IBTIDA': ***Media Komunikasi Hasil Penelitian Pendidikan Guru Madrasah Ibtidaiyah*** p-ISSN: 2722-8452 (Print)
Volume 07, No. 02, November 2026, Hal. 179-187 e-ISSN: 2722-8290 (Online)

DOI: https://doi.org/10.37850/ibtida'
https://journal.faibillfath.ac.id/index.php/ibtida

UJI INSTRUMEN TES SEBAGAI ALAT UKUR HASIL BELAJAR SISWA KELAS 2 PELAJARAN PENDIDIKAN AGAMA ISLAM

Noor Alfi Fajriyani[1], Ummi Kulsum[2], Atika Rahmah[3], Nurhayati[4], Abdul Qodir[5], Miatul Khairo[6]
[1,2,3,4,5,6] Sekolah Tinggi Ilmu Tarbiyah Aqidah Usymuni, Sumenep
email : *nooralfifajriyani@stitasumenep.ac.id*

Received 31 October 2026; Received in revised form 21 November 2026; Accepted 23 November 2026

Abstrak

Guru Pendidikan Agama Islam (PAI) seusai pembelajaran berlangsung, kebanyakan hanya menggunakan penilaian tes yang ada di buku. Tidak memiliki kreatifitas untuk membuat instrumen penilaian yang sesuai dengan karater peserta didik di kelas. Penelitian ini bertujuan untuk menghasilkan produk instrumen penilaian tes objektif pilihan ganda yang layak untuk mengukur hasil belajar siswa pelajaran Pendidikan Agama Islam (PAI). Penelitian ini merupakan penelitian kuantitatif deskripitif. Untuk menggambarkan reabilitas, daya kesukaran, daya pembeda, dan pengecohan pada instrument tes PAI. Subjek penelitian pada uji coba ialah seluruh siswa kelas 2 SD Al-Fatah Wiyung Surabaya yang berjumlah 30 siswa. Hasil penelitian menunjukkan instrumen penilaian tes objektif layak untuk mengukur hasil belajar siswa pada materi "Mari Mengenal Allah" mata pelajaran PAI. Dengan kriteria reabilitas 0,741. Tingkat kesukaran sulit 15%, Sedang 70%, mudah 15%. Daya Pembeda kriteria buruk 35%, cukup 15%, baik 50%. Dan Tingkat pengecohan instrument kriteria baik 70% kurang baik 30%.
***Kata kunci**: Hasil Belajar ; Instrumen Tes ; Pendidikan Agama Islam*

Abstract

Islamic Religious Education (PAI) teachers, after the learning process, mostly only use the assessment tests found in the book. They lack creativity to create assessment instruments that match the character of the students in the classroom. This study aims to produce a suitable multiple-choice objective test assessment instrument product to measure students' learning outcomes in Islamic Religious Education (PAI). This research is a descriptive quantitative study. To describe the reliablity, difficulty, discriminating power, and distortion of the PAI test instrument. The research subjects in the trial were all 2nd grade students at SD Al-Fatah Wiyung Surabaya, totaling 30 students. The results showed that the objective test assessment instrument was feasible to measure student learning outcomes on the material "Let's Get to Know Allah" in the PAI subject. With a reliability criterion of 0.741. The difficulty level was 15% difficult, 70% moderate, and 15% easy. Discrimination power for poor criteria was 35%, fair 15%, good 50%. And instrument distortion level was 70%, poor was 30%.
***Keywords**: Learning Outcomes ; Test Instruments ; Islamic Religious Education.*

© 2026 by the authors. This work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License.

PENDAHULUAN

Penilaian merupakan salah satu kegiatan penting yang harus tercanangkan dalam dunia Pendidikan (Hidayah, 2020). Baik mulai dari tingkatan pendidikan sekolah dasar hingga tingkatan pendidikan perguran tinggi (Primasari et al., 2021). Penilaian dalam dunia pendidikan mempunyai kedudukan sebagai penentu kualitas dari pendidikan itu sendiri (Setiadi, 2016). Sehingga perlu adanya kegiatan penilaian yang sistematis dan berkesinambungan sesuai dengan standar penilaian pada pendidikan (Hafidhoh & Rifa'i, 2021). Dengan demikian, penilaian yang kredibel dan berkualitas, akan menghasilkan juga mencerminkan lulusan yang berkualitas (Santiyadnya, 2011)).

Penilaian pembelajaran merupakan suatu bentuk kegiatan pengukuran hasil belajar yang sebelumnya telah melalui proses kegiatan pengajaran demi mencapai tujuan belajar (Imania & Bariah, 2019). Terdapat prinsip yang wajib terpenuhi pada siswa ketika penilaian dilakukan, diantaranya; mendemontrasikan, menunjukkan, dan memahami (Efendi et al., 2021). Dan hasil dari penilaian itu sendiri digunakan guru sebagai acuan evaluasi pembelajaran. Sehingga guru bisa meningkatan mutu pendidikan serta memetakan kemampuan kognitif dari setiap siswa (Agisna et al., 2023).

Menilai dan meng-evaluasi pengajaran ialah merupakan suatu kegiatan fundamental yang harus di kerjakan oleh guru (Irawati et al., 2017). Karena melalui penilaian guru bukan hanya bisa mengukur keberhasilan belajar, namun juga bisa mengukur mutu pengajaran yang telah di terapkan oleh guru kala itu (Muslim, 2018). Penilaian sendiri tergolong menjadi beberapa bentuk penilaian. Pertama ada penilaian berupa tes dan kedua berupa penilaian non tes (Rizqiyah, 2018). Penilaian tes terdiri dari tes lisan dan tes tulis. Penilaian tes tulis terdiri tes tulis berupa uraian dan pilihan ganda (Maulidia & Pahlevi, 2020). Sedangkan penilaian non tes yaitu bisa berupa penilaian sikap, proyek, portopolio, penilaian diri, dll. Penilaian non tes biasanya dikumpulkan menggunakan lembar observasi atau berupa angket (Rusilowati, 2013). Ranah dari instrument non tes sendiri yaitu aspek afektif dan psikomotorik (Bahari, 2023).

Penilaian yang digunakan oleh guru Indonesia seringkali tidak membantu siswa secara optimal dalam menghadapi masalah kontekstual. Karena masih ada beberapa diantara guru yang belum siap dalam merencanakan, meng-implemtasikan serta mengevaluasi pembelajaran. Sehingga hasil belajar siswa menjadi kurang optimal (Kartomo & Slameto, 2016). Masih banyak guru yang membuat soal tes tidak berdasarkan kisi-kisi tes dengan indicator materi pelajaran, tetapi cenderung hanya menggunakan soal pada buku yang disediakan. Kecenderungan guru untuk membuat pertanyaan yang tidak dibimbing oleh kisi-kisi tes inilah yang membuat siswa tidak terlatih dalam bepikir kritis (Setiawan et al., 2021).

© 2026 by the authors. This work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License.

Masih banyakya guru yang belum memiliki pengetahuan terkait evaluasi pembelajaran, informasi hasil beajar kurang akurat. Sedangkan, hasil belajar siswa sangat terpengaruh dengan kualitas instrument penilaian yang telah dibuat oleh guru. Oleh sebab itu guru harus memiliki keterampilan dalam membuat instrument penilaian, agar mendapatkan kualitas instrument yang baik (Ramadhan et al., 2019).

Penelitian yang dilakukan wijayanti,dkk. juga menyatakan bahwa keterampilan dan inovasi guru dalam membuat instrument digital itu penting. Penelitian menggunakan metode pengembangan dan kunatitatif deskriptif. Menghasilkan kelayakan dalam pengembangan instrument digital, sehingga dapat mengukur aspek kognitif pada Pelajaran Ips kelas IV (Wijayanti et al., 2023).

Hal yang sama juga dilakukan pengembangan instrument tes untuk menilai semua aspek komptensi siswa. Dengan menggunakan metode pengembangan dan menguji instrument dengan menggunakan validitas konstruk. Didapatkan guru kelompok ppg dapat mengembangkan instrument dengan validitas kategori sangat baik di 84,,76% pada Pelajaran IPA (Destiana et al., 2020).

Sedangkan guru Pendidikan Agama Islam (PAI) di Sekolah Dasar Islam al-Fatah Wiyung Surbaya, mengutarakan bahwa selama kegiatan penilaian di akhir materi pelajaran, tidak membuat instrument tes. Guru hanya menggunakan instrument tes yang terdapat di buku siswa. Sehingga ada kemungkinan instrument tes yang ada di buku siswa sudah terlebih dahulu di kerjakan di rumah bersama orang tua atau lainnya. Dengan begitu informasi kognitif yang didapat dari hasil pengerjaan instrument oleh siswa, tidak bisa dijadikan landasan sebagai bentuk evaluasi atau memetakan kemampuan kognitif setiap siswa. Karena instrument tes yang ada di buku siswa belum tentu reliabel.

Dengan demikian guru PAI di sekolah dasar wiyung, belum memiliki inovasi dan keterampilan dalam mengembangkan instrument. Berangkat dari permasalahan diatas, dirasa peneliti perlu mengembangkan instrument tes tipe objektif dengan memenuhi tiga standart prinsip penilaian mendemostrasikan, menunjukkan dan memahami (Efendi et al., 2021).

Penetilitian ini, berbeda dengan penelitian sebelumnya. Dimana penelitian tidak hanya mengukur validitas reabilitas, melainkan juga ,mrngukur daya beda dan kesukaran insturmen. Selain itu, instrument yang dikembangkan ialah Pelajaran PAI. Dimana masih belum banyak peneliti mengembangkan dan mengukur instrument tes pada Pelajaran PAI.

Tujuan dari penelitian ini ialah untuk menguji instrument pada Pelajaran PAI yang telah dikembangkan berbentuk tipe objektif pilihan ganda. Dan mengukur reabilitas, daya beda, kesukaran, dan daya pengecoh instrument dengan berbantuan ITEMAN 4.3..

© 2026 by the authors. This work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License.

https://journal.faibillfath.ac.id/index.php/ibtida

METODE PENELITIAN

Kulitatif deskriptif digunakan sebagai metode penelitian ini. Metode kualitatif dimana menginterpretasikan temuan-temuan dengan menggunakan sebuah kata-kata (Sugiyono, 2013). Tujuan dari metode ini ialah untuk menginterpretasikan data secara menyeluruh dan efektif. Sehingga dapat menggambarkan fenomena yang factual dan sistematis (Sari et al., 2024).

Tempat penelitian di Sekolah Dasar Al-Fatah Wiyung Surabaya. Dengan seluruh siswa kelas 2 SD Al-Fatah Wiyung Surabaya menjadi subjek dari penelitian. Siswa yang di jadikan subjek penelitin ialah siswa yang telah menerima materi ajar "Mari Mengenal Allah"Pelajaran Pendidikan Agama Islam (PAI). Total keseluruhan siswa kelas 2, dari rombel "A" dan rombel "B" berjumlah 30 siswa.

Teknik Pengumpulan data menggunakan instrument tes objektif tipe pilihan ganda. Dengan empat pilihan jawaban "A,B,C,D" dan jumlah soal berjumlah 20 pertanyaan. Hasil jawaban siswa kemudian di kumpulkan sebagai bentuk data emperis.

Data yang telah didapatkan dari hasil jawaban siswa, kemudian dianalisis menggunakan program software ITEMAN 4.3 untuk teori uji klasik, yang meliputi: uji reabilitas, uji kesukaran item, daya pembeda, dan pengecoh soal (Ramadhan et al., 2019). Uji ini digunakan untuk mengukur kualitas instrument sebagai bentuk pengukur hasil belajar siswa.

HASIL DAN PEMBAHASAN

Penilaian merupakan bentuk tolak ukur seorang guru dalam menentukan kualitas hasil belajar siswa. Hasil belajar yang baik akan mencerminkan kualitas pemahaman terhadap materi yang telah diajarkan. Oleh sebabnya penting bagi guru untuk mengembangkan instrument tes dengan kuliatas reabilitas yang baik. Disamping itu guru juga harus memperhatikan Tingkat kesukaran, daya pembeda, dan pengecohan dalam mengembangkan instrument. Agar tidak menimbulkan heteregoen hasil belajar.

Guru PAI SD Al-Fatah Wiyung Surabaya telah menerapkan uji coba instrument test pada instrument Pelajaran PAI yang telah dikembangkan. Uji coba dilakukan di SD Al-Fatah Wiyung Surabaya dengan melibatkan siswa kelas 2 yang berjumlah 30 siswa terdiri dari dua rombel kelas. Kualitas instrument tes objektif pilihan ganda berdasarkan karakteristik teori uji klasik meliputi: reabilitas, tingkat kesukaran, daya pembeda, dan pengecohan.

A. Reabilitas Instrumen

Analisis butir tes pilihan ganda mari mengenal Allah pelajaran Pendidikan Agama Islam (PAI) bagi siswa kelas 2 dilakukan dengan support software statistic ITEMAN versi 4.3.0.3.. Dari hasil analisis, didapat hasil analisis sebagai berikut:

© 2026 by the authors. This work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License.

Tabel 1. Hasil Reabilitas Instrumen

Score	Alpha	SEM	Split-Half (Random)	Split-Half (First-Last)	Split-Half (Odd-Even)	S-B Random	S-B First -Last	S-B Odd-Even
Scored items	0,741	1,936	0,567	0,548	0,713	0,723	0,708	0,833

Berdasarkan dari hasil analisis dengan menggunakan ITEMAN, didapatkan nilai alpha reabilitas adalah 0,741. Sehingga dapat diartikan bahwa instrument tes objektif materi "mari mengenal Allah" pelajaran PAI memiliki reabilitas soal yang baik. Sebagaimana dengan ketentuan, apabila nilai reabilitas ≥0,7 maka dinyatakan instrument tersebut masuk pada kategori baik (Machali, 2021). Dengan demikian soal tersebut bisa dijadikan sebagai alat ukur untuk menilai kemampuan aspek kognitif siswa (Ramadhan et al., 2019).

B. Tingkat Kesukaran Instrumen

Ketika guru mengembangkan instrument test, perlu diperhatikan instrument tersebut harus memiliki tingkat kesukaran yang baik. Semakin baik Tingkat kesukaran yang dikandung oleh instrument, maka guru akan dengan mudah untuk mengukur dan membedakan kemampuan dari siswa (Agisna et al., 2023). Kesukaran atau kesulitan Item didefinisikan sebagai proporsi peserta test pada respon jawaban yang benar dalam item tes tertentu dan nilainya sekitar 0 sampai 1 (Ramadhan et al., 2020).

Berikut hasil analisis tingkat kesukaran instrument pada materi "mari mengenal Allah" Pelajaran PAI, dengan menggunakan software ITEMAN 4.3, sebagai berikut:

Tabel 2. Hasil Tingkat Kesukaran Instrumen

No.	Indeks Kesukaran	Kategori	Nomor Item	Total	Persentase
1	0.00 – 0.30	Sulit	4, 15, 18	3	15%
2	0.31 – 0.70	Sedang	3, 5, 6, 7, 8, 9, 10, 12, 13, 14, 16, 17, 19, 20	14	70%
3	0.71 – 1.00	Mudah	1, 2, 11	3	15%

© 2026 by the authors. This work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License.

IBTIDA': ***Media Komunikasi Hasil Penelitian Pendidikan Guru Madrasah Ibtidaiyah***
Volume 07, No. 02, November 2026, Hal. 179-187
p-ISSN: 2722-8452 (Print)
e-ISSN: 2722-8290 (Online)

DOI: https://doi.org/10.37850/ibtida'
https://journal.faibillfath.ac.id/index.php/ibtida

Tingkat kesukaran instrument test, diperlukan untuk memeriksa item yang termasuk dalam ketegori sulit, sedang, atau mudah. Hasil yang tercatat dalam table menunjukkan bahwa 3 soal termasuk kategori sulit, 14 soal termasuk kategori sedang dan 3 soal lainnya termasuk kategori mudah. Hal ini selarasa dengan kriteria 0.00-0.30 "sulit", 0.31-0.70 "sedang", dan 0.71-1.00 "mudah" (Sudijono, 2012).

C. Daya Pembeda Instrumen

Daya pembeda merupakan salah satu indicator penting yang harus terkandung dalam instrument test. Karena, daya pembeda adalah kemampuan tes untuk membedakan antara siswa yang memiliki kemampuan kognitif tinggi dengan kemampuan kognitif rendah (Setiawan et al., 2021).

Hasil analisis daya pembeda item instrument test materi "mari mengenal Allah" Pelajaran PAI sebagai berikut:

Tabel 3. Hasil Daya Pembeda Instrumen

No.	Indeks Daya Pembeda	Item	Total	Persentase
1.	0.00-0.20 (buruk)	5,7,1,16,17,18,20	7	35 %
2.	0.21-0.40 (cukup)	1,6,8	3	15 %
3.	0.41-0.70 (baik)	2,3,4,9,10,11,12,13,14,19	10	50 %
4.	0.71-1.00 (sangat baik)	-	0	0 %

Berdasarkan dari hasil uji table di atas, terdapat 7 soal yang memiliki indeks daya pembedanya buruk. 3 soal lainnya berindeks-kan cukup dan 10 lainnya memiliki indeks daya pembeda baik. Jika soal test lebih banyak berindeks buruk, guru bisa membuang soal dengan menggantinya atau bisa merevisi soal tersebut (Qadir et al., 2024).

D. Tingkat Pengecoh Istrumen

Daya pengecoh mempunyai peran penting pada bentuk isntrumen test objektif atau pilihan ganda. Karena, dengan adanya daya pengecoh dapat menetukan, apakah pilihan jawaban yang salah berkerja sebagai penguji pemahaman materi, atau hanya sekedar pelengkap yang jawabannya juga mudah ditebak (Santosa & Badawi, 2022).

Dalam hal bentuk pilihan ganda ada jawaban alternatif yang menipu, menipu di sini adalah daya pengeceoh untuk jawaban atas pertanyaan yang dapat menipu jawaban yang sebenarnya (Setiawan et al., 2021).

Hasil analisis indeks tingkat pengecoh instrument tes materi "mari mengenal Allah" Pelajaran PAI dengan menggunakan software iteman, sebagai berikut:

© 2026 by the authors. This work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License.

Tabel 4. Hasil Tingkat Pengecoh Instrumen

No	Indeks Pengecoh	Item	Total	Presentase
1.	Baik (≥5%)	1,3,4,7,9,10,12,13,14,15,16,17,18,20	14	70 %
2.	Kurang Baik (≤5%)	2,5,6,8,11,19	6	30 %

Dari hasil table diatas dapat diketahui bahwa terdapat 14 soal instrumen yang menduduki indeks pengecoh baik, yaitu 70%. Dan 6 soal pertanyaan lainnya menduduki indeks kurang baik, yaitu 30%. Pengecoh yang baik berindeks > dari 5%. Dan dari keenam soal yang memiliki indeks kurang baik kurang dari < 5% bisa diganti atau revisi (Warju et al., 2020).

KESIMPULAN DAN SARAN

Hasil analisis uji instrument test objektif atau pilihan ganda dengan menggunakan bantuan software ITEMAN 4.3., menunjukkan instrumen penilaian objektif layak digunakan. Hasil karakteristik item pilihan ganda menunjukkan bahwa dari 20 item tersebut, mendapatkan rata-rata tingkat reliabilitas adalah 0,7 (kategori baik). Daya kesukaran berada pada 15% kategori sulit, 70% sedang, 15% mudah. Daya pembeda berada pada 35% kategori buruk, 15% cukup, 50% baik, 0% sangat baik. Daya pengecoh didapatkan pada kategori indeks 70% baik dan 30% kurang baik. Terdapat beberapa item soal instrument yang menduduki kategori mudah, buruk, dan kurang baik. item ini bisa diganti atau direvisi oleh guru. Sehingga instrument penilaian objektif layak untuk digunakan sebagai pengukuran hasil belajar kognitif materi "Mari Mengenal Allah" pelajaran Pendidikan Agama Islam siswa kelas 2.

Saran bagi peneliti selanjutnya ialah untuk menguji instrument test siswa dapat menambahkan uji validitas konstruk pada instrument Pelajaran PAI sebagai penelitian pembaharuan.

© 2026 by the authors. This work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License.

DAFTAR PUSTAKA

Agisna, R., Jauhari, Z. A., Zuar, M. S., Sholihin, M., & I, A. K. (2023). Evaluasi Pembelajaran sebagai Upaya Peningkatan Kompetensi Guru dan Pencapaian Kompetensi Belajar Siswa. *Social Science Academic*, *1*(2), 353–362. https://doi.org/10.37680/ssa.v1i2.3582

Bahari, J. I. (2023). EVALUATION OF COGNITIVE, PSYCHOMOTOR, AND AFFECTIVE ASPECTS IN THE SUBJECT OF ISLAMIC EDUCATION. *INCARE, International Journal of Educational Resources*, *4*(2), 137–163. https://doi.org/10.59689/incare.v4i2.702

Destiana, D., Suchyadi, Y., & Anjaswuri, F. (2020). PENGEMBANGAN INSTRUMEN PENILAIAN UNTUK MENINGKATKAN KUALITAS PEMBELAJARAN PRODUKTIF DI SEKOLAH DASAR. *Jurnal Pendidikan Dan Pengajaran Guru Sekolah Dasar (JPPGuseda)*, *3*(2), 119. https://doi.org/10.55215/jppguseda.v3i2.2720

Efendi, I., Prawitasari, M., & Susanto, H. (2021). Implementasi Penilaian Pembelajaran Pada Kurikulum 2013 Mata Pelajaran Sejarah. *Prabayaksa: Journal of History Education*, *1*(1), 21. https://doi.org/10.20527/prb.v1i1.3081

Hafidhoh, N., & Rifa'i, M. R. (2021). Karakteristik Penilaian Pembelajaran Pada Kurikulum 2013 Di MI. *Awwaliyah: Jurnal Pendidikan Guru Madrasah Ibtidaiyah*, *4*(1), 11–18. https://doi.org/10.58518/awwaliyah.v4i1.673

Hidayah, I. (2020). Analisis Standar Penilaian Pendidikan di Indonesia (Telaah atas Peraturan Menteri Pendidikan Nasional Nomor 20 Tahun 2007, Peraturan Menteri Pendidikan dan Kebudayaan Nomor 66 Tahun 2013, dan Peraturan Menteri Pendidikan dan Kebudayaan Nomor 23 Tahun 201. *Al Iman: Jurnal Keislaman Dan Kemasyarakatan*, *4*(1), 85–105.

Imania, K. A. N., & Bariah, S. H. (2019). Rancangan Pengembangan Instrumen Penilaian Pembelajaran Berbasis Daring. *Petik: Jurnal Pendidikan Teknologi Informasi Dan Komunikasi*, *5*(1), 31–47. https://doi.org/10.31980/petik.v5i1.1534

Irawati, H., Saifuddin, M. F., & Ma'rifah, D. R. (2017). PENGEMBANGAN INSTRUMEN TES DAN NON TES DALAM RANGKA MENYIAPKAN PENILAIAN AUTENTIK PADA KURIKULUM 2013 DI SMP/MTS MUHAMMADIYAH SE-KABUPATEN BANTUL. *Jurnal Pemberdayaan: Publikasi Hasil Pengabdian Kepada Masyarakat*, *1*(2), 503–506. https://doi.org/10.12928/jp.v1i2.362

Kartomo, A. I., & Slameto, S. (2016). EVALUASI KINERJA GURU BERSERTIFIKASI. *Kelola: Jurnal Manajemen Pendidikan*, *3*(2), 219–229. https://doi.org/10.24246/j.jk.2016.v3.i2.p219-229

Machali, I. (2021). *Metode Penelitian Kuantitatif: Panduan Praktis Merencanakan, Melaksanakan dan Analisis dalam Penelitian Kuantitatif*. Fakultas Ilmu Tarbiyah dan Keguruan Universitas Islam Negeri Sunan Kalijaga Yogyakarta.

Maulidia, F., & Pahlevi, T. (2020). Pengembangan Instrumen Penilaian Tes Soal Pilihan Ganda Berbasis HOTS Pada Mata Pelajaran Administrasi Umum Jurusan OTKP SMK Negeri 1 Lamongan. *Jurnal Pendidikan Administrasi Perkantoran (JPAP)*, *8*(1), 136–145.

© 2026 by the authors. This work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License.

https://doi.org/10.26740/jpap.v8n1.p136-145

Muslim, A. Q. (2018). *MODEL PENILAIAN KINERJA GURU*. *11*(1).

Primasari, I. F. N. D., Marini, A., & Sumantri, M. S. (2021). Analisis Kebijakan dan Pengelolaan Pendidikan Terkait Standar Penilaian di Sekolah Dasar. *Jurnal Basicedu*, *5*(3), 1479–1491. https://doi.org/10.31004/basicedu.v5i3.956

Qadir, A., Huda, N., & Hermina, D. (2024). ANALISIS BUTIR TES: TINGKAT KESUKARAN, DAYA PEMBEDA DAN EFEKTIVITAS PENGECOH. *Al-Furqan : Jurnal Agama, Sosial, Dan Budaya*, *3*(3), 1450–1467.

Ramadhan, S., Mardapi, D., Prasetyo, Z. K., & Utomo, H. B. (2019). The Development of an Instrument to Measure the Higher Order Thinking Skill in Physics. *European Journal of Educational Research*, *8*(3), 743–751. https://doi.org/10.12973/eu-jer.8.3.743

Rizqiyah, L. (2018). *Teknik tes dan nontes sebagai alat evaluasi hasil belajar* (Zr32w_v1). OSF Preprints. https://doi.org/10.31219/osf.io/zr32w

Rusilowati, A. (2013). *PENGEMBANGAN INSTRUMEN NONTES*.

Santiyadnya, N. (2011). IMPLEMENTASI UJI KOMPETENSI DAN PENGARUHNYA TERHADAP KUALITAS LULUSAN SMK NEGERI BIDANG TEKNOLOGI DI PROVINSI BALI. *Jurnal Pendidikan Teknologi Dan Kejuruan*, *8*(1). https://doi.org/10.23887/jptk-undiksha.v8i1.2889

Santosa, S., & Badawi, J. A. (2022). Analisis Butir Soal Pilihan Ganda Tema Pertumbuhan dan Perkembangan Makhluk Hidup Kelas III Madrasah Ibtidaiyah. *Jurnal Basicedu*, *6*(2), 1678–1686. https://doi.org/10.31004/basicedu.v6i2.2206

Sari, M. N., Abdillah, L. A., Mappanyompa, Asmarany, A. I., Rakhmawati, I., Pattiasina, P. J., Kusnadi, I. H., Hasanuddin, R., Pradana, I. P. Y. B., Rela, I. Z., isma, A., Darman, & Hadikusumo, R. A. (2024). *Metode Penelitian Kualitatif (Konsep & Aplikasi)*. Mega Press Nusantara.

Setiadi, H. (2016). Pelaksanaan penilaian pada Kurikulum 2013. *Jurnal Penelitian Dan Evaluasi Pendidikan*, *20*(2), 166–178. https://doi.org/10.21831/pep.v20i2.7173

Setiawan, J., Ajat Sudrajat, A., Aman, A., & Kumalasari, D. (2021). Development of higher order thinking skill assessment instruments in learning Indonesian history. *International Journal of Evaluation and Research in Education (IJERE)*, *10*(2), 545. https://doi.org/10.11591/ijere.v10i2.20796

Sudijono, S. (2012). *Pengantar Evaluasi*. Rajawali Pres.

Sugiyono, S. (2013). *Metode Penelitian Kuantitatif, Kualitatif, dan R&D*. Alfabeta.

Warju, W., Ariyanto, S. R., Soeryanto, S., & Trisna, R. A. (2020). ANALISIS KUALITAS BUTIR SOAL TIPE HOTS PADA KOMPETENSI SISTEM REM DI SEKOLAH MENENGAH KEJURUAN. *Jurnal Pendidikan Teknologi Dan Kejuruan*, *17*(1), 95–104. https://doi.org/10.23887/jptk-undiksha.v17i1.22914

Wijayanti, R. W., Riyanto, Y., & Subroto, W. T. (2023). Pengembangan Instrumen Penilaian Digital Untuk Mengukur Hasil Belajar IPS Kelas IV Sekolah Dasar. *EDUKASIA Jurnal Pendidikan Dan Pembelajaran*, *4*(1), 127–136. https://doi.org/10.62775/edukasia.v4i1.230

© 2026 by the authors. This work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License.