Vol. 22, No. 2, November 2026 *p-ISSN: 2087-3476*
Jurnal Pepatuzdu: Media Pendidikan dan Sosial Kemasyarakat *e-ISSN: 2541-5700*
DOI: http://dx.doi.org/10.35329/fkip.v22i2.7198

ANALISIS TINGKAT KESUKARAN BUTIR SOAL UJIAN SEKOLAH MATEMATIKA KELAS VI SD DI KABUPATEN KEPULAUAN SELAYAR

Bau Ajir[1], Patahuddin Patahuddin[1], Suryadi Ishak[1]*, Shabrina Syntha Dewi[2]
[1]Program Studi Penelitian dan Evaluasi Pendidikan, Universitas Negeri Makassar
[2]Program Studi Pendidikan Teknik Elektro, Universitas Negeri Makassar
***Email: suryadi.ishak@unm.ac.id**

ABSTRACT

This study focuses on measuring the difficulty level of items on the School Mathematics Exam for sixth-grade elementary school students in Selayar Islands Regency. This study employs a descriptive quantitative approach, analyzing the exam responses of 113 students to 15 mathematics questions. Data sources were obtained from exam documents, including the test booklet, answer key, test blueprint, and students' answer sheets. Item analysis was conducted based on Classical Test Theory (CTT) by calculating the difficulty index based on the proportion of students who answered each item correctly. The interpretation criteria used were: difficult if $P < 0.30$, moderate if $0.30 \leq P \leq 0.70$, and easy if $P > 0.70$. The results of the study show that 2 test items (13.33%) fall into the easy category, 12 test items (80.00%) fall into the moderate category, and 1 test item (6.67%) falls into the difficult category. The average difficulty index of 0.514 indicates that the test items fall into the moderate category. These results indicate that the Grade 6 Elementary School Mathematics Exam questions in Selayar Islands Regency are sufficiently representative for measuring students' abilities within the average ability range. However, the distribution of difficulty levels still needs to be expanded to obtain a more comprehensive picture of students' abilities.

Keywords: item analysis; item difficulty; elementary schools; mathematics; school examinations.

PENDAHULUAN

Pendidikan dasar memiliki peran penting dalam membangun fondasi kemampuan akademik peserta didik. Pada tingkat sekolah dasar, matematika menempati posisi penting sebagai sarana untuk melatih proses berpikir peserta didik, termasuk penalaran yang logis, kemampuan mengevaluasi secara kritis, berpikir secara sistematis, serta melakukan analisis (Ndiung & Jediut, 2020). Kemampuan tersebut diperlukan peserta didik tidak hanya untuk menyelesaikan persoalan matematika, tetapi juga membantu peserta didik menghadapi berbagai

permasalahan yang dijumpai dalam kehidupan sehari-hari. Kemampuan matematika mencakup kemampuan untuk bernalar, mengonstruksi, menerapkan, dan memaknai konsep matematika ketika berhadapan dengan situasi kehidupan nyata. mengonstruksi, menerapkan, dan memaknai konsep matematika ketika berhadapan dengan situasi kehidupan nyata (OECD, 2023). Oleh karena itu, pencapaian kompetensi matematika perlu diukur melalui instrumen evaluasi yang tepat, objektif, valid, dan reliabel sehingga mampu menggambarkan tingkat kemampuan peserta didik secara akurat (OECD, 2023; Weintrop et al., 2016).

Ujian sekolah digunakan untuk mengetahui sejauh mana peserta didik telah menguasai kompetensi yang dipelajari selama menempuh pendidikan. Pada kelas VI sekolah dasar, hasil ujian sekolah tidak hanya menggambarkan pemahaman peserta didik terhadap penguasaan materi, melainkan juga menjadi salah satu dasar dalam menentukan kesiapan peserta didik melanjutkan pendidikan ke jenjang berikutnya (Nafs et al., 2023). Oleh karena itu, soal yang digunakan dalam ujian perlu memiliki kualitas yang baik agar hasil penilaian benar-benar mencerminkan kemampuan peserta didik (Bahar et al., 2021).

Tingkat kesukaran menjadi salah satu komponen yang perlu diperhatikan dalam evaluasi butir soal. Tingkat kesukaran dapat dipahami sebagai ukuran yang menunjukkan proporsi siswa yang mengikuti tes dalam memberikan jawaban yang benar terhadap suatu soal (Arikunto, 2021) Butir soal yang terlalu mudah kurang mampu memberikan tantangan dan dapat menimbulkan kesan bahwa peserta didik telah menguasai materi secara memadai. Analisis tingkat kesukaran penting dilakukan karena indeks ini dapat mengindikasikan apakah sebuah butir soal termasuk dalam kategori mudah, sedang, atau sukar, sekaligus menjadi pertimbangan dalam menentukan kesesuaian soal sebagai bagian dari proses penilaian pembelajaran (Fauzie et al., 2021). Sebaliknya, butir soal dengan Tingkat kesukaran yang tinggi dapat mempersulit peserta didik dalam menunjukkan kemampuan yang dimilikinya, terutama Ketika tingkat kesukaran tidak seimbang dengan kemampuan soal dalam membedakan peserta didik serta fungsi pengecoh yang digunakan (Rezigalla et al., 2024). Dalam evaluasi formatif, kualitas butir soal perlu diperiksa melalui beberapa karakteristik, seperti validitas, reliabilitas, tingkat kesukaran, daya pembeda, serta kinerja pengecoh. Hal ini diperlukan agar soal dapat digunakan untuk mengukur kemampuan berpikir peserta didik dengan lebih akurat (Trianingsih, 2023). Di samping itu, analisis tingkat kesukaran butir soal juga berperan dalam mengidentifikasi pemahaman dan miskonsepsi peserta didik terhadap berbagai konsep pembelajaran (Soeharto & Csapó, 2021). Oleh karena itu, penyusunan soal perlu memperhatikan keseimbangan tingkat kesukaran agar tes mampu mengukur kemampuan peserta didik pada berbagai rentang kemampuan secara lebih objektif dan representatif (Konakcı, 2024).

Pada pendekatan *Classical Test Theory* (CTT), indeks kesukaran merupakan informasi dasar yang penting untuk menilai karakteristik butir soal (Quaigrain & Arhin, 2017). Hubungan antara indeks tersebut dan tingkat kemudahan soal bersifat berbanding lurus: semakin besar nilai yang diperoleh, semakin banyak peserta tes yang mampu menjawab butir soal dengan benar. Sebaliknya, apabila nilai indeks menurun, butir soal cenderung semakin sulit untuk dijawab (Alim et al., 2025). Informasi mengenai tingkat kesukaran dapat dimanfaatkan guru maupun penyusun soal sebagai bahan pertimbangan untuk menilai kesesuaian butir soal terhadap karakteristik kemampuan peserta didik serta sasaran evaluasi pembelajaran.

Konteks Kabupaten Kepulauan Selayar penting untuk dikaji karena analisis empirik terhadap soal Ujian Sekolah Matematika kelas VI SD masih perlu diperkuat. Praktik penyusunan soal di sekolah sering kali lebih menekankan kesesuaian dengan kisi-kisi, tetapi belum selalu diikuti dengan analisis butir soal secara sistematis. Padahal, tanpa analisis empirik, guru akan kesulitan menentukan apakah soal yang digunakan memiliki tingkat kesukaran yang sesuai dengan kemampuan peserta didik dan tersusun secara proporsional. Penelitian ini diarahkan untuk menjawab kebutuhan tersebut dengan memfokuskan kajian pada tingkat kesukaran butir soal Ujian Sekolah Matematika Kelas VI SD di Kabupaten Kepulauan Selayar.

Berdasarkan uraian tersebut, rumusan masalah dalam artikel ini adalah: bagaimanakah tingkat kesukaran soal Ujian Sekolah Matematika Kelas VI SD di Kabupaten Kepulauan Selayar? Sejalan dengan rumusan masalah tersebut, penelitian ini diarahkan untuk menggambarkan tingkat kesukaran butir soal Ujian Sekolah Matematika Kelas VI SD di Kabupaten Kepulauan Selayar berdasarkan kategori mudah, sedang, dan sukar.

METODE PENELITIAN

Pendekatan kuantitatif deskriptif diterapkan dalam penelitian ini karena data yang dianalisis berupa skor dan respons peserta didik terhadap setiap butir soal Ujian Sekolah. Penelitian tidak memberikan perlakuan khusus kepada subjek, tetapi mendeskripsikan karakteristik butir soal berdasarkan hasil analisis empirik. Fokus utama penelitian adalah tingkat kesukaran butir soal Ujian Sekolah Matematika Kelas VI SD di Kabupaten Kepulauan Selayar.

Objek penelitian berupa 15 butir soal Ujian Sekolah Matematika beserta respons jawaban peserta didik. Penelitiann ini melibatkan 113 peserta didik kelas VI sekolah dasar yang berada di Kabupaten Kepulauan Selayar sebagai subjek penelitian. Seluruh peserta tersebut mengikuti Ujian Sekolah Matematika pada tahun ajaran 2024/2025. Pengumpulan data dilakukan melalui teknik dokumentasi terhadap beberapa dokumen penelitian, meliputi naskah soal, kunci jawaban dan kisi-kisi, serta jawaban peserta didik pada lembar ujian.

Data dianalisis menggunakan pendekatan *Classical Test Theory* (CTT) sebagaimana lazim digunakan dalam analisis karakteristik butir tes Indeks tingkat kesukaran dihitung dengan rumus:

$$P = B / N$$

Keterangan:
P = indeks tingkat kesukaran butir soal;
B = jumlah peserta didik yang menjawab benar pada suatu butir;
N = jumlah seluruh peserta tes.

Kriteria interpretasi indeks tingkat kesukaranmengacu pada kategori umum dalam evaluasi pendidikan, yaitu mudah, sedang, dan sukar (Sudjiono, 2006).

Tabel 1. Kriteria interpretasi indeks tingkat kesukaran butir soal

Nilai P	Kategori
P < 0,30	Sukar
0,30 ≤ P ≤ 0,70	Sedang
P > 0,70	Mudah

Kriteria tersebut digunakan dalam menentukan klasifikasi setiap soal berdasarkan tingkat kesukarannya, yaitu mudah, sedang, atau sukar. Hasil pengolahan data kemudian disajikan dalam bentuk frekuensi, persentase, nilai rata-rata indeks kesukaran, dan penjelasan deskriptif sehingga distribusi tingkat kesukaran soal dapat terlihat secara lebih komprehensif.

HASIL DAN PEMBAHASAN

Berdasarkan hasil pengujian, instrumen Ujian Sekolah Matematika Kelas VI SD di Kabupaten Kepulauan Selayar memiliki validitas dan reliabilitas yang baik. Validitas isi dianalisis menggunakan koefisien Aiken's V dengan melibatkan lima orang ahli sebagai validator. Dari hasil analisis diketahui 15 butir soal memperoleh koefisien Aiken's V pada rentang 0,85 sampai 0,95, lebih tinggi daripada nilai kritis 0,75, sehingga seluruh butir soal dinyatakan valid dan memiliki kesesuaian isi yang tinggi dengan kompetensi yang diukur. Selanjutnya, hasil pengujian reliabilitas menggunakan koefisien Kuder–Richardson 20 (KR-20) memperlihatkan konsistensi internal instrumen yang kuat, sehingga dapat dinyatakan reliabel sebagai alat ukur. Dengan demikian, berdasarkan aspek validitas dan reliabilitas, instrumen Ujian Sekolah Matematika dapat digunakan

sebagai alat pengukuran kompetensi matematika peserta didik kelas VI SD di Kabupaten Kepulauan Selayar

Hasil analisis indeks tingkat kesukaran terhadap 15 butir soal Ujian Sekolah Matematika Kelas VI SD di Kabupaten Kepulauan Selayar menunjukkan bahwa nilai indeks kesukaran berada pada rentang 0,19 sampai 0,88. Nilai tersebut menunjukkan adanya variasi indeks tingkat kesukaran antar butir, meskipun sebagian besar butir termasuk dalam kategori sedang. Nilai rata-rata indeks kesukaran yang diperoleh sebesar 0,514, sehingga secara umum butir soal ujian berada pada kategori sedang.

Tabel 2. Indeks kesukaran setiap butir soal

Butir Soal	Indeks Kesukaran (P)	Kategori
B1	0,82	Mudah
B2	0,54	Sedang
B3	0,47	Sedang
B4	0,19	Sukar
B5	0,88	Mudah
B6	0,48	Sedang
B7	0,50	Sedang
B8	0,53	Sedang
B9	0,49	Sedang
B10	0,41	Sedang
B11	0,45	Sedang
B12	0,57	Sedang
B13	0,43	Sedang
B14	0,44	Sedang
B15	0,51	Sedang

Merujuk pada data pada tabel, sebanyak dua butir soal berada dalam kategori mudah, yakni B1 yang memperoleh indeks 0,82 dan B5 dengan indeks 0,88. Kondisi tersebut menunjukkan bahwa mayoritas peserta didik dapat memberikan jawaban yang benar pada kedua soal tersebut. Selanjutnya, terdapat 12 butir yang berada pada kategori sedang, meliputi B2, B3, B6, B7, B8, B9, B10, B11, B12, B13, B14, serta B15. Adapun kategori sukar hanya ditemukan pada satu butir, yaitu B4 yang memiliki indeks kesukaran sebesar 0,19.

Vol. 22, No. 2, November 2026 *p-ISSN: 2087-3476*
Jurnal Pepatuzdu: Media Pendidikan dan Sosial Kemasyarakat *e-ISSN: 2541-5700*
DOI: http://dx.doi.org/10.35329/fkip.v22i2.7198

Tabel 3. Distribusi indeks tangka kesukaran butir soal

Kategori	Jumlah Butir	Persentase
Mudah	2	13,33%
Sedang	12	80,00%
Sukar	1	6,67%
Jumlah	15	100,00%

Distribusi tersebut memperlihatkan bahwa sebagian besar butir soal terkonsentrasi dalam kategori sedang. Dominasi kategori sedang mengindikasikan bahwa soal Ujian Sekolah Matematika Kelas VI SD di Kabupaten Kepulauan Selayar relatif mampu mengukur kemampuan peserta didik pada rentang kemampuan rata-rata. Namun, proporsi soal mudah dan sukar masih relatif kecil sehingga butir soal ujian belum sepenuhnya mempresentasikan keragaman tingkat kemampuan peserta didik secara luas.

Temuan penelitian memperlihatkan bahwa tingkat kesukaran soal Ujian Sekolah Matematika Kelas VI SD di Kabupaten Kepulauan Selayar didominasi oleh kategori sedang. Dari 15 soal, 12 diantaranya atau 80,00% berada pada kategori sedang. Temuan ini mengindikasikan bahwa mayoritas soal memiliki tingkat kesulitan yang relatif proporsional bagi peserta didik. Dalam evaluasi pembelajaran, soal yang berada pada tingkat kesukaran sedang umumnya lebih efektif digunakan sebagai bagian dari instrumen penilaian, sebab hasil respons peserta tes dapat memberikan gambaran yang lebih informatif mengenai kemampuan mereka dibandingkan dengan butir yang berada pada tingkat kesukaran ekstrem.

Rata-rata indeks kesukaran sebesar 0,514 memperkuat temuan bahwa butir soal ujian secara keseluruhan termasuk dalam kategori sedang. Nilai ini menggambarkan bahwa sekitar separuh peserta tes mampu menjawab soal dengan benar secara rata-rata. Dalam konteks ujian sekolah, kondisi tersebut dapat dipandang cukup baik karena soal mampu memberikan tantangan yang wajar bagi peserta didik dan tidak menimbulkan tingkat kesulitan yang terlalu ekstrem. Soal dengan tingkat kesukaran sedang juga dapat membantu guru memperoleh gambaran umum mengenai penguasaan kompetensi matematika peserta didik.

Butir B1 dan B5 termasuk kategori mudah. B1 memiliki indeks kesukaran 0,82, sedangkan B5 memiliki indeks 0,88. Berdasarkan kisi-kisi, B1 berkaitan dengan operasi pecahan, sedangkan B5 berkaitan dengan pola bilangan. Banyaknya peserta didik yang berhasil menjawab kedua butir tersebut dengan benar mengindikasikan bahwa materi atau bentuk soal pada kedua butir relatif mudah dipahami oleh peserta didik. Butir soal yang mudah tetap diperlukan dalam sebuah tes karena dapat memberikan informasi mengenai penguasaan

kompetensi dasar sekaligus memberi ruang bagi peserta didik berkemampuan rendah untuk menunjukkan kemampuan minimal mereka. Akan tetapi, jumlah butir mudah yang terlalu banyak dapat mengakibatkan tes menjadi kurang mampu menggambarkan keragaman kemampuan peserta didik. Dalam penelitian ini, proporsi butir mudah sebesar 13,33%, sehingga jumlahnya tidak mendominasi butir soal ujian.

Butir B4 merupakan satu-satunya soal dalam kategori sukar dengan nilai indeks kesukaran sebesar 0,19. Berdasarkan kisi-kisi, butir ini berkaitan dengan soal cerita pembagian desimal. Rendahnya proporsi peserta didik yang menjawab benar dapat mengindikasikan bahwa materi pembagian desimal dalam konteks soal cerita masih menjadi tantangan bagi sebagian besar peserta didik. Kesulitan tersebut dapat disebabkan oleh beberapa kemungkinan, seperti belum kuatnya pemahaman konsep desimal, kesulitan memahami konteks soal cerita, atau kurangnya ketelitian dalam menerjemahkan informasi verbal ke dalam operasi matematika. Oleh sebab itu, hal ini harus diperhatikan secara khusus dalam proses belajar dan ketika menyusun soal di masa mendatang.

Dominasi butir berkategori sedang menunjukkan bahwa soal cukup sesuai dengan kemampuan mayoritas peserta didik. Namun, dari sudut pandang pengembangan instrumen, distribusi tingkat kesukaran masih dapat ditingkatkan. Komposisi yang terlalu terpusat pada kategori sedang membuat tes lebih sensitif untuk mengukur peserta didik dengan kemampuan rata-rata, tetapi kurang optimal untuk mengidentifikasi peserta didik yang berada pada kemampuan sangat rendah atau sangat tinggi. Peserta didik berkemampuan tinggi membutuhkan sejumlah butir yang lebih menantang agar kemampuan puncaknya dapat terukur, sedangkan peserta didik berkemampuan rendah membutuhkan butir yang lebih mudah untuk mengukur penguasaan kompetensi dasar.

Temuan ini sejalan dengan prinsip analisis butir dalam CTT yang menempatkan tingkat kesukaran sebagai parameter penting dalam menilai mutu soal (Ashraf & Jaseem, 2020). Tingkat kesukaran yang seimbang berperan dalam meningkatkan kualitas informasi yang diperoleh dari tes. Soal yang terlalu mudah cenderung dijawab benar oleh hampir semua peserta, sehingga kurang memberi informasi mengenai perbedaan kemampuan. Sebaliknya, soal yang terlalu sukar cenderung dijawab salah oleh hampir semua peserta, sehingga juga kurang informatif. Analisis tingkat kesulitan sangat penting untuk menentukan apakah kualitas soal sudah sesuai dan dapat mendukung tujuan evaluasi hasil belajar (Hidayati & Nisa', 2023). Karena itu, keberadaan soal pada kategori sedang sangat penting, tetapi tetap perlu dilengkapi dengan proporsi soal mudah dan sukar yang memadai. Pemerataan tingkat kesukaran juga perlu dipertimbangkan bersama daya pembeda agar tes mampu menggambarkan kemampuan peserta didik secara lebih tepat (Dias et al., 2024; Alemu et al., 2024).

Dibandingkan dengan distribusi ideal yang sering direkomendasikan dalam pengembangan tes, yaitu variasi antara soal mudah, sedang, dan sukar, hasil penelitian ini menunjukkan bahwa instrumen masih perlu penguatan pada kategori mudah dan sukar (Rezigalla et al., 2024). Penambahan butir mudah dapat membantu mengukur pencapaian kompetensi dasar, sedangkan penambahan butir sukar dapat membantu mengukur kemampuan berpikir lebih tinggi. Pengembangan instrumen penilaian perlu memperhatikan keseimbangan karakteristik butir, termasuk tingkat kesukaran, sehingga instrumen yang dihasilkan dapat memberikan gambaran kemampuan peserta didik secara lebih akurat (Rahman et al., 2021). Di samping itu, penyusunan soal yang ditujukan untuk mengukur higher-order thinking skills hendaknya mengacu pada indikator berpikir tingkat tinggi, seperti analisis, evaluasi, dan pemecahan masalah (Gradini et al., 2022). Dengan demikian, pengembangan soal di masa mendatang perlu mempertimbangkan pemerataan tingkat kesukaran agar tes dapat memberikan gambaran kemampuan peserta didik secara lebih komprehensif. Hal ini juga sejalan dengan pengembangan instrumen asesmen HOTS berbasis literasi numerasi yang menekankan pentingnya validitas, daya pembeda, dan tingkat kesukaran butir dalam menghasilkan instrumen penilaian yang layak digunakan (Wardani & Siregar, 2023).

Implikasi praktis dari temuan ini adalah pentingnya pelaksanaan analisis butir secara rutin setelah ujian sekolah. Analisis butir dapat digunakan untuk menilai mutu soal berdasarkan sejumlah karakteristik, antara lain tingkat kesukaran, daya pembeda, reliabilitas, serta kinerja pengecoh (Rezigalla et al., 2024; Trianingsih, 2023). Hasil analisis tingkat kesukaran dapat menjadi dasar bagi guru maupun penyusun soal dalam menentukan butir yang masih layak digunakan, perlu diperbaiki, atau memerlukan pengembangan lebih lanjut. Butir yang berada pada kategori sedang dan sesuai indikator dapat dipertimbangkan untuk menjadi bagian dari bank soal, sedangkan butir dengan tingkat kesukaran terlalu rendah atau telalu tinggi perlu ditelaah kembali dari aspek materi, konstruksi, bahasa, dan kesesuaian dengan indikator. Pandangan tersebut sejalan dengan kajian Fauzie et al. (2021) yang menunjukkan bahwa analisis indeks kesukaran pada bank soal berdasarkan aspek kognitif Taksonomi Bloom. Dengan cara ini, sekolah dapat membangun bank soal yang lebih terstandar dan bermutu.

Selain itu, temuan penelitian ini juga dapat digunakan sebagai bahan pertimbangan dalam pelaksanaan pembelajaran matematika di kelas VI. Butir sukar pada materi pembagian desimal berbentuk soal cerita menunjukkan perlunya penguatan pembelajaran pada materi tersebut. Analisis tingkat kesukaran tidak terbatas pada penilaian terhadap mutu instrumen, tetapi juga dapat dimanfaatkan untuk memperoleh gambaran mengenai materi yang masih menjadi kendala bagi peserta didik. Soeharto dan Csapó (2021) menjelaskan bahwa pola tingkat kesukaran butir dapat membantu guru memahami

pemahaman dan miskonsepsi yang dialami peserta didik selama proses pembelajaran. Berdasarkan informasi tersebut, guru dapat memberikan latihan yang lebih beragam, menggunakan konteks kehidupan sehari-hari, dan membimbing peserta didik dalam memahami informasi soal cerita sebelum melakukan operasi hitung. Dengan demikian, informasi dari analisis tingkat kesukaran dapat dimanfaatkan sebagai salah satu dasar untuk mengevaluasi instrumen sekaligus mendukung perbaikan pembelajaran matematika (Konakcı, 2024; Soeharto & Csapó, 2021).

SIMPULAN

Berdasarkan analisis terhadap 15 soal Ujian Sekolah Matematika untuk Kelas VI SD di Kabupaten Kepulauan Selayar, ditemukan bahwa soal-soalnya sebagian besar memiliki tingkat kesukaran sedang. Dari keseluruhan soal yang dianalisis, sebanyak 2 soal (13,33%) tergolong mudah, 12 soal (80,00%) termasuk dalam kelompok sedang, sedangkan 1 soal (6,67%) dalam kategori sukar. Rata-rata indeks kesukaran yang diperoleh adalah 0,514. Nilai tersebut menggambarkan bahwa secara umum, soal yang digunakan dalam ujian memiliki tingkat kesukaran yang berada dalam kategori sedang.

Dominasi butir berkategori sedang menunjukkan bahwa soal cukup representatif untuk mengukur kemampuan peserta didik pada rentang kemampuan rata-rata. Namun, komposisi butir pada kategori mudah dan sukar masih perlu ditingkatkan sehingga instrumen dapat memberikan gambaran kemampuan peserta didik secara lebih komprehensif. baik pada kelompok berkemampuan rendah maupun tinggi. Dengan demikian, soal Ujian Sekolah Matematika Kelas VI SD di Kabupaten Kepulauan Selayar dapat digunakan sebagai alat evaluasi, tetapi pengembangan variasi tingkat kesukaran tetap perlu dilakukan sebagai upaya memperbaiki mutu instrumen yang digunakan pada penyelenggaraan ujian selanjutnya..

SARAN DAN REKOMENDASI

Guru dan tim penyusun soal disarankan melakukan analisis tingkat kesukaran secara rutin setiap selesai pelaksanaan ujian sekolah. Hasil analisis tersebut dapat digunakan untuk memperbaiki butir soal, menyusun bank soal, dan meningkatkan mutu evaluasi pembelajaran matematika. Butir yang berkategori sedang dan sesuai indikator dapat dipertahankan, sedangkan butir yang terlalu mudah atau terlalu sukar perlu ditelaah kembali dari sisi materi, konstruksi, bahasa, dan kesesuaiannya dengan kompetensi yang diukur.

Sekolah juga perlu mendorong pengembangan soal dengan distribusi tingkat kesukaran yang lebih bervariasi. Penambahan butir mudah dan sukar yang proporsional penting dilakukan agar tes tidak hanya mengukur kemampuan rata-rata peserta didik, tetapi juga mampu mengidentifikasi peserta didik yang

memerlukan penguatan serta peserta didik yang memiliki kemampuan tinggi. Penelitian berikutnya dapat memperluas analisis dengan memasukkan aspek daya pembeda, efektivitas pengecoh, validitas empiris, atau menggunakan pendekatan *Item Response Theory* (IRT) untuk memperoleh informasi karakteristik butir yang lebih mendalam.

DAFTAR PUSTAKA

Alemu, A. T., Tesfa, H., Mulugeta, A., Fenta, E. T., & Belay, M. A. (2024). Quality of multiple choice question items: Item analysis. *International Journal of Scientific Reports, 10*(6), 195–199. https://doi.org/10.18203/issn.2454-2156.IntJSciRep20241316

Arikunto, S. (2021). *Dasar-dasar evaluasi pendidikan edisi 3*. Bumi aksara.

Ashraf, Z. A., & Jaseem, K. (2020). Classical and modern methods in item analysis of test tools. *International Journal of Research and Review*, *7*(5), 397–403.

Bahar, R., Istiyono, E., Widihastuti, W., Munadi, S., Nuryana, Z., & Fajaruddin, S. (2021). Analisis karakteristik soal ujian sekolah hasil musyawarah guru matematika di Tasikmalaya. *AKSIOMA: Jurnal Program Studi Pendidikan Matematika*, *10*(4), 2660–2674. https://doi.org/http://dx.doi.org/10.24127/ajpm.v10i4.4359

Dias, B. M., Silva, L. M. G. da, Hamamoto Filho, P. T., Bollela, V. R., & Gabriel, C. S. (2024). Development of qualified items for nursing education assessment: The progress testing experience. *Nurse Education in Practice, 81*, Article 104199. https://doi.org/10.1016/j.nepr.2024.104199

Fauzie, M., Pada, A. U. T., & Supriatno, S. (2021). Analysis of the difficulty index of item bank according to cognitive aspects during the Covid-19 pandemic. *Jurnal Penelitian dan Evaluasi Pendidikan, 25*(2), 162–169. https://doi.org/10.21831/pep.v25i2.42603

Gradini, E., Khairunnisak, C., & Noviani, J. (2022). Development of higher-order thinking skill (HOTS) test on mathematics in secondary school. *AKSIOMA: Jurnal Program Studi Pendidikan Matematika, 11*(1), 319–330. https://doi.org/10.24127/ajpm.v11i1.4649

Hidayati, K., & Nisa', Z. E. K. (2023). Analisis butir soal penilaian akhir semester matematika. *AKSIOMA: Jurnal Program Studi Pendidikan Matematika, 12*(3).

Konakcı, S. (2024). Item analysis in multiple choice questions: A study on question difficulty and authors' evaluation. *Journal of Basic and Clinical Health Sciences, 8*(2). https://doi.org/10.30621/jbachs.1427464

Nafs, H., Sridana, N., Hikmah, N., & Soeprianto, H. (2023). Analisis kualitas butir soal ulangan akhir semester genap mata pelajaran matematika kelas vii smpn 6 mataram tahun ajaran 2022/2023. *Jurnal Ilmiah Profesi Pendidikan*, *8*(4), 2324–2331.

https://doi.org/https://doi.org/10.29303/jipp.v8i4.1785

Ndiung, S., & Jediut, M. (2020). Pengembangan instrumen tes hasil belajar matematika peserta didik sekolah dasar berorientasi pada berpikir tingkat tinggi. *Premiere Educandum: Jurnal Pendidikan Dasar Dan Pembelajaran*, *10*(1), 94. https://doi.org/10.25273/pe.v10i1.6274

OECD. (2023). *PISA 2022 assessment and analytical framework: Mathematics, reading, science and financial literacy*. OECD Publishing. https://doi.org/10.1787/dfe0bf9c-en.

Rahman, A., Rusnayati, H., & Muslim, M. (2021). Analysis of the characteristics of higher order thinking skills (HOTS) test on momentum and impulse for senior high school student using item response theory. *Jurnal Penelitian Fisika dan Aplikasinya*, *11*(2), 127–137. https://doi.org/10.26740/jpfa.v11n2.p127-137

Rezigalla, A. A., Eleragi, A. M. E. S. A., Elhussein, A. B., Alfaifi, J., ALGhamdi, M. A., Ameer, A. Y. Al, Mohammed, O. A., & Adam, M. I. E. (2024). Item analysis: the impact of distractor efficiency on the difficulty index and discrimination power of multiple-choice items. *BMC Medical Education*, *24*. https://doi.org/https://doi.org/10.1186/s12909-024-05433-y

Soeharto, S., & Csapó, B. (2021). Evaluating item difficulty patterns for assessing student misconceptions in science across physics, chemistry, and biology concepts. *Heliyon*, *7*(11), Article e08352. https://doi.org/10.1016/j.heliyon.2021.e08352

Trianingsih, L. (2023). Analisis kualitas tes dan butir soal integral pada evaluasi formatif matematika teknik. *Indonesian Journal of Civil Engineering Education*, *9*(2), 44–57.

Wardani, N., & Siregar, T. J. (2023). The development of HOTS (Higher-Order Thinking Skills) assessment instrument based on numeration literacy in mathematics for junior high school. *Indonesian Journal of Science and Mathematics Education*, *6*(2), 238–252.

Weintrop, D., Beheshti, E., Horn, M., Orton, K., Jona, K., Trouille, L., & Wilensky, U. (2016). Defining computational thinking for mathematics and science classrooms. *Journal of Science Education and Technology*, *25*(1), 127–147. https://doi.org/10.1007/s10956-015-9581-5.