



Prosedur Pengembangan Tes Hasil Belajar Standar: Dari Kisi-Kisi Hingga Perakitan Soal

Aprilia Mengga¹, Angreni Tamal Nenohai²

Pendidikan Agama Kristen, Institut Agama Kristen Negeri Kupang, Indonesia^{1,2}

*Email: apriliamengga@gmail.com, ninynenohai@gmail.com,

Diterima: 13-06-2026 | Disetujui: 17-06-2026 | Diterbitkan: 19-06-2026

ABSTRACT

The procedure for developing standardized learning achievement tests is a systematic process aimed at producing evaluation instruments that are valid, reliable, and objective. The development stages begin with the preparation of a test blueprint as a guideline for determining competencies, materials, indicators, and levels of item difficulty. This is followed by item writing and review to ensure the appropriateness of the content, construction, language, and alignment with learning objectives and Bloom's Taxonomy. The test is then piloted and analyzed to determine the level of difficulty, discriminating power, distractor effectiveness, validity, and reliability of the test items. Based on the analysis results, the items are revised, selected, and assembled into a test instrument ready for use. Through a structured procedure, standardized learning achievement tests can serve as accurate measurement tools for assessing students' competency achievement.

Keywords: Learning achievement test, test blueprint, validity, reliability, item analysis, educational evaluation.

ABSTRAK

Prosedur pengembangan tes hasil belajar standar merupakan proses sistematis yang bertujuan menghasilkan instrumen evaluasi yang valid, reliabel, dan objektif. Tahapan pengembangan dimulai dari penyusunan kisi-kisi sebagai pedoman dalam menentukan kompetensi, materi, indikator, dan tingkat kesulitan soal. Selanjutnya dilakukan penulisan dan penelaahan soal untuk memastikan kesesuaian isi, konstruksi, bahasa, serta keterkaitannya dengan tujuan pembelajaran dan Taksonomi Bloom. Tes kemudian diujicobakan dan dianalisis untuk mengetahui tingkat kesukaran, daya pembeda, efektivitas pengecoh, validitas, dan reliabilitas soal. Berdasarkan hasil analisis, soal diperbaiki, diseleksi, dan dirakit menjadi instrumen tes yang siap digunakan. Dengan prosedur yang terstruktur, tes hasil belajar standar mampu menjadi alat ukur yang akurat dalam menilai pencapaian kompetensi peserta didik.

Katakunci: Tes hasil belajar, kisi-kisi, validitas, reliabilitas, analisis butir soal, evaluasi pendidikan.

Bagaimana Cara Sitasi Artikel ini:

Mengga, A., & Nenohai, A. T. . (2026). Prosedur Pengembangan Tes Hasil Belajar Standar: Dari Kisi-Kisi Hingga Perakitan Soal. *Educational Journal*, 1(4), 2072-2079. <https://doi.org/10.63822/e5g4vb11>

PENDAHULUAN

Menurut Undang-Undang Republik Indonesia Nomor 20 Pasal 1 tahun 2003 tentang Sistem Pendidikan Nasional bahwa Pendidikan adalah usaha sadar dan terencana untuk mewujudkan suasana belajar dan proses pembelajaran agar peserta didik secara aktif mengembangkan potensi dirinya untuk memiliki kekuatan spiritual keagamaan, pengendalian diri, kepribadian, kecerdasan, akhlak mulia, serta keterampilan yang diperlukan dirinya, masyarakat, bangsa dan Negara (Abdullah, 2022).

Sehubungan dengan hal itu, telah dilakukan berbagai upaya oleh pemerintah untuk meningkatkan mutu pendidikan antara lain diadakan perubahan dan penyesuaian kurikulum di semua jenjang pendidikan, perbaikan mutu pendidikan seperti penataran guru-guru, pengadaan buku paket dan penambahan sarana dan prasarana untuk kegiatan belajar mengajar di kelas termasuk pengelolaan proses pembelajaran setiap mata pelajaran. Pengelolaan proses pembelajaran guru dituntut untuk mampu mengembangkan instrumen penilaian dan evaluasi proses dan hasil belajar (Simanjuntak et al., 2022).

Melalui evaluasi, kita dapat mengetahui dan mengambil keputusan dari hasil pengukuran yang dilakukan. Evaluasi pembelajaran merupakan proses penentuan nilai berdasarkan kriteria untuk memperoleh informasi atau data dari tingkat kemampuan dan pemahaman siswa baik dari segi kognitif, psikomotor, maupun afektif nantinya akan digunakan untuk memperbaiki hal-hal yang memang perlu diperbaiki pada proses pembelajaran.

Menurut Purwanto, dalam arti luas evaluasi adalah proses merencanakan, memperoleh dan menyediakan informasi yang sangat diperlukan untuk membuat alternative-alternatif keputusan. Hal yang sangat mempengaruhi kualitas penilaian, evaluasi proses dan hasil belajar adalah instrumen yang akan digunakan. Instrumen merupakan alat yang digunakan untuk mengukur kegiatan evaluasi pembelajaran. Salah satunya adalah tes. Tes yang digunakan di sekolah biasanya untuk mengukur tingkat kemampuan atau prestasi siswa dalam bidang kognitif, seperti pengetahuan, pemahaman, analisis, sintesis, dan evaluasi. Sebagaimana di jelaskan bahwa tes merupakan alat ukur prosedur yang digunakan untuk mengetahui atau mengukur sesuatu dalam suasana, dengan cara dan aturan-aturan yang sudah ditentukan. Tes dapat juga di jadikan untuk mengukur banyaknya pengetahuan yang di peroleh individu dari suatu bahan pepelajaran yang terbatas pada tingkat tertentu. (Magdalena, Hifziyah, Aeni, & Rahayu, 2020; Ramadani, Yennita, & Ernidawati, 2024)

METODE PENELITIAN

Penelitian ini menggunakan metode studi pustaka (library research) dengan pendekatan deskriptif kualitatif untuk mengkaji prosedur pengembangan tes hasil belajar standar. Data penelitian dikumpulkan dari berbagai literatur terpercaya seperti buku, jurnal ilmiah, dan peraturan perundang-undangan yang berkaitan dengan evaluasi pembelajaran. Fokus kajiannya mencakup seluruh tahapan pengembangan tes, mulai dari penyusunan kisi-kisi, penulisan dan penelaahan soal, uji coba, analisis butir soal, hingga perakitan tes menjadi instrumen yang valid, reliabel, dan objektif (Faujah, Mulyani, Ananda, & Witarsa, 2022).

Proses analisis data dilakukan menggunakan teknik analisis isi (content analysis) dengan cara mengidentifikasi konsep utama, membandingkan pendapat para ahli, dan mengelompokkan informasi

berdasarkan tema. Hasil analisis tersebut kemudian disajikan secara deskriptif untuk memberikan gambaran komprehensif mengenai langkah-langkah penyusunan tes. Melalui pendekatan ini, penelitian diharapkan dapat memberikan pemahaman mendalam sekaligus menjadi referensi bagi para pendidik dalam menyusun instrumen evaluasi pembelajaran yang berkualitas.

HASIL PENELITIAN

Prosedur pengembangan tes hasil belajar standar merupakan proses sistematis untuk menghasilkan alat ukur yang valid, reliabel, dan objektif dalam menilai pencapaian kompetensi siswa. Berikut adalah tahapan lengkap prosedur pengembangan tes hasil belajar standar, mulai dari penyusunan kisi-kisi hingga perakitan soal:

1. Menyusun Spesifikasi Tes (Kisi-Kisi)

□ Menyusun spesifikasi tes adalah menentukan tujuan tes, menyusun kisi-kisi, memilih bentuk tes, dan tes sumatif. Hal ini dilakukan agar mempermudah dalam menulis soal dan siapa saja yang menulis soal akan menghasilkan tingkat kesulitan yang relatif sama (Magdalena et al., 2020).

Dalam dunia evaluasi pendidikan, yang dimaksud dengan tes adalah cara atau prosedur pengukuran dan penilaian dalam bidang pendidikan, yang berbentuk pemberian tugas atau serangkaian tugas, baik berupa pertanyaan yang harus dijawab, atau perintah oleh testee, sehingga dapat dihasilkan nilai yang melambangkan tingkah laku atau prestasi, nilai yang dapat dibandingkan dengan nilai yang dicapai oleh testee lainnya, atau dibandingkan dengan nilai standar tertentu. Dalam membuat tes, harus dibuat secara logis dan rasional mengenai pokok-pokok materi apa saja yang patut ditanyakan sebagai bahan pengetahuan penting untuk diketahui dan dipahami oleh peserta didik, instrumen berbentuk tes digunakan untuk mengukur pencapaian kompetensi kognitis peserta didik (Phafiandita, Permadani, Pradani, & Wahyudi, 2022).

1. Menyusun Kisi-kisi (Blueprints)

Kisi-kisi merupakan sebuah struktur atau matriks yang mencakup kriteria mengenai soal-soal yang dibutuhkan atau yang akan dibuat. Kisi-kisi juga dikenal sebagai cetak biru yang menggambarkan kompetensi dan materi yang akan diujikan. Bentuknya adalah tabel yang menjelaskan rincian materi dan perilaku, serta proporsi yang diinginkan oleh penilai. Setiap kotak di tabel diisi dengan angka yang menunjukkan jumlah soal (Amanullah, 2020).

Kisi-kisi juga merupakan kerangka yang sangat krusial dalam merancang alat pembelajaran yang efektif. Kisi-kisi dapat diartikan sebagai panduan yang mengatur dan mengorganisasi materi yang akan dievaluasi, sehingga alat yang dihasilkan dapat mencerminkan sasaran pembelajaran yang telah ditentukan, dapat dianggap sah dan dapat dipercaya. Kisi-kisi tidak hanya sebagai sarana pendukung, tetapi juga merupakan komponen vital dalam proses evaluasi pendidikan yang terstruktur dan terencana.

2. Tujuan merancang kisi-kisi

Tujuan utama dalam merancang kisi-kisi adalah memberikan kerangka yang jelas dalam pengembangan alat evaluasi, membantu pengajar dalam merumuskan pertanyaan atau tugas yang terkait dengan sasaran pembelajaran, dan juga memastikan bahwa semua kompetensi yang diharapkan dapat dinilai dengan adil. Dengan adanya kisi-kisi, pengajar dapat lebih mudah menentukan materi yang harus

diuji dan memastikan bahwa alat evaluasi tidak hanya menilai pengetahuan siswa, tetapi juga keterampilan dan sikap yang diharapkan (Muhadi, Jarir, Khairina, Rajuna, & Prasetyo, 2025).

Melibatkanmurid dalam proses pembuatan kisi-kisi, mereka akan merasa lebih memiliki tanggung jawab terhadap pembelajaran mereka, yang dapat memperbaiki motivasi dan partisipasi mereka. Kerja sama ini juga berpotensi membangun rasa kepemilikan di kalangan murid yangpada akhirnya bisa memperkuat komitmen mereka dalam menjalani proses belajar. Kisi-kisi pembelajaran harus mencakup isi yang relevan dengan masalah terbaru dan kebutuhan masyarakat. Namun, sangat penting untuk memastikan bahwa referensi yang digunakan untuk mendukung pernyataan ini relevan dan menguatkan klaim tersebut. Kualitas materi ajar berpengaruh besar terhadap minat dan motivasi belajar siswa, sehingga relevansi isi tidak boleh diabaikan. Standar kisi-kisi pembelajaran juga dapat ditingkatkan melalui umpan balik yang positif dari rekan sejawat dan para ahli. Proses penilaian dan revisi yang terus-menerus akan membantu memastikan bahwa kisi-kisi tetap efektif dan sesuai dengan sasaran pembelajaran (Sati, 2026)

Kerangka kerja matriks yang menjadi pedoman penulisan soal. Komponen utamanya mencakup:Standar Kompetensi (SK) dan Kompetensi Dasar (KD), Materi pokok dan indikator soal. Ranah kognitif (C1–C6), bentuk soal (Pilihan Ganda atau Uraian), alokasi nomor soal dan proporsi tingkat kesulitan (mudah, sedang, sukar) dan distribusi ranah kognitif (misalnya mengacu pada Taksonomi Bloom) serta menentukan panjang tes yakni menghitung jumlah butir soal yang proporsional sesuai bobot materi pembelajaran.

2. Menulis Soal Tes (Item Writing)

Menulis soal tes merupakan penjabaran dari indikator menjadi pertanyaan-pertanyaan yang karakteristiknya sesuai dengan perincian pada kisi-kisi yang telah dibuat dengan memperhatikan hal-hal berikut; (1) Kaidah Konstruksi: Penulisan harus mematuhi kaidah penulisan soal yang baik, seperti kesesuaian bahasa, tidak mengandung bias budaya atau SARA, dan hanya memiliki satu kunci jawaban yang paling benar (untuk soal objektif). (2) Menggunakan Kartu Soal: Setiap butir soal ditulis pada kartu soal yang mencantumkan identitas materi dan kunci jawaban (Resita, 2022).

3. Menelaah Soal Tes (Validasi Kualitatif)

Menelaah soal tes merupakan hal yang dilakukan untuk meminimalisir kesalahan atau kekurangan. Salah satu tahapan untuk menghasilkan tes yang baik adalah dengan melakukan penelaahan tes dengan melibatkan pakar pendidikan yaitu Dr. Muhammad Japar, M.Si (Magdalena et al., 2020).

Penelaahan butir tes didahului dengan penetapan level tes berdasarkan perjenjangan Taksonomi Bloom edisi revisi dan kesesuaian antara instrumen tes tertulis, bentuk pilihan ganda dan uraian dengan materi pembelajaran. Selain itu pakar memberikan masukan lain seperti kesesuaian konten, konstruksi, dan bahasa dengan konsep Taksonomi Bloom. Beberapa hal yang harus diperhatikan: (1) Aspek Telaah: Reviewer (dosen atau guru senior) memeriksa kesesuaian soal dengan kisi-kisi, kejelasan stimulus, dan ketepatan pilihan pengecoh. (2) Revisi Awal: Memperbaiki kalimat yang ambigu atau mengganti soal yang tidak valid secara logis. Sebelum dirakit, draf soal wajib ditelaah ulang untuk meminimalkan kesalahan atau cacat pada butir soal. (3) Validasi Kualitatif:Melibatkan expert judgment (pendapat para ahli/rekan sejawat) untuk menilai kesesuaian materi, konstruksi penulisan, dan bahasa.Pada tahap ini, soal yang tidak memenuhi syarat akan direvisiatau didrop (Magdalena et al., 2020).

4. Melakukan Uji Coba Tes

Uji coba tes dilakukan sebagai sarana memperoleh data empiris untuk membedakan tes standar dengan tes biasa tentang tingkat kebaikan soal yang telah disusun dengan tujuan memperoleh respons jawaban siswa untuk dianalisis secara statistik. Perangkat tes yang telah dirakit diujicobakan pada sampel populasi yang representatif (Kurniawan, 2022).

5. Menganalisis Butir Soal

Analisis butir soal merupakan aktivitas seorang pengajar untuk mengevaluasi pencapaian belajar siswa dan memperbaiki kualitas soal yang telah dibuat. Aktivitas ini mencakup pengumpulan, penyusunan, dan pemanfaatan informasi dari jawaban siswa untuk menciptakan. Tujuannya untuk menilai setiap item soal supaya bisa melahirkan soal-soal yang bermutu saat dilaksanakan ujian (YUSUF, 2024).

Proses analisis ini juga berfungsi untuk memperbaiki atau mengganti soal-soal yang tidak memenuhi kriteria untuk diujikan. Soal yang baik adalah soal yang mampu menunjukkan apakah siswa telah menguasai materi yang diberikan atau tidak. Analisis butir-butir soal tes juga dilakukan agar dapat mengetahui tingkat kesulitan butir soal, daya pembeda, dan efektivitas pengecoh. Data hasil uji coba diolah menggunakan program statistik (seperti SPSS, Anates, atau IteMan) untuk mengetahui karakteristik psikometrik setiap soal (Savika & Zuhriyah, 2024).

Tingkat Kesukaran (Item difficulty)

Hal ini dilakukan untuk menilai proporsi siswa yang menjawab benar (mudah, sedang, sukar) dan menilai tingkat kesulitan soal-soal dalam tes. Tujuannya agar bisa ditentukan bobot yang sesuai untuk soal-soal untuk mencerminkan tingkat kesukarannya. Tingkat kesukaran soal harus diidentifikasi, dan hal ini dapat diketahui berdasarkan hasil jawaban siswa. Jika jumlah siswa yang menjawab dengan benar sedikit, maka soal tersebut dianggap memiliki tingkat kesulitan yang tinggi. Indeks Tingkat Kesulitan (ITK), yang juga dikenal sebagai ukuran yang menilai tingkat kesulitan butir soal dalam pengujian siswa, menunjukkan seberapa mudah atau sulit soal tersebut. Skor yang digunakan memiliki rentang antara 0,00 hingga 1,00 (Rajagukguk & Naibaho, 2023).

Kesulitan soal dikategorikan menjadi tiga kelompok: mudah (0,71-1,00), sedang (0,31-0,70), dan sulit (0,00-0,30). Proses untuk mengukur tingkat kesulitan ini dilakukan dengan menggunakan perangkat lunak IteMan. Tingkat kesukaran suatu soal merupakan peluang untuk mendapatkan jawaban yang benar berdasarkan tingkat kemampuan tertentu, biasanya dinyatakan dalam bentuk indeks. Semakin tinggi angka indeks tingkat kesulitan yang diperoleh dari perhitungan, semakin mudah soal tersebut (Magdalena, Fauziah, Fазiah, & Nupus, 2021).

Daya Pembeda (discrimination power)

Mengukur kemampuan soal dalam membedakan kelompok siswa berkemampuan tinggi dan rendah. Analisis kemampuan membedakan dilakukan dengan menilai item-item, pertanyaan dalam ujian untuk menentukan seberapa efektif mereka dalam mengukur keberhasilan siswa (Apriliyanti, Saputri, Hakim, & Pratiwi, 2024).

Efektivitas Pengecoh (distractor)

Memastikan semua opsi salah (distraktor) berfungsi dengan baik. Opsi pengecoh yang telah teruji keberhasilannya bisa disimpan dalam bank soal karena mampu menipu siswa dengan efektif. Namun, pilihan pengecoh yang belum terbukti berhasil harus diperbaiki atau dihapus (Magdalena, 2025).

Validitas & Reliabilitas

Menilai keabsahan empiris butir soal dan keajekan skor instrumen. Validitas dan reliabilitas dilakukan agar dapat dijadikan alat ukur yang betul-betul berkualitas dan sah sesuai kemampuan siswa. Tes yang valid adalah tes yang mengukur dengan tepat keadaan yang ingin diukur. Sebaliknya, tes dikatakan tidak valid bila digunakan untuk mengukur suatu keadaan yang tidak tepat diukur dengan tes tersebut. Maka validitas berhubungan dengan tes apa yang mesti diukurinya dan seberapa baik melakukannya (KENDAL, n.d.).

Begitu pun dengan uji reliabilitasnya yang berhubungan pada tingkat akurasi tes apa yang mesti diukur. Menurut Thorndike dan Hagen Reliabilitas berhubungan pada akurasi instrumen atau tes untuk menghasilkan kecermatan hasil ukur. Jadi, Syarat yang harus dipenuhi untuk menjadi alat ukur hasil belajar yang baik berhubungan dengan validitas dan reliabilitas. sehingga Tes yang memenuhi syarat alat ukur yang baik dapat menghasilkan hasil ukur yang akurat (Magdalena et al., 2020).

6. Memperbaiki dan Menyeleksi Soal

Memperbaiki tes adalah Langkah yang biasanya dilakukan tes butir soal, yaitu memperbaiki masing-masing butir soal yang ternyata masih belum baik. Setelah proses penelaahan oleh pakar pendidikan, maka selanjutnya instrumen direvisi berdasarkan saran yang telah diberikan (Ida & Musyarofah, 2021). Hasil revisi tersebut diuraikan berdasarkan penjabaran indikator menghasilkan 10 butir soal pilihan ganda dan 5 butir soal uraian. Berdasarkan hasil analisis empiris, butir-butir soal dikelompokkan ke dalam tiga kategori:

- Diterima: Soal yang memenuhi standar statistik langsung disimpan.
- Direvisi: Soal yang memiliki kelemahan kecil pada kalimat atau pilihan pengecoh diperbaiki.
- Ditolak: Soal yang tidak valid atau memiliki daya pembeda buruk dibuang.

7. Perakitan Soal (Test Assembly)

Dalam merakit soal, hal-hal yang dapat memengaruhi validitas soal seperti nomor urut soal, pengelompokan bentuk soal, layout, dan sebagainya harus diperhatikan karena walaupun butir-butir soal yang disusun sudah baik tetapi jika penyusunannya sembarang dapat menyebabkan soal tersebut menjadi tidak baik. Banyaknya butir soal tes untuk setiap indikator minimal satu untuk setiap butir soal (MARENGKE, n.d.).

Penyusunan butir soal disesuaikan dengan tingkatan kognitif yang sesuai dengan Taksonomi Bloom edisi revisi yaitu level Mengingat (C1), memahami (C2), mengaplikasikan (C3), menganalisis (C4), Mengevaluasi (C5), dan mengkreasi (C6).

Tahap akhir penyusunan instrumen siap pakai (Izzah, Azizah, Ekawati, & Rufiana, 2025).

- Pengurutan Soal: Menyusun soal terpilih berdasarkan tingkat kesukaran (dari mudah ke sukar) atau berdasarkan kelompok materi sejenis agar tidak mematahkan semangat siswa. Misalnya: Pilihan Ganda dikelompokkan dengan Pilihan Ganda.

□ Kelengkapan Dokumen: Menambahkan petunjuk pengerjaan yang jelas dan tegas kepada peserta tes mengenai cara menjawab dan alokasi waktu, lembar jawaban, serta rubrik penskoran baku.

SIMPULAN

Prosedur pengembangan tes hasil belajar standar merupakan proses yang sistematis untuk menghasilkan alat evaluasi yang valid, reliabel, dan objektif. Proses ini dimulai dari penyusunan spesifikasi tes atau kisi-kisi yang berfungsi sebagai pedoman dalam menentukan kompetensi, materi, indikator, serta tingkat kesulitan soal. Dengan adanya kisi-kisi, penyusunan soal menjadi lebih terarah dan mampu mencerminkan tujuan pembelajaran secara menyeluruh.

Adapun tahap penulisan dan penelaahan soal. Penulisan soal harus mengikuti kaidah konstruksi yang baik agar tidak menimbulkan bias dan memiliki tingkat kejelasan yang tinggi. Setelah itu, soal ditelaah oleh ahli atau reviewer untuk memastikan kesesuaian materi, bahasa, dan konstruksi dengan indikator pembelajaran serta Taksonomi Bloom. Tahap ini penting untuk meminimalkan kesalahan dan meningkatkan kualitas soal sebelum diujicobakan.

Selanjutnya, tes yang telah disusun perlu diuji coba dan dianalisis secara empiris. Analisis butir soal dilakukan untuk mengetahui tingkat kesukaran, daya pembeda, efektivitas pengecoh, validitas, dan reliabilitas soal. Melalui proses ini dapat diketahui apakah suatu soal layak digunakan, perlu direvisi, atau harus dibuang. Dengan demikian, kualitas instrumen evaluasi dapat terus diperbaiki agar benar-benar mampu mengukur kemampuan siswa secara tepat.

Tahap akhir adalah memperbaiki, menyeleksi, dan merakit soal menjadi instrumen tes yang siap digunakan. Penyusunan soal dilakukan secara sistematis berdasarkan tingkat kesulitan dan ranah kognitif agar peserta didik dapat mengerjakan tes dengan baik. Selain itu, kelengkapan seperti petunjuk pengerjaan, lembar jawaban, dan rubrik penskoran juga harus disiapkan secara jelas. Dengan prosedur yang terencana dan terstruktur, tes hasil belajar standar dapat menjadi alat ukur yang akurat dalam menilai pencapaian kompetensi peserta didik.

DAFTAR PUSTAKA

- Abdullah, M. (2022). Lembaga Pendidikan Sebagai Suatu Sistem Sosial (Studi Tentang Peran Lembaga Pendidikan di Indonesia Dalam Perspektif Undang-Undang Nomor 20 Tahun 2003 Tentang Sistem Pendidikan Nasional). *Mamba'ul'Ulum*, 38–48.
- Amanullah, R. (2020). Analisis Kisi-kisi Soal Ujian Akhir Semester Ganjil Mata Pelajaran Bahasa Indonesia Kelas VIII SMP YLPI Pekanbaru Tahun Ajaran 2017/2018. Universitas Islam Riau.
- Apriliyanti, K. D., Saputri, A. K. A., Hakim, L., & Pratiwi, V. (2024). Analisis Butir Soal HOTS Pilihan Ganda Pada Elemen Komputer Akuntansi Di SMKN 10 Surabaya Menggunakan Aplikasi Anates. *PESHUM: Jurnal Pendidikan, Sosial Dan Humaniora*, 4(1), 727–738.
- Faujah, H., Mulyani, R. D., Ananda, R., & Witarso, R. (2022). Analisis standar penilaian pendidikan dasar: Studi literatur review. *JPDI (Jurnal Pendidikan Dasar Indonesia)*, 7(3), 90.
- Ida, F. F., & Musyarofah, A. (2021). Validitas dan reliabilitas dalam analisis butir soal. *Al-Muarrib Journal of Arabic Education*, 1(1), 34–44.

- Izzah, A. N., Azizah, S. S. L., Ekawati, R., & Rufiana, I. S. (2025). Eksplorasi Analisis Butir Soal Bahasa Indonesia di Kelas II Sekolah Dasar Berdasarkan Taksonomi Bloom. *Jurnal Pendidikan Dan Pembelajaran Indonesia (JPPI)*, 5(2), 1031–1043.
- KENDAL, B. (n.d.). VALIDITAS DAN RELIABILITAS INSTRUMEN TES BUATAN GURU PENDIDIKAN AGAMA ISLAM MATA PELAJARAN Fiqih Kelas VII di MTS. AL-MA'ARIF NU 02.
- Kurniawan, H. (2022). Pengantar praktis penyusunan instrumen penelitian. Deepublish.
- Magdalena, I. (2025). Implementasi Pembelajaran Abad 21 dan Bank Soal. Fahmi Karya.
- Magdalena, I., Fauziah, S. N., Fiazah, S. N., & Nupus, F. S. (2021). Analisis validitas, reliabilitas, tingkat kesulitan dan daya beda butir soal ujian akhir semester tema 7 kelas III SDN Karet 1 Sepatan. *Bintang*, 3(2), 198–214.
- Magdalena, I., Hifziah, M., Aeni, V. N., & Rahayu, R. P. (2020). Pengembangan Instrumen Tes Siswa Tingkat Sekolah Dasar Kabupaten Tangerang. *Nusantara*, 2(2), 227–237.
- MARENGKE, M. (n.d.). PERSEPSI KONSEPTUALITAS PERAKITAN BUTIR SOAL OLEH.
- Muhadi, M., Jarir, J., Khairina, K., Rajuna, R., & Prasetyo, E. (2025). Evaluasi perencanaan desain pembelajaran, pelaksanaan proses kegiatan pembelajaran, dan evaluasi instrumen hasil pembelajaran. *Edu Society: Jurnal Pendidikan, Ilmu Sosial Dan Pengabdian Kepada Masyarakat*, 5(2), 156–165.
- 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.
- Rajagukguk, M. J. T., & Naibaho, D. (2023). Mampu Memilih Soal Berdasarkan Tingkat Kesukaran. *Jurnal Pendidikan Sosial Dan Humaniora*, 2(4), 12736–12747.
- Ramadani, R. D., Yennita, Y., & Ernidawati, E. (2024). Pengembangan instrumen tes berbasis literasi sains siswa smp pada materi getaran dan gelombang. *Silampari Jurnal Pendidikan Ilmu Fisika*, 6(1), 25–34.
- Resita, A. (2022). PENGEMBANGAN INSTRUMEN TES PEMBELAJARAN TEKS PROSEDUR KELAS XI SMA/MA TAHUN AJARAN 2020/2021.
- Sati, N. L. (2026). Pengembangan Media Pembelajaran Berbasis Problem Based Learning Pada Materi Sistem Periodik Unsur Berorientasi Minat Belajar Murid. UNIVERSITAS JAMBI.
- Savika, H. I., & Zuhriyah, I. A. (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.
- Simanjuntak, H., Bakti Tonni Endaryono, M. M., Sinaga, D., Siagian, B. A., Saragih, E. L. L., SS M, H. U. M., & Siagian, H. (2022). Mutu pendidikan untuk jenjang sekolah dasar. Penerbit Qiara Media.
- YUSUF, F. W. (2024). Analisis butir soal asesmen sumatif biologi materi perubahan lingkungan dengan menggunakan anates pada kelas x sma. *LEARNING*, 4(1).