

Analisis Butir Uji Validitas, Reliabilitas Tingkat Kesukaran, Daya Pembeda Soal Sumatif Pendidikan Pancasila

Atika Nur Rohmah* & Fina Fakhriyah

Program Studi Magister Pendidikan Dasar, Fakultas Keguruan dan Ilmu Pendidikan,
Universitas Muria Kudus, Indonesia

*Corresponding Author: 202403076@std.umk.ac.id

Article History

Received : January 20th, 2026

Revised : January 28th, 2026

Accepted : February 23th, 2026

Abstract: Evaluasi pembelajaran merupakan salah satu komponen penting dalam memastikan ketercapaian tujuan pendidikan secara optimal. Oleh karena itu, instrumen penilaian, khususnya soal sumatif, perlu dianalisis secara sistematis agar memenuhi kriteria validitas dan reliabilitas sehingga mampu memberikan hasil pengukuran yang konsisten dan akurat. Penelitian ini bertujuan untuk menganalisis butir soal sumatif Pendidikan Pancasila MI An Nashriyah materi Keberagaman Budaya Indonesia. Penelitian ini menggunakan pendekatan kuantitatif dengan metode deskriptif. Sampel penelitian ini adalah kelas VI MI An Nashriyah tahun ajaran 2025/2026 dengan jumlah 30 siswa. Metode analisa data yang digunakan adalah analisis validitas, reliabilitas, tingkat kesukaran dan daya pembeda. Berdasarkan hasil analisis yang dilakukan, maka didapatkan dari 25 soal terdapat 14 soal valid sedangkan 11 soal lainnya dinyatakan tidak valid. Uji reliabilitas soal dilakukan menggunakan perhitungan nilai *Cronbach's Alpha* yang didapatkan dengan nilai 0,701. Sebanyak 12 soal memiliki tingkat kesukaran mudah dan 13 soal lainnya memiliki tingkat kesukaran sedang. Berdasarkan perhitungan daya pembeda, didapatkan 6 soal berkategori daya pembeda baik, 12 soal berkategori daya pembeda cukup baik, dan 6 soal berkategori daya pembeda kurang baik/buruk. Berdasarkan hasil penelitian, dapat disimpulkan bahwa kualitas butir soal sumatif Pendidikan Pancasila di MI An Nashriyah secara umum telah memenuhi kriteria reliabilitas dengan kategori cukup, namun masih memerlukan perbaikan pada aspek validitas dan daya pembeda. Dengan demikian, hasil penelitian ini dapat dijadikan sebagai bahan evaluasi bagi guru dalam menyusun dan memperbaiki kualitas butir soal sumatif agar lebih valid, reliabel, serta sesuai dengan karakteristik peserta didik dan tujuan pembelajaran Pendidikan Pancasila.

Keywords: *Assessment Questions, Item Analysis, Pancasila Education*

PENDAHULUAN

Pendidikan merupakan proses fundamental dalam membentuk kualitas sumber daya manusia yang berilmu, berkarakter, dan berdaya saing. Melalui pendidikan, siswa tidak hanya dibekali pengetahuan dan keterampilan, tetapi juga nilai-nilai moral dan sosial yang menjadi dasar dalam kehidupan bermasyarakat dan berbangsa. Di era globalisasi yang ditandai dengan pesatnya perkembangan teknologi dan arus informasi lintas budaya, pendidikan moral dan karakter menjadi aspek yang sangat penting untuk membentuk peserta didik yang tidak hanya cerdas secara intelektual, tetapi juga memiliki integritas, etika, dan jati diri bangsa yang kuat sebagai bekal dalam menghadapi berbagai tantangan global. Salah satu mata pelajaran yang

memiliki peran penting dalam pembentukan karakter dan jati diri bangsa adalah Pendidikan Pancasila.

Pendidikan Pancasila merupakan salah satu komponen pendidikan kewarganegaraan yang secara sistematis diarahkan pada internalisasi nilai-nilai ideologi Pancasila melalui proses pembelajaran, sehingga peserta didik diharapkan mampu berkembang menjadi warga negara yang berkarakter, beretika, dan berlandaskan nilai-nilai kebangsaan (Lestari & Kurnia, 2022). Sebagai sebuah negara dengan keanekaragaman budaya, suku, agama, dan bahasa, Indonesia membutuhkan suatu fondasi yang kuat untuk mempersatukan seluruh elemen masyarakatnya. Dalam konteks tersebut, Pendidikan Pancasila memiliki peran strategis sebagai sarana internalisasi nilai-nilai persatuan,

toleransi, dan kebhinekaan kepada peserta didik sejak dini (Amelia et al., 2023). Salah satu materi penting dalam mata pelajaran ini adalah Keberagaman Budaya Indonesia, yang bertujuan menanamkan pemahaman sejak dini mengenai nilai-nilai kebhinekaan dalam kehidupan bermasyarakat. Untuk mengetahui sejauh mana pemahaman dan penguasaan siswa terhadap materi tersebut, diperlukan proses penilaian yang sistematis dan berkelanjutan melalui instrumen evaluasi yang berkualitas.

Penilaian merupakan komponen penting dalam proses pembelajaran untuk memotivasi siswa meningkatkan pemahamannya dan untuk memenuhi harapan standar ketercapaiannya keberhasilan belajar (Handayani et al., 2022). Untuk mengukur keberhasilan pembelajaran secara komprehensif, guru perlu menerapkan penilaian formatif sebagai sarana perbaikan proses pembelajaran dan penilaian sumatif sebagai alat evaluasi capaian akhir siswa. Dalam pelaksanaan penilaian formatif dan sumatif, diperlukan instrumen penilaian yang disusun secara sistematis dan terstandar agar hasil penilaian yang diperoleh dapat memberikan gambaran yang akurat mengenai capaian belajar peserta didik. Oleh karena itu, instrumen penilaian yang baik harus memenuhi kriteria validitas, reliabilitas, tingkat kesukaran, dan daya pembeda agar hasil penilaian yang diperoleh benar-benar mencerminkan kemampuan siswa secara objektif.

Kualitas butir soal merupakan aspek yang sangat penting karena soal yang disusun dengan baik mampu memberikan gambaran yang akurat mengenai kondisi dan tingkat kemampuan peserta didik. Sebaliknya, butir soal yang memiliki kualitas rendah tidak dapat mencerminkan kemampuan peserta didik secara tepat, sehingga hasil penilaian yang diperoleh menjadi kurang akurat (Tarmizi et al., 2021). Butir soal yang berkualitas ditandai dengan tingkat validitas yang tinggi, reliabilitas yang baik, tingkat kesukaran yang berada pada kategori sedang, serta daya pembeda yang efektif pada setiap nomor soal (Saputra et al., 2022). Dengan terpenuhinya kriteria tersebut, instrumen penilaian yang digunakan diharapkan mampu mengukur kemampuan siswa secara objektif, akurat, dan adil, sehingga hasil penilaian dapat dijadikan dasar yang valid dalam pengambilan keputusan pembelajaran serta perbaikan proses evaluasi di masa selanjutnya.

Penelitian ini diperkuat oleh beberapa relevansi penelitian tentang uji validitas, reliabilitas, daya pembeda dan tingkat kesukaran soal yang telah dilakukan sebelumnya. Penelitian tersebut diantaranya oleh Dianova & Anwar (2024) yang berjudul “*Analisis Butir Uji Validitas, Reliabilitas, Tingkat Kesukaran, dan Daya Pembeda Soal Sumatif Bahasa Arab SD Islam*” dengan hasil penelitian bahwa dari 20 butir soal yang valid diperoleh nilai reliabilitas Cronbach’s Alpha sebesar 0,885. Sebagian besar soal memiliki tingkat kesukaran mudah (17 soal) dan sisanya berkategori sedang (3 soal). Berdasarkan daya pembeda, 11 soal tergolong baik, 5 soal cukup baik, 3 soal sedang, dan 6 soal berkategori buruk. Penelitian lainnya oleh (Herlina et al., 2023) yang berjudul “*Analytical Review of Final Examination Questions for Fifth Grade at SDN 03 East Pontianak*” dengan hasil penelitian terhadap 30 butir soal pilihan ganda menunjukkan bahwa 73% soal dinyatakan valid dan 27% tidak valid, diperoleh koefisien reliabilitas sebesar 0,899 yang termasuk kategori tinggi, tingkat kesukaran, sekitar 13% soal berkategori sukar, 73% berkategori sedang, dan 13% berkategori mudah, dan ditinjau dari daya pembeda, sekitar 39% soal termasuk kategori baik, 42% cukup, dan 19% tidak baik, serta dari aspek pengecoh, sekitar 89% soal diterima, 8% direvisi, dan 3% ditolak. Penelitian relevan lainnya oleh Nanda et al., (2025) yang berjudul “*Item analysis of mathematical literacy test and self efficacy of learners*” dengan hasil penelitian menunjukkan bahwa seluruh instrumen penilaian yang digunakan, baik soal esai tingkat tinggi, soal pilihan ganda, maupun kuesioner, dinyatakan valid (100%). Namun, reliabilitas soal ujian esai dan pilihan ganda masih berada pada kategori sedang. Analisis lebih lanjut terhadap soal pilihan ganda menunjukkan bahwa 92% butir soal berada pada tingkat kesukaran mudah dan 25% butir soal memiliki daya pembeda yang rendah.

Berdasarkan uraian di atas, peneliti ingin menggali lebih mendalam tentang uji validitas, reliabilitas, tingkat kesukaran dan daya beda soal sumatif Pendidikan Pancasila materi Keberagaman Budaya Indonesia agar instrumen penilaian yang digunakan mampu mengukur kemampuan siswa secara objektif, akurat, dan adil, sehingga hasil penilaian dapat dijadikan dasar yang valid dalam pengambilan keputusan pembelajaran serta perbaikan proses evaluasi di masa selanjutnya.

METODE

Penelitian ini menggunakan pendekatan kuantitatif dengan metode deskriptif, yang bertujuan untuk mendeskripsikan secara kuantitatif karakteristik data penelitian, meliputi tingkat validitas dan reliabilitas instrumen, tingkat kesulitan butir soal, serta daya pembeda soal yang digunakan pada muatan pembelajaran Pendidikan Pancasila, sehingga diperoleh gambaran objektif mengenai kualitas instrumen penilaian yang dikembangkan. Penelitian dilaksanakan di MI An Nashriyah pada semester ganjil tahun ajaran 2025/2026 tepatnya pada 29 November 2025. Subjek penelitian adalah siswa kelas VI MI An Nashriyah yang berjumlah 30 siswa. Pemilihan sampel dilakukan dengan menggunakan teknik *purposive random sampling*, yaitu metode pengambilan sampel secara acak yang tetap memperhatikan kriteria atau pertimbangan tertentu sesuai dengan tujuan penelitian (Farhana Maulida et al., 2023). Pertimbangan yang dimaksud adalah kelas tersebut telah mempelajari materi keberagaman budaya Indonesia. Instrumen penelitian yang digunakan berupa soal sumatif Pendidikan Pancasila yang terdiri atas 25 butir soal pilihan ganda dan mencakup level kognitif C1-C6. Instrumen tersebut digunakan untuk mengukur pencapaian hasil belajar peserta didik serta ketercapaian kompetensi pada berbagai tingkat kognitif.

Teknik analisis data pada penelitian ini terdiri atas analisis kualitas soal sumatif yang meliputi uji validitas, uji reliabilitas, daya pembeda, dan tingkat kesukaran. Uji validitas merupakan prosedur untuk memastikan bahwa data yang diperoleh benar-benar mengukur aspek yang menjadi tujuan penelitian (Anggraini et al., 2022). Dalam penelitian ini, pengujian validitas dilakukan dengan mengacu pada kriteria pengambilan keputusan pada uji statistik korelasi, yaitu suatu indikator dalam kuesioner dinyatakan valid apabila nilai koefisien korelasi yang diperoleh (r hitung) lebih besar daripada nilai r tabel (r hitung > r tabel). Setelah proses pengujian validitas selesai, instrumen selanjutnya dianalisis tingkat reliabilitasnya untuk memastikan konsistensi pengukuran. Pengujian reliabilitas bertujuan untuk mengetahui sejauh mana instrumen penelitian mampu memberikan hasil yang konsisten dalam pengukuran (Al Hakim et al., 2021). Instrumen dikatakan reliabel apabila menghasilkan data

yang relatif sama ketika digunakan pada waktu yang berbeda atau pada kondisi yang setara, sehingga dapat dipercaya sebagai alat ukur yang stabil dalam mengumpulkan data penelitian. Uji reliabilitas dalam penelitian ini dilakukan menggunakan koefisien Cronbach's Alpha untuk mengetahui tingkat konsistensi internal instrumen penelitian. Instrumen dinyatakan reliabel apabila nilai koefisien Cronbach's Alpha yang diperoleh lebih besar dari 0,70 dan menunjukkan bahwa instrumen memiliki tingkat reliabilitas yang memadai dan dapat digunakan sebagai alat ukur yang konsisten dalam penelitian. Pengujian daya beda instrumen dilakukan untuk mengetahui kemampuan setiap butir soal dalam membedakan peserta didik yang telah menguasai materi dengan peserta didik yang belum menguasainya (Saputri et al., 2023). Daya beda yang baik menunjukkan bahwa butir soal mampu mengidentifikasi perbedaan kemampuan belajar peserta didik secara efektif, sehingga layak digunakan sebagai alat evaluasi hasil belajar. Daya beda pada penelitian ini dihitung menggunakan rumus Kusaeri-Supranto dengan cara menentukan selisih proporsi jawaban benar antara kelompok atas dan kelompok bawah. Perhitungan daya beda dinyatakan dengan rumus $D = P_A - P_B$, di mana P_A merupakan proporsi jawaban benar pada kelompok atas dan P_B merupakan proporsi jawaban benar pada kelompok bawah.

Tabel 1. Interpretasi Indeks Daya Pembeda Butir

Daya Beda	Kalsifikasi	Interpretasi
0,70-1,00	Excellent	Baik Sekali
0,40-0,69	Baik	Baik
0,20-0,39	Memuaskan	Cukup
0,00-0,19	Lemah	Kurang Baik
Bertanda Negatif		Jelek

(Magdalena et al., 2021)

Tahap selanjutnya adalah melakukan analisis tingkat kesukaran butir soal. Analisis tingkat kesukaran butir soal dilakukan untuk mengetahui tingkat kesulitan soal yang digunakan (Nurhalimah et al., 2022). Selain memenuhi kriteria validitas dan reliabilitas, kualitas butir soal yang baik juga ditandai oleh adanya keseimbangan tingkat kesukaran, sehingga instrumen tidak didominasi oleh soal yang terlalu mudah maupun terlalu sukar.

Tabel 2. Interpretasi Tingkat Kesukaran Butir Tes

Nilai P	Kategori
0,00 – 0,30	Sukar
0,31 – 0,70	Sedang
0,71 – 1,00	Mudah

Sudijono (Magdalena et al., 2021)

HASIL DAN PEMBAHASAN

Uji Validitas

Penelitian telah dilaksanakan di MI An Nashriyah tahun ajaran 2025/2026. Sampel berjumlah 30 siswa dari kelas VI. Dalam penelitian ini, diperoleh hasil perhitungan dari uji validitas, reliabilitas, daya pembeda, dan tingkat kesukaran soal sumatif Pendidikan Pancasila. Soal yang digunakan dalam penelitian ini sebanyak 25 butir soal dengan 6 indikator pencapaian kompetensi. Adapun indikator pencapaian kompetensi yang dimaksud yaitu: (1)

menjelaskan faktor penyebab keberagaman, (2) menyebutkan jenis keberagaman di Indonesia, (3) menentukan contoh keberagaman daerah, (4) menyebutkan semboyan persatuan dan kesatuan, (5) mengkritik pertunjukan festival budaya, dan (6) menciptakan karya tentang keberagaman. Instrumen soal yang telah disusun selanjutnya dianalisis validitasnya menggunakan bantuan Microsoft Excel melalui uji korelasi statistik, dengan cara membandingkan nilai r hitung dan r tabel berdasarkan distribusi nilai r pada taraf signifikansi 5%. Dari 25 soal yang telah diujikan, sebanyak 44% atau 11 soal dinyatakan tidak valid sedangkan 56% atau 14 soal lainnya dinyatakan valid. Temuan tersebut mengindikasikan bahwa terdapat 11 butir soal yang perlu direvisi agar layak digunakan pada pelaksanaan tes selanjutnya. Berikut adalah hasil uji validitas butir soal:

Tabel 3. Hasil Uji Validitas Butir Soal

No. Soal	R Hitung	R Tabel	Keterangan
1	0,466	0,361	Valid
2	0,252	0,361	Tidak valid
3	0,249	0,361	Tidak valid
4	0,403	0,361	Valid
5	0,154	0,361	Tidak valid
6	0,072	0,361	Tidak valid
7	0,292	0,361	Tidak valid
8	0,101	0,361	Tidak valid
9	0,352	0,361	Tidak valid
10	0,442	0,361	Valid
11	0,664	0,361	Valid
12	0,390	0,361	Valid
13	0,204	0,361	Tidak valid
14	0,280	0,361	Tidak valid
15	0,557	0,361	Valid
16	0,428	0,361	Valid
17	0,394	0,361	Valid
18	0,059	0,361	Tidak valid
19	0,428	0,361	Valid
20	0,390	0,361	Valid
21	0,172	0,361	Tidak valid
22	0,389	0,361	Valid
23	0,461	0,361	Valid
24	0,478	0,361	Valid
25	0,507	0,361	Valid

Tabel 4. Persentase Hasil Uji Validitas Soal

Indeks Validitas	Nomor Soal	Jumlah	Persentase
Valid (r hitung > 0,361)	1,4,10,11,12,15,16,17,19,20,22,23,24,25	14	56%
Tidak Valid (r hitung < 0,361)	2,3,5,6,7,8,9,13,14,18,21	11	44%
Total		25	100%

Beberapa butir soal yang tidak valid, dapat diidentifikasi penyebabnya diantaranya adalah tingkat kesukaran soal yang kurang proporsional, yakni soal yang diberikan terlalu mudah sehingga banyak siswa yang menjawab betul dengan

perhitungan $r_{hitung} = 0,059$ sedangkan $r_{tabel} = 0,361$ sehingga $r_{hitung} < r_{tabel}$ artinya tidak valid. Soal yang dimaksudkan adalah sebagai berikut:

18. Setiap suku bangsa juga memiliki Bahasa sebagai alat komunikasi yang berbeda beda, salah satunya adalah Bahasa Sunda. Asal daerah dari Bahasa Sunda adalah....
- Jawa Tengah
 - Jawa Timur
 - Jawa Barat
 - DKI Jakarta

Gambar 1. Contoh Soal Tidak Valid

Penyebab lain ketidakvalidan soal adalah konteks soal tidak familiar bagi siswa. Soal yang diberikan jauh dari pengalaman siswa, atau tidak berada pada lingkungan siswa. Ini menyebabkan

anak akan cenderung menebak jawaban dari soal yang diberikan. Adapun contoh soal tersebut sebagai berikut:

9. Tarian data merupakan budaya yang diwariskan secara turun menurun untuk menggambarkan legenda, atau nilai adat istiadat Masyarakat setempat. Di bawah ini yang merupakan contoh tarian adat dari Kalimantan yang paling tepat adalah....

A.		Kalimantan utara
B.		Kalimantan selatan
C.		Kalimantan timur
D.		Kalimantan tengah

Gambar 2. Contoh Soal Tidak Valid

Uji Reliabilitas

Tahapan setelah dilakukan uji validitas adalah melakukan uji reliabilitas untuk mengetahui tingkat konsistensi hasil pengukuran. Uji reliabilitas dilakukan dengan menghitung nilai *Cronbach's Alpha* dengan batas acuan sebesar 0,70 yang termasuk dalam kategori

reliabilitas cukup. Hasil perhitungan menunjukkan nilai *Cronbach's Alpha* sebesar 0,701, sehingga dapat disimpulkan bahwa instrumen soal tersebut memiliki tingkat reliabilitas yang baik atau reliabel. Adapun hasil uji reliabilitas soal sebagai berikut:

Tabel 5. Hasil Uji Reliabilitas Butir Soal

Kriteria Pengukuran		
Nilai Acuan	Nilai Cronbach's Alpha	Keterangan
0,70	0,701	Reliabel

Uji Daya Pembeda

Soal yang telah memenuhi kriteria validitas dan reliabilitas selanjutnya dianalisis tingkat daya pembedanya. Analisis ini bertujuan untuk mengetahui sejauh mana setiap butir soal

mampu membedakan siswa yang berkemampuan tinggi dan siswa yang berkemampuan rendah. Hasil analisis daya pembeda pada 25 butir soal sumatif Pendidikan Pancasila materi Keberagaman Budaya Indonesia menunjukkan

bahwa sebesar 24% atau 6 butir soal berada pada kategori daya pembeda baik, 52% atau 13 butir soal termasuk dalam kategori daya beda cukup baik, dan 24% lainnya atau 6 soal lainnya berada pada kategori daya pembeda kurang baik. Temuan tersebut dapat diartikan bahwa sebagian besar soal telah mampu membedakan peserta didik yang memiliki kemampuan tinggi dan rendah. Secara rinci, 6 butir soal memiliki daya pembeda yang baik sehingga sangat efektif dalam mengidentifikasi perbedaan kemampuan

siswa, 13 butir soal berada pada kategori cukup baik yang berarti masih mampu membedakan kemampuan siswa meskipun belum optimal, sedangkan 6 butir soal dengan daya pembeda kurang baik menunjukkan kemampuan yang terbatas dalam membedakan siswa berkemampuan tinggi dan rendah sehingga perlu dilakukan perbaikan atau revisi sebelum digunakan kembali. Berikut disajikan hasil uji daya pembeda soal:

Tabel 6. Hasil Uji Daya Pembeda Soal

Kategori	Nomor Soal	Jumlah	Persentase
Baik	10,11,12,17,23,26	6	24%
Cukup	1,2,4,7,9,13,14,15,16,19,20,22	13	52%
Kurang Baik	3,5,6,18,21	6	24%
Total		25	100%

Adapun contoh soal yang memiliki daya pembeda kurang baik sebagai berikut:

21. Sekolah Nusantara mengadakan Festival Budaya Indonesia. Setiap kelas menampilkan pertunjukan dari daerah yang berbeda. Kelas V menampilkan Tari Saman dari Aceh. Penari sudah mengenakan kostum yang hampir sesuai dan menampilkan gerakan secara kompak. Namun, selama pertunjukan berlangsung, musik pengiring terdengar terlalu pelan, sehingga beberapa gerakan tidak selaras dengan irama. Selain itu, ekspresi wajah penari terlihat kurang bersemangat, meskipun gerakan tangan sudah cukup rapi. Niken Salindri diminta menjadi juri saat pertunjukan tersebut. Kritik membangun yang harus disampaikan Niken Salindri sebagai dewan juri yang objektif adalah...
- Pertunjukan sebaiknya diganti dengan tarian dari daerah lain agar lebih menarik.
 - Gerakan tari tidak perlu diperhatikan karena yang terpenting adalah kostumnya.
 - Musik pengiring perlu diperjelas dan penari sebaiknya menampilkan ekspresi yang lebih bersemangat agar pertunjukan lebih hidup.
 - Pertunjukan sudah sempurna sehingga tidak memerlukan perbaikan apa pun.

Gambar 3. Contoh Soal Daya Pembeda Kurang Baik

Soal tersebut menggunakan redaksi soal yang terlalu panjang. Hal ini bisa menjadikan siswa salah memahami soal yang diberikan. Siswa akan merasa malas membaca soal yang terlalu panjang hingga akhirnya mereka akan cenderung menebak jawabannya. Hal tersebut dapat menyebabkan daya pembeda soal rendah sehingga memerlukan revisi atau perbaikan agar mampu membedakan kemampuan siswa lebih akurat.

Uji Tingkat Kesukaran

Analisis tingkat kesukaran ini bertujuan untuk mengetahui apakah soal yang diberikan termasuk dalam kategori mudah, sedang, atau sulit. Hasil analisis tingkat kesukaran pada 25 butir soal sumatif Pendidikan Pancasila materi

Keberagaman Budaya Indonesia menunjukkan bahwa sebesar 48% atau 12 soal memiliki kategori tingkat kesukaran mudah, sedangkan 52% atau 13 soal lainnya memiliki kategori tingkat kesukaran sedang. Temuan ini mengindikasikan bahwa sebagian besar soal telah disusun dengan tingkat kesukaran yang proporsional, sehingga mampu mengukur pemahaman siswa secara bertahap. Namun demikian, dominasi soal dengan kategori mudah dan sedang menunjukkan bahwa soal dengan tingkat kesukaran tinggi masih perlu ditambahkan agar instrumen penilaian dapat mengukur kemampuan berpikir tingkat tinggi siswa secara lebih optimal. Berikut dilampirkan hasil perhitungan tingkat kesukaran soal:

Tabel 7. Hasil Uji Tingkat Kesukaran Soal

Kategori	Nomor Soal	Jumlah	Persentase
Mudah	1,2,3,5,6,7,8,9,13,18,19,23	12	48%
Sedang	4,10,11,12,14,15,16,17,20,21,22,24,25	13	52%
Sulit	-	0	0%

Total	25	100
--------------	----	-----

Hasil Hasil analisis tingkat kesukaran menunjukkan bahwa tidak terdapat butir soal yang termasuk dalam kategori sulit, meskipun secara konseptual dan konstruksi soal telah dirancang pada level berpikir tingkat tinggi (*Higher Order Thinking Skills/HOTS*). Kondisi ini dapat dijelaskan melalui beberapa faktor berikut:

- 1) Tinggi-rendahnya IQ siswa. Tingkat kecerdasan atau tinggi rendahnya IQ seseorang siswa menjadi sangat berpengaruh dalam penangkapan dari setiap yang dibacanya untuk mendapatkan hasilnya (Gunarwati et al., 2021)
- 2) Persepsi kesulitan soal yang bersifat subjektif. Soal yang tampak kompleks, baik dari segi panjang wacana maupun tuntutan penalaran, belum tentu sulit bagi siswa apabila mereka telah memiliki pemahaman konsep yang memadai. Dengan pemahaman yang matang, siswa memiliki pola pikir yang terstruktur sehingga mampu menyelesaikan permasalahan secara mandiri tanpa mengalami hambatan yang berarti, bahkan ketika dihadapkan pada masalah yang muncul secara tiba-tiba (Rohmawati, et al., 2023).
- 3) Kejelasan petunjuk dan redaksi soal. Soal yang disusun dengan bahasa yang komunikatif, konteks yang familiar, serta pilihan jawaban yang logis dan tidak ambigu akan memudahkan siswa dalam memahami permasalahan yang disajikan. Kejelasan ini membantu siswa memfokuskan pada proses berpikir yang dituntut tanpa terhambat oleh kesalahan interpretasi soal.

Dengan demikian, dapat disimpulkan bahwa soal yang disusun telah memenuhi kriteria kualitas yang baik dan berorientasi pada pengembangan HOTS, namun hasil uji empiris menunjukkan tidak adanya butir soal yang tergolong sulit. Hal ini menegaskan bahwa tingkat kesukaran soal ditentukan oleh respons siswa secara nyata, bukan semata-mata oleh asumsi atau kompleksitas konstruksi soal.

KESIMPULAN

Berdasarkan hasil penelitian yang telah dilaksanakan di MI An Nashriyah pada tahun ajaran 2025/2026 dengan melibatkan 30 siswa kelas VI, dapat disimpulkan bahwa kualitas soal sumatif Pendidikan Pancasila materi Keberagaman Budaya Indonesia secara umum

berada pada kategori cukup baik, namun masih memerlukan perbaikan pada beberapa aspek. Dari 25 butir soal yang dikembangkan berdasarkan enam indikator pencapaian kompetensi, hasil uji validitas menunjukkan bahwa 14 soal dinyatakan valid, sedangkan 11 soal lainnya belum memenuhi kriteria validitas sehingga perlu direvisi sebelum digunakan kembali. Hasil uji reliabilitas menunjukkan nilai Cronbach's Alpha sebesar 0,701, yang menandakan bahwa instrumen soal memiliki tingkat konsistensi yang cukup baik dan dapat digunakan sebagai alat ukur yang reliabel. Selanjutnya, analisis daya pembeda menunjukkan bahwa sebagian besar soal telah mampu membedakan siswa dengan kemampuan tinggi dan rendah, meskipun masih terdapat beberapa butir soal dengan daya pembeda kurang baik yang memerlukan perbaikan. Ditinjau dari tingkat kesukaran, soal yang digunakan didominasi oleh kategori mudah dan sedang, sehingga secara proporsional telah mampu mengukur pemahaman siswa. Namun demikian, masih diperlukan penambahan soal dengan tingkat kesukaran tinggi agar instrumen penilaian dapat mengukur kemampuan berpikir tingkat tinggi siswa secara lebih optimal. Secara keseluruhan, instrumen soal yang disusun dapat digunakan sebagai alat evaluasi pembelajaran, dengan catatan dilakukan revisi dan penyempurnaan pada butir soal tertentu agar kualitas penilaian menjadi lebih baik.

UCAPAN TERIMA KASIH

Penulis menyampaikan apresiasi dan terima kasih kepada rekan-rekan satu kelompok yang telah menjalin kerja sama secara optimal, saling bertukar gagasan, serta memberikan dukungan selama pelaksanaan penelitian ini. Ucapan terima kasih juga ditujukan kepada dosen pembimbing atas arahan, saran, dan bimbingan yang diberikan dengan penuh perhatian sepanjang proses penelitian. Penulis sangat menghargai kesabaran, komitmen, dan waktu yang telah dicurahkan. Selain itu, penulis mengucapkan terima kasih kepada seluruh pihak terkait yang telah memberikan bantuan berupa dukungan teknis, fasilitas, serta sumber daya lain sehingga penelitian ini dapat terlaksana dengan baik. Seluruh kontribusi dan dukungan yang diberikan menjadi bagian penting dalam keberhasilan penelitian ini. Semoga hasil

penelitian ini dapat memberikan manfaat bagi pengembangan ilmu pengetahuan dan masyarakat.

REFERENSI

- Amelia, R., Nur, P., Linashar, A., Truvadi, R., Trinita, A., Fauzi, I., & Salam, B. (2023). *Peran pendidikan Pancasila dalam membentuk karakter bangsa Indonesia: Tinjauan dan implikasi*. *Advances in Social Humanities Research*, 1. <https://asrpublishing.org/index.php/ashr>
- Angraini, F. D. P., Aprianti, A., Setyawati, V. A. V., & Hartanto, A. A. (2022). Pembelajaran statistika menggunakan software SPSS untuk uji validitas dan reliabilitas. *Jurnal Basicedu*, 6(4), 6491–6504. <https://doi.org/10.31004/basicedu.v6i4.3206>
- Dianova, F. R., & Anwar, N. (2024). Analisis butir uji validitas, reliabilitas, tingkat kesukaran, dan daya pembeda soal sumatif Bahasa Arab SD Islam. *Jurnal Bahasa Daerah Indonesia*, 1(3), 1–13. <https://doi.org/10.47134/jbdi.v1i3.2863>
- Farhana Maulida, A., Zulkarnain, I., & Hidayanto, T. (2023). Analisis miskonsepsi peserta didik menggunakan tes diagnostik four-tier pada materi aljabar. *Mandalika Mathematics and Education Journal*, 5(1), 152–160. <https://doi.org/10.29303/jm.v5i1.4930>
- Gunarwati, R., Maula, L. H., & Nurasiah, I. (2021). Analisis kemampuan membaca pemahaman berbasis daring pada siswa sekolah dasar. *Janacitta: Journal of Primary and Children's Education*, 4(1). <http://jurnal.unw.ac.id/index.php/janacitta>
- Handayani, F., Maharani, R. A., & Fitria, Y. (2022). Penilaian dan jenis tes yang dibuat oleh guru di tingkat sekolah dasar. *Jurnal Basicedu*, 6(1), 726–737. <https://doi.org/10.31004/basicedu.v6i1.2033>
- Herlina, S., Sahrani, S., Ainin, M., & Nugraha, M. T. (2023). Analytical review of final examination questions for fifth grade at SDN 03 East Pontianak. *J-PAI: Jurnal Pendidikan Agama Islam*, 10(1). <https://doi.org/10.18860/jpai.v10i1.25892>
- Lestari, S. O., & Kurnia, H. (2022). Peran pendidikan Pancasila dalam pembentukan karakter. *Jurnal Citizenship: Media Publikasi Pendidikan Pancasila dan Kewarganegaraan*, 5(1), 1–25. <https://doi.org/10.12928/citizenship.v5i2.23179>
- Magdalena, I., Fauziah, S. N., Faziah, 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: Jurnal Pendidikan dan Sains*, 3(2). <https://ejournal.stitpn.ac.id/index.php/bintang>
- Nanda, E. M., Kismiantini, K., & Arliani, E. (2025). Item analysis of mathematical literacy test and self-efficacy of learners. *Jurnal Math Educator Nusantara*, 11(2), 307–318. <https://doi.org/10.29407/jmen.v11i2.27281>
- Nurhalimah, S., Hidayati, Y., Rosidi, I., Dwi, W., & Hadi, P. (2022). Hubungan antara validitas item dengan daya pembeda dan tingkat kesukaran soal pilihan ganda PAS. *Jurnal Natural Science Educational Research*, 4(3). <https://journal.trunojoyo.ac.id/nser>
- Rohmawati, A. A., & Fadly, W. (2023). Analisis kemampuan membuat solusi melalui kegiatan eksperimen. *Jurnal Tadris IPA Indonesia*. <http://ejournal.iainponorogo.ac.id/index.php/jtii>
- Saputra, H. D., Purwanto, W., Setiawan, D., Fernandez, D., & Putra, R. (2022). Hasil belajar mahasiswa: Analisis butir soal tes. *Edukasi: Jurnal Pendidikan*, 20(1), 15–27. <https://doi.org/10.31571/edukasi.v20i1.3432>
- Saputri, H. A., Zulhijrah, Larasati, N. J., & Shaleh. (2023). Analisis instrumen asesmen: Validitas, reliabilitas, tingkat kesukaran dan daya beda butir soal. *Jurnal Pendidikan*. <https://ejournal.uin-suska.ac.id>
- Tarmizi, P., Setiono, P., Amaliyah, Y., & Agrian, A. (2021). Analisis butir soal pilihan ganda tema “Sehat Itu Penting” kelas V SD Negeri 04 Kota Bengkulu. *ELSE (Elementary School). Education Journal*, 4(2), 124–132. <https://doi.org/10.30651/else.v4i2.7090>