenis validitas yang berasal dari hasil tes belajar terbagi menjadi dua, yaitu pengujian validitas secara logis dan secara empiris.

Validitas Logis

Yang dimaksud dengan "validitas logis" mencakup kata "logis" yang memiliki arti penalaran. Oleh karena itu, jenis validitas logis merupakan suatu instrumen evaluasi yang mengacu kepada instrumen yang berdasarkan hasil penalaran. item dianggap valid apabila pengujian dipersiapkan dengan menyesuaikan teori dan peraturan agar menjadi instrumen yang benar. Validitas logis memiliki dua aspek, yakni validitas isi dan validitas konstruk.

Validitas Isi

Setelah dirancang dengan mempertimbangkan materi yang dievaluasi, instrumen tersebut memperoleh validitas isi yang menunjukkan bahwa tes tersebut mencakup contoh-contoh hasil belajar yang sejalan dengan tujuan pembelajaran yang ditetapkan. Dengan kata lain, instrumen tersebut dapat diandalkan untuk mengukur pemahaman siswa terhadap materi pelajaran dengan memperhatikan kesesuaian antara materi ujian dan tujuan pembelajaran yang diinginkan. Validitas isi tes ditentukan oleh pendapat juri ahli yang disebut validator pada saat proses validasi awal. Dengan menggunakan spesifikasi tes yang dikembangkan, analisis logis dilakukan untuk menentukan apakah pertanyaan yang dikembangkan benar-benar mewakili apa yang diukur.

Validitas Konstruk

Suatu instrumen dikatakan mencapai validitas konstruk apabila dibangun di atas konstruk dimensi psikologis yang akan dievaluasi atau diukur adalah aspek berpikir yang disebutkan dalam indikator pembelajaran. Validitas konstruk teoritis merujuk pada sejauh mana hasil yang diperoleh dari pengukuran menggunakan instrumen yang bersangkutan mencerminkan konstruksi teoritis yang menjadi dasar pembuatan instrumen pengukuran tersebut. Validitas konstruk berdasarkan indikator-indikator (juga disebut variabel yang diamati) menggunakan ilmu statistik dengan sangat kompleks: Analisis faktor dengan SPSS atau validitas konvergen dan diskriminan dapat diuji melalui dua pendekatan berikut: a) Melakukan analisis faktor statistik (konfirmasi). b) Menggunakan asumsi dasar verifikasi konvergensi dan diskriminan sebagai berikut: Pengujian seharusnya menunjukkan korelasi yang tinggi dengan variabel yang seharusnya memiliki korelasi tinggi secara teoritis (untuk validasi konvergen). Pada saat yang sama, tes tersebut tidak boleh dikorelasikan dengan variabel lain yang secara teoritis harus dikorelasikan (Solichin, 2017).

Validitas Empiris

Jenis validitas selanjutnya bertajuk validitas empiris, diksi empiris sendiri memiliki makna pengalaman. Maka dari itu suatu instrumen dapat dikatakan valid melalui sisi pengalaman, dalam artian telah teruji dari pengalaman. Didalam jenis empiris ini didapati beberapa perinciannya, yakni ; validitas prediksi (*predictive validity*) dan validitas sama saat (*concurrent validity*) (Riyani et al., 2017).

Validitas Prediksi/Ramalan (*predictive validity*)

Keberadaan validitas ini terjadi ketika sebuah instrumen mampu dengan tepat memprediksi hasil atau peristiwa yang akan terjadi di masa depan. Dengan kata lain, instrumen tersebut memiliki kemampuan yang handal untuk meramalkan atau memperkirakan apa yang akan terjadi berdasarkan informasi yang diberikan. Hal ini menunjukkan bahwa instrumen tersebut dapat diandalkan untuk mengukur atau mengidentifikasi variabel yang relevan dalam memprediksi hasil di masa yang akan datang, maka instrumen tersebut telah memenuhi aspek prediksi dalam konteks validitas. Sedangkan tolak ukur keakuratan tersebut dapat dicermati melalui korelasi antara hasil tes dan alat ukur di masa yang mendatang (Realita, 2015).

Salah satu contohnya adalah ketika diadakannya tes masuk dalam suatu universitas, dengan tes tersebut membuahkan dua hasil yakni antara lulus dan tidak lulus. Yang dapat lulus tentunya diprediksi sanggup dalam mengikuti perkuliahan secara maksimal sedangkan yang tidak lulus diprediksi tidak sanggup dalam mengikuti perkuliahan yang akan mendatang.

Atau secara singkatnya standar tes yang memenuhi unsur ini jika ditemukan hubungan diantara hasil tes yang diuji dan kriteria yang sudah ditetapkan. Salah satu teknik untuk mengetahui ada tidaknya validitas prediksi ini adalah melalui teknik *product moment* karya Karl Pearson (Pramana et al., 2012).

Validitas Sama Saat (*concurrent validity*).

Validitas konkrue ini dapat tercipta jika adanya hubungan yang harmonis antara antara skor tes dan kinerja saat ini. Dengan demikian, hasil tes menjadi perbandingan yang berarti. Validitas ini juga dapat disebut sebagai validitas empiris. Sebagai contoh:

Seorang pendidik ingin memastikan validitas tes sumatif yang telah mereka buat. Untuk melakukannya, diperlukan data historis sebagai pembanding, seperti hasil ulangan harian dan nilai total ujian sebelumnya yang tersedia saat ini. Dengan membandingkan hasil tes sumatif dengan data historis ini, guru dapat mengevaluasi sejauh mana tes tersebut mencerminkan kemajuan siswa secara akurat dan valid.

Pengertian Uji Reliabilitas

Uji reliabilitas seringkali dipergunakan dalam menghitung konsistensi suatu alat pengukur, seperti kuesioner yang didalamnya terdiri dari indikator-indikator berdasarkan variabel atau konstruk tertentu (Ghozali, 2018). Tujuan dari adanya uji reliabilitas ialah memastikan bahwa alat pengukur tersebut dapat

memberikan hasil yang sama secara konsisten (berkesinambungan) pada percobaan atau pengukuran yang berulang, dan dapat diandalkan. Dalam konteks uji reliabilitas, Suatu instrumen pengukuran dianggap memiliki reliabilitas apabila memberikan hasil yang konsisten saat diukur berulang kali. Dengan kata lain, alat pengukur dapat diandalkan untuk memberikan hasil yang serupa dalam situasi pengukuran yang berbeda-beda. Ini menunjukkan bahwa instrumen tersebut memberikan tingkat kestabilan dan konsistensi yang diperlukan untuk evaluasi yang akurat dan dapat diandalkan.

Menurut Nursalam, uji reliabilitas mengacu pada kesamaan hasil pengukuran atau pengamatan saat dilakukan dalam rentang waktu yang berbeda. Penting untuk memperhatikan peran alat dan metode yang digunakan dalam mengukur atau mengamati, karena keduanya berkontribusi dalam mencapai kesamaan hasil pada waktu yang bersamaan. Dalam hal kuesioner, reliabilitas mengindikasikan sejauh mana jawaban seseorang terhadap *statement* (pernyataan) dalam tes tersebut konsisten. apabila jawabannya relatif sama atau konsisten dalam pengukuran yang berulang, maka kuesioner dianggap reliabel dan dapat dipercaya (Nursalam, 2003). Uji reliabilitas digunakan sebagai alat atau tes untuk mengevaluasi ketepatan atau konsistensi suatu tes. Artinya, uji reliabilitas memastikan bahwa tes yang telah dilakukan tersebut menunjukkan hasil yang tidak berubah atau relatif sama setiap kali digunakan.

Berdasarkan teori yang termasyhur, reliabilitas terkait pada kecocokan sebuah instrumen tes dalam pengukuran. Reliabilitas mengacu pada stabilitas nilai yang didapatkan oleh individu yang sama pada saat diuji lagi bersama dengan instrumen tes yang sama pula dalam ruang waktu yang tidak sama. Maka, uji reliabilitas mengukur bagaimana hasil dari pengukuran tersebut apakah tetap konsisten ketika pengukuran dilakukan kembali pada instrumen yang sama. Harapannya adalah bahwa hasil pengukuran akan sama jika pengukuran tersebut diulang (Surapranata, 2006).

Ada beberapa uji reliabilitas yang dapat digunakan untuk menguji reliabilitas suatu instrumen. Beberapa uji reliabilitas yang umum digunakan meliputi uji *test-retest*, uji ekuivalen, dan uji *internal consistency*. Uji *internal consistency* sendiri memiliki beberapa teknik yang berbeda, termasuk uji KR 21, KR 20, Alfa Cronbach dan split half. Setiap teknik uji reliabilitas memiliki kriteria instrumen yang berbeda untuk dapat diuji menggunakan teknik tersebut (Purba & Octobe, 2021). Dengan menggunakan berbagai uji reliabilitas ini, peneliti dapat menguji reliabilitas instrumen dan memastikan bahwa instrumen tersebut memberikan hasil yang konsisten dan dapat diandalkan.

Faktor yang Mempengaruhi Reliabilitas

Jumlah butir tes (panjang tes)

Seringkali reliabilitas tes cenderung meningkat seiring dengan peningkatan jumlah butir soal tes. Hal ini disebabkan oleh fakta bahwa apabila instrumen tes yang digunakan semakin banyak butir soal, maka semakin besar yang dapat diukur dengan lebih tepat. Uji reliabilitas akan menaik dan itu dijadikan sebagai fungsi dari panjang tes dapat dipaparkan melalui rumus Spearman-Brown, yaitu:

$$r_n = \frac{nr}{1 + (n - 1)r}$$

r_n : indeks reliabilitas setelah ditambahkan soal

n : perkalian penambahan awal

r : indeks reliabilitas awal

Apabila sebuah instrumen tes mempunyai butir soal yang banyak, kontinuitasnya akan mengalami peningkatan asalkan instrumen soal tersebut didasarkan pada penjelasan yang terperinci mengenai variabel yang sedang diukur. Oleh karena itu, menyusun tes harus mempertimbangkan seberapa banyak butir soal yang tepat. Meskipun penambahan butir soal dapat meningkatkan reliabilitas, bukan berarti bahwa harus

disiapkan sebanyak-banyaknya butir soal. Tetap penting untuk mempertimbangkan batas ukuran dan indikator yang relevan dengan variabel.

Variabilitas kelompok atau heterogenitas kelompok

Salah satu faktor yang mempengaruhi uji reliabilitas instrumen ialah variabilitas kelompok. Variabilitas kelompok mengacu pada variasi nilai mata pelajaran pada variabel yang diukur pada instrumen tes tersebut. Jika tes tersebut mengukur kecakapan dalam bahasa Inggris, maka variabilitas kelompok akan tercermin dalam ragam nilai kecakapan berbahasa Inggris pada sejumlah peserta didik yang diuji. Maka relevan dengan rumus reliabilitas, yakni:

$$r_{xx} = \frac{\sigma_T^2}{\sigma_x^2}$$

Koefisien reliabilitas menggambarkan sejauh mana tes secara konsisten menempatkan siswa pada posisi yang tepat di dalam kelompok mereka. Koefisien reliabilitas dipengaruhi secara langsung oleh nilai yang tersebar dalam kelompok yang akan diukur. Jika penyebaran skor semakin besar (menunjukkan variasi yang lebih tinggi dalam kelompok), maka indeks reliabilitas akan semakin tinggi. Perbedaan kemampuan siswa ketika diukur akan lebih sulit pada kelompok yang memiliki banyak kesamaan (homogen) dibanding dengan kelompok yang bermacam-macam (heterogen).

Variabilitas dalam kelompok memiliki dampak pada nilai yang tersebar. Koefisien reliabilitas secara langsung dapat dipengaruhi oleh nilai yang tersebar pada kelompok yang diukur. Dengan asumsi faktor-faktor lain tetap konstan, semakin besar nilai yang tersebar dalam kelompok, maka akan semakin tinggi pula perolehan dari indeks reliabilitas itu sendiri.

Objektivitas penskoran

Reliabilitas pemberi nilai mengacu pada tingkat kesepakatan antara observer atau evaluator yang berbeda dalam memberikan penilaian pada instrumen serupa. Semakin besar kesesuaian antara penilai, maka keandalan reliabilitas penilaian juga akan meningkat. Sebaliknya, jika terdapat perbedaan penilaian yang signifikan antara evaluator, reliabilitas pemberi skor akan menurun. Banyak tes bakat dan tes akademik standar menunjukkan tingkat objektivitas tinggi. Hal ini mengindikasikan bahwa nilai-nilai yang diberikan dalam tes, seperti pada soal pilihan ganda, tidak dipengaruhi oleh subjektivitas atau pendapat pemberi penilaian. Dalam konteks ini, objektivitas menunjukkan bahwa evaluasi dilakukan dengan konsistensi dan tidak dipengaruhi oleh faktor-faktor yang tidak relevan atau keputusan subjektif dari pemberi penilaian. Dalam tes objektif, skor yang dihasilkan memiliki tingkat objektivitas yang tinggi karena dapat diukur secara konsisten oleh evaluator yang berbeda. Semakin tinggi tingkat objektivitas tes, maka reliabilitasnya juga cenderung lebih tinggi.

Dalam rangka meningkatkan reliabilitas, penting untuk memastikan adanya kesepakatan yang tinggi antara evaluator dalam memberikan penilaian. Menggunakan tes yang memiliki tingkat objektivitas yang tinggi juga dapat membantu meningkatkan reliabilitas instrumen.

Metode estimasi reliabilitas

Pengujian koefisien reliabilitas tes standar, penting untuk memilih langkah apa atau metode yang akan dipakai dalam memutuskan besarnya koefisien reliabilitas. Metode yang digunakan secara langsung mempengaruhi nilai koefisien reliabilitas yang diestimasi. Berikut ini beberapa cara untuk mengestimasi koefisien reliabilitas:

Metode tes ulang (*Test Retest Method*): Hasil koefisien reliabilitas dapat lebih tinggi jika interval waktu antara tes pertama dan tes ulangnya pendek. Hal ini karena semakin pendek interval waktu, semakin stabil dan konsisten respons siswa terhadap tes tersebut. Dalam kasus ini, koefisien reliabilitas yang diestimasi dapat lebih tinggi dibandingkan dengan metode belah dua. Namun, jika interval waktu antara tes pertama dan tes ulangnya ditingkatkan, koefisien reliabilitas yang dihasilkan cenderung lebih rendah. Hal ini disebabkan oleh

kemungkinan adanya perubahan dalam kemampuan atau faktor-faktor lain yang dapat memengaruhi respons siswa antara dua tes yang berbeda. Dalam situasi ini, reliabilitas tes dapat dipengaruhi negatif oleh perubahan tersebut, sehingga koefisien reliabilitas yang diestimasi menjadi lebih rendah.

Level kelompok dan tingkat kesulitan tes

Selain variasi dalam kelompok, karakteristik kelompok juga bisa mempengaruhi koefisien reliabilitas. Akurasi pengukuran dari instrumen tes dapat berkaitan dengan tingkat kemampuan individu yang diuji. Namun, tidak ada aturan baku yang dapat menjelaskan korelasi ini secara pasti. Hubungan ini sangat tergantung pada desain tes itu sendiri. Ketika seseorang menganggap tes terlalu sulit, mereka mungkin cenderung menebak dalam menjawab, yang berpotensi mengurangi keakuratan tes. Di sisi lain, jika tes terlalu mudah bagi kelompok tertentu sehingga mereka mampu menjawab sebagian besar item dengan benar, maka tes tersebut mungkin tidak efektif dalam membedakan antara anggota kelompok tersebut. Semua ini bergantung pada tingkat kesulitan tes itu sendiri. Tes yang terlalu sulit atau terlalu mudah mungkin tidak dapat mengukur perbedaan antarindividu dengan baik.

Homogenitas tes

Homogenitas tes juga mempengaruhi reliabilitas. Sebagai contoh, ujian PAI dengan 100 item untuk kelas X akan lebih reliabel daripada tes PAI dengan 100 item yang mengukur tingkat SMA secara keseluruhan. Prinsip yang serupa berlaku untuk bidang studi lainnya seperti bahasa Arab atau matematika, yang membutuhkan pendekatan yang lebih terstruktur dengan penekanan yang lebih besar pada logika, aturan, keterampilan, dan hasil, berbeda dengan mata pelajaran seperti sejarah. Ini merupakan dimensi tambahan dari konsistensi dalam tes yang memungkinkan reliabilitas yang lebih tinggi lebih mudah tercapai dalam pengujian bahasa asing dan matematika dibandingkan dengan evaluasi hasil belajar dalam mata pelajaran lainnya.

Hasil Uji Validitas Butir Soal

Peneliti melakukan uji pada tes sumatif akhir semester ganjil yang terdiri dari 50 butir soal pilihan ganda mata pelajaran PAI di tiga kelas, yaitu XA, XB, dan XC, dengan total 62 siswa. Dalam penelitian ini, validitas empiris dari setiap butir soal akan dievaluasi melalui analisis korelasi. Suatu soal dianggap valid jika koefisien korelasi yang dihitung (r_{hitung}) lebih besar atau sama dengan nilai korelasi yang tercantum dalam tabel (r_{tabel}) pada tingkat signifikansi 5%. Berikut adalah hasil pengujian:

Tabel 1. Hasil Validitas Butir Soal

Butir Soal	Hasil Uji		Keputusan	Butir Soal	Hasil Uji		Keputusan
	r_{hitung}	r_{tabel}			r_{hitung}	r_{tabel}	
1	0.329	0.250	Valid	26	0.333	0.250	Valid
2	0.635	0.250	Valid	27	0.523	0.250	Valid
3	0.598	0.250	Valid	28	0.424	0.250	Valid
4	0.322	0.250	Valid	29	0.424	0.250	Valid
5	0.263	0.250	Valid	30	0.130	0.250	Tidak Valid
6	0.154	0.250	Tidak Valid	31	0.192	0.250	Tidak Valid
7	0.413	0.250	Valid	32	0.236	0.250	Tidak Valid
8	0.005	0.250	Tidak Valid	33	0.386	0.250	Valid
9	0.049	0.250	Tidak Valid	34	0.357	0.250	Valid
10	0.486	0.250	Valid	35	0.460	0.250	Valid
11	0.421	0.250	Valid	36	0.147	0.250	Tidak Valid
12	0.572	0.250	Valid	37	0.421	0.250	Valid
13	0.176	0.250	Tidak Valid	38	0.597	0.250	Valid
14	0.353	0.250	Valid	39	0.278	0.250	Valid
15	0.268	0.250	Valid	40	0.319	0.250	Valid

16	0.174	0.250	Tidak Valid	41	0.409	0.250	Valid
17	0.549	0.250	Valid	42	0.357	0.250	Valid
18	0.140	0.250	Tidak Valid	43	0.477	0.250	Valid
19	0.279	0.250	Valid	44	0.741	0.250	Valid
20	0.403	0.250	Valid	45	0.554	0.250	Valid
21	0.550	0.250	Valid	46	0.242	0.250	Tidak Valid
22	0.169	0.250	Tidak Valid	47	0.609	0.250	Valid
23	0.317	0.250	Valid	48	0.430	0.250	Valid
24	0.132	0.250	Tidak Valid	49	0.567	0.250	Valid
25	0.222	0.250	Tidak Valid	50	0.265	0.250	Valid

Tabel 2. Distribusi Butir Soal Berdasarkan Validitas

No	Indeks Validitas	Butir Soal
1	≥ 0.250 (Valid)	1, 2, 3, 4, 5, 7, 10, 11, 12, 14
		15, 17, 19, 20, 21, 23, 26, 27
		28, 29, 33, 34, 35, 37, 38, 39
		40, 41, 42, 43, 44, 45, 47
		48, 49, 50
2	< 0.250 (Tidak Valid)	6, 8, 9, 13, 16, 18, 22, 24, 25
		30, 31, 32, 36, 46

Berdasarkan hasil analisis 50 butir soal pilihan ganda ujian Sumatif akhir semester ganjil SMA Muhammadiyah Tanah Grogot dapat diketahui bahwa 36 butir soal dinyatakan valid yaitu soal nomor 1, 2, 3, 4, 5, 7, 10, 11, 12, 14, 15, 17, 19, 20, 21, 23, 26, 27, 28, 29, 33, 34, 35, 37, 38, 39, 40, 41, 42, 43, 44, 45, 47, 48, 49 dan 50 sedangkan 14 butir soal dinyatakan tidak valid yaitu soal nomor 6, 8, 9, 13, 16, 18, 22, 24, 25, 30, 31, 32, 36 dan 46.

Dari paparan di atas maka yang menunjukkan valid berjumlah 36 butir soal dan menggambarkan keberhasilan sejumlah besar pertanyaan dalam mengukur pemahaman siswa terhadap materi PAI. Butir soal yang valid menunjukkan bahwa pertanyaan tersebut mampu memberikan gambaran yang akurat tentang tingkat pemahaman siswa terkait dengan tujuan pembelajaran yang telah ditetapkan. Validitas butir soal yang tinggi juga mencerminkan upaya guru dalam merancang instrumen evaluasi yang sesuai dengan kurikulum dan materi pembelajaran yang diimplementasikan di SMA Muhammadiyah Tanah Grogot.

Namun, perlu diakui bahwa adanya 14 butir soal yang dinyatakan tidak valid memberikan indikasi terhadap potensi perbaikan yang dapat dilakukan. Analisis mendalam terhadap butir-butir soal yang tidak valid menjadi penting untuk mengevaluasi sejauh mana butir soal tersebut memenuhi kriteria kevalidan dan mengidentifikasi faktor-faktor yang mungkin menyebabkan ketidakvalidan. Proses ini dapat memberikan wawasan yang berharga bagi guru dan penyusun kurikulum untuk menyempurnakan desain butir soal dan meningkatkan kualitas instrumen evaluasi yang digunakan di sekolah.

Selanjutnya, perlu diperhatikan bahwa validitas butir soal tidak hanya bersifat statis tetapi dapat berubah seiring waktu dan perkembangan kurikulum. Oleh karena itu, penelitian ini seharusnya menjadi dasar untuk melakukan pemantauan terus-menerus terhadap kualitas butir soal yang digunakan dalam ujian sumatif akhir semester. Dengan melakukan pembaruan dan penyesuaian secara berkala, guru dapat memastikan bahwa

instrumen evaluasi mereka selalu relevan dengan perubahan kurikulum dan memenuhi standar kevalidan yang diperlukan.

Hasil Uji Reliabilitas Butir Soal

Peneliti menjalani uji pada tes sumatif akhir semester ganjil yang terdiri dari 50 soal pilihan ganda mata pelajaran PAI di tiga kelas, yakni XA, XB, dan XC, dengan jumlah total siswa sebanyak 62 orang. Untuk mengetahui reliabilitas instrumen penelitian, peneliti melakukan perhitungan menggunakan rumus Alpha Cronbach. Instrumen penelitian diklasifikasikan sebagai reliabel jika nilai hasil perhitungan reliabilitas melebihi atau sama dengan 0,65. Berikut adalah hasil uji reliabilitas:

Tabel 3. Hasil Uji Reliabilitas Instrumen Soal

Reliability Statistics	
Cronbach's Alpha	Jumlah Butir Soal
0,855	50

Pada penelitian ini, nilai Cronbach's Alpha sebesar 0,855 menunjukkan tingkat reliabilitas yang sangat baik, mengindikasikan bahwa tes pilihan ganda PAI yang digunakan dalam penelitian ini dapat diandalkan untuk mengukur pemahaman siswa terhadap materi PAI. Nilai reliabilitas yang tinggi memberikan keyakinan kepada peneliti bahwa instrumen yang digunakan memberikan hasil yang konsisten dan dapat dipercaya, sehingga hasil penelitian dapat dianggap representatif.

Perlu diperhatikan bahwa nilai 0,855 yang diperoleh jauh melampaui ambang batas minimal reliabilitas sebesar 0,65. Hal ini menegaskan bahwa instrumen tes pilihan ganda memiliki tingkat konsistensi yang sangat tinggi, memberikan keyakinan bahwa hasil evaluasi yang diperoleh dari tes tersebut dapat diandalkan dan akurat. Hasil uji reliabilitas yang positif ini memberikan dukungan kuat terhadap kevalidan instrumen penelitian dan memperkuat kesimpulan yang dapat diambil dari hasil analisis butir soal sebelumnya.

Dengan adanya nilai reliabilitas yang tinggi, dapat diasumsikan bahwa tes pilihan ganda ini memiliki konsistensi yang baik dalam mengukur apa yang seharusnya diukur, yaitu pemahaman siswa terhadap mata pelajaran PAI. Tingkat reliabilitas yang tinggi juga dapat meminimalkan kemungkinan adanya kesalahan pengukuran dan meningkatkan keakuratan hasil penelitian. Oleh karena itu, keberhasilan dalam mendapatkan nilai Cronbach's Alpha sebesar 0,855 menegaskan bahwa instrumen penelitian ini dapat memberikan kontribusi yang signifikan dalam evaluasi pemahaman siswa terhadap materi PAI di SMA Muhammadiyah Tanah Grogot.

Dalam konteks evaluasi pembelajaran, nilai reliabilitas yang tinggi pada tes pilihan ganda ini dapat menjadi dasar untuk pembaruan dan penyesuaian selanjutnya. Guru dapat menggunakan informasi ini untuk merancang strategi pengajaran yang lebih efektif, dan sekolah dapat mempertimbangkan penerapan instrumen ini dalam ujian dan evaluasi lainnya. Dengan demikian, hasil reliabilitas yang baik tidak hanya memberikan dukungan terhadap penelitian ini tetapi juga memiliki dampak positif dalam perbaikan sistem evaluasi pendidikan.

Terakhir, pemahaman terhadap hasil analisis butir soal ini tidak hanya bermanfaat bagi guru, tetapi juga dapat memberikan informasi berharga kepada pihak-pihak terkait dalam proses pengambilan keputusan di tingkat sekolah. Hasil penelitian ini dapat menjadi landasan untuk pengembangan strategi evaluasi dan pengajaran yang lebih baik, serta dapat memberikan sumbangan positif terhadap evaluasi pendidikan secara keseluruhan. Oleh karena itu, penting untuk mendiskusikan dan memanfaatkan temuan ini secara kolaboratif di antara staf pengajar, penyusun kurikulum, dan pimpinan sekolah guna mencapai tujuan bersama dalam meningkatkan kualitas pendidikan.

SIMPULAN

Berdasarkan analisis validitas dan reliabilitas tes ujian sumatif akhir semester ganjil mata pelajaran PAI tahun ajaran 2022/2023 di SMA Muhammadiyah Tanah Grogot, disimpulkan bahwa soal pilihan ganda untuk kelas X dianggap valid dan reliabel. Ini menandakan bahwa instrumen ini efektif untuk digunakan sebagai alat evaluasi dalam jangka waktu yang panjang. Oleh karena itu, peneliti merekomendasikan kepada guru untuk melakukan uji coba terhadap butir soal sebelum diberikan kepada siswa, guna menjamin kualitas soal yang akan diujikan. Dukungan langsung dari pihak sekolah juga penting untuk membantu meningkatkan kemampuan guru dalam menganalisis butir soal.

DAFTAR PUSTAKA

- Alfani, M. F. (2022). *Analisis Butir Soal Ulangan Akhir Semester Ganjil Dalam Mata Pelajaran Pancasila Kurikulum Merdeka Kelas Iv Sd/Mi*. 1(1), 35–44.
- Cahyaningrum, I. Y., Fuady, A., & Sunismi. (2023). Analisis Butir Soal Sumatif Akhir Semester Ganjil Mata Pelajaran Matematika Kelas Vii Dengan Berbantuan Aplikasi Software Anates. *Mathema Journal*, 5(2), 67–81.
- Fadli, R., Hidayati, S., Cholifah, M., Siroj, R. A., & Afgani, M. W. (2023). Validitas Dan Reliabilitas Pada Penelitian Motivasi Belajar Pendidikan Agama Islam Menggunakan Product Moment. *Jiip-Jurnal Ilmiah Ilmu Pendidikan*, 6(3).
- Farida, & Musyarofah, A. (2021). Validitas Dan Reliabilitas Dalam Analisis Butir Soal. *Al-Mu'arrib: Jurnal Pendidikan Bahasa Arab*, 1(1), 34–44.
- Ghozali, I. (2018). *Aplikasi Analisis Multivariate Dengan Program Ibm Spss 25 Edisi 9*. Badan Penerbit Universitas Diponegoro.
- Hayati, S., & Lailatussaadah, L. (2016). Validitas Dan Reliabilitas Instrumen Pengetahuan Pembelajaran Aktif, Kreatif Dan Menyenangkan (Pakem) Menggunakan Model Rasch. *Jurnal Ilmiah Didaktika*, 16(2).
- Hikmah, & Muslimah. (2021). Validitas Dan Reliabilitas Tes Dalam Menunjang Hasil Belajar Pai. *Proceedings*, 1(1), 345–356.
- Indah, M., Karoma, K., & Rusdi, A. (2021). Analisis Tes Butir Soal Guru Dalam Mata Pelajaran Pendidikan Agama Islam (Pai) Di Sekolah Menengah Pertama Negeri 8 Palembang. *Muaddib: Islamic Education Journal*, 4(1), 22.
- Khafidin, Z. (2014). Analisis Validitas Dan Realibilitas Tes Mata Pelajaran Pendidikan Agama Islam Tingkat Sma. *Edukasia : Jurnal Penelitian Pendidikan Islam*, 9(2), 260.
- Magdalena, I., Fauziah, S. N., Faziha, S. N., & Nupus, F. S. (2021). Analisis Validitas, Reliabilitas, Tingkat Kesulitan Dan Data Beda Butir Sial Ujian Akhir Semester Tema 7 Kelas Iii Sdb Jaret 1 Sepatan. *Bintang : Jurnal Pendidikan Dan Sains*, 3(2), 198–214.
- Muzaffir, A. (2016). Validitas Tes Dan Kualitas Butir Soal. *Lisanuna: Jurnal Bahasa Arab Dan Pembelajaran*, 05(01).
- Nursalam. (2003). *Konsep Dan Penerapan Metodologi Penelitian Ilmu Keperawatan (Pedoman Skripsi, Tesis Dan Insrumen Penelitian Keperawatan*. Salemba Medika.
- Pahri, E. D. (2021). *Validitas Dan Tingkat Reliabilitas Soal Buatan Pilihan Ganda Tes Sumatif Siswa Kelas Xi Akl 3 Smk Negeri 4*.
- Pramana, I. N. D., Putra, N. P. S. W., Phalguna Bg, K. W., & Nugraha, K. Y. (2012). Evaluasi Pendidikan. In *Jakarta: Pt Bumi Aksara*. Undhiksa.
- Purba, & Octobe, Y. (2021). *Teknik Uji Instrumen Penelitian Pendidikan*. Widina Bhakti Persada.

- 977 *Analisis Validitas dan Reliabilitas Butir Soal Sumatif Akhir Semester Ganjil Mata Pelajaran PAI - Muhammad Isa Anshari, Rodiah Nasution, Muhammad Irsyad, Alifia Zuhriatul Alifa, Indah Aminatus Zuhriyah*
DOI : <https://doi.org/10.31004/edukatif.v6i1.5931>
- Realita. (2015). Uji Validitas Tes Buatan Dosen Mata Kuliah Fiqh. *Jurnal Mudarrisuna*, 5(1), 87–90.
- 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), 62.
- Shodiyah, N., & Septi, B. S. (2014). *Asesmen Pembelajaran*. Umsida Press.
- 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), 195.
- Surapranata, S. (2006). *Analisis, Validitas, Reliabilitas, Dan Interpretasi Hasil Tes*. Remaja Rosdakarya.