

Analisis Instrumen Tes Penilaian Pendidikan Agama Islam Menggunakan Tes Teori Klasik

Boini ^{a,1,*}, Sahrul ^{b,2}, Asmanah ^{c,3}, Sumin ^{d,4}

^{a,b,c,d} Pascasarjana Institut Agama Islam Negeri Pontianak, Indonesia

¹boiniboy20@gmail.com; ²sahrul273@guru.sma.belajar.id; ³asmanahpptq144@gmail.com; ⁴amien.ptk@gmail.com

*Correspondent Author

KATAKUNCI

Karakteristik, Instrumen penilaian, Pendidikan Agama Islam

KEYWORDS

Characteristics, assessment instruments, Islamic religious education

ABSTRAK

Tujuan dari penelitian untuk mengetahui dan mendeskripsikan tentang Validitas, Reliabilitas, dan Daya beda. Tingkat kesukaran merupakan sebuah instrument yang penting dalam sebuah pendidikan agar data yang di hasilkan dapat di gunakan dengan baik. Validasi terbagi menjadi validasi suatu isi, konstruk. Validasi isi adalah sebuah insstrumen yang dapat dikatakan valid jika sesuai dengan kurikulum, materi dan tujuan dalam sebuah pembelajaran. Sedangkan validitas konstruk adalah validitas yang menggunakan pola untuk berpikir seberapa tercapainya butir tes dapat mengukur sesuatu yang akan di ukur dan disesuaikan dengan konsep yang telah di tetapkan. Reliabilitas ditentukan dengan nilai cronbach's alpha, apabila nilai cronbach's alpha sebesar 0,5 dapat digunakan untuk tujuan penelitian. Daya pembeda kemampuan membedakan kelompok atas dan kelompok bawah dalam bentuk soal. Tingkat kesukaran membedakan soal dari segi kesulitan.

Analysis of Islamic Religious Education Assessment Test Instruments Using Classical Theory Tests

The aim of the research is to find out and describe Validity, Reliability and Distinction. The level of difficulty is an important instrument in education so that the resulting data can be used well. Validation is divided into validation of a content, construct. Content validation is an instrument that can be said to be valid if it is in accordance with the curriculum, material and objectives of a lesson. Meanwhile, construct validity is validity that uses patterns to think about how achieved test items can measure something that will be measured and adjusted to the concept that has been determined. Reliability is determined by the Cronbach's alpha value, if the Cronbach's alpha value is 0.5 it can be used for research purposes. Discriminating power is the ability to differentiate between upper and lower groups in the form of questions. The level of difficulty differentiates questions in terms of difficulty.

This is an open-access article under the CC-BY-SA license.



Pendahuluan

Sistem Pendidikan nasional yang tercantum dalam undang-undang nomor 20 tahun 2003 menjelaskan bahwa pendidikan tidak hanya membentuk peserta didik yang cerdas tetapi membentuk karakter (Ratnawatiningsih & Hastuti, 2022) atau kepribadian siswa sehingga lahir generasi yang berakhlak serta berakarakter (Tri & Sofiyatul, 2022) yang selaras dengan nilai-nilai bangsa dan agama (Zakarya, Hafidz, Martaputu, 2023). Menurut (Ali, 2022)

sebagai tenaga profesional, guru berkewajiban dalam Merencanakan pembelajaran, melaksanakan proses pembelajaran yang bermutu, serta menilai dan mengevaluasi hasil penilaian.

Sedangkan (Afifatun, 2018) berpendapat Evaluasi merupakan bagian penting dan tak terpisahkan dalam sistem pendidikan saat ini. Peningkatan kualitas pendidikan dapat dilihat dari nilai-nilai yang diperoleh siswa (Rhain et al., 2023). Tentu saja untuk itu diperlukan sistem penilaian yang baik dan tidak bias. Adanya hubungan antara pelajaran juga engagement siswa sangat dibutuhkan agar proses pembelajaran lebih efektif (Sunawan, 2017).

Dalam Alqur'an surah Al-ankabut ayat 2-3 menjelaskan tentang bahwa perlunya mengadakan evaluasi.

أَحْسِبِ النَّاسُ أَنْ يُتْرَكُوا أَنْ يَقُولُوا آمَنَّا وَهُمْ لَا يُفْتَنُونَ وَلَقَدْ فَتَنَّا الَّذِينَ مِنْ قَبْلِهِمْ فَلَيَعْلَمَنَّ اللَّهُ الَّذِينَ صَدَقُوا وَلَيَعْلَمَنَّ الْكَاذِبِينَ .

Terjemahan Kemenag (Depag, 2007) "Apakah manusia mengira bahwa mereka akan dibiarkan (hanya dengan) berkata, "Kami telah beriman," sedangkan mereka tidak diuji? Sungguh, Kami benar-benar telah menguji orang-orang sebelum mereka. Allah pasti mengetahui orang-orang yang benar dan pasti mengetahui para pendusta".

Berdasarkan ayat Alqur'an diatas sang penciptaka memulai ujian, latihan atau evaluasi kepada semua makhluknya (Rhain et al., 2023). Mengukur potensi keimanan mereka dalam khazanah knowledge (Ngabdul Shodikin et al., 2023), evaluasi tersebut diadakan untuk mengukur sampai mana pencapaian yang telah di capai siswa dalam pemebelajaran dan melihat kompon tersebut (Robbaniyah, 2022). (Daryanto, 2013) memberikan suatu komentar mengenai TP (Tujuan Penilaian) : (a) Mendeskripsikan kemampuan belajar para peserta didik sehingga dapat diketahui positif dan negatifnya. (2) Melihat keberhasilan proses pendidikan dan pengajaran disekolah, yakni seberapa jauh keefektifannya dalam mengubah tingkah laku para peserta didik (Novita et al., 2022). Alat ukur yang disebut instrument adalah alat ukur yang telah melalui uji validitas dan reliabilitas data. Beberapa faktor yang mempengaruhi validitas dan reliabilitas suatu data tergantung pada baik tidaknya instrumen pengumpulan data atau pengukur objek dari suatu variable penelitian. Baik tidaknya suatu instrumen ditentukan oleh validitas dan reliabilitasnya (Zogara, 2021). Senada dengan yang di ungkapkan (Zogara, 2021) kelayakan status validitas dan rehabilitas memiliki beberapa data dan syarat tertentu.

Kenyataannya di lapangan, instrumen penilaian yang digunakan selama ini belum diuji validitas dan reliabilitasnya instrumen tersebut sudah layak standarilisasi, Sehingga validitas instrument sangatlah penting (Hariyati T, 2019). Secara konstruk atau validasi konstruk dan secara kriteria dengan validasi kriteria (Yusup, 2018). Dalam prosedur Pendidikan serta pembelajaran, sistem kualitas bahan mengajar mengutamakan profesionalisme pendidik dan prestasi siswa.

Proses Pendidikan dan pembelajaran membutuhkan penggabungan antara komponen-komponen pembelajaran (Pratama et al., 2022) yang besinergi untuk membentuk kompleksitas disiplin ilmu baik dalam hal kognisi, emosional dan psikomotorik (Jatmiko Wibisono, Hafidz, Iffat Abdul Ghalib, 2023). Penerapan pembelajaran PAI menitikberatkan pada keistimewaan dan pada aspek keimanan, akhlak dan ibadah. Semua ranah penelitian PAI didasarkan pada AlQur'an dan hadis (Robbaniyah et al., 2022). Selanjutnya, pendidik perlu memiliki interpretasi yang baik tentang metode pembelajaran PAI yang tepat untuk memastikan pendidik tidak melakukan kesalahan dalam melaksanakannya (Sulaiman, 2017). Rangkaian pembelajaran di hubungkan sesuai dengan konteks pembelajaran dan didasarkan hasil akhir. Kemudian, hasil siswa-siswa sebagai bahan evaluasi untuk mengintegrasikan interpretasi siswa dalam materi diajarkan. Pendidikan yang nyata, numerik adalah hasil belajar siswa. Maka jawabannya ialah dengan cara uji validitas karna menghasilkan pengukuran, berbeda dengan uji realiabilitas yang mengukur sejauh mana instumremn yang daapat dipercaya.

Metode

Metode pada artikel ini menggunakan studi pustaka (library research) (Syaiful Anam, 2023) yaitu metode dengan pengumpulan data dengan cara memahami dan mempelajari teori dari literatur yang berhubungan dengan penelitian tersebut. Pengumpulan data tersebut menggunakan cara mencari sumber dan mengkontruksikan dari berbagai sumber contohnya seperti buku, jurnal, riset yang sudah pernah dilakukan. Selanjutnya peneliti menggunakan Aplikasi Software yaitu JASP untuk mendapatkan hasil penelitian yang valid sesuai arah dan tujuan penelitian ini.

Hasil dan Pembahasan

1. Validitas

Validitas isi

Validitas isi ditentukan menggunakan kesepakatan ahli. Kesepakatan ahli bidang studi atau sering digunakan dalam menentukan validitas isi (Retnawati, 2014) digunakan atau konstruk psikologi yang diukur.

Table 1. Kriteria Penilaian Butir Instrumen oleh Validator

Nilai	Keterangan
1.	Tidak Relevan
2.	Kurang Relevan
3.	Cukup
4.	Relevan
5.	Sangat Relevan

Tabel 2. Kriteria Validatas Isi

Koefisien Validas Isi	Kriteria
0.00-0.19	Sangat Rendah
0.20-0.39	Rendah
0.40-0.59	Sedang
0.60-0.79	Tinggi
0.80-1.10	Sangat Tinggi

Dari dua teori diatas peneliti menyimpulkan bahwa validitas isi adalah sebuah instrumen yang dapat dikatakan valid jika sesuai dengan kurikulum, materi dan tujuan.

2. Validitas konstruk

Menurut (Sugiyono, 2019) membutuhkan sekitar 30 anggota sampel yang digunakan , untuk pengumpulan data, menganalisis dan mengkorelasikan skor item untuk melakukan validitas konstruk dengan mamakai rumus produks momen: (Sugiyono, 2019).

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

Yaitu:

r = koefesien korelasi

n = total skor item instrument

$\sum X_i$ = keseluruhan skor jawaban

$\sum Y_i$ = hasil kuadrat skor item

$\sum Y_i^2$ = hasil kuadrat skor total jawaban

$\sum X_i Y_i$ = berapa kali skor item jawaban di kalikan dengan skor total

jika $r > 0,3$, persyaratan minimum dianggap dapat diterima jika korelasi kurang dari 0,3, item dalam perangkat tersebut dikatakan tidak valid (Sugiono, 2016). Dari dua teori diatas peneliti menyimpulkan bahwa validitas konstruk adalah validitas yang mempermasalahkan seberapa jauh butir-butir tes mampu mengukur apa yang benar-benar hendak diukur sesuai dengan konsep khusus atau definisi konseptual yang telah ditetapkan.

3. Reliabilitas

Reliabilitas instrumen

Reliabilitas mengacu pada suatu pengertian bahwa instrumen cukup dapat dipercaya untuk digunakan sebagai alat pengumpul data karena instrumen tersebut sudah baik. Reliabilitas juga mengacu pada tingkat kredibilitas sesuatu dan dapat dipercaya (Arikunto, 2013).

TABEL 3. DERAJAT RELIABILITAS

Koefisien Reliabilitas	Interpretasi
$0,90 < \leq 1,00$	Sangat tinggi
$0,70 < \leq 0,90$	Tinggi
$0,40 < \leq 0,70$	Sedang
$0,20 < \leq 0,40$	Rendah
$R \leq 0,20$	Sangat rendah

Menurut (Basuki, 2014) reliabel instrumen adalah jika nilai Cronbach's alpha 0,60 dan kurang dari 1, maka instrumen tersebut memiliki korelasi tinggi atau reliabel, sedangkan jika nilai Cronbach's alpha di bawah 0,5 ke bawah, maka instrumen tersebut berkecenderungan rendah atau tidak reliabel. Dan hal ini juga didukung oleh (Surapranata, 2004) bahwa koefisien reliabilitas sebesar 0,5 dapat digunakan untuk tujuan penelitian. Dari dua teori diatas peneliti menyimpulkan bahwa reliabilitas instrumen adalah ditentukan dengan nilai Cronbach's alpha, apabila nilai Cronbach's

4. Daya pembeda

Daya pembeda soal

Daya beda butir soal merupakan suatu indeks yang menunjukkan tingkat kemampuan butir soal membedakan kelompok atas (peserta tes yang berprestasi tinggi) dari kelompok bawah (peserta tes yang berprestasi rendah). (Elviana, 2020) Dari teori diatas peneliti menyimpulkan bahwa daya pembeda soal adalah kemampuan soal untuk membedakan kelompok peserta tes yang paham dan kelompok peserta tes yang kurang paham. Menurut (Arikunto, 2013) dalam buku Evaluasi Pembelajaran (Widiyanto, 2018) Klasifikasi Daya Beda Instrumen.

TABEL 4. KLASIFIKASI DAYA BEDA

NILAI D	KLASIFIKASI
0,00-0,20	Jelek (<i>poor</i>)
0,20-0,40	Cukup (<i>satisfactory</i>)
0,40 - 0,70	Baik (<i>good</i>)
0,70 - 1,00	Baik sekali (<i>excellent</i>)
0,00	Soal tidak bisa dipakai

Sejalan dengan pendapat (Elviana, 2020) Koefisien daya beda berkisar antara -1,00 sampai dengan +1,00, nilai positif saja yang mempunyai arti semakin besar indeks daya beda, maka semakin mendekati 1 maka butir tersebut mampu membedakan kelompok atas dan kelompok bawah. Butir yang dianggap mempunyai daya beda yang baik apabila butir tersebut $> 0,5$.

5. Tingkat kesukaran

Tingkat kesukaran butir soal

Tingkat kesukaran butir soal adalah persentase atau proporsi dari peserta tes untuk menjawab benar suatu butir soal. Besarnya tingkat kesukaran berkisar 0,00 - 1,00. Semakin besar tingkat kesukaran yang diperoleh dari hasil hitungan, berarti semakin mudah soal itu dan soal itu harus direvisi. Suatu soal memiliki $p = 0,00$ artinya bahwa tidak ada siswa yang menjawab benar dan bila memiliki $p = 1,00$ artinya bahwa semua siswa menjawab benar (Arifin, 2017). Soal yang sukar akan membuat murid tidak bersemangat karena diluar jangkauannya, sedangkan jika soal yang terlalu gampang akan membuat murid tidak tertarik untuk mengerjakan. Menurut ketentuan yang sering diikuti, indeks kesukaran sering diklasifikasikan sebagai berikut:

TABEL 5. PREDIKAT TINGKAT KESUKARAN INSTRUMEN

NILAI K	Predikat
0,00 – 0,30	Sukar
0,30 – 0,70	Sedang
0,71 – 1,00	Mudah

Walaupun demikian ada yang berpendapat bahwa soal-soal yang dianggap baik, yaitu soal-soal sedang yaitu soal-soal yang mempunyai indeks kesukaran 0,30 – 0,70. Perlu diketahui bahwa soal-soal yang terlalu mudah atau terlalu sukar tidak berarti tidak boleh digunakan. Hal ini tergantung kegunaannya.

Simpulan

Dalam hal ini (*validitas, Reliabilitas, Daya beda, Tingkat kesukaran*) merupakan sebuah instrument yang sangat penting dalam sebuah pendidikan supaya data yang di hasilkan dapat di gunakan dengan baik.

- 1) Validasi terbagi menjadi validasi isi dan konstruk. Validasi isi merupakan sebuah instrumen yang dapat dikatakan valid jika sesuai dengan kurikulum, materi dan tujuan. Sedangkan validitas konstruk Adalah Validitas yang berpikir
- 2) seberapa tercapainya butir tes dapat mengukur sesuatu yang akan di ukur dan disesuaikan dengan konsep khusus yang telah di tetapkan.
- 3) Reliabilitas instrument ditentukan dengan nilai cronbach's alpha, apabila nilai cronbach's alpha sebesar 0,5 dapat digunakan untuk tujuan penelitian.
- 4) Daya pembeda kemampuan soal untuk membedakan kelompok peserta tes yang faham dan kelompok peserta tes yang kurang paham.
- 5) Tingkat kesukaran menyimpulkan bahwa soal yang sukar akan membuat murid tidak bersemanagat karena diluar jangkauannya, sedangkan jika soal yang terlalu gampang akan membuat murid tidak tertarik untuk mengerjakan.

Jadi jika instrumen penilaian dari ke empat konsep didalam masuk kriteria baik, jadi instrumens yang dianalisa oleh pengkaji mempunyai kapasitas elok.

Daftar Pustaka

- Afifatun. (2018). *Perencanaan Sistem Evaluasi Pendidikan*. CV. LADUNY ALIFATAMA.
- Ali, M. (2022). Optimalisasi Kompetensi Kepribadian Dan Profesionalisme Guru Pendidikan Agama Islam dalam Mengajar. *Ar-Rusy*, 100-121.
- Allen. (1979). *Introduction to measurement theory*. Mexico: Cole Publishing Company.
- Arikunto. (2013). *Dasar-Dasar Evaluasi Pendidikan*. Jakarta: Bumi Aksara.
- Basuki, H. (2014). *Asesemen Pembelajaran*. Bandung: PT.Remaja Rosdakarya.
- Daryanto, A. (2013). *Evaluasi dan Penilaian Pembelajaran Kurikulum*. Yogyakarta: Gava Media.
- Depag. (2007). *Mushaf Alquran dan Terjemahan*. Jakarta: Alquran Pentashihan Departemen Agama RI.
- Elviana. (2020). ANALISIS BUTIR SOAL EVALUASI PEMBELAJARAN PAI PROGRAM ANATES. *Jurnal Mudarrisuna*, 58-74.
- Erlinawati, M. (2021). Test Validity and Reliability In Learning Evaluation. *Bulletin of Community Engagment*, 1-7.
- Hariyati T, Z. (2019). VALIDITY AND RELIABILITY OF PERFORMANCE ASSESSMENT. *JURNAL LITERASIOLOGI*, 100-107.

- Hikmah, M. (2021). Validitas dan Realibitas Tes dalam menunjang Hasil Belajar PAI. Vol I.
- Jatmiko Wibisono, Hafidz, Iffat Abdul Ghalib, H. N. (2023). Konsep Pemikiran Pembaharuan Muhammadiyah Bidang Pendidikan (Studi Pemikiran Muhammad Abduh). *Attractive: Innovative Education Journal*, 5(2), 514–522.
- Khafidin. (2014). Analisis Validitas Dan Reliabilitas Tes Mata Pelajaran Pendidikan Agama Islam Tingkat SMA. *Jurnal Penelitian Pendidikan Islam*, 253–266.
- Ngabdul Shodikin, E., Sucipto, E., Wasith Achadi, M., Muzaky, F., Wahyu Laras Pertiwi, R., & Tinggi Ilmu Tarbiyah Madani Yogyakarta, S. (2023). *Implementation of the Sabaq, Sabqi, Manzil Methods in Improving the Quality of Memorizing Qur'an Learning in Class V Salafiyah Ula Islamic Center Bin Baz Bantul*. 1(1), 34–44. <https://journal.amorfati.id/index.php/postaxial>
- Novita, M., Zakki, M., & Inayati, N. L. (2022). Implementasi Pendidikan Moral Dalam Membina Perilaku Siswa Di Sekolah Menengah Atas Islam Terpadu Al Huda. *JIPSI: Jurnal Ilmu Pendidikan Dan Sains Islam Interdisipliner*, 2(1), 95–105.
- Mardapi. (2016). *Pengukuran Penilaian Dan Evaluasi Pendidikan*. Yogyakarta: UIN Yogyakarta.
- Pratama, N., Syafii, M., Qur, A.-, Arab, B., Arab, B., & Umar, S. (2022). Problematika Pembelajaran Mahasiswa Lulusan Sekolah Umum pada Program Studi Pendidikan Bahasa Arab di Sekolah Tinggi Ilmu Tarbiyah Madani Yogyakarta. *JIPSI: Jurnal Ilmu Pendidikan Dan Sains Islam Interdisipliner*, 1(2), 117–124.
- Ratnawatiningsih, E., & Hastuti, A. P. (2022). Penanaman Karakter Anak Usia Dini Melalui Program Pembiasaan Pola Hidup Bersih dan Sehat di RA Miftahul Falah Gondosuli. *JIPSI: Jurnal Ilmu Pendidikan Dan Sains Islam Interdisipliner*, 1(1), 44–52.
- Rhain, A., Hafidz, Nashihin, H., Srihananto, T. H., & Hermawati, T. (2023). Tahsin Reading Assistance for Islamic Boarding School Tahfidz Qur'an Muhammadiyah Daarul Arqom Sawahan Ngemplak Boyolali. *Jurnal Pengabdian Masyarakat Bestari*, 2(1), 27–44. <https://doi.org/10.55927/jpmb.v2i1.2729>
- Robbaniyah, Q. (2022). Eksplorasi Strategi Kontra Radikalisme pada Santri di Pondok Pesantren Islamic Center Bin Baz Yogyakarta. *AMORTI: Jurnal Studi Islam Interdisipliner*, X(X), 1–10.
- Robbaniyah, Q., Lina, R., Ustadz, S., Rofiq, A., Islami, F. Al, & Faiz, A. (2022). Kontribusi Pemikiran Abu Nida ` dalam Pengembangan Pendidikan Islam Pondok Pesantren di Indonesia. *JIPSI: Jurnal Ilmu Pendidikan Dan Sains Islam Interdisipliner*, 1(1), 23–34.
- Retnawati. (2014). *Membuktikan validitas Instrumen Pengukuran*. Jakarta: UI Press.
- Sugiyono. (2019). *Statistika Untuk Penelitian*. Bandung: Alfabete.
- Sulaiman. (2017). *Metodologi Pembelajaran PAI*. Banda Aceh: Yayasan Pena Banda Aceh.
- Sunawan, D. H. (2017). Classroom Engagement. *Atlantis Press*, 50-82.
- Sundayana. (2016). *Statistika Penelitian Pendidikan*. Bandung: Alfabeta.
- Surapranata. (2004). *Analisis, validitas, reliabilitas, dan interpretasi hasil tes implementasi kurikulum 2004*. Bandung: PT. Remaja Rosdakarya.
- Syaiful Anam, H. N. (2023). *Metode Penelitian (Kualitatif, Kuantitatif, Eksperimen, dan R&D)*. PT GLOBAL EKSEKUTIF TEKNOLOGI. https://www.google.co.id/books/edition/Metode_Penelitian_Kualitatif_Kuantitatif/w-bFEAAAQBAJ?hl=id&gbpv=0
- Suryadi, A. (2020). *Evaluasi Pembelajaran*. Bandung: CV. Jejak.
- Tri, S., & Sofiyatul, A. (2022). Implementasi Pendidikan Karakter Religius Remaja melalui Kegiatan Rutin Pembacaan Kitab Maulid Diba ' di Desa Dangkel Kecamatan Parakan Kabupaten Temanggung. *JIPSI: Jurnal Ilmu Pendidikan Dan Sains Islam Interdisipliner*, 1(2), 84–94.
- Widiyanto. (2018). *Evaluasi Pembelajaran*. Madiun: UNIPMA PRESS.
- Yusup. (2018). Uji Validitas Dan Reliabilitas Instrumen Penelitian. *Jurnal Ilmiah Kependidikan*, 1-7.
- Zakarya, Hafidz, Martaputu, H. N. (2023). Peran Guru Pendidikan Agama Islam dalam Meningkatkan Motivasi Belajar peserta didik di SMA Muhammadiyah 1 Surakarta. *Pendidikan Agama Islam (PAI) Dalam Kurikulum Nasional*, 2(2), 1–13.
- Zogara, A. u. (2021). *MEMAHAMI DASAR-DASAR PENELITIAN KESEHATAN*. In Haura Publishing.