



## Evaluasi Hasil Belajar Siswa Melalui Analisis Tingkat Kesulitan Butir Soal pada Mata Pelajaran IPA

Fenisa Putri Meliani<sup>1\*</sup>, Karolin Nazwa Liya<sup>2</sup>, Syaula Abid Faiziya<sup>3</sup>, Rahmawati Sholihah<sup>4</sup>, Dika Meilia<sup>5</sup>, Nazwa Wulan Fiantika<sup>6</sup>

<sup>1,2,3,4,5,6</sup>Universitas Muhammadiyah Prof. Dr. Hamka Jakarta, Indonesia

\*Penulis Korespondensi: [fenisaputrim@gmail.com](mailto:fenisaputrim@gmail.com), [karolinnazwaliya@gmail.com](mailto:karolinnazwaliya@gmail.com), [syaulaabid@gmail.com](mailto:syaulaabid@gmail.com), [rahmawatisholihah6@gmail.com](mailto:rahmawatisholihah6@gmail.com), [dikamelia21@gmail.com](mailto:dikamelia21@gmail.com), [nazwawulan13@gmail.com](mailto:nazwawulan13@gmail.com)

**Abstract.** *Learning outcome evaluation is an essential component of the educational process that functions to measure students' competency achievement and assess the effectiveness of instruction. The quality of evaluation results is highly influenced by the quality of the assessment instruments used. One method to determine the quality of a test instrument is through item difficulty analysis. This study aimed to analyze the difficulty level of test items in Science (IPA) subjects as part of student learning outcome evaluation and to assess the quality of the instrument used in the learning process. The study employed a quantitative approach with a descriptive method. Data were collected through documentation, including test items, answer keys, and students' answer sheets, and were analyzed using the difficulty index. The findings revealed that the distribution of item difficulty consisted of easy, moderate, and difficult categories, with moderate items dominating the test. This indicates that the instrument had a relatively balanced level of difficulty and was capable of measuring students' abilities more objectively. The results further demonstrate that item difficulty analysis is useful not only for identifying the quality of test items but also for improving assessment instruments and enhancing the effectiveness of the learning process. Therefore, item analysis should be conducted regularly to ensure the development of valid, objective, and high-quality assessment instruments.*

**Keywords:** *Learning Outcome Evaluation, Item Analysis, Item Difficulty, Assessment Instrument, Science Education.*

**Abstrak.** Evaluasi hasil belajar merupakan salah satu komponen penting dalam proses pembelajaran yang berfungsi untuk mengukur tingkat pencapaian kompetensi peserta didik serta menilai efektivitas pembelajaran yang telah dilaksanakan. Kualitas hasil evaluasi sangat dipengaruhi oleh kualitas instrumen penilaian yang digunakan. Salah satu cara untuk mengetahui kualitas instrumen tes adalah melalui analisis tingkat kesulitan butir soal. Penelitian ini bertujuan untuk menganalisis tingkat kesulitan butir soal pada mata pelajaran Ilmu Pengetahuan Alam (IPA) sebagai bagian dari evaluasi hasil belajar siswa serta mengevaluasi kualitas instrumen yang digunakan dalam proses pembelajaran. Penelitian menggunakan pendekatan kuantitatif dengan metode deskriptif. Data diperoleh melalui dokumentasi berupa naskah soal, kunci jawaban, dan lembar jawaban peserta didik, kemudian dianalisis menggunakan indeks tingkat kesulitan. Hasil penelitian menunjukkan bahwa distribusi tingkat kesulitan soal terdiri atas kategori mudah, sedang, dan sukar, dengan dominasi soal berkategori sedang. Temuan tersebut menunjukkan bahwa instrumen yang digunakan memiliki tingkat kesulitan yang relatif proporsional sehingga mampu mengukur kemampuan peserta didik secara lebih objektif. Analisis tingkat kesulitan tidak hanya bermanfaat untuk mengetahui kualitas butir soal, tetapi juga dapat digunakan sebagai dasar dalam memperbaiki instrumen evaluasi dan meningkatkan efektivitas proses pembelajaran. Oleh karena itu, analisis butir soal perlu dilakukan secara berkala guna menghasilkan instrumen penilaian yang valid, objektif, dan berkualitas.

**Kata kunci:** Evaluasi Hasil Belajar, Analisis Butir Soal, Tingkat Kesulitan Soal, Instrumen Penilaian, IPA.

### 1. LATAR BELAKANG

Evaluasi hasil belajar merupakan salah satu komponen penting dalam proses pendidikan yang berfungsi untuk mengetahui tingkat ketercapaian tujuan pembelajaran

serta efektivitas proses pembelajaran yang telah dilaksanakan. Melalui kegiatan evaluasi, guru dapat memperoleh informasi mengenai tingkat penguasaan kompetensi peserta didik, keberhasilan strategi pembelajaran yang digunakan, serta berbagai faktor yang memengaruhi hasil belajar. Informasi tersebut selanjutnya digunakan sebagai dasar dalam pengambilan keputusan untuk memperbaiki proses pembelajaran dan meningkatkan kualitas pendidikan. Arikunto (2018) menyatakan bahwa evaluasi merupakan proses sistematis dalam mengumpulkan, menganalisis, dan menafsirkan data guna menentukan tingkat keberhasilan suatu program atau kegiatan pendidikan. Sejalan dengan pendapat tersebut, Mardapi (2012) menjelaskan bahwa evaluasi pembelajaran bertujuan untuk memperoleh informasi yang akurat mengenai pencapaian belajar peserta didik sehingga dapat dijadikan dasar dalam perencanaan dan perbaikan pembelajaran.

Keberhasilan evaluasi hasil belajar sangat dipengaruhi oleh kualitas instrumen penilaian yang digunakan. Instrumen yang baik harus mampu mengukur kompetensi peserta didik secara objektif, akurat, dan konsisten. Oleh karena itu, suatu instrumen penilaian harus memenuhi berbagai persyaratan psikometrik, seperti validitas, reliabilitas, daya pembeda, tingkat kesulitan, dan efektivitas pengecoh. Apabila instrumen yang digunakan tidak memenuhi kriteria tersebut, maka hasil evaluasi yang diperoleh berpotensi tidak menggambarkan kemampuan peserta didik yang sebenarnya. Dalam konteks ini, analisis butir soal menjadi salah satu langkah penting yang harus dilakukan sebelum maupun sesudah instrumen digunakan dalam kegiatan evaluasi. Sudijono (2015) menegaskan bahwa analisis butir soal bertujuan untuk mengidentifikasi kualitas setiap soal sehingga instrumen yang digunakan benar-benar mampu mengukur hasil belajar secara tepat dan objektif.

Salah satu aspek penting dalam analisis butir soal adalah tingkat kesulitan soal. Tingkat kesulitan merupakan indikator yang menunjukkan proporsi peserta didik yang mampu menjawab suatu soal dengan benar. Analisis tingkat kesulitan diperlukan untuk mengetahui apakah suatu soal tergolong mudah, sedang, atau sukar. Informasi tersebut sangat penting karena tingkat kesulitan yang tidak proporsional dapat memengaruhi ketepatan hasil pengukuran. Soal yang terlalu mudah cenderung tidak mampu membedakan kemampuan peserta didik karena sebagian besar siswa dapat menjawab dengan benar. Sebaliknya, soal yang terlalu sukar dapat menyebabkan peserta didik

mengalami kesulitan sehingga hasil tes tidak mencerminkan kemampuan yang sesungguhnya. Menurut Kusaeri dan Suprananto (2012), tingkat kesulitan merupakan salah satu karakteristik utama yang menentukan kualitas suatu instrumen penilaian karena berpengaruh terhadap kemampuan tes dalam mengukur hasil belajar secara akurat.

Dalam pembelajaran Ilmu Pengetahuan Alam (IPA) di sekolah dasar, penggunaan tes pilihan ganda masih menjadi salah satu bentuk evaluasi yang paling banyak digunakan. Bentuk tes ini dinilai efektif untuk mengukur penguasaan konsep, pemahaman materi, serta kemampuan berpikir peserta didik dalam cakupan materi yang luas. Namun demikian, kualitas tes pilihan ganda sangat bergantung pada kualitas setiap butir soal yang disusun. Penyusunan soal yang tidak memperhatikan tingkat kesulitan dapat menghasilkan instrumen yang kurang mampu menggambarkan kemampuan peserta didik secara komprehensif. Oleh karena itu, analisis tingkat kesulitan menjadi langkah yang penting dalam memastikan bahwa instrumen evaluasi IPA mampu memberikan informasi yang valid mengenai pencapaian hasil belajar siswa.

Berbagai penelitian sebelumnya menunjukkan bahwa analisis tingkat kesulitan memiliki peran yang signifikan dalam meningkatkan kualitas instrumen evaluasi. Hadmar, Ali, dan Yurfiah (2024) menemukan bahwa distribusi tingkat kesukaran yang tidak proporsional dapat menurunkan kualitas hasil evaluasi pembelajaran IPA di sekolah dasar. Penelitian Valen dan Satria (2021) juga menunjukkan bahwa instrumen dengan komposisi tingkat kesulitan yang seimbang mampu menghasilkan informasi yang lebih akurat mengenai kemampuan peserta didik. Selain itu, Paskalin dan Susanti (2020) menjelaskan bahwa soal dengan kategori sedang merupakan jenis soal yang paling efektif dalam mengukur penguasaan konsep karena mampu membedakan peserta didik yang telah memahami materi dengan peserta didik yang masih mengalami kesulitan belajar.

Meskipun penelitian mengenai analisis tingkat kesulitan butir soal telah banyak dilakukan, sebagian besar penelitian masih berfokus pada pengelompokan soal ke dalam kategori mudah, sedang, dan sukar tanpa mengaitkan hasil analisis tersebut dengan evaluasi hasil belajar sebagai dasar perbaikan pembelajaran. Padahal, hasil analisis tingkat kesulitan tidak hanya memberikan informasi mengenai kualitas instrumen penilaian, tetapi juga dapat digunakan untuk mengidentifikasi materi yang belum dikuasai peserta didik, mengevaluasi efektivitas strategi pembelajaran, serta menyusun tindak

lanjut pembelajaran yang lebih tepat. Dengan demikian, analisis tingkat kesulitan memiliki fungsi yang lebih luas dibandingkan sekadar menentukan kualitas soal.

Berdasarkan uraian tersebut, penelitian ini penting dilakukan untuk mengkaji evaluasi hasil belajar siswa melalui analisis tingkat kesulitan butir soal pada mata pelajaran IPA. Hasil penelitian diharapkan dapat memberikan gambaran mengenai kualitas instrumen evaluasi yang digunakan sekaligus menjadi dasar dalam memperbaiki proses pembelajaran dan pengembangan instrumen penilaian pada masa mendatang. Adapun tujuan penelitian ini adalah menganalisis tingkat kesulitan butir soal pada mata pelajaran IPA sebagai bagian dari evaluasi hasil belajar siswa sehingga dapat diketahui kualitas instrumen penilaian yang digunakan dalam proses pembelajaran.

## **2. METODE PENELITIAN**

Penelitian ini menggunakan pendekatan kuantitatif dengan metode deskriptif. Pendekatan kuantitatif dipilih karena penelitian berfokus pada pengumpulan dan analisis data numerik untuk menggambarkan karakteristik butir soal berdasarkan tingkat kesulitannya. Menurut Sugiyono (2023), penelitian kuantitatif merupakan metode penelitian yang berlandaskan pada filsafat positivisme dan digunakan untuk mengkaji fenomena melalui data yang dapat diukur secara objektif. Sementara itu, metode deskriptif digunakan untuk memberikan gambaran secara sistematis mengenai kondisi objek penelitian tanpa memberikan perlakuan atau manipulasi terhadap variabel yang diteliti.

Penelitian dilaksanakan pada instrumen evaluasi hasil belajar mata pelajaran Ilmu Pengetahuan Alam (IPA) di sekolah dasar. Objek penelitian berupa butir-butir soal pilihan ganda yang digunakan oleh guru dalam kegiatan evaluasi pembelajaran, sedangkan subjek penelitian adalah seluruh peserta didik yang mengikuti tes pada materi yang telah diajarkan. Data penelitian diperoleh melalui teknik dokumentasi dengan mengumpulkan naskah soal, kunci jawaban, serta lembar jawaban peserta didik yang telah digunakan dalam proses evaluasi pembelajaran.

Analisis data dilakukan menggunakan analisis tingkat kesulitan butir soal (*difficulty index*). Tingkat kesulitan merupakan proporsi peserta didik yang mampu menjawab benar suatu butir soal. Menurut Arikunto (2018), semakin tinggi nilai indeks kesukaran,

semakin mudah soal tersebut, sedangkan semakin rendah nilai indeks menunjukkan bahwa soal semakin sukar. Analisis ini bertujuan untuk mengetahui kualitas setiap butir soal dalam mengukur kemampuan peserta didik.

Perhitungan indeks tingkat kesulitan dilakukan menggunakan rumus:

$$P = B / JS$$

Keterangan:

P = indeks tingkat kesulitan;

B = jumlah peserta didik yang menjawab benar;

JS = jumlah seluruh peserta didik yang mengikuti tes.

Berdasarkan nilai indeks yang diperoleh, setiap butir soal diklasifikasikan ke dalam kategori tingkat kesulitan menurut kriteria Arikunto (2018), yaitu: (1) 0,00–0,30 termasuk kategori sukar, (2) 0,31–0,70 termasuk kategori sedang, dan (3) 0,71–1,00 termasuk kategori mudah.

Tahapan analisis data diawali dengan menghitung indeks tingkat kesulitan untuk setiap butir soal. Selanjutnya, hasil perhitungan digunakan untuk mengelompokkan soal ke dalam kategori mudah, sedang, dan sukar. Data yang diperoleh kemudian disajikan dalam bentuk tabel dan dianalisis secara deskriptif untuk memberikan gambaran mengenai distribusi tingkat kesulitan soal. Hasil analisis tersebut digunakan sebagai dasar untuk mengevaluasi kualitas instrumen penilaian serta menentukan butir soal yang layak dipertahankan, direvisi, atau diganti sehingga instrumen evaluasi yang digunakan pada pembelajaran berikutnya memiliki kualitas yang lebih baik.

### **3. HASIL DAN PEMBAHASAN**

Berdasarkan hasil analisis tingkat kesulitan terhadap seluruh butir soal yang digunakan dalam evaluasi pembelajaran IPA, diperoleh distribusi tingkat kesulitan yang terdiri atas kategori mudah, sedang, dan sukar. Hasil analisis tersebut menunjukkan bahwa sebagian besar butir soal berada pada kategori sedang. Kondisi ini menunjukkan bahwa instrumen evaluasi yang digunakan telah memiliki tingkat kesulitan yang relatif proporsional sehingga mampu mengukur kemampuan peserta didik secara lebih objektif.

Tabel 1. Distribusi Tingkat Kesulitan Butir Soal

| Kategori | Jumlah Soal | Persentase |
|----------|-------------|------------|
| Mudah    | 6           | 20%        |
| Sedang   | 18          | 60%        |
| Sukar    | 6           | 20%        |
| Total    | 30          | 100%       |

Berdasarkan Tabel 1 diketahui bahwa sebanyak 18 butir soal atau sebesar 60% berada pada kategori sedang. Selanjutnya, sebanyak 6 butir soal atau sebesar 20% termasuk kategori mudah, sedangkan 6 butir soal lainnya atau sebesar 20% berada pada kategori sukar. Komposisi tersebut menunjukkan bahwa instrumen evaluasi telah memiliki keseimbangan tingkat kesulitan yang cukup baik karena didominasi oleh soal berkategori sedang. Keberadaan soal dengan kategori mudah dan sukar tetap diperlukan sebagai variasi dalam mengukur kemampuan peserta didik yang memiliki tingkat penguasaan materi yang berbeda-beda.

Hasil analisis menunjukkan bahwa butir soal pada mata pelajaran IPA memiliki tingkat kesulitan yang bervariasi, yaitu terdiri atas kategori mudah, sedang, dan sukar. Perbedaan tingkat kesulitan tersebut menunjukkan bahwa setiap butir soal memiliki karakteristik yang berbeda dalam mengukur kemampuan peserta didik. Secara umum, sebagian besar butir soal berada pada kategori sedang, sedangkan sebagian lainnya termasuk kategori mudah dan sukar.

Soal dengan kategori sedang merupakan kategori yang paling ideal karena mampu mengukur kemampuan siswa secara lebih objektif. Butir soal dengan tingkat kesulitan sedang memberikan kesempatan kepada peserta didik yang memiliki kemampuan tinggi maupun rendah untuk menunjukkan tingkat penguasaan materi secara optimal. Sebaliknya, soal yang terlalu mudah cenderung tidak mampu membedakan kemampuan peserta didik karena hampir seluruh siswa dapat menjawab dengan benar. Di sisi lain, soal yang terlalu sukar juga kurang efektif karena hanya dapat dijawab oleh sebagian kecil peserta didik.

Hasil penelitian ini sejalan dengan penelitian Valen dan Satria (2021) yang menyatakan bahwa distribusi tingkat kesulitan yang proporsional menghasilkan instrumen evaluasi yang lebih berkualitas. Instrumen yang memiliki keseimbangan antara

soal mudah, sedang, dan sukar mampu memberikan informasi yang lebih akurat mengenai pencapaian hasil belajar peserta didik.

Temuan penelitian ini juga didukung oleh penelitian Paskalin dan Susanti (2020) yang menjelaskan bahwa soal dengan kategori sedang memberikan informasi yang lebih baik mengenai kemampuan siswa dibandingkan soal yang terlalu mudah maupun terlalu sukar. Oleh karena itu, penyusunan instrumen evaluasi perlu memperhatikan distribusi tingkat kesulitan agar kualitas tes tetap terjaga.

Selain digunakan untuk mengetahui kualitas instrumen penilaian, analisis tingkat kesulitan juga dapat dimanfaatkan sebagai bahan evaluasi terhadap proses pembelajaran. Butir soal yang dijawab salah oleh sebagian besar siswa menunjukkan bahwa materi tertentu masih sulit dipahami sehingga guru perlu melakukan perbaikan strategi pembelajaran. Sebaliknya, apabila hampir seluruh siswa mampu menjawab suatu soal dengan benar, guru dapat mempertimbangkan apakah materi telah dikuasai dengan baik atau tingkat kesulitan soal terlalu rendah.

Hasil penelitian menunjukkan bahwa sebagian besar butir soal berada pada kategori sedang. Temuan tersebut mengindikasikan bahwa instrumen evaluasi yang digunakan telah memenuhi salah satu karakteristik tes yang baik, yaitu mampu mengukur kemampuan peserta didik secara optimal. Soal dengan kategori sedang dianggap lebih efektif karena tidak terlalu mudah sehingga seluruh peserta didik dapat menjawab dengan benar, tetapi juga tidak terlalu sukar sehingga masih dapat dikerjakan oleh sebagian besar peserta didik yang telah memahami materi pembelajaran.

Keberadaan soal dengan tingkat kesulitan sedang memungkinkan guru memperoleh informasi yang lebih akurat mengenai kemampuan peserta didik. Melalui hasil tersebut, guru dapat mengidentifikasi siswa yang telah menguasai materi pembelajaran maupun siswa yang masih memerlukan bimbingan lebih lanjut. Oleh karena itu, soal dengan kategori sedang memiliki kemampuan yang lebih baik dalam membedakan tingkat penguasaan materi antar peserta didik dibandingkan soal yang terlalu mudah ataupun terlalu sukar.

Selain itu, hasil analisis tingkat kesulitan juga dapat dimanfaatkan sebagai bahan evaluasi terhadap pelaksanaan proses pembelajaran. Apabila terdapat sejumlah butir soal

yang termasuk kategori sukar karena hanya mampu dijawab oleh sebagian kecil peserta didik, maka kondisi tersebut dapat menjadi indikator bahwa materi pembelajaran masih belum dipahami secara optimal. Sebaliknya, apabila sebagian besar peserta didik mampu menjawab suatu butir soal dengan benar, guru perlu meninjau kembali apakah tingkat kesulitan soal yang digunakan terlalu rendah atau memang materi tersebut telah benar-benar dikuasai oleh peserta didik.

Melalui analisis tersebut, guru tidak hanya memperoleh informasi mengenai kualitas instrumen penilaian, tetapi juga memperoleh gambaran mengenai efektivitas strategi pembelajaran yang telah diterapkan di kelas. Dengan demikian, hasil analisis tingkat kesulitan dapat dijadikan sebagai dasar dalam menyusun instrumen evaluasi yang lebih baik sekaligus memperbaiki proses pembelajaran pada materi yang masih dianggap sulit oleh peserta didik. Analisis butir soal yang dilakukan secara berkala diharapkan mampu meningkatkan mutu evaluasi pembelajaran sehingga hasil yang diperoleh benar-benar mencerminkan kemampuan peserta didik secara objektif dan akurat.

Dominasi butir soal pada kategori sedang menunjukkan bahwa instrumen evaluasi yang digunakan telah mampu mengukur kemampuan peserta didik secara lebih proporsional. Butir soal dengan tingkat kesulitan sedang umumnya memiliki kemampuan yang lebih baik dalam membedakan siswa yang telah menguasai materi dengan siswa yang masih mengalami kesulitan. Oleh karena itu, keberadaan soal kategori sedang menjadi indikator bahwa instrumen penilaian telah mendukung proses evaluasi yang objektif dan mampu memberikan informasi yang lebih akurat mengenai pencapaian hasil belajar siswa. Hal tersebut sejalan dengan konsep evaluasi pendidikan yang menyatakan bahwa instrumen yang baik tidak hanya mengukur hasil belajar, tetapi juga mampu memberikan gambaran nyata mengenai tingkat penguasaan kompetensi peserta didik (Arikunto, 2021).

Temuan penelitian ini juga dapat dijelaskan melalui Taksonomi Bloom Revisi yang dikembangkan oleh Anderson dan Krathwohl. Dalam taksonomi tersebut dijelaskan bahwa proses kognitif peserta didik berkembang mulai dari kemampuan mengingat (*remember*), memahami (*understand*), menerapkan (*apply*), menganalisis (*analyze*), mengevaluasi (*evaluate*), hingga mencipta (*create*). Instrumen evaluasi yang memiliki variasi tingkat kesulitan memungkinkan guru mengukur kemampuan peserta didik pada

berbagai jenjang proses kognitif tersebut sehingga hasil penilaian menjadi lebih komprehensif. Dengan demikian, penyusunan soal sebaiknya tidak hanya berfokus pada kemampuan mengingat, tetapi juga memberikan ruang bagi peserta didik untuk menunjukkan kemampuan berpikir tingkat tinggi.

Selain dipengaruhi oleh kualitas penyusunan soal, tingkat kesulitan butir soal juga dipengaruhi oleh proses pembelajaran yang berlangsung di kelas. Apabila sebagian besar siswa mampu menjawab soal kategori sedang dengan benar, hal tersebut mengindikasikan bahwa strategi pembelajaran yang diterapkan guru telah sesuai dengan karakteristik peserta didik serta tujuan pembelajaran yang ingin dicapai. Sebaliknya, apabila banyak soal berada pada kategori sukar, kondisi tersebut dapat menjadi indikator bahwa materi belum dipahami secara optimal, metode pembelajaran kurang efektif, atau terdapat ketidaksesuaian antara indikator pembelajaran dengan soal yang disusun. Oleh karena itu, hasil analisis tingkat kesulitan tidak hanya digunakan untuk mengevaluasi kualitas instrumen, tetapi juga menjadi bahan refleksi terhadap efektivitas proses pembelajaran.

Keberadaan soal yang termasuk kategori mudah maupun sukar tetap memiliki fungsi dalam suatu instrumen evaluasi. Soal kategori mudah diperlukan untuk mengukur penguasaan kompetensi dasar yang harus dimiliki seluruh peserta didik, sedangkan soal kategori sukar digunakan untuk mengidentifikasi peserta didik yang memiliki kemampuan akademik lebih tinggi. Namun demikian, proporsi kedua kategori tersebut tidak boleh mendominasi keseluruhan instrumen karena dapat mengurangi ketepatan hasil pengukuran. Oleh sebab itu, penyusunan instrumen evaluasi perlu mempertimbangkan keseimbangan tingkat kesulitan agar seluruh indikator pembelajaran dapat diukur secara optimal (Arikunto, 2021).

Implikasi dari hasil penelitian ini adalah pentingnya pelaksanaan analisis butir soal secara rutin setelah kegiatan evaluasi selesai dilaksanakan. Analisis tersebut membantu guru menentukan butir soal yang layak dipertahankan, direvisi, maupun diganti sebelum digunakan kembali pada evaluasi berikutnya. Dengan adanya analisis yang dilakukan secara berkelanjutan, kualitas bank soal akan semakin meningkat sehingga instrumen penilaian yang digunakan benar-benar memenuhi karakteristik tes yang baik, yaitu valid, reliabel, memiliki tingkat kesulitan yang proporsional, serta mampu memberikan

informasi yang akurat mengenai kemampuan peserta didik. Melalui langkah tersebut, evaluasi pembelajaran tidak hanya berfungsi sebagai alat pemberian nilai, tetapi juga sebagai dasar dalam meningkatkan mutu pembelajaran IPA di sekolah.

Hasil penelitian ini sejalan dengan pendapat Ningrum et al. (2023) yang menyatakan bahwa analisis butir soal merupakan langkah penting dalam meningkatkan kualitas instrumen evaluasi. Hartono et al. (2024) juga menjelaskan bahwa analisis tingkat kesulitan dapat membantu guru dalam memperbaiki kualitas pembelajaran karena memberikan informasi mengenai materi yang masih memerlukan penguatan.

Dengan demikian, analisis tingkat kesulitan tidak hanya berfungsi sebagai alat untuk menilai kualitas instrumen evaluasi, tetapi juga sebagai dasar dalam meningkatkan mutu pembelajaran. Guru disarankan melakukan analisis butir soal secara berkala sehingga instrumen yang digunakan benar-benar mampu mengukur hasil belajar peserta didik secara valid, objektif, dan akurat.

#### **4. KESIMPULAN DAN SARAN**

Berdasarkan hasil penelitian, dapat disimpulkan bahwa analisis tingkat kesulitan butir soal merupakan salah satu cara yang efektif untuk mengevaluasi kualitas instrumen penilaian pada mata pelajaran IPA. Hasil analisis menunjukkan bahwa butir soal memiliki tingkat kesulitan yang bervariasi, yaitu mudah, sedang, dan sukar. Soal dengan kategori sedang merupakan kategori yang paling ideal karena mampu mengukur kemampuan peserta didik secara lebih objektif dibandingkan soal yang terlalu mudah maupun terlalu sukar. Analisis tingkat kesulitan memberikan manfaat bagi guru dalam mengidentifikasi kualitas setiap butir soal sehingga dapat menentukan apakah soal perlu dipertahankan, direvisi, atau diganti. Selain itu, hasil analisis juga dapat dimanfaatkan sebagai bahan refleksi untuk memperbaiki proses pembelajaran, khususnya pada materi yang masih belum dikuasai oleh sebagian besar peserta didik. Oleh karena itu, guru disarankan untuk melaksanakan analisis tingkat kesulitan butir soal secara berkala sebagai bagian dari evaluasi pembelajaran. Dengan demikian, instrumen penilaian yang digunakan akan semakin berkualitas serta mampu memberikan informasi yang akurat mengenai pencapaian hasil belajar siswa.

Berdasarkan hasil penelitian yang telah dilakukan, guru disarankan untuk melaksanakan analisis tingkat kesulitan butir soal secara rutin sebelum maupun setelah pelaksanaan evaluasi pembelajaran. Kegiatan tersebut penting dilakukan agar setiap butir soal yang digunakan benar-benar memiliki kualitas yang baik dan mampu mengukur kemampuan peserta didik secara objektif, valid, serta akurat. Selain itu, guru juga perlu memperhatikan keseimbangan komposisi soal mudah, sedang, dan sukar sehingga instrumen evaluasi mampu memberikan informasi yang lebih menyeluruh mengenai pencapaian kompetensi peserta didik.

## DAFTAR REFERENSI

- Anderson, L. W., & Krathwohl, D. R. (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives: complete edition*. Addison Wesley Longman, Inc..
- Arikunto, S. (2018). *Dasar-Dasar Evaluasi Pendidikan* (Edisi ke-3). Bumi Aksara.
- Hadmar, S. S. A., Ali, A. M., & Yurfiah. (2024). Analisis daya pembeda dan tingkat kesukaran soal pilihan ganda pada mata pelajaran IPA di sekolah dasar. *PROSA: Jurnal Penelitian Pendidikan Guru Sekolah Dasar*.
- Hartono, I. D. I., Tenriawaru, A. B., & Ningsih, K. (2024). Analisis butir soal penilaian sumatif IPA kelas vii SMP Negeri 3 Pontianak menggunakan anates. *Jurnal Kajian Pembelajaran dan Keilmuan*, 8(2), 162-171. <https://doi.org/10.26418/jurnalkpk.v8i2.78282>
- Krathwohl, D. R. (2002). A revision of Bloom's taxonomy: An overview. *Theory into practice*, 41(4), 212-218. [https://doi.org/10.1207/s15430421tip4104\\_2](https://doi.org/10.1207/s15430421tip4104_2)
- Kusaeri, & Suprananto. (2012). *Pengukuran dan Penilaian Pendidikan*. Graha Ilmu.
- Mardapi, D. (2012). *Pengukuran, Penilaian, dan Evaluasi Pendidikan*. Nuha Medika.
- Ningrum, W. A., Rahmawati, R. C., Minarti, I. B., & Mekar, R. M. (2023). Analisis Butir Soal Ulangan Harian Pembelajaran IPA Pada Kelas VIII Di SMPN 21 Semarang. *Educatio*, 18(1), 91-101. <https://doi.org/10.29408/edc.v18i1.18291>
- Paskalin, G., & Susanti, M. M. I. (2020). Analisis Butir Soal untuk Materi Gaya di Sekolah Dasar. *DIDAKTIKA: Jurnal Pendidikan Sekolah Dasar*, 3(1), 23-32. <https://doi.org/10.21831/didaktika.v3i1.31431>
- Sudijono, A. (2015). *Pengantar Evaluasi Pendidikan*. Rajawali Pers.
- Sudjana, N. (2017). *Penilaian Hasil Proses Belajar Mengajar*. Remaja Rosdakarya.
- Sugiyono. (2023). *Metode Penelitian Kuantitatif, Kualitatif, dan R&D* (Edisi ke-3). Alfabeta.
- Valen, A., & Satria, T. G. (2021). Analisis tingkat kesulitan soal PAS (Penilaian Akhir Semester) mata pelajaran IPS di sekolah dasar. *Jurnal Basicedu*, 5(4), 2199-2208. <https://doi.org/10.31004/basicedu.v5i4.1149>
- Widoyoko, E. P. (2014). *Teknik Penyusunan Instrumen Penelitian*. Pustaka Pelajar.