

Research Article

Analisis Butir Soal Tes Formatif Mata Pelajaran Akidah Akhlak Kelas VI MIN 2 Kubu Raya 2022

Nurhayati¹, Sukino², Zulkifli³

1. Mahasiswa prodi Magister Pendidikan Agama Islam Pascasarjana IAIN Pontianak, umynurhayati76@gmail.com
2. Prodi Magister PAI Pascasarjana IAIN Pontianak, ariefsukino@yahoo.co.id
3. FTIK IAIN Pontianak, zulkifli.abdillah@ymail.com

Copyright © 2023 by Authors, Published by Risalah: Jurnal Pendidikan dan Studi Islam. This is an open access article under the CC BY License (<https://creativecommons.org/licenses/by-sa/4.0>).

Received : June 30, 2023

Revised : August 7, 2023

Accepted : August 16, 2023

Available online : September 3, 2023

How to Cite: Nurhayati, Sukino, and Zulkifli. 2023. "Analisis Butir Soal Tes Formatif Mata Pelajaran Akidah Akhlak Kelas VI MIN 2 Kubu Raya 2022". *Risalah, Jurnal Pendidikan Dan Studi Islam* 9 (3):997-1006. https://doi.org/10.31943/jurnal_risalah.v9i3.584.

Abstract. The purpose of this study was to analyze the quality level of the formative test items in the Akidah Akhlak lesson which consisted of the levels of item validity, reliability, difficulty, and item discrimination. This study uses a quantitative descriptive research method. The research was conducted at MIN 2 Kubu Raya, especially in class VI, using 10 multiple-choice questions. Data collection in this study used Excel and SPSS to analyze the data obtained. Based on the results of the analysis, you can find that the difficulty level of the questions is very difficult 0%, difficult 0%, moderate 0%, easy 2 questions 20%, and very easy 8 questions 80%. The next step is a revision in order to measure students' abilities. The level of validity of the questions is that there are 7 items out of 10 questions that are declared valid and 3 questions that are declared invalid. The results of this instrument reliability test were carried out on 49 students with a significance level of 5%, and degrees of freedom (df) = n-2 = 49 - 2 = 47, so it was found that the value of $r_{11} = 0.5145$ and smaller than the value of r_{11} according to the provisions proposed by Heale and Twycross, namely 0.70 or in other words, r_{11} count is less than 0.70 ($0.659 > 0.70$), which means that the question is not reliable and must be corrected before it is used to measure students' ability to understand the subject matter of Akidah Akhlak.

Keywords : Validity; Reliability; Level of Difficulty, Power of Distinction, Akidah Akhlak.

Abstrak. Tujuan penelitian ini adalah untuk menganalisis tingkat kualitas soal tes formatif pada pembelajaran Akidah Akhlak yang terdiri dari tingkat validitas soal, reliabilitas, kesukaran, dan daya diskriminasi soal. Penelitian ini menggunakan metode penelitian deskriptif kuantitatif. Penelitian ini dilaksanakan di MIN 2 Kubu Raya khususnya kelas VI dengan menggunakan 10 soal pilihan ganda. Pengumpulan data pada penelitian ini menggunakan Excel dan SPSS untuk menganalisis data yang

diperoleh. Berdasarkan hasil analisis diketahui tingkat kesukaran soal sangat sulit 0%, sukar 0%, sedang 0%, mudah 2 soal 20%, dan sangat mudah 8 soal 80%. Langkah selanjutnya adalah revisi guna mengukur kemampuan siswa. Tingkat validitas soal yaitu terdapat 7 butir pertanyaan dari 10 soal yang dinyatakan valid dan 3 soal yang dinyatakan tidak valid. Hasil uji reliabilitas instrumen ini dilakukan terhadap 49 siswa dengan taraf signifikansi 5%, dan derajat kebebasan ($df = n - 2 = 49 - 2 = 47$) sehingga diperoleh nilai $r_{11} = 0,5145$ dan lebih kecil dari nilai r_{11} menurut ketentuan yang dikemukakan oleh Heale dan Twycross yaitu 0,70 atau dengan kata lain r_{11} hitung kurang dari 0,70 ($0,659 > 0,70$) yang berarti pertanyaan tersebut tidak reliabel dan harus diperbaiki terlebih dahulu. digunakan untuk mengukur kemampuan siswa dalam memahami materi pelajaran Akidah Akhlak.

Kata Kunci: Validitas; Keandalan; Tingkat Kesulitan, Daya Pembeda, Akidah Akhlak.

PENDAHULUAN

Krisis moral yang terjadi di bangsa ini menjadi suatu kegelisahan untuk semua kalangan. Maraknya berbagai kasus kriminal seperti korupsi yang tidak pernah surut dan terus mengalami peningkatan dari waktu ke waktu. Misalnya banyaknya krisis yang kompleks dari berbagai kejadian yang sangat memilukan seperti tawuran pelajar, penyalahgunaan obat terlarang, pergaulan bebas, aborsi, penganiayaan dan pembunuhan. Kejadian ini sungguh sangatlah bertentangan dengan keadaan keagamaan serta kepribadian bangsa Indonesia. Apabila krisis ini terus dibiarkan maka akan terus berlarut dan dianggap suatu hal yang biasa dan akan membudaya di era saat ini. Sekecil apapun krisis moral secara tidak langsung bisa memberikan kerapuhan untuk nilai hidup berbangsa serta bernegara (M. Turmudi 2018).

Manusia hidup selalu memiliki rencana untuk mencapai suatu yang lebih baik. Dan untuk mengetahui tingkat keberhasilan suatu rencana maka diperlukan suatu instrument sebagai alat ukur (Sardiyanah, 2020). Begitu juga dalam pembelajaran guru memiliki rencana capaian pembelajaran atas kinerjanya pada tiap pertemuan hingga dilakukan tes sumatif dan formatif. Pengukuran keberhasilan dapat dilakukan dalam berbagai bentuk sesuai dengan tujuan atau kebutuhan masing-masing. Adapun prasyarat yang penting dilakukan dalam melakukan pengukuran hasil kerja individu maupun kelompok adalah menguji validitas dan reliabilitas instrument yang digunakan, dalam hal ini telah dilakukan penelitian oleh (Rahmaini & Taufiq, 2018).

Pengukuran merupakan bagian yang penting dalam proses pendidikan, pengukuran menjadi dasar mengambil keputusan yang mampu merubah sikap peserta didik dalam proses pembelajaran (Sukino, 2015). sehingga setiap pendidik (guru di sekolah formal), harus memiliki kompetensi pedagogik seperti kemampuan dalam menyusun instrument tes yang valid dan reliabel (Lester et al., 2014). Kemampuan tersebut tidak saja untuk memenuhi administrasi sekolah seperti tersedianya soal, namun jauh lebih penting bahwa dengan instrument soal yang bermutu akan berdampak signifikan terhadap kemampuan berpikir kritis siswa dalam menyelesaikan masalah.

Menyusun instrument tes perlu mengikuti ketentuan, jadi tidak boleh sesuai dengan keinginan pembuatnya, instrument tes harus menggunakan Bahasa yang baku dan jumlahnya juga proporsional. Dalam menyusun instrument tes juga berdasarkan pada tujuan dari pembelajaran, baik bersifat kognitif, afektif maupun psikomotorik. Berdasarkan pengalaman di lapangan masih terdapat kesalahan penyusunan tes oleh guru bidang studi, termasuk guru Pendidikan agama Islam. Menurut pendapat Mardapi kesalahan yang sering ditemukan adalah kesalahan acak dan kesalahan

sistematis hal ini telah dibuktikan oleh penelitian (Cahyaningrum, 2019) dan (Purwanto, 2010).

Penyusunan soal tes dalam satu kali pengukuran harus mampu mengukur keseluruhan kemampuan siswa, dan soal tes tersebut dapat memberikan informasi tentang perbedaan individu. Karena tujuan pengukuran adalah melihat derajat ketuntasan dalam belajar, melalui hasil tes tersebut pihak yang berkepentingan seperti guru kelas, guru mata pelajaran, dan juga orang tua sangat membutuhkan informasi ketuntasan belajar anak didiknya.

Ketuntasan belajar memang diharapkan oleh guru dan orang tua, namun dalam pencapaiannya harus melalui proses yang bermutu, salah satunya adalah dengan menggunakan soal yang berkualitas. Tidak terlalu sukar dan tidak terlalu mudah, karena apabila para peserta didik diberikan tes yang terlalu sukar maka mereka akan menyerah dalam mengerjakannya sehingga membuat tes itu tidak efektif diberikan kepada peserta didik, sebaliknya apabila tes yang diberikan guru terlalu mudah maka hasil dari tes tersebut tidak dapat mengukur kemampuan peserta didik secara efektif.

Maka dari itu perlu diadakannya uji terhadap bentuk soal yang akan diberikan kepada peserta didik. Di antara uji terhadap bentuk soal tersebut dapat dilakukan dengan cara uji tingkat kesukaran butir soal, Uji Daya Beda Butir Soal, Uji Validitas dan Uji Reabilitas pada setiap soal yang akan diberikan kepada peserta didik. Tujuannya adalah supaya soal yang diberikan kepada peserta didik itu dapat benar-benar mengukur kemampuan siswa dalam mata pelajaran tersebut, dan juga agar tercapainya tujuan dalam pendidikan di Indonesia.

Tujuan dari kegiatan analisis soal adalah sebagai berikut: Jawaban atas pertanyaan-pertanyaan dan merupakan data diagnostik yang dapat digunakan untuk menganalisis pembelajaran di kelas, memberikan informasi akurat tentang kelemahan dan kekuatan soal. Menurut Thorndik dan Hagnr dikutip oleh (Purwanto, 2010) analisis butir soal memberi masukan pada guru dalam menyusun strategi meningkatkan hasil belajar dalam berbagai aspek pada pembelajaran berikutnya. Mengarahkan persiapan yang lebih baik untuk ujian tahun depan dengan meninjau kembali pertanyaan dan jawaban ujian tahun sebelumnya dan merevisi pertanyaan-pertanyaan bermasalah berdasarkan tinjauan tersebut.

Berdasarkan paparan di atas maka setiap soal sebaiknya dilakukan analisis agar memberikan dampak terhadap kemampuan mereka dan dapat mengembangkan pengetahuan mereka dalam memahami mata pelajaran tersebut. Atas dasar itu, peneliti dalam hal ini akan membahas tentang kajian isi tes yang berkaitan dengan analisis kualitas butir soal yang akan ditinjau dari tingkat validitas, reliabilitas, kesukaran soal, tingkat daya beda.

METODE PENELITIAN

Penelitian ini merupakan penelitian analisis deskriptif, dengan menggunakan metode kuantitatif. Penelitian ini dilakukan di MIN 2 Kubu Raya pada kelas VI. Penelitian ini bertujuan untuk mendeskripsikan karakteristik butir soal dan besaran estimasi kesalahan dalam pengukuran pada saat dilakukan ujian akhir semester ganjil pada mata pelajaran Pendidikan Akidah Akhlak. Teknik analisis data pada penelitian ini menggunakan bantuan Excel dan juga SPSS versi 23. Penelitian ini dilakukan secara objektif untuk mengukur tingkat 1). Tingkat Kesukaran, 2). Daya

Beda, 3). Validitas Soal, 4). Reabilitas Soal. Soal yang dianalisis berupa soal tes formatif dengan 10 butir soal dan berbetuk pilihan ganda. Subjek penelitian ini adalah 49 siswa kelas VI di MIN 2 Kubu Raya. Objek penelitian adalah dokumen soal yang dibuat oleh Guru dan lembar jawaban siswa serta kunci jawabannya.

HASIL DAN PEMBAHASAN

Validitas Butir Soal

Validitas adalah analisis yang dilakukan untuk mengukur kelayakan soal yang akan diujikan atau diberikan kepada peserta didik, validitas juga dapat dikatakan sebagai kecermatan suatu instrumen dalam pengukuran kevalidan butir soal yang akan diujikan. Validitas memiliki tujuan untuk mengukur instrumen soal tersebut apakah layak atau tidak layak untuk diujikan (Purwati et al., 2021).

Jadi betapa penting melakukan validitas terhadap soal yang akan diujikan atau diberikan kepada peserta didik agar soal yang diberikan kepada peserta didik benar-benar dapat mengukur kemampuannya dalam mengukur kemampuannya memahami materi tersebut. Berdasarkan output dari SPSS 23. Dapat sajikan data sebagai berikut:

Tabel 1: Hasil Analisis Validitas Butir Soal Akidah Akhlak Kelas VI MIN Kubu Raya

No Soal	Person Corelation	Nilai Sig	Kesimpulan	Interpretasi
Soal_1	0.039	0.791	Tidak valid	
Soal_2	0.721	0.000	Valid	Tinggi
Soal_3	0.625	0.000	Valid	Tinggi
Soal_4	0.414	0.003	Valid	Rendah
Soal_5	0.039	0.791	Tidak Valid	
Soal_6	0.332	0.020	Valid	Rendah
Soal_7	0.000	0.000	Valid	Sangat Rendah
Soal_8	0.185	0.202	Tidak Valid	
Soal_9	0.684	0.000	Valid	Sangat Rendah
Soal_10	0.587	0.000	Valid	Sangat Rendah

Setelah dilakukan analisis validitas butir soal, maka dapat disimpulkan jika nilai sig. <0.05, maka instrumen soal tersebut dapat dikatakan Valid. Adapun penjelasan tentang interpretasi nilai validitas yaitu sebagai berikut :

- 0,800-1,000 : Sangat Tinggi
- 0,600-0,799 : Tinggi
- 0,400-0,599 : Cukup
- 0,200-0,399 : Rendah
- 0,000-0,199 : Sangat Rendah

Berdasarkan hasil uji validitas terhadap data analisis kevalidan butir soal yang telah dibuat guru Akidah Akhlak di MIN 2 Kubu Raya, maka dapat disimpulkan bahwa dari 10 soal tes formatif Akidah Akhlak Kelas VI MIN 2 Kubu Raya terdapat 7 soal yang dinyatakan valid dan 3 soal yang dinyatakan tidak valid, dan hasil analisis kevaliditasan masuk kategori cukup. Berdasarkan data tersebut maka dapat dianalisis bahwa tiga soal yang tidak valid dapat direvisi. Soal tidak valid biasanya dipengaruhi

oleh struktur bahasa yang digunakan sulit dipahami siswa, tingkat kesulitab butir tidak sesuai dengan materi yang diajarkan dan jawaban masing-masing butir soal dapat diprediksi oleh siswa (Purwanto, 2010) Sebaiknya guru perlu melakukan uji coba terlebih dahulu sebelum digunakan. Jadi tes yang validitasnya tinggi adalah tes yang dapat menjalankan fungsi pengukuran dan memberikan hasil ukur yang sesuai dengan tujuan melakukan tes. (Purwanto, 2010).

Reabilitas Butir Soal

Reabilitas adalah perhitungan estimasi yang dilakuakn untuk menganalisis konsistensi didalam soal yang telah dibuat. Reabilitas juga dapat mengukur apakah soal tersebut layak digunakan ataukah perlu diulang (Purwati et al., 2021).

Maka dari itu uji reabilitas ini sangat menentukan kualitas soal yang telah dibuat guru, apakah soal tersebut konsisiten apa tidak. Soal yang memiliki kekonsistenan yang baik maka soal tersebut dapat dikatan sebagai soal yang layak diberikan kepada peserta didik, sebaliknya soal yang memiliki kekonsistenan yang kurang baik maka soal tersebut perlu dilakukan pengulangan dalam pembuatannya. Berdasarkan hasil uji dengan SPSS versi 23. Didapatkan hasil sebagai berikut:

Tabel 3 Hasil Analisis Validitas Butir Soal Akidah Akhlak Kelas VI MIN Kubu Raya

Cronbach's Alpha	N of Items
.659	11

Dari tabel diatas dapat dirujuk dengan klasifikasi reliabilitas sebagai berikut:

Tabel 2 Nilai Koefisien Korelasi

Tes Reliabilitas	Kategori	Koefisien Korelasi
Sangat Tinggi		0,800 – 1,000
Tinggi		0,600 – 0,799
Cukup		0,400 – 0,599
Rendah		0,200 – 0,399
Sangat Rendah		0,000- 0,199

Pengujian reabilitas instrumen ini dilakukan terhadap 49 orang siswa kelas VI di MIN 2 Kubu Raya dengan tingkat signifikansi 5%, dan derajat kebebasan (df) = $n-2=49-2=47$, sehingga didapatkan bahwa nilai $r_{11} = 0,5145$ dan lebih kecil dari nilai r_{11} menurut ketentuan yang dikemukakan Heale and Twycross yaitu sebesar 0,70 atau dengan kata lain r_{11} hitung lebih kecil dari 0,70 ($0,659 > 0,70$) yang berarti soal tersebut tidak reliabel, dan harus dilakukan perbaikan sebelum digunakan untuk mengukur kemampuan siswa dalam memahami mata pelajaran Akidah Akhlak.

Menurut Azwar hasil pengukuran dapat dipercaya apabila dalam beberapa kali penggunaan pengukuran terhadap subjek yang sama akan diperoleh hasil yang relatif sama, selama aspek yang diukur dalam diri subjek tidak atau belum berubah (Azwar, 2013) dengan demikian dapat ditegaskan bahwa derajat kepercayaan instrument tes

ditentukan oleh keajegan atau konsistensi hasil yang didapatkan meskipun dengan subjek yang berbeda. (Secolsky, 2018).

Tingkat Kesukaran Butir Soal

Tingkat kesukaran soal dalam sebuah tes adalah presentase jumlah siswa yang menjawab benar atau salah (Arikunto, 2010). Tingkat kesukaran soal adalah peluang menjawab benar atau salah suatu soal pada tingkatan kemampuan atau bisa dikatakan sebagai alat ukur untuk dapat mengetahui sebuah soal tersebut apakah tergolong mudah ataupun sukar, biasanya tingkat kesukaran soal dapat dilakukan dalam bentuk indeks (Purwati et al., 2021).

Untuk dapat mengetahui tingkat kesukaran butir soal yang akan diberikan kepada siswa maka diperlukan adanya analisis soal. Analisis tingkat kesukaran butir soal dilakukan agar soal yang diberikan kepada siswa dapat mengukur kemampuannya dalam memahami mata pelajaran tersebut. Pada tingkat kesukaran butir soal itu terbagi kedalam beberapa tingkatan yaitu: Mudah, Sedang, dan Sulit. Berdasarkan hasil jawaban yang telah dilakukan siswa, maka dapat dianalisis tingkat kesukaran pada setiap soal tes formatif Akidah Akhlak yang telah dibuat oleh guru yaitu sebagai berikut:

Tabel 4 Analisis Tingkat Kesukaran Pada Butir Soal Akidah Akhlak Kelas VI

Soal	Jumlah Soal Benar	Jumlah Siswa	Indeks Kesukaran	Kategori Soal
1	48	49	0,979	Mudah Sekali
2	35	49	0,714	Mudah
3	48	49	0,979	Mudah Sekali
4	46	49	0,938	Mudah Sekali
5	48	49	0,979	Mudah Sekali
6	48	49	0,979	Mudah Sekali
7	49	49	1	Mudah Sekali
8	48	49	0,979	Mudah Sekali
9	47	49	0,959	Mudah Sekali
10	37	49	0,755	Mudah

Dari hasil analisis yang dilakukan peneliti terhadap tingkat kesukaran butir soal tes formatif Akidah Akhlak pada siswa kelas VI MIN 2 Kubu Raya dapat disimpulkan sebagai berikut:

Tabel 5 Hasil Analisis Tingkat Kesukaran Pada Butir Soal Akidah Akhlak Kelas VI

No	Kategori	Butir Soal	Jumlah	Presentase
1	Sangat Sukar	0	0	0%
2	Sukar	0	0	0%
3	Sedang	0	0	0%
4	Mudah	2 dan 10	2	20%
5	Mudah Sekali	1,3,4,5,6,7,8, dan 9	8	80%

Dari hasil perhitungan didapatkan bahwa dari 49 soal tes formatif yang telah dibuat oleh guru Akidah Akhlak di MIN 2 Kubu Raya yaitu 2 soal tergolong mudah, dan 8 soal tergolong sangat mudah. Maka seharusnya soal seperti ini harus diperbaiki, karena menurut indeks kesulitan soal, dalam menyusun soal tes harus seimbang karena apabila soal terlalu sukar maka akan menghilangkan semangat peserta didik untuk menyelesaikan soal tersebut, namun apabila bentuk soal terlalu mudah maka soal tersebut tidak menggairahkan peserta didik dalam mengasah dan menguji kemampuannya.

Butir soal dikatakan baik apabila butir soal yang dibuat tidak terlalu mudah dan juga terlalu sukar. Dari pernyataan tersebut dapat diartikan bahwa sebuah soal yang baik harus memiliki tingkat kesukaran yang sedang atau cukup. Ada tiga kemungkinan untuk tindak lanjut dari butir soal yang sukar pertama dapat dianulir tidak digunakan lagi untuk masa mendatang kedua dilakukan telaah ulang sehingga dapat diketahui factor penyebab butir soal tersebut sulit dijawab siswa, ketiga soal tersebut dapat digunakan untuk tes yang bersifat seleksi yang ketat.

Daya Beda Butir Soal

Daya pembeda ialah kemampuan suatu instrumen soal dalam mengklasifikasikan siswa berdasarkan tingkat kemampuannya, membedakan hasil tes dengan hasil jawaban setiap siswa. Bukan hanya itu daya beda juga dapat mengukur kemampuan setiap peserta didik dalam memahami materi yang telah diajarkan oleh gurunya (Purwati et al., 2021). Dengan adanya daya beda pada setiap soal yang dikerjakan oleh peserta didik maka dapat terlihat kemampuan setiap peserta didik yang menjawab soal tersebut. Sehingga dapat menjadikan evaluasi bagi guru dalam mengajarkan dan mendidik para peserta didiknya. Dari hasil data berikut maka dapat disimpulkan bahwa analisis daya beda soal tes formatif yang telah dibuat oleh guru Akidah Akhlak MIN 2 Kubu Raya Adalah Sebagai Berikut.

Tabel 6: Hasil Analisis Daya Beda Butir Soal Akidah Akhlak Kelas VI

No	Kategori	Butir Soal	Jumlah	Presentasi
1	Sangat Kuat	0	0	0%
2	Kuat	0	0	0%
3	Lemah	1,3,4,5,6,7,8, dan 9	8	80%
4	Sangat Lemah	2 dan 10	2	20%

Dari hasil analisis yang dilakukan peneliti terhadap tingkat daya beda butir soal yang dilakukan peneliti didapatkan bahwa dari 10 soal yang dianalisis terdapat 0 soal yang dikatakan sangat kuat, 0 soal yang dikategorikan kuat, 8 soal dikategorikan lemah, dan 2 soal yang dikategorikan sangat lemah. Berdasarkan hasil ini maka bentuk soal seperti ini diperlukan adanya revisi pada soal tersebut, tujuannya agar soal yang diberikan kepada peserta didik benar-benar dapat mengukur kemampuan mereka dalam memahami mata pelajaran Akidah Akhlak yang telah diajarkan oleh gurunya.

KESIMPULAN

Berdasarkan hasil analisis butir soal yang dilakukan oleh penelitian yang berupa mengukur tingkat kesukaran, daya beda, validitas dan reabilitas butir soal.

Maka dapat disimpulkan yakni tingkat kesukaran dari beberapa soal yang telah dianalisis didapatkan bahwa tingkat kesukaran soal sangat sukar 0%, sukar 0%, sedang 0%, mudah 2 soal 20%, dan mudah sekali 8 soal 80% sehingga didapatkan bahwa bentuk soal seperti ini harus diadakan revisi mayor agar dapat mengukur kemampuan siswa. Sedangkan tingkat kevalidan soal yaitu terdapat 7 butir soal dari 10 soal yang dinyatakan valid dan 3 soal yang dinyatakan tidak valid.

Sedangkan hasil Pengujian reabilitas instrumen ini dilakukan terhadap 49 orang siswa dengan tingkat signifikansi 5%, dan derajat kebebasan (df) = $n-2=49-2=47$, sehingga didapatkan bahwa nilai $r_{11} = 0,5145$ dan lebih kecil dari nilai r_{11} menurut ketentuan yang dikemukakan Heale and Twycross yaitu sebesar 0,70 atau dengan kata lain r_{11} hitung lebih kecil dari 0,70 ($0,659 > 0,70$) yang berarti soal tersebut tersebut tidak reliabel, dan harus dilakukan perbaikan sebelum digunakan untuk mengukur kemampuan siswa dalam memahami mata pelajaran Akidah Akhlak.

DAFTAR PUSTAKA

- Agis, M., Batani, R. A., Inda, D., Susanti, S., & (2021). Analisis Validitas Media Pembelajaran LKPD dan Video Pembelajaran Hukum Bernoulli. *Jurnal Siliwangi*:
<http://jurnal.unsil.ac.id/index.php/jspendidikan/article/view/421>
- Ali Hamzah. (2014). *Evaluasi Pembelajaran Matematika*. rajawali pers.
- Ambarwati, Y. F., & Ismiyati, I. (2022). Analisis butir soal pilihan ganda ulangan akhir semester genap mata pelajaran kearsipan. *Measurement In Educational Research*
<http://ejournal.ressi.id/index.php/meter/article/view/144>
- Amiruddin, K., Mania, S., & ... (2020). Analisis Butir Soal Ujian Akhir Sekolah (UAS) Mata Pelajaran Matematika. *Alauddin Journal of*. <https://journal3.uin-alauddin.ac.id/index.php/ajme/article/view/17364>
- Arifin, Y. Y. (2017). Mutu Pendidikan Pada Sekolah Standar Nasional (Studi Kasus pada Sekolah Dasar di Samarinda). In *IN^{TEL}EGENSIA: Jurnal Pendidikan dan Pembelajaran*.
- Basuki, B. (2020). Evaluasi dan Supervisi Standar Kepala Sekolah dalam Peningkatan Proses Pembelajaran. *Jurnal Pendidikan Guru*.
<https://jurnal.literasikitaindonesia.com/index.php/jurpendigu/article/view/166>
- Bukhari. (2005). *Teknik-teknik Dasar Evaluasi Pendidikan*. Jemmars.
- Djuanda, I. (2020). Implementasi Evaluasi Program Pendidikan Karakter Model CIPP (Context, Input, Process dan Output). *Al Amin: Jurnal Kajian Ilmu Dan Budaya Islam*. <http://stitalamin.ac.id/jurnal/index.php/alaman/article/view/39>
- Fitrianawati, M. (2017). *Peran analisis butir soal guna meningkatkan kualitas butir soal, kompetensi guru dan hasil belajar peserta didik*. publikasiilmiah.ums.ac.id.
<https://publikasiilmiah.ums.ac.id/xmlui/handle/u617/9117>
- Gimo, G., & Nugrahani, F. (2020). Analisis Butir Soal Ulangan Akhir Semester (UAS) Mata Pelajaran Bahasa Indonesia Kelas XI Sekolah Menengah Kejuruan Negeri 1 Wonogiri Tahun Pelajaran 2015 *Stilistika: Kajian Bahasa, Sastra*
<http://journal.univetbantara.ac.id/index.php/stilistika/article/view/609>
- Halik, A. S., Mania, S., & Nur, F. (2019). Analisis butir soal Ujian Akhir Sekolah (UAS) mata pelajaran matematika pada tahun ajaran 2015/2016 SMP Negeri 36

- Makassar. *Al Asma: Journal of Islamic* <https://journal3.uin-alauddin.ac.id/index.php/almasma/article/view/11249>
- Hamimi, L., Zamharirah, R., & ... (2020). Analisis butir soal ujian Matematika kelas VII semester ganjil tahun pelajaran 2017/2018. *Mathema: Jurnal* <https://ejurnal.teknokrat.ac.id/index.php/jurnalmathema/article/view/459>
- Hidayati, A. N. (2020). *Evaluasi pembelajaran pada mata pelajaran pendidikan agama Islam di sekolah inklusi SD Anak Saleh Malang*. *etheses.uin-malang.ac.id*. <http://etheses.uin-malang.ac.id/id/eprint/19423>
- Huda, N., & Wahyuni, T. S. (2019). Analisis butir soal IPA Try Out USBN Tahun Ajaran 2018/2019 dalam kaitannya dengan level kognitif. *Madrasah: Jurnal Pendidikan* ... <http://repository.uin-malang.ac.id/5469/>
- Ida, F. F. (2021). Validitas dan Reliabilitas dalam Analisis Butir Soal. *AL-MU'ARRIB: JOURNAL OF ARABIC EDUCATION*. <https://lp2msasbabel.ac.id/jurnal/index.php/AL-MUARRIB/article/view/2100>
- Latifah, N. (2018). Pengembangan Bahan Ajar Berbasis Budaya Lokal Subtema Indahnya Keragaman Budaya Negeriku di Kelas IV Sekolah Dasar. *Budaya Negeriku Di Kelas IV Sekolah Dasar*. <https://repository.unja.ac.id/4369/>
- Nasir, M. (2019). Analisis empirik program analisis butir soal dalam rangka menghasilkan soal yang baik dan bermutu sebagai alat evaluasi pembelajaran fisika. *SEMIRATA* 2015. <https://jurnal.untan.ac.id/index.php/semirata2015/article/view/13534/0>
- Novia, T., Wardani, A., Canda, C., Nurdi, N., & ... (2020). Analisis Validitas dan Reliabilitas Butir Soal UTS Fisika Kelas X SMA Swasta Muhammadiyah 4 Langsa. ... *Fisika Dan Sains*. <https://ejournalunsam.id/index.php/JDFS/article/view/2256>
- Pangestu, D. (2019). Evaluative study relevance of curriculum development model of elementary school teacher education with elementary school curriculum. ... *Jurnal Pendidikan Dan Pembelajaran Dasar*. <http://ejournal.radenintan.ac.id/index.php/terampil/article/view/3827>
- Purwati, L. M., Arianty, R., Syakilah, D. M., Ridlo, S., & Susilaningsih, E. (2021). Analisis Soal Tes Pilihan Ganda Berbasis Higher Order Thinking Skill menggunakan Aplikasi Anates Windows Versi 4.0.9 For Windows. *Jurnal Pendidikan UNIGA*, 15(2), 460. <https://doi.org/10.52434/jp.v15i2.1287>
- Santosa, S., & Badawi, J. A. (2022). Analisis Butir Soal Pilihan Ganda Tema Pertumbuhan dan Perkembangan Makhluk Hidup Kelas III Madrasah Ibtidaiyah. *Jurnal Basicedu*. <https://jbasic.org/index.php/basicedu/article/view/2206>
- Suharsimi Arikunto. (2007). *Dasar-Dasar Evaluasi Pendidikan*. Bumi Aksara.
- Azwar, S. (2013). *Sikap Manusia: Teori dan Pengukurannya*. Yogyakarta: Pustaka Pelajar, 1(69).
- Cahyaningrum, I. (2019). Analisis Kualitas Butir Soal Pendidikan Agama Islam Pada Penilaian Kahir Semester Gasal Kelas VII SMP Negeri 1 Kembaran Tahun Pelajaran 2018/2019. *Institut Agama Islam Negeri Purwokerto*.
- Lester, P. E., Inman, D., & Bishop, L. K. (2014). *Handbook of tests and measurement in education and the social sciences*.

- Purwanto, N. (2010). Prinsip-Prinsip dan Teknik Evaluasi Pengajaran. In *PT Remaja Rosdakarya*.
- Purwati, L. M., Arianty, R., Syakilah, D. M., Ridlo, S., & Susilaningsih, E. (2021). Analisis Soal Tes Pilihan Ganda Berbasis Higher Order Thinking Skill menggunakan Aplikasi Anates Windows Versi 4.0.9 For Windows. *Jurnal Pendidikan UNIGA*, 15(2), 460. <https://doi.org/10.52434/jp.v15i2.1287>
- Rahmaini, A., & Taufiq, A. N. (2018). Analisis Butir Soal Pendidikan Agama Islam Di SMK N 1 Sedayu Tahun Ajaran 2017/2018 (Analisis Tingkat Kesukaran, Daya Pembeda dan Fungsi Distraktor pada Soal Pilihan Ganda Kelas XI). *Jurnal MUDARRISUNA*, 8(1).
- Sardiyannah, S. (2020). KONSEP EVALUASI DALAM PENDIDIKAN. *Jurnal Al-Qalam: Jurnal Kajian Islam & Pendidikan*, 8(1). <https://doi.org/10.47435/al-qalam.v8i1.199>
- Secolsky, C. (2018). 1Measurement, Assessment, and Evaluation in Higher Education. *Handbook on Measurement, Assessment, and Evaluation in Higher Education*, 1–31. <https://doi.org/10.4324/9781315709307-1>
- Sukino, S. (2015). The effect of Application an Examination for Student's Science Attitude and Learning Motivation. *Tarbawi*, 1 nomor 2, 63–80. <https://jurnal.uinbanten.ac.id/index.php/tarbawi/article/view/1997/1657>
- Tarmizi, P., Setiono, P., Amaliyah, Y., & ... (2021). Analisis Butir Soal Pilihan Ganda Tema Sehat Itu Penting Kelas V SD Negeri 04 Kota Bengkulu. *ELSE (Elementary)* <http://journal.um-surabaya.ac.id/index.php/pgsd/article/view/7090>
- Warju, W., Ariyanto, S. R., Soeryanto, S., & ... (2020). Analisis Kualitas Butir Soal Tipe Hots pada Kompetensi Sistem Rem di Sekolah Menengah Kejuruan. *Jurnal Pendidikan* <https://ejournal.undiksha.ac.id/index.php/JPTK/article/view/22914>
- Widayanti, W., Bistari, B., & ... (2021). Analisis Butir Soal Pilihan Ganda Penilaian Tengah Semester Pada Pembelajaran Tematik Kelas V Sekolah Dasar Negeri 39 Pontianak Kota. *Jurnal DIDIKA* <http://e-journal.hamzanwadi.ac.id/index.php/didika/article/view/4370>