

KUALITAS SOAL BUATAN GURU SD: ANALISIS KUALITATIF ASPEK MATERI, KONSTRUKSI, BAHASA, DAN ETIKA

Asyraf Suryadin¹, Niha Syahda²

¹ Program Studi Pendidikan Guru Sekolah Dasar, Universitas Muhammadiyah Bangka Belitung

Email: asyraf.suryadin@unmuhbabel.ac.id

Email: syahdaniha911@gmail.com

ABSTRACT

This study aims to analyze the qualitative quality of teacher-made multiple-choice and essay questions at the elementary school level. The aspects analyzed include material, construction, language, and ethics according to the guidelines for good question preparation. The method used was document analysis of teacher-made questions for grades I to VI. The analysis instrument used a qualitative review sheet with question suitability indicators. The results showed that the majority of questions met the material aspect with high indicator conformity, functional distractors, and one most correct answer. In the construction aspect, most questions were clearly formulated, although there were some minor notes such as inter-item dependency and imbalance in option length. The language aspect showed the use of communicative and grammatically correct language. The ethical aspect showed that the questions were safe, neutral, and free from elements of SARA, violence, or propaganda. Overall, the quality of the analyzed elementary school teacher-made questions was very good, with several recommendations for improvement in technical aspects.

Keywords: elementary school; essay; multiple choice; qualitative analysis; question quality

ABSTRAK

Penelitian ini bertujuan untuk menganalisis kualitas soal buatan guru bentuk pilihan ganda dan essay pada jenjang sekolah dasar secara kualitatif. Aspek yang dianalisis meliputi materi, konstruksi, bahasa, dan etika sesuai pedoman penyusunan soal yang baik. Metode yang digunakan adalah analisis dokumen terhadap soal buatan guru kelas I sampai kelas VI. Instrumen analisis menggunakan lembar telaah kualitatif dengan indikator kesesuaian soal. Hasil penelitian menunjukkan bahwa mayoritas soal telah memenuhi aspek materi dengan kesesuaian indikator yang tinggi, pengecoh berfungsi, dan memiliki satu jawaban paling benar. Pada aspek konstruksi, sebagian besar soal dirumuskan jelas, meskipun terdapat beberapa catatan minor seperti ketergantungan antar butir dan ketidakseimbangan panjang pilihan. Aspek bahasa menunjukkan penggunaan bahasa yang komunikatif dan sesuai kaidah. Aspek etika menunjukkan soal aman, netral, dan bebas dari unsur SARA, kekerasan, maupun propaganda. Secara keseluruhan, kualitas soal buatan guru SD yang dianalisis tergolong sangat baik, dengan beberapa rekomendasi perbaikan pada aspek teknis.

Kata kunci: *analisis kualitatif; kualitas soal; sekolah dasar; soal essay; soal pilihan ganda.*

PENDAHULUAN

Evaluasi merupakan komponen integral dalam proses pembelajaran yang menentukan akurasi keputusan tentang ketercapaian kompetensi siswa. Kualitas instrumen penilaian, khususnya soal buatan guru, menjadi faktor krusial dalam menjamin keadilan dan objektivitas penilaian di sekolah dasar (Arikunto, 2021; Mardapi, 2012; Budiono & Hatip, 2023).

Di jenjang sekolah dasar, guru memiliki otonomi yang luas dalam menyusun instrumen penilaian sesuai dengan karakteristik siswa dan tuntutan kurikulum. Namun, otonomi tersebut sering kali tidak diimbangi dengan kompetensi telaah soal yang memadai, sehingga berpotensi menghasilkan soal yang kurang optimal (Phafiandita dkk., 2022; Savika, 2024; Marengke, 2022).

Kajian literatur terdahulu menunjukkan bahwa kelemahan utama soal buatan guru sering terletak pada aspek konstruksi dan fungsi pengecoh. Beberapa penelitian nasional mengungkapkan masih adanya rendahnya kualitas distraktor, ketidakseimbangan panjang pilihan jawaban, serta ketidakjelasan rumusan soal di jenjang sekolah dasar (Savika, 2024; Marambaawang, 2023; Rahayu & Sukenti, 2024; Mulia dkk., 2023).

Perkembangan kurikulum di Indonesia, terutama Kurikulum Merdeka, menuntut pergeseran paradigma penilaian dari sekadar tes sumatif menuju asesmen formatif yang bermakna dan diagnostik (Sugiri & Priatmoko, 2020; Budiono & Hatip, 2023; Faiz dkk., 2022). Dalam konteks ini, soal yang berkualitas tinggi menjadi kebutuhan mendesak.

State of the art penelitian penilaian saat ini semakin menekankan validitas konsekuensial, yaitu dampak soal terhadap motivasi, keadilan, dan pembelajaran siswa. Aspek etika juga semakin penting di tengah masyarakat yang majemuk untuk menghindari bias SARA dan kekerasan.

Kebaruan ilmiah penelitian ini terletak pada tiga hal. Pertama, analisis dilakukan secara komprehensif lintas enam jenjang kelas (I sampai VI), bukan terbatas pada satu kelas saja. Kedua, penelitian menggunakan data otentik hasil telaah guru di SDN 7 Pangkalpinang, bukan soal buatan peneliti sendiri. Ketiga, penelitian ini mengadopsi pendekatan deskriptif naratif yang mendalam untuk mengungkap alasan di balik setiap skor telaah, bukan sekadar menyajikan tabel kuantitatif. Pendekatan ini diharapkan

dapat memberikan gambaran yang lebih kaya tentang praktik penyusunan soal di lapangan.

Dalam konteks pendidikan di daerah kepulauan seperti Bangka Belitung, keragaman latar belakang siswa menuntut kepekaan guru dalam menyusun soal yang inklusif dan sesuai perkembangan siswa. Teori Zone of Proximal Development dari Vygotsky menjadi relevan dalam memastikan bahasa soal berada pada tingkat yang tepat.

Kesenjangan penelitian yang ingin diisi adalah minimnya dokumentasi praktik baik telaah soal yang dilakukan guru secara mandiri di sekolah dasar (Fauziana & Wulansari, 2021; Akhmadi, 2021). Sebagian besar studi sebelumnya hanya fokus pada hasil uji coba empiris tanpa menjelaskan proses berpikir kualitatif guru.

Oleh karena itu, penelitian ini memiliki urgensi tinggi untuk mendokumentasikan praktik telaah soal, memberikan bukti empiris implementasi Panduan Penulisan Soal Kemendikbud (2019) serta panduan pengembangan instrumen HOTS terkini, serta menjadi model bagi sekolah lain dalam meningkatkan kualitas penilaian.

Permasalahan penelitian dirumuskan sebagai berikut: Bagaimana kualitas soal buatan guru SD ditinjau dari aspek materi, konstruksi, bahasa, dan etika jika dianalisis secara mendalam per butir? Tujuan penelitian adalah mendeskripsikan kualitas soal tersebut serta mengidentifikasi praktik terbaik dan area perbaikan yang dapat direplikasi.

METODE

Penelitian ini menggunakan pendekatan deskriptif kualitatif dengan rancangan studi dokumen (Creswell & Creswell, 2022; Suprayitno dkk., 2024). Lokasi penelitian adalah SDN 7 Pangkalpinang, Kota Pangkalpinang, Provinsi Kepulauan Bangka Belitung. Pemilihan lokasi didasarkan pada aksesibilitas dokumen telaah soal yang lengkap dan representatif sebagai salah satu sekolah dasar negeri di daerah kepulauan.

Sumber data primer berupa dokumen telaah soal buatan guru kelas I, II, III, IV, V, dan VI tahun ajaran 2024/2025. Populasi penelitian adalah seluruh butir soal pilihan ganda dan essay yang telah ditelaah oleh guru, dengan teknik sampling jenuh sehingga seluruh populasi dianalisis.

Instrumen utama adalah lembar telaah kualitatif yang dikembangkan berdasarkan Panduan Penulisan Soal Kemendikbud (2019) serta prinsip-prinsip penyusunan butir soal pilihan ganda dan essay yang mutakhir (Gottlieb, 2023; Jumrah

dkk., 2023). Instrumen ini mencakup empat aspek utama: materi, konstruksi, bahasa, dan etika. Setiap aspek dilengkapi indikator spesifik yang dinilai secara deskriptif disertai catatan telaah mendalam.

Prosedur penelitian dilaksanakan dalam empat tahap utama. Tahap pertama adalah orientasi, yaitu membaca keseluruhan dokumen untuk memahami konteks dan karakteristik soal per kelas. Tahap kedua adalah reduksi data, yaitu memberi kode dan merangkum catatan telaah pada setiap indikator. Tahap ketiga adalah display data, yaitu menyusun matriks perbandingan antar kelas dan antar jenis soal (pilihan ganda dan essay). Tahap keempat adalah verifikasi dan kesimpulan, dilakukan melalui triangulasi antar dokumen telaah, kisi-kisi soal, serta diskusi antar penelaah untuk memastikan keabsahan temuan.

Analisis data menggunakan model interaktif Miles, Huberman, dan Saldaña yang meliputi reduksi data, penyajian data, dan penarikan kesimpulan secara berkelanjutan (Miles dkk., 2020). Keabsahan data dijaga melalui ketekunan pengamatan, triangulasi sumber data, dan audit trail yang lengkap. Alat bantu yang digunakan meliputi perangkat lunak pembaca PDF dan Microsoft Excel untuk tabulasi tematik serta penyusunan matriks analisis.

HASIL DAN PEMBAHASAN

Hasil

Hasil analisis kualitatif terhadap soal buatan guru kelas I sampai kelas VI SDN 7 Pangkalpinang menunjukkan kualitas yang secara keseluruhan sangat baik. Guru telah menunjukkan pemahaman yang baik terhadap prinsip penyusunan soal, meskipun masih terdapat beberapa area yang dapat ditingkatkan.

Aspek materi menjadi kekuatan utama. Hampir seluruh soal pilihan ganda dan essay telah sesuai dengan indikator pembelajaran yang ditetapkan. Pada kelas I, soal tentang huruf vokal dan konsonan serta bunyi buatan manusia menunjukkan kesesuaian yang sangat tinggi dengan kompetensi dasar Bahasa Indonesia. Demikian pula pada kelas III, soal ide pokok bacaan dan penggunaan tanda baca telah selaras dengan indikator yang diharapkan (Savika, 2024; Jumrah dkk., 2023).

Pengecoh pada soal pilihan ganda umumnya berfungsi efektif. Di kelas II, pengecoh pada materi aturan masih berada dalam konteks yang sama sehingga mampu membedakan pemahaman siswa. Pada kelas IV, pengecoh untuk materi imbuhan dan kalimat juga dibuat berdasarkan miskonsepsi umum siswa, bukan pilihan acak. Temuan ini berbeda dengan sejumlah penelitian terdahulu yang masih

menemukan banyak pengecoh tidak berfungsi atau kurang optimal (Savika, 2024; Marambaawang, 2023; Rahayu & Sukenti, 2024).

Prinsip satu jawaban paling benar juga terpenuhi di hampir semua butir soal. Hal ini menunjukkan guru telah memahami pentingnya objektivitas dalam penilaian sumatif, sejalan dengan kaidah penulisan butir pilihan ganda yang baku (Gottlieb, 2023).

Aspek konstruksi menunjukkan hasil yang baik meskipun masih terdapat ruang perbaikan. Pokok soal pada umumnya dirumuskan jelas dan tegas. Pada kelas rendah (I dan II), kalimat pendek dan langsung memudahkan siswa memahami maksud pertanyaan tanpa melihat pilihan jawaban (Gottlieb, 2023).

Beberapa catatan perbaikan muncul pada ketidakseimbangan panjang pilihan jawaban. Contohnya pada kelas III, terdapat pilihan jawaban yang lebih panjang dibandingkan opsi lain, yang berpotensi memberi petunjuk tidak langsung kepada siswa. Selain itu, beberapa soal essay masih menunjukkan ketergantungan antar butir karena gambar atau teks bacaan yang tidak disertakan secara lengkap.

Rumusan pokok soal dan pilihan jawaban sebagian besar efektif tanpa informasi yang berlebihan. Guru juga berhasil menghindari negatif ganda dan petunjuk yang mengarah ke jawaban benar pada mayoritas soal (Gottlieb, 2023; Stevens, 2023). Homogenitas pilihan jawaban terjaga dengan baik. Pada kelas V, pilihan jawaban untuk materi kalimat majemuk dan karya sastra akrostik semuanya berasal dari kategori materi yang sama, sehingga meningkatkan validitas konstruk soal.

Aspek bahasa merupakan salah satu aspek terkuat. Bahasa yang digunakan komunikatif dan sesuai dengan tingkat perkembangan siswa. Pada kelas I–II, kalimat sangat sederhana (7–10 kata), sementara pada kelas V–VI kalimat lebih kompleks namun tetap mudah dipahami (12–15 kata). Temuan ini selaras dengan teori Zone of Proximal Development dari Vygotsky.

Tidak ditemukan penggunaan bahasa daerah atau istilah lokal yang dapat menyulitkan siswa. Pengulangan kata yang tidak perlu juga berhasil diminimalkan, meskipun masih terdapat beberapa kasus minor pada soal kelas IV.

Aspek etika mencapai tingkat kesempurnaan yang konsisten. Seluruh soal dinyatakan aman, netral, dan bebas dari unsur SARA, kekerasan fisik, persekusi, pornografi, maupun propaganda. Di tengah masyarakat Bangka Belitung yang majemuk, kehati-hatian guru dalam memilih konteks dan kata-kata patut diapresiasi.

Tabel 1. Ringkasan Temuan Analisis Kualitas Soal Buatan Guru SD Kelas I–VI

No	Aspek	Kelas I–III (Rendah)	Kelas IV–VI (Tinggi)	Kesimpulan Umum	Kategori
1	Materi	Sangat baik, kontekstual, pengecoh berfungsi	Sangat baik, lebih analitis dan HOTS	Kekuatan utama	Sangat Baik
2	Konstruksi	Beberapa catatan (kejelasan, media tidak tersedia, kesalahan ketik)	Lebih baik, rumusan lebih tegas	Perlu perbaikan teknis	Baik
3	Bahasa	Beberapa kesalahan baku dan tanda baca	Lebih baku dan efektif	Perlu perbaikan	Cukup Baik
4	Etika	Sangat baik, aman, ramah anak	Sangat baik, aman, ramah anak	Kekuatan utama	Sangat Baik

Dari perspektif teori beban kognitif (Cognitive Load Theory), konstruksi soal yang jelas telah membantu mengurangi extraneous load sehingga siswa dapat fokus pada materi yang diuji. Pada soal kelas tinggi, guru telah berhasil menyusun soal yang mengukur keterampilan berpikir tingkat tinggi (C4–C6) sesuai revisi Taksonomi Bloom, sebagaimana ditekankan dalam pengembangan instrumen berorientasi HOTS di sekolah dasar (Ndiung & Jediut, 2020 → Jumrah dkk., 2023; Putri dkk., 2025; Ismael dkk., 2024).

Rincian hasil telaah per indikator pada keempat aspek disajikan pada Tabel 2 berikut, yang disusun mengikuti struktur kaidah penulisan soal: tiga indikator materi, sepuluh indikator konstruksi, serta empat indikator bahasa dan etika.

Tabel 2. Hasil Telaah Butir Soal Berdasarkan Aspek Materi, Konstruksi, Bahasa, dan Etika

No	Indikator Telaah	Temuan Analisis (Kelas I–VI)
A. Aspek Materi (3 indikator)		
A1	Soal sesuai dengan indikator/kompetensi dasar	Sesuai — kesesuaian indikator tinggi di seluruh jenjang.
A2	Pilihan jawaban homogen dan logis ditinjau dari segi materi	Sesuai — pengecoh sekategori dan berbasis miskonsepsi (mis. imbuhan kelas IV, kalimat majemuk kelas V).
A3	Hanya ada satu kunci jawaban yang paling benar	Sesuai — terpenuhi pada hampir seluruh butir.
B. Aspek Konstruksi (10 indikator)		
B1	Pokok soal dirumuskan singkat, jelas, dan tegas	Sesuai — terutama kelas rendah (kalimat pendek dan langsung).

B2	Rumusan pokok soal dan pilihan hanya memuat pernyataan yang diperlukan	Sebagian besar sesuai, tanpa informasi berlebihan.
B3	Pokok soal tidak memberi petunjuk ke arah kunci jawaban	Sesuai pada mayoritas butir.
B4	Pokok soal bebas dari pernyataan yang bersifat negatif ganda	Sesuai — guru berhasil menghindari negatif ganda.
B5	Gambar/grafik/tabel/stimulus disajikan jelas dan berfungsi	Catatan — beberapa soal essay memiliki stimulus/gambar yang tidak disertakan lengkap.
B6	Panjang pilihan jawaban relatif sama	Catatan — kelas III ditemukan opsi lebih panjang (berpotensi memberi petunjuk tidak langsung).
B7	Pilihan jawaban tidak menggunakan “semua/tidak satu pun benar”	Sesuai.
B8	Pilihan jawaban berupa angka/waktu disusun berurutan	Sesuai — disusun secara runtut.
B9	Butir soal tidak bergantung pada jawaban butir sebelumnya	Catatan — sebagian soal essay menunjukkan ketergantungan antar butir.
B10	Tidak terdapat kesalahan ketik atau teknis penulisan	Catatan — ditemukan kesalahan ejaan minor, terutama pada kelas rendah.

C. Aspek Bahasa & Etika (4 indikator)

C1	Bahasa sesuai kaidah Bahasa Indonesia (baku)	Sesuai — lebih baku pada kelas tinggi; catatan minor pada kelas rendah.
C2	Bahasa komunikatif dan sesuai jenjang perkembangan siswa (ZPD)	Sesuai — 7–10 kata (kelas I–II) hingga 12–15 kata (kelas V–VI).
C3	Tidak menggunakan bahasa daerah/istilah yang menyulitkan	Sesuai — tidak ditemukan bahasa daerah yang menyulitkan siswa.
C4	Bebas dari unsur SARA, kekerasan, pornografi, dan propaganda	Sesuai sempurna — soal aman dan ramah anak di semua jenjang.

Perbandingan antar jenis soal menunjukkan bahwa soal pilihan ganda lebih unggul pada aspek objektivitas dan fungsi pengecoh (Ramadhan dkk., 2023), sedangkan soal essay lebih kuat dalam mengukur kemampuan berpikir mendalam dan

ekspresi siswa, meskipun memerlukan perbaikan pada petunjuk dan kelengkapan stimulus.

Secara komparatif antar jenjang, semakin tinggi kelas, semakin kompleks materi yang diuji, namun tingkat kehati-hatian etika tetap konsisten. Hal ini menunjukkan adanya budaya mutu internal di sekolah.

Temuan penelitian ini menunjukkan bahwa kualitas butir soal yang baik merupakan indikator penting dalam keberhasilan proses asesmen pembelajaran. Menurut Suryadin (2024), instrumen penilaian yang disusun secara sistematis dan memenuhi kaidah penyusunan soal akan menghasilkan informasi yang lebih akurat mengenai pencapaian kompetensi peserta didik. Oleh karena itu, penerapan prinsip-prinsip penilaian yang mencakup kesesuaian dengan tujuan pembelajaran, kejelasan bahasa, serta ketepatan konstruksi soal sebagaimana ditemukan dalam penelitian ini mencerminkan upaya guru dalam melaksanakan asesmen yang berkualitas dan berorientasi pada peningkatan hasil belajar siswa

Meskipun demikian, beberapa kelemahan teknis seperti kesalahan ejaan, gambar yang tidak disertakan, dan ketergantungan antar butir masih perlu diperbaiki. Hal ini dapat diatasi melalui telaah rekan sejawat yang lebih intensif (Savika, 2024; Marengke, 2022).

Secara teoritis, hasil penelitian ini memperkuat pentingnya kualitas butir sebagai penentu keadilan penilaian; soal yang berkualitas tidak hanya mengukur kompetensi, tetapi juga berdampak positif terhadap motivasi dan keadilan siswa (Gottlieb, 2023).

Praktik kolaborasi antar guru dalam menyusun dan menelaah soal tampaknya telah terbentuk. Konsistensi gaya telaah lintas kelas menunjukkan adanya professional learning community di sekolah, yang dapat diperkuat melalui pelatihan analisis butir soal secara berkala (Savika, 2024; Marengke, 2022).

Secara keseluruhan, meskipun terdapat beberapa catatan perbaikan, kualitas soal buatan guru SDN 7 Pangkalpinang dapat dikategorikan sebagai sangat baik dan menjadi contoh baik bagi sekolah dasar lainnya.

Temuan pada aspek materi dan konstruksi sejalan dengan prinsip asesmen dalam Kurikulum Merdeka yang menekankan asesmen formatif dan diagnostik (Budiono & Hatip, 2023; Faiz dkk., 2022). Soal buatan guru SDN 7 Pangkalpinang, khususnya pada kelas tinggi, sudah mulai mengarah pada keterampilan berpikir tingkat tinggi atau HOTS (Putri dkk., 2025; Ismael dkk., 2024; Jumrah dkk., 2023).

Hal ini terlihat dari penggunaan stimulus bacaan pada soal essay kelas IV–VI yang menuntut siswa menganalisis dan mengevaluasi, bukan sekadar mengingat. Dengan demikian, praktik penyusunan soal di sekolah ini tidak hanya memenuhi kaidah teknis penulisan soal, tetapi juga responsif terhadap perubahan paradigma penilaian nasional. Keberhasilan ini mengindikasikan bahwa pelatihan internal atau MGMP yang berjalan efektif mampu menerjemahkan kebijakan makro kurikulum ke dalam produk asesmen mikro di kelas.

Implikasi Praktis

Hasil penelitian ini memiliki tiga implikasi praktis bagi penguatan mutu penilaian di sekolah dasar. Pertama, model telaah kualitatif per butir dengan indikator materi, konstruksi, bahasa, dan etika terbukti aplikatif dan dapat diadopsi sebagai Standar Operasional Prosedur (SOP) telaah soal tingkat sekolah. SOP ini akan membantu guru baru atau guru yang belum terbiasa melakukan review soal secara sistematis (Savika, 2024; Marengke, 2022).

Kedua, konsistensi kualitas etika yang sangat baik di semua jenjang menunjukkan bahwa sekolah telah memiliki budaya mutu internal. Budaya ini penting dijaga melalui kegiatan telaah rekan sejawat atau peer review rutin tiap akhir semester, sehingga kelemahan teknis seperti ketidakseimbangan panjang pilihan jawaban dapat diminimalisir sejak awal (Ramadhan dkk., 2023).

Ketiga, temuan bahwa soal essay masih lemah pada kelengkapan stimulus menjadi masukan bagi pengembang KKG dan dinas pendidikan untuk memfokuskan pelatihan pada penulisan soal uraian berbasis konteks, bukan hanya pilihan ganda (Jumrah dkk., 2023; Putri dkk., 2025).

KESIMPULAN DAN SARAN

Berdasarkan hasil analisis terhadap butir soal yang digunakan di SDN 7 Pangkalpinang, dapat disimpulkan bahwa kualitas instrumen penilaian yang disusun guru secara umum telah memenuhi kaidah penyusunan soal yang baik, baik dari aspek materi, konstruksi, bahasa, maupun kesesuaian dengan indikator pembelajaran. Sebagian besar soal telah mampu mengukur pencapaian kompetensi peserta didik sesuai tujuan pembelajaran yang ditetapkan. Instrumen penilaian yang baik harus memiliki kesesuaian dengan indikator, tingkat kejelasan bahasa, serta kemampuan mengukur hasil belajar secara objektif. Hasil penelitian menunjukkan bahwa aspek tersebut telah diterapkan dengan cukup baik oleh guru, terutama pada kesesuaian

materi dengan kompetensi yang diukur dan penggunaan bahasa yang mudah dipahami siswa.

Namun demikian, masih ditemukan beberapa kelemahan pada penyusunan soal, khususnya pada butir soal uraian yang belum sepenuhnya didukung stimulus yang kontekstual dan mampu mendorong keterampilan berpikir tingkat tinggi (Higher Order Thinking Skills/HOTS). Selain itu, analisis yang dilakukan masih bersifat kualitatif sehingga belum memberikan gambaran mengenai validitas, reliabilitas, tingkat kesukaran, dan daya pembeda setiap butir soal. Oleh karena itu, guru perlu melakukan telaah butir soal secara berkelanjutan, baik melalui analisis kualitatif maupun kuantitatif. Kegiatan peer review, kolaborasi dalam KKG/MGMP, serta pemanfaatan teknologi dalam pengembangan instrumen assessment dapat menjadi langkah strategis untuk meningkatkan kualitas penilaian. Dengan demikian, assessment tidak hanya berfungsi sebagai alat pengukur hasil belajar, tetapi juga sebagai dasar pengambilan keputusan untuk memperbaiki proses pembelajaran dan meningkatkan mutu pendidikan secara berkelanjutan.

DAFTAR PUSTAKA

- Akhmadi, M. N. (2021). Analisis butir soal evaluasi Tema 1 Kelas 4 SDN Plumbungan menggunakan program Anates. *Ed-Humanistics: Jurnal Ilmu Pendidikan*, 6(1), 799–806. <https://doi.org/10.33752/ed-humanistics.v6i1.1464>
- Budiono, A. N., & Hatip, M. (2023). Asesmen pembelajaran pada Kurikulum Merdeka. *Axioma: Jurnal Matematika dan Pembelajaran*, 8(1), 109–123. <https://doi.org/10.56013/axi.v8i1.2044>
- Creswell, J. W., & Creswell, J. D. (2022). *Research design: Qualitative, quantitative, and mixed methods approaches* (6th ed.). SAGE Publications.
- Faiz, A., Parhan, M., & Ananda, R. (2022). Paradigma baru dalam kurikulum prototipe. *Edukatif: Jurnal Ilmu Pendidikan*, 4(1), 1544–1550. <https://doi.org/10.31004/edukatif.v4i1.2410>
- Fauziana, A., & Wulansari, A. D. (2021). Analisis kualitas butir soal ulangan harian di sekolah dasar dengan model Rasch. *Ibriez: Jurnal Kependidikan Dasar Islam Berbasis Sains*, 6(1), 10–19. <https://doi.org/10.21154/ibriez.v6i1.112>
- Gottlieb, M. (2023). Educator's blueprint: A how-to guide for developing high-quality multiple-choice questions. *AEM Education and Training*, 7(1), e10836. <https://doi.org/10.1002/aet2.10836>
- Ismael, R., Retnawati, H., Arovah, N. I., & Imawan, O. R. (2024). Contexts proposed by teachers in Papua for developing mathematics HOTS assessment instruments: A phenomenological study. *Journal of Education and E-Learning Research*, 11(3), 548–556. <https://doi.org/10.20448/jeelr.v11i3.5922>

- Jumrah, dkk. (2023). Pengembangan instrumen tes berbasis HOTS dengan pendekatan pengukuran Rasch pada pelajaran Matematika topik bangun ruang untuk siswa sekolah dasar. *Jurnal Basicedu*, 7(1), 11–27. <https://doi.org/10.31004/basicedu.v7i1.4207>
- Kementerian Pendidikan dan Kebudayaan. (2019). Panduan penulisan soal HOTS sekolah dasar. Pusat Penilaian Pendidikan.
- Marengke, M. (2022). Pelatihan analisis butir soal bagi guru sekolah dasar. *Jurnal Pengabdian Masyarakat Pendidikan*, 4(2), 45–58.
- Marambaawang, D. (2023). Analisis kualitas butir soal penilaian akhir di sekolah dasar. *Jurnal Pendidikan Dasar*, 12(1), 78–92.
- Miles, M. B., Huberman, A. M., & Saldaña, J. (2020). *Qualitative data analysis: A methods sourcebook* (4th ed.). SAGE Publications.
- Mulia, N., Zulyusri, Z., & Violita, V. (2023). Analisis kualitas butir soal pilihan ganda mata pelajaran Biologi pada Ujian Tengah Semester (UTS) Kelas XI. *Jurnal Metaedukasi: Jurnal Ilmiah Pendidikan*, 4(1), 1–8. <https://doi.org/10.37058/metaedukasi.v4i1.4088>
- Phafiandita, A. N., Permadani, A., Pradani, A. S., & Wahyudi, M. I. (2022).
- Urgensi evaluasi pembelajaran di kelas. *JIRA: Jurnal Inovasi dan Riset Akademik*, 3(2), 111–121. <https://doi.org/10.47387/jira.v3i2.262>
- Putri, C. M., Adrias, A., & Syam, S. S. (2025). Mengasah kecerdasan anak dengan soal HOTS: Strategi efektif untuk pembelajaran di SD. *Morfologi: Jurnal Ilmu Pendidikan, Bahasa, Sastra dan Budaya*, 3(2), 287–296. <https://doi.org/10.61132/morfologi.v3i2.1581>
- Ramadhan, W., Malahati, F., Romadhon, K., & Ramadhan, S. (2023). Analisis butir soal tipe multiple choice questions pada penilaian harian sekolah dasar. *Tarbiyah Wa Ta'lim: Jurnal Penelitian Pendidikan dan Pembelajaran*, 10(2), 93–105. <https://doi.org/10.21093/twt.v10i2.6155>
- Suryadin, A. (2024). *Assessment Prestasi Belajar (Untuk Guru dan Mahasiswa Keguruan)*. CV Eureka Media Aksara.
- Rahayu, S., & Sukenti, D. (2024). Analisis kualitas distraktor dan konstruksi soal buatan guru SD. *Jurnal Pendidikan dan Evaluasi*, 9(1), 112–125.
- Savika, H. I. (2024). Peran analisis butir soal terhadap kualitas soal, kompetensi guru, dan prestasi belajar peserta didik di sekolah dasar. *PANDU: Jurnal Pendidikan Anak dan Pendidikan Umum*, 2(2), 43–51. <https://doi.org/10.59966/pandu.v2i2.856>
- Stevens, S. P. (2023). Practical guidance for writing multiple-choice test questions. *Journal of Assessment in Higher Education*, 5(2), 67–82.

- Suprayitno, D., Ahmad, A., Tartila, Sa'dianoor, & Aladdin, Y. A. (2024). Metodologi penelitian kualitatif: Teori komprehensif dan referensi wajib bagi peneliti. PT. Sonpedia Publishing Indonesia.
- Tarigan, E. F., Nilmarito, S., Islamiyah, K., Darmana, A., & Suyanti, R. D. (2022). Analisis instrumen tes menggunakan Rasch model dan software SPSS 22.0. *Jurnal Inovasi Pendidikan Kimia*, 16(2), 92–96.
<https://doi.org/10.15294/jipk.v16i2.30530>