



Analisis butir soal sumatif Pendidikan Pancasila dengan pendekatan Rasch di SMP Negeri 1 Pleret

Agusti Puji Lestari ^{a,1}, Suyato ^{b,2}, Dian Satria Charismana ^{c,3}

¹ (agustipuji.2021@student.uny.ac.id) Mahasiswa Departemen PKnH, Fishipol, UNY

² (suyato@uny.ac.id) Dosen Fakultas Ilmu Sosial, Hukum, dan Politik, UNY

³ (dian_satria@uny.ac.id) Dosen Fakultas Ilmu Sosial, Hukum, dan Politik, UNY

^a Agusti Puji Lestari (Departemen Pendidikan Kewarganegaraan dan Hukum FISHIPOL UNY), Yogyakarta Indonesia

^b Suyato (Departemen Pendidikan Kewarganegaraan dan Hukum FISHIPOL UNY), Yogyakarta Indonesia

^c Dian Satria Charismana (Departemen Pendidikan Kewarganegaraan dan Hukum FISHIPOL UNY), Yogyakarta Indonesia

ABSTRAK

Analisis butir soal dilakukan untuk meningkatkan kualitas butir soal melalui revisi atau membuang soal yang tidak efektif. Penelitian ini bertujuan untuk mengidentifikasi karakteristik butir soal pilihan ganda asesmen sumatif akhir semester (ASAS) Tahun Ajaran 2024/2025 kelas VIII mata pelajaran Pendidikan Pancasila di SMP Negeri 1 Pleret, Kabupaten Bantul. Penelitian ini merupakan penelitian deskriptif dengan menggunakan pendekatan kuantitatif. Teknik pengumpulan data yang digunakan yaitu dokumentasi. Data dianalisis menggunakan *software* Winsetp dengan Rasch model yang mengukur satu parameter yaitu tingkat kesulitan. Hasil penelitian menunjukkan bahwa butir soal pilihan ganda dalam ASAS memiliki tingkat kesulitan yang bervariasi. Terdapat 1 soal (2%) yang tergolong sangat sulit, 11 soal (22%) yang tergolong sulit, 25 soal (50%) tergolong sedang, 7 soal (14%) tergolong mudah, dan 6 soal (12%) tergolong sangat mudah. Berdasarkan temuan tersebut, tingkat kesulitan soal ASAS memiliki variasi yang cukup baik.

ABSTRACT

Analysis of question items is carried out to improve the quality of question items through revision or discarding ineffective questions. This study aims to identify the characteristics of multiple-choice question items for the summative assessment at the end of the semester (ASAS) for the 2024/2025 Academic Year grade VIII Pancasila Education subject at SMP Negeri 1 Pleret, Bantul Regency. This research is a descriptive research using a quantitative approach. The data collection technique used is documentation. The data was analyzed using Winstep software with a Rasch model that measures one parameter, namely the level of difficulty. The results of the study show that the multiple-choice question items in ASAS have varying levels of difficulty. There was 1 question (2%) that was classified as very difficult, 11 questions (22%) that were classified as difficult, 25 questions (50%) classified as moderate, 7 questions (14%) classified as easy, and 6 questions (12%) classified as very easy. Based on these findings, the difficulty level of ASAS questions varies quite well.

Sejarah Artikel

Diterima: 14 April 2025

Disetujui: 15 April 2025

Kata kunci:

Analisis butir soal, rasch model, soal pendidikan pancasila

Keywords:

Analysis of question, rasch model, pancasila education questions

Pendahuluan

Pembelajaran adalah suatu rangkaian program yang mencakup tiga elemen utama, yakni tujuan pembelajaran, proses pembelajaran, dan evaluasi pembelajaran. Evaluasi dalam pembelajaran menjadi tahap penting karena berperan untuk mengukur pencapaian peserta didik terhadap tujuan pembelajaran yang telah ditetapkan serta memberikan umpan balik bagi guru dan peserta didik. Hasil evaluasi dapat digunakan sebagai masukan untuk memperbaiki dan meningkatkan program serta kegiatan pembelajaran (Ismail, 2020). Evaluasi dilakukan untuk mengetahui apakah tujuan pembelajaran tercapai atau tidak dan sejauh mana kualitas soal yang dibuat (Priowuntato, 2020).

Penilaian memegang peran signifikan dalam menentukan tingkat pencapaian kompetensi peserta didik selama proses pembelajaran. Penilaian bertujuan mengumpulkan informasi mengenai kemajuan belajar peserta didik yang menjadi dasar dalam pengambilan keputusan yang didasarkan pada standar serta berbagai faktor yang relevan (Asrul, Saragih, & Mukhtar, 2022, p.16). Hasil penilaian mencerminkan kualitas pembelajaran yang telah berlangsung. Oleh karena itu, guru perlu memiliki keterampilan yang baik dalam mengevaluasi pembelajaran, mulai dari perencanaan, penyusunan instrumen penilaian, analisis, hingga interpretasi (Jumini, Madnasri, Cahyono, & Parmin, 2023, p.759).

Instrumen penilaian dapat dikatakan baik jika mampu memberikan data yang akurat. Instrumen penilaian terbagi menjadi dua kategori, yaitu tes dan nontes. Tes adalah bagian terpenting dari pembelajaran, dimana tes sebagai suatu alat yang dilakukan untuk mengevaluasi kemampuan peserta didik (Utomo, 2022). Kualitas tes secara keseluruhan dipengaruhi oleh kualitas setiap butir soal. Oleh karena itu, tes harus diujicobakan di lapangan sebagai tahap penilaian untuk menentukan apakah ada elemen yang perlu direvisi, disempurnakan, atau bahkan dihilangkan. Menurut Retnawati (2014, p.62), tes berkualitas tersusun dari pertanyaan-pertanyaan yang dirancang dengan baik. Untuk memperoleh butir soal yang berkualitas, setiap butir soal dalam perangkat tes perlu di analisis secara mendalam. Sebab, jika instrumen tes kurang baik, peserta didik dapat mengalami kesulitan atau melakukan kesalahan dalam memahami soal, terutama pada pertanyaan yang bersifat HOTS (Anggarani & Hidayat, 2022).

Analisis butir soal adalah langkah penting dalam mengevaluasi kualitas sebuah tes, baik secara keseluruhan maupun pada masing-masing butir soal yang ada (Arifin, 2017, p.246). Sebuah soal yang berkualitas adalah soal yang dapat menyajikan informasi yang akurat sesuai dengan tujuan yang ditetapkan, misalnya untuk mengidentifikasi peserta didik yang sudah atau belum memahami materi. Menurut Farida dan Musyarofah (2021, p.38), analisis umumnya dilakukan dengan dua pendekatan yaitu analisis kualitatif (*qualitative control*) dan analisis kuantitatif (*quantitative control*). Analisis kualitatif butir soal mengacu pada kaidah penyusunan soal dan dilakukan sebelum diujikan, mencakup aspek materi, konstruksi, bahasa, dan kunci jawaban. Sementara itu, analisis kuantitatif didasarkan pada bukti empiris setelah soal diujikan.

Pada kenyataannya masih banyak guru yang belum menyadari betapa pentingnya kualitas pertanyaan yang digunakan, sehingga guru cenderung melakukan penilaian tanpa memperhatikan efektivitas soal-soal yang diujikan dalam mengukur kemampuan kognitif peserta didik. Guru hanya cenderung melakukan analisis subjektif berupa validitas logis berdasarkan kemampuan peserta didik menjawab soal, padahal itu tidak cukup untuk menilai kualitas soal. Akibatnya, kualitas butir soal yang akan diujikan tersebut belum dapat dipastikan (Aeni & Suyato, 2020). Oleh karena itu, soal yang akan diujikan kepada peserta didik perlu diteliti dan dianalisis untuk memastikan bahwa soal tersebut memenuhi kualitas soal yang baik. Menganalisis butir soal merupakan aktivitas yang harus dilakukan oleh guru dengan tujuan meningkatkan kualitas soal yang telah disusun. Selain itu, kegiatan ini juga bertujuan untuk mengidentifikasi soal-soal yang tergolong baik, kurang baik, maupun tidak layak digunakan (Ida & Musyarofah, 2021).

Menurut Nese Guler et al. sebagaimana dikutip Purnama dan Alfarisa (2020, p.38), terdapat dua teori utama yang digunakan untuk menilai kualitas soal, yaitu teori tes klasik dan teori respon butir. Karakteristik teori tes klasik bergantung pada kelompok (*group dependent*) dan item (*item dependent*). Hal ini berarti bahwa selain dipengaruhi oleh pertanyaan atau butir soal, peserta tes juga mempengaruhi koefisien reliabilitas, tingkat kesulitan, dan indeks daya pembeda tes. Sebaliknya, teori respon butir tidak dipengaruhi oleh peserta tes, kelompok peserta yang mengikuti ujian tidak menjadi masalah menurut teori ini asalkan peserta tes kompeten.

Pada penelitian ini, analisis butir-butir soal dilakukan dengan menggunakan teori respons butir. Model karakteristik butir yang digunakan adalah model *Rasch*, yang sering dikenal dengan butir soal satu parameter (1PL) (Harmurni, 2019, p.11). Menurut model *Rasch*, peluang seorang peserta tes menjawab suatu butir soal dengan benar dipengaruhi oleh satu faktor utama, yaitu tingkat kesulitan soal (Ayub, Istiyono, Munadi, Permadi, Patiserlihun, & Sudjipto, 2020, p.47). Model *Rasch* digunakan dalam penelitian ini untuk menganalisis hasil tes berbentuk pilihan ganda. Model ini mengasumsikan bahwa kemampuan peserta didik dapat diukur secara akurat dan objektif, serta setiap butir soal memiliki tingkat kesulitan berbeda tanpa bergantung pada kumpulan soal tertentu (Pudjiati & Madani, 2023, p.4). Menurut Sumintono dan Widhiarso (2015), tujuan utama model *Rasch* adalah menghasilkan skala pengukuran dengan interval yang setara, sehingga dapat memberikan informasi yang akurat mengenai peserta tes serta kualitas soal yang dikerjakan (Tarigan et al., 2022, p. 93). Parameter tingkat kesulitan dalam *Rasch* model, dianggap baik jika indeksinya berada dalam rentang -2 hingga +2 (DeMars, 2018). Dalam *Rasch* model suatu butir soal dapat dikatakan baik apabila memenuhi kriteria kecocokan butir dengan memperhatikan hasil analisis *outfit mean square*, *outfit z-standard*, dan *point measure correlation*. Butir soal dianggap valid jika terpenuhi dua dari tiga kriteria tersebut (Abdullah, Jahja, & Setiawan, 2022, p.46).

Hasil observasi dan wawancara dengan guru mata pelajaran Pendidikan Pancasila di SMP Negeri 1 Pleret menyatakan bahwa, guru belum menganalisis butir soal secara teknik respon per butir soal. Guru hanya berfokus pada ketuntasan peserta didik yang mengikuti remedial untuk memperbaiki nilai. Hampir semua guru tidak melakukan analisis terhadap butir soal yang dibuatnya, karena kegiatan tersebut memerlukan waktu yang cukup lama. Guru tersebut mengakui bahwa analisis butir soal seharusnya dilakukan untuk menilai kualitas soal yang telah disusun. Pada dasarnya, analisis butir soal merupakan langkah penting yang harus dilakukan guru untuk meningkatkan kualitas soal dan memahami karakteristiknya secara jelas, serta memberikan masukan sebagai dasar dalam pengambilan keputusan.

Peneliti menemukan sejumlah penelitian sebelumnya yang membahas analisis butir soal dengan menggunakan *Rasch* model. Pertama penelitian yang dilakukan oleh Ikhwa Nudin dan Rachmad Syarifudin Hidayatullah tahun 2023. Penelitian tersebut menganalisis butir soal menggunakan model *Rasch* yang dibantu dengan program *Winstep*. Hasil analisis menunjukkan bahwa butir soal pada penelitian ini memiliki nilai reliabilitas logit person sebesar 0,49, nilai reliabilitas butir soal sebesar 0,94, dan nilai logit alpha cronbach sebesar 0,81 (Nudin & Hidayatullah, 2023). Penelitian ini memiliki persamaan dan perbedaan dengan penelitian yang dilakukan oleh peneliti. Persamaan penelitian tersebut dengan penelitian ini adalah menganalisis butir soal dengan menggunakan model *Rasch* yang dibantu dengan program *Winstep*. Sedangkan perbedaannya terletak pada soal yang dianalisis, pencarian indeks daya pembeda, dan jenjang kelas yang berbeda serta tingkatan kelas yang berbeda.

Kedua penelitian dari Wiwin Fajriah Hanna dan Heri Retnawati tahun 2022. Hasil penelitian menunjukkan bahwa sepuluh item memiliki nilai INFIT MNSQ berkisar antara 0,77 hingga 1,33, sedangkan dua item tidak dapat masuk ke dalam model *Rasch* karena nilai INFIT MNSQ yang berada di luar jangkauan. Jika prediksi OUTFIT menunjukkan bahwa t kurang dari atau sama dengan 2,00, maka item tersebut dapat digunakan. Analisis OUTFIT t digunakan untuk mengidentifikasi sepuluh item yang memenuhi syarat dan dua item yang tidak memenuhi syarat, yang menunjukkan bahwa item yang sesuai dengan persyaratan dapat digunakan. Estimasi item memiliki nilai reliabilitas sebesar 0,97, sedangkan estimasi kasus memiliki nilai reliabilitas sebesar 0,25 (Hanna & Retnawati, 2022). Terdapat persamaan dan perbedaan antara penelitian tersebut dengan penelitian peneliti. Persamaannya terletak pada penggunaan model *Rasch* dalam menganalisis butir soal, sementara perbedaannya terdapat pada perangkat lunak yang digunakan dalam proses analisis butir soal.

Ketiga penelitian dari Debbi Silvia Aqida tahun 2024. Penelitian tersebut menganalisis butir soal untuk menentukan kesesuaian dan tingkat kesulitannya, dan temuan menunjukkan bahwa, berdasarkan output program Winstep, 15 butir soal sesuai dengan model Rasch, dengan nilai Outfit MNSQ 1,03 dan nilai Outfit ZSTD 0,0; 11 butir soal memiliki nilai Pt-measure Corr yang positif, sementara 4 butir soal memiliki nilai negatif, yang menunjukkan bahwa butir soal tersebut tidak sesuai dengan model. Butir soal matematika memenuhi setidaknya satu dari tiga kriteria yang telah ditetapkan dan sesuai dengan model, namun satu butir soal dinilai mudah berdasarkan tingkat kesulitannya (Aqida, 2024). Terdapat persamaan dan perbedaan antara penelitian ini dengan penelitian yang dilakukan oleh peneliti. Persamaannya yaitu menganalisis butir soal berdasarkan kesesuaian tingkat kesulitan butir soal dan perangkat lunak analisis butir soal yaitu Winstep. Namun, topik yang akan diteliti menjadi pembeda penelitian ini dengan peneliti.

Model *Rasch* belum banyak mewarnai jagad penelitian di Indonesia. Masih terbatas penelitian yang menganalisis butir soal menggunakan Model *Rasch* pada mata pelajaran Pendidikan Pancasila di SMP. Sehingga jelas bahwa penelitian ini masih jarang dilakukan dan memiliki unsur kebaruan. Penelitian ini tidak hanya berkontribusi dalam pengembangan kajian akademik, tetapi juga memberikan manfaat praktis bagi dunia pendidikan. Sehingga penelitian ini sangat penting agar guru mampu mengetahui kualitas soal yang disusunnya dan juga mampu mengukur kelebihan dan kelemahan peserta didiknya.

Tujuan utama penelitian ini adalah untuk mengevaluasi serta meningkatkan perangkat tes, khususnya pada mata pelajaran Pendidikan Pancasila untuk kelas VIII di SMP Negeri 1 Pleret. Hasil analisis terhadap perangkat tes dapat dijadikan sebagai pijakan awal dalam proses persiapan bank soal untuk mata pelajaran Pendidikan Pancasila SMP Negeri 1 Pleret. Guru diharapkan dapat menggunakan temuan dari analisis kualitas butir soal penelitian ini sebagai panduan untuk memperbaiki dan menyempurnakan perangkat tes yang digunakan.

Metode

Penelitian ini menggunakan penelitian deskriptif kuantitatif. Data yang dikumpulkan berupa angka untuk kemudian dianalisis dan ditarik kesimpulannya menggunakan pendekatan *Rasch*. Penelitian ini dilaksanakan di SMP N 1 Pleret Kabupaten Bantul. Teknik pengumpulan data pada penelitian ini adalah dokumentasi. Menurut Sugiyono (2019, p.329), dokumentasi adalah mengumpulkan informasi dan data yang telah tersimpan dalam berbagai bentuk atau format. Instrumen yang digunakan pada penelitian ini adalah soal ASAS Pendidikan Pancasila tahun ajaran 2024/2025 dan lembar jawaban peserta didik. Teknik pengambilan data digunakan pada tes jenis, pilihan ganda (*multiple choice*) yang dibuat oleh guru mata pelajaran Pendidikan Pancasila berjumlah 50 butir soal. Setelah diujikan kepada peserta didik butir soal dianalisis menggunakan teori modern dengan analisis pemodelan *Rasch* digunakan *software Winstep*.

Hasil dan Pembahasan

Validitas

Validitas yang digunakan untuk mengukur perangkat soal ASAS Pendidikan Pancasila kelas VIII yaitu validitas isi. Validitas isi digunakan untuk memastikan bahwa instrumen soal ASAS Pendidikan Pancasila benar-benar mengukur kompetensi dan pemahaman peserta didik sesuai dengan tujuan pembelajaran yang telah ditetapkan. Penelitian ini menilai validitas isi menggunakan rumus *Aiken's*. Validitas ini didasarkan pada penilaian para ahli terhadap sebuah butir untuk menentukan sejauh mana butir tersebut mempresentasikan konstruk yang diukur. Analisis validitas isi formula *Aiken's*

bertujuan menilai sejauh mana ketepatan setiap item dalam perangkat tes assessmen sumatif akhir semester (ASAS) dalam menjalankan fungsi pengukurannya. Dalam penelitian ini, validitas isi setiap item dianalisis dengan melibatkan tiga ahli. Berikut hasil validitas isi data asesmen sumatif akhir semester (ASAS) kelas VIII SMP Negeri 1 Pleret menggunakan formula Aiken's.

Gambar 1. Hasil Analisis Validitas Butir Soal Formula Aiken's

Soal	V	Ket	Soal	V	Ket
Soal 1	0,8333333	Tinggi	Soal 26	0,8333333	Tinggi
Soal 2	0,1666667	Rendah	Soal 27	0,9166667	Tinggi
Soal 3	0,75	Tinggi	Soal 28	0,4166667	Rendah
Soal 4	0,6666667	Sedang	Soal 29	0,75	Tinggi
Soal 5	0,75	Tinggi	Soal 30	0,75	Tinggi
Soal 6	1	Tinggi	Soal 31	0,9166667	Tinggi
Soal 7	0,75	Tinggi	Soal 32	0,8333333	Tinggi
Soal 8	0,6666667	Sedang	Soal 33	0,8333333	Tinggi
Soal 9	0,6666667	Sedang	Soal 34	0,8333333	Tinggi
Soal 10	0,9166667	Tinggi	Soal 35	0,6666667	Sedang
Soal 11	0,75	Tinggi	Soal 36	0,75	Tinggi
Soal 12	0,6666667	Sedang	Soal 37	0,8333333	Tinggi
Soal 13	0,8333333	Tinggi	Soal 38	0,75	Tinggi
Soal 14	0,9166667	Tinggi	Soal 39	0,8333333	Tinggi
Soal 15	0,9166667	Tinggi	Soal 40	0,9166667	Tinggi
Soal 16	0,75	Tinggi	Soal 41	0,9166667	Tinggi
Soal 17	0,9166667	Tinggi	Soal 42	0,9166667	Tinggi
Soal 18	0,8333333	Tinggi	Soal 43	0,9166667	Tinggi
Soal 19	1	Tinggi	Soal 44	0,9166667	Tinggi
Soal 20	0,9166667	Tinggi	Soal 45	0,9166667	Tinggi
Soal 21	0,75	Tinggi	Soal 46	0,8333333	Tinggi
Soal 22	0,8333333	Tinggi	Soal 47	0,9166667	Tinggi
Soal 23	0,9166667	Tinggi	Soal 48	0,9166667	Tinggi
Soal 24	0,9166667	Tinggi	Soal 49	0,9166667	Tinggi
Soal 25	0,75	Tinggi	Soal 50	0,9166667	Tinggi

Sumber: Data yang sudah diolah, 2025

Berdasarkan gambar tersebut, keputusan mengenai validitas butir soal ditentukan berdasarkan nilai validitasnya. Butir soal dianggap valid jika memiliki nilai lebih dari 0,7, cukup valid jika nilainya berada dalam rentang 0,5 hingga 0,7, dan tidak valid jika kurang dari 0,5. Hasil analisis pada soal ASAS menunjukkan bahwa 43 yaitu nomor 1, 3, 5, 6, 7, 10, 11, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 29, 30, 31, 32, 33, 34, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50 butir soal tergolong valid (tinggi), 5 butir soal yaitu nomor 4, 8, 9, 12, 35 dinyatakan cukup valid (sedang), dan 2 butir soal yaitu nomor 2 dan 28 dinyatakan tidak valid (rendah).

Uji validitas isi menggunakan formula Aiken's dalam penelitian ini, dari 50 butir soal ASAS Pendidikan Pancasila kelas VIII, sebanyak 48 butir dinyatakan valid. Suatu butir soal dianggap valid jika mampu mengukur secara tepat aspek yang seharusnya diukur sesuai dengan tujuan pembelajaran yang telah ditetapkan (Sumintono & Widhiarso, 2014, p.34). Sementara itu, 2 butir soal dianggap tidak valid karena memiliki nilai $V < 0,5$, tergolong dalam kategori rendah atau tidak valid. Ketidakvalidan kedua soal tersebut disebabkan oleh hasil penilaian dari tiga ahli, karena menurut validator butir soal tersebut tidak memenuhi kriteria yang ditetapkan, terutama dalam kesesuaian dengan indikator pencapaian kompetensi serta kesesuaian isi soal yang baik. Oleh karena itu, kedua butir tersebut memperoleh nilai rendah dan dinyatakan tidak valid dalam uji validitas isi formula Aiken's.

Dua butir soal yang tidak valid, tidak dapat dilakukan uji reliabilitas. Hal ini disebabkan oleh fakta bahwa reliabilitas hanya dapat diukur pada instrumen yang telah terbukti valid, karena pengukuran yang tidak akurat tidak akan memberikan hasil yang konsisten. Jika suatu soal tidak valid, maka pengukurannya tidak akurat, sehingga tidak dapat menghasilkan hasil yang konsisten. Dengan

demikian, hanya butir soal yang valid yang akan diuji untuk memastikan konsistensi dan keandalan hasil pengukuran.

Reliabilitas

Uji reliabilitas dalam penelitian ini dilakukan menggunakan koefisien *Cronbach's Alpha*. Menurut Sumintono dan Widhiarso (2014, p.31), reliabilitas merujuk pada sejauh mana suatu pengukuran yang dilakukan berulang kali dapat menghasilkan informasi yang konsisten. Untuk mengukur reliabilitas, penelitian ini memanfaatkan perangkat lunak SPSS. Berikut merupakan hasil uji reliabilitas soal ASAS Pendidikan Pancasila kelas VIII menggunakan SPSS.

Tabel 1. Hasil Uji Reliabilitas

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of items
.892	.886	48

Sumber: Data yang sudah diolah, 2025

Hasil dari SPSS menunjukkan bahwa soal yang diujikan reliabel, karena hasil dari pengukuran SPSS *Cronbach's Alpha* menunjukkan 0,892. Nilai *Cronbach's Alpha* yang diperoleh kemudian dibandingkan dengan nilai r product moment yang terdapat dalam tabel. Ketentuan yang digunakan adalah jika nilai r hitung lebih besar dari r tabel, maka instrumen tersebut dianggap reliabel. Dalam uji coba yang dilakukan, terdapat 48 soal yang diuji berdasarkan keterangan valid dengan $\alpha = 0,05$ dan $n = 48$, sehingga diperoleh nilai r tabel sebesar 0,284. Hasil perhitungan menunjukkan bahwa nilai *Cronbach's Alpha* adalah 0,892. Karena nilai r hitung lebih besar dari r tabel. Sedangkan jika dihitung dari perbandingan nilai *Cronbach Alpha* maka dapat disimpulkan bahwa hasil reliabilitas butir tes asesmen sumatif akhir semester (ASAS) mata pelajaran Pendidikan Pancasila dinyatakan reliabel, karena nilai *Cronbach Alpha* $>0,60$ maka dinyatakan reliabel. Sebuah soal dianggap reliabel jika pengukuran yang dilakukan secara berulang menghasilkan informasi yang konsisten (Sumintono & Widhiarso, 2014, p.31).

Hasil uji reliabilitas dengan nilai Cronbach's Alpha tersebut menunjukkan bahwa instrumen yang digunakan dalam asesmen memiliki tingkat konsistensi yang tinggi dalam mengukur kemampuan peserta didik. Artinya, soal ASAS tersebut dapat diandalkan dan memberikan hasil yang stabil serta akurat, sehingga mampu mencerminkan pemahaman peserta didik terhadap materi yang telah diajarkan. Tingkat reliabilitas yang baik berperan penting dalam evaluasi pendidikan, karena memastikan bahwa hasil tes dapat dipercaya dan dijadikan dasar dalam pengambilan keputusan terkait pembelajaran serta perkembangan peserta didik. Oleh karena itu, penggunaan butir soal ASAS dapat memberikan gambaran yang lebih akurat mengenai pencapaian kompetensi peserta didik dalam mata pelajaran Pendidikan Pancasila.

Tingkat Kesesuaian Butir Soal (Item Fit)

Uji kesesuaian butir dalam penelitian ini didasarkan pada asumsi model *Rasch*. Proses uji kesesuaian butir dengan pendekatan model *Rasch* dilakukan dengan mengevaluasi sejauh mana butir tersebut sesuai atau tidak dengan model yang digunakan. Analisis uji ini dilakukan menggunakan program *Winstep*. Kriteria agar butir dapat dianggap sesuai dengan model dalam program *Winstep* apabila memenuhi kriteria kesesuaian butir dengan memperhatikan hasil analisis dari tiga kriteria yang digunakan untuk memeriksa kesesuaian butir dengan model. Suatu butir dikatakan sesuai dengan model jika nilai Outfit MNSQ berada dalam rentang 0,5 hingga 1,5, nilai Outfit ZSTD berkisar antara -2 hingga 2, serta nilai Pt Measure Corr berada dalam kisaran 0,4 hingga 0,85 (Sumintono &

Widhiarso, 2015, p.72). Berikut adalah hasil analisis kesesuaian butir menggunakan program *Winstep* disajikan pada gambar berikut.

Gambar 2. Hasil Output Kesesuaian Butir

File Edit Format View Help

TABLE 10.1 DataTAS 20099305.TXT Feb 17 20:09 2025

INPUT: 221 Person 50 Item REPORTED: 221 Person 50 Item 2 CATS WINSTEPS 3.73

Person: REAL SEP.: 2.59 REL.: .87 ... Item: REAL SEP.: 4.54 REL.: .95

Item STATISTICS MISFIT ORDER

ENTRY	TOTAL	TOTAL	MEASURE	S.E.	INPIT	OUTFIT	PT-MEASURE	EXACT MATCH	
NUMBER	SCORE	COUNT			INPIT	OUTFIT	EXP.	QBSN	EXP.
4	212	221	-2.34	.95	1.08	4.3.74	3.21A-.02	.17	95.9 95.9
3	200	221	-1.37	.24	1.03	2.1.54	1.31B-.21	.25	90.0 90.5
6	208	221	-1.93	.80	1.12	0.1.49	1.01C-.10	.20	94.1 94.1
24	97	221	1.68	.16	1.21	2.01.47	3.91D-.34	.98	69.1 72.7
42	169	221	-.14	.17	1.10	1.21.38	1.61E-.20	.37	70.6 77.8
50	111	221	1.34	.16	1.14	2.21.36	3.01F-.38	.49	65.9 71.1
9	99	221	1.63	.16	1.22	3.01.36	3.11G-.35	.50	62.3 72.4
7	39	221	3.46	.21	1.27	2.11.32	1.31H-.32	.48	82.3 85.3
39	98	221	1.66	.16	1.19	2.71.28	2.81I-.37	.58	67.7 72.8
27	139	221	-.88	.16	1.12	1.91.21	1.61J-.37	.46	65.0 70.8
20	109	221	-.14	.17	1.13	1.51.17	.01K-.28	.37	75.9 77.8
16	128	221	1.05	.15	1.13	2.01.14	1.21L-.39	.47	67.3 70.7
5	196	221	-1.15	.22	1.12	.81.07	.31M-.20	.27	87.7 88.7
86	137	221	.71	.16	1.11	1.81.12	.91N-.37	.45	69.1 71.2
44	121	221	.86	.16	1.09	1.51.01	.11O-.44	.46	62.7 70.6
46	121	221	1.10	.15	1.01	1.11.00	.01P-.47	.48	70.0 70.7
12	212	221	-2.34	.95	1.08	1.11.08	.31Q-.36	.17	95.9 95.9
41	136	221	.76	.16	1.08	1.31.06	.51R-.40	.45	68.2 71.1
10	187	221	-.75	.20	1.93	-.61.07	.51S-.36	.31	85.5 85.0
11	168	221	.17	.16	1.07	1.01.06	-.21T-.37	.40	70.0 74.4
23	184	221	-.64	.19	1.06	-.61.01	-.21U-.38	.32	83.6 83.7
40	164	221	.91	.17	1.09	-.41.08	-.11V-.37	.39	77.3 76.2
42	168	221	-.11	.17	1.03	-.41.09	-.51W-.37	.38	75.5 77.5
13	139	221	.66	.16	1.00	.11.88	-.81X-.46	.44	70.9 71.2
30	122	221	1.08	.15	.99	-.11.93	-.61Y-.48	.47	71.4 70.7
33	209	221	-2.02	.31	.99	-.11.83	-.21X-.20	.20	94.5 94.6
28	128	221	1.05	.15	.98	-.11.99	-.11W-.48	.47	69.1 70.7
37	158	221	.17	.16	.91	-.11.91	-.91V-.45	.40	80.0 74.4
38	207	221	-1.85	.29	.92	-.11.98	1.11U-.20	.21	93.6 93.6
47	203	221	-1.56	.26	.97	-.11.82	-.31T-.26	.24	91.8 91.8
48	147	221	.46	.16	.97	-.51.87	-.81S-.46	.43	72.2 72.1
21	176	221	-.36	.10	.96	-.41.85	-.51P-.38	.35	80.9 80.4
15	99	221	1.63	.16	.93	-.51.95	-.51Q-.54	.50	74.1 72.4
14	144	221	.54	.16	.93	-.41.90	-.41P-.49	.43	71.8 71.7
29	192	221	-.97	.21	.92	-.51.78	-.61O-.34	.29	88.6 87.1
25	173	221	-.20	.18	.92	-.61.84	-.61N-.42	.37	81.4 78.5
18	140	221	.64	.16	.92	-.11.79	-.11M-.51	.44	72.7 71.3
45	114	221	1.27	.15	.90	-.11.80	-.11L-.55	.49	75.5 70.9
6	205	221	-1.69	.27	.90	-.41.45	-.11K-.33	.23	92.7 92.7
35	215	221	-2.78	.42	.88	-.21.42	-.11J-.24	.14	97.3 97.3
22	150	221	.39	.16	.87	-.21.04	-.11I-.51	.42	78.2 72.6
49	143	221	.56	.16	.87	-.21.77	-.11H-.53	.44	76.8 71.6
19	188	221	-.79	.20	.86	-.11.67	-.11G-.42	.31	85.9 85.4
34	203	221	-1.56	.26	.86	-.71.54	-.11F-.35	.24	91.8 91.8
1	131	221	.86	.16	.85	-.21.78	-.11E-.56	.46	78.2 70.9
26	212	221	-2.34	.95	.85	-.41.55	-.81D-.28	.17	95.9 95.9
32	150	221	.39	.16	.82	-.21.67	-.21C-.55	.42	79.1 72.6
17	126	221	.98	.15	.80	-.31.74	-.21B-.60	.47	82.3 70.8
31	124	221	1.03	.15	.74	-.41.69	-.21A-.64	.47	89.2 70.7
MEAN	157.2	221.0	-.12	.23	1.00	.01.02	.01		79.4 79.2
S.D.	40.4	.0	1.95	.24	.12	1.61.47	1.51		10.8 9.3

Sumber: Data yang sudah diolah, 2025

Analisis tingkat kesesuaian butir soal (item fit) ASAS Pendidikan Pancasila kelas VIII di SMP Negeri 1 Pleret menggunakan *software Winstep* menunjukkan bahwa dari 50 butir soal yang dianalisis, sebanyak 41 butir (82%) dinyatakan sesuai atau *fit*. Kesimpulan ini diambil karena 41 butir tersebut telah memenuhi minimal dua dari tiga kriteria yang ditetapkan. Sementara itu, terdapat 9 butir soal yang tidak sesuai dengan model (*misfit*) karena tidak memenuhi setidaknya dua kriteria yang telah ditetapkan. Dengan demikian, butir-butir soal tersebut perlu diperbaiki agar lebih sesuai dengan standar yang diharapkan.

Soal yang paling atas yaitu soal nomor 4, tidak *fit*. Jika dilihat dari tiga kriteria, butir soal nomor 4 tidak memenuhi ketiga kriteria kesesuaian butir soal. Sedangkan soal nomor 2 tidak memenuhi kriteria dan tidak memiliki nilai pengukuran, karena butir soal tersebut tidak memenuhi kriteria (*maximum measure*) dan berada di luar batas yang ditentukan (*outlier*). Hal ini menunjukkan bahwa soal tersebut perlu disusun ulang dengan mengganti dengan soal yang baru. Dengan demikian, soal nomor 2 dan 4 harus dibuang. Untuk butir soal lainnya yaitu soal nomor 3, 6, 9, 24, 35, 39, 50 hanya tidak memenuhi dua dari tiga kriteria sehingga soal-soal tersebut perlu dilakukan perbaikan. Dalam penyusunan soal yang telah diujikan pada peserta didik, keberadaan soal yang tidak sesuai menunjukkan adanya kesalahan dalam soal tersebut. Oleh karena itu, guru dapat merevisi soal yang telah disusun serta meningkatkan metode pembelajaran yang diterapkan (Sumintono & Widhiarso, 2015, p.72).

Dalam penyusunan soal yang telah diujikan pada peserta didik, keberadaan soal yang tidak sesuai menunjukkan adanya kesalahan dalam soal tersebut. Oleh karena itu, guru dapat merevisi soal yang telah disusun serta meningkatkan metode pembelajaran yang diterapkan (Sumintono & Widhiarso, 2015, p.72). Perbaikan terhadap butir-butir yang tidak *fit* ini sangat diperlukan untuk meningkatkan efektivitas dan akurasi asesmen. Proses perbaikan ini tidak hanya akan memperbaiki kualitas soal, tetapi juga akan berkontribusi pada peningkatan hasil belajar peserta didik. Dengan demikian, langkah selanjutnya adalah melakukan analisis mendalam terhadap butir-soal yang tidak *fit* untuk mengidentifikasi penyebab ketidaksesuaian dan merumuskan strategi perbaikan yang tepat.

Tingkat Kesulitan Butir Soal

Setelah menganalisis kesesuaian butir atau *model fit*, langkah berikutnya adalah mengevaluasi karakteristik butir dalam instrumen yang telah dikembangkan. Pada tahap ini, karakteristik butir soal dianalisis berdasarkan tingkat kesulitannya, karena penelitian ini menggunakan Pemodelan *Rasch* dalam pengembangannya. Berdasarkan *Rasch* model, Item dikatakan baik jika indeks kesulitan lebih dari -2,0 atau kurang dari 2,0 (Hambleton et al., 1991). Analisis karakteristik butir soal pada tahap ini menggunakan program *Winstep*, dimana tingkat kesulitan dilihat pada kolom *measure*. Berikut adalah hasil analisis karakteristik butir soal pada masing-masing instrument. Butir yang dianalisis pada instrumen Penilaian sebanyak 50 butir. Adapun output hasil analisis karakteristik butir dapat dilihat pada gambar berikut.

Gambar 3. Hasil Output Tingkat Kesulitan

File Edit Format View Help 200772045.TXT Feb 16 12:31 2025
 INPUT: 221 Person 88 Item REPORTED: 221 Person 88 Item 2 CATS WINSTEPS 2.73
 REPORT: REAL SEP.: 2.99 REL.: .87 ... Item: REAL SEP.: 4.94 REL.: .95

Item STATISTICS: MEASURE ORDER

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	MEASURE	MODEL S.E.	INFIT RMSQ	OUTFIT RMSQ	OUTFIT CORR.	PT-MEASURE	EXACT MATCH	EXP. OBS.	EXP. EXP.	Item
7	39	221	3.46	.23	1.37	2.11	1.32	1.3	.32	.48	85.3	87
24	87	221	1.68	.16	1.21	2.81	1.47	3.0	.34	.50	69.1	72.7
39	98	221	1.60	.16	1.19	2.71	1.28	2.8	.37	.50	67.7	72.6
9	99	221	1.63	.16	1.22	3.01	1.36	3.1	.35	.50	62.3	72.4
15	99	221	1.63	.16	.99	.0	.98	.5	.34	.50	74.1	72.4
68	111	221	1.34	.16	1.14	2.21	1.26	3.0	.38	.49	65.9	71.1
48	114	221	1.27	.15	.98	-1.8	.88	-1.1	.38	.49	76.6	70.9
46	123	221	1.10	.15	1.01	1.1	.98	.6	.47	.48	70.6	70.7
38	122	221	.99	.15	.99	.1	.99	.6	.47	.71	4	70.7
18	123	221	1.05	.15	1.13	2.0	1.14	1.2	.39	.47	67.3	70.7
29	123	221	1.05	.15	.98	-.2	.99	-.1	.48	.47	69.1	70.7
11	124	221	1.03	.15	.74	-.6	.99	-.9	.64	.47	83.2	70.7
17	126	221	.98	.15	.80	-3.6	.74	-2.3	.60	.47	82.3	70.8
27	138	221	.88	.16	1.12	1.9	1.22	1.6	.37	.46	88.0	70.8
1	131	221	.86	.16	.85	2.1	.76	1.8	.40	.46	69.2	70.9
44	131	221	.86	.16	1.09	1.9	1.01	1.1	.41	.46	62.7	70.9
41	135	221	.76	.16	1.08	1.3	1.00	.5	.40	.45	88.2	71.1
36	137	221	.71	.16	1.11	1.0	1.12	.9	.37	.45	89.1	71.2
33	139	221	.66	.16	1.00	.1	.88	-.8	.40	.44	70.9	71.2
10	140	221	.64	.16	.92	-1.4	.79	-1.5	.51	.44	72.7	71.3
49	143	221	.58	.16	.87	-2.2	.77	-3.7	.44	.44	76.8	71.8
14	144	221	.54	.16	.93	1.2	.80	1.4	.49	.43	71.0	71.7
48	147	221	.46	.16	.97	-.8	.87	-.8	.48	.43	78.2	72.1

Data Tingkat Kesulitan - Notepad

File Edit Format View Help

48	147	221	.46	.16	.97	-.8	.87	-.8	.48	.43	78.2	72.1
22	150	221	.36	.16	.87	-2.0	.64	-1.0	.51	.42	78.2	72.6
13	158	221	.17	.16	1.07	1.0	.98	.2	.37	.40	70.9	74.4
37	158	221	.17	.16	.91	-1.2	.98	.0	.45	.40	88.0	74.4
60	164	221	.03	.17	1.03	.0	.88	-.1	.37	.39	77.1	74.2
43	160	221	-.11	.17	1.03	.4	.89	-.5	.37	.38	75.5	77.5
26	169	221	-.14	.17	1.13	1.5	1.17	.8	.28	.37	76.9	77.8
42	189	221	-.14	.17	1.10	1.1	.88	.6	.28	.37	76.4	77.8
25	172	221	-.20	.16	.92	-.9	.84	-.6	.42	.37	81.4	78.5
21	176	221	-.36	.16	.96	-.4	.88	-.5	.38	.35	88.9	80.4
33	184	221	-.64	.19	1.06	.6	.91	.2	.30	.33	83.6	83.7
10	187	221	-.75	.20	.93	-.0	1.07	.2	.30	.31	89.0	85.0
19	188	221	-.79	.20	.86	-1.1	.97	-1.5	.42	.31	85.9	85.4
29	192	221	-.97	.21	.92	-.5	.78	-.6	.34	.29	88.0	87.1
6	196	221	-1.15	.22	1.12	.8	1.07	.3	.28	.27	87.7	88.7
3	200	221	-1.37	.24	1.03	2.1	.94	1.3	.21	.25	90.9	90.5
34	202	221	-1.56	.26	.86	-.7	.84	-1.2	.35	.28	91.9	92.8
47	203	221	-1.56	.26	.97	-.1	.82	-.8	.26	.24	91.8	93.8
6	205	221	-1.69	.27	.90	-.4	.45	-1.4	.53	.23	92.7	92.7
38	207	221	-1.85	.28	.92	.9	.88	.1	.26	.21	92.6	93.6
8	208	221	-1.93	.30	1.12	.6	1.49	1.0	.10	.20	94.1	94.1
33	209	221	-2.02	.31	.99	.1	.83	-.2	.20	.20	94.5	94.6
4	212	221	-2.34	.36	1.08	-0.8	.74	8.2	-.02	.17	95.9	95.9
12	213	221	-2.34	.35	.99	1.1	.88	.3	.16	.17	95.9	95.9
28	212	221	-2.34	.35	.88	-.4	.85	-.8	.28	.17	88.9	85.9
15	215	221	-2.76	.42	.88	-.2	.42	1.1	.24	.14	97.3	97.3
2	221	221	-5.84	1.03					.00	.00	100.0	100.0

MEASURE ORDER

MEAN: 157.2 221.0 -1.12 -231.00 -011.02 .01 79.4 79.2
 S.D.: 40.4 .0 1.95 .28 .12 1.6 .47 1.5 10.0 9.2

Sumber: Data yang sudah diolah, 2025

Output tingkat kesulitan butir pada gambar 2 sudah berurutan, mulai dari tingkat kesulitan tertinggi sampai terendah. Tingkat kesulitan butir dilihat berdasarkan hasil pada kolom *measure*. Berdasarkan gambar 2 diatas menunjukan bahwa butir dengan tingkat kesulitan tertinggi adalah butir ke 7 yaitu dengan indeks tingkat kesulitan sebesar 3,46, sedangkan untuk tingkat kesulitan terendah adalah butir ke 2 yaitu dengan indeks tingkat kesulitan sebesar -5,84. Butir soal nomor 7 adalah butir yang sangat sulit dikerjakan oleh peserta didik, dibuktikan jumlah peserta didik yang menjawab yaitu dari 221 peserta didik hanya 39 peserta didik yang mampu menjawab butir nomor 7. Sedangkan butir

soal nomor 2 adalah butir soal yang sangat mudah dikerjakan oleh peserta didik, dibuktikan dari 221 peserta didik yang menjawab benar yaitu 221 peserta didik.

Nilai *logit* yang sama pada setiap butir menunjukkan tingkat kesulitan yang tidak jauh berbeda, contohnya pada butir 39 dan 9. Semakin besar nilai *logit*, maka tingkat kesulitan soal semakin tinggi, sedangkan semakin kecil nilai *logit*, maka soal tersebut menjadi semakin mudah. Arikunto (2018, p.232), sebuah soal dikategorikan baik jika tidak terlalu mudah maupun terlalu sulit. Soal yang dianggap efektif memiliki tingkat kesulitan sedang, karena kategori ini dinilai ideal dalam mengukur kemampuan peserta tes secara optimal. Soal dengan tingkat kesulitan yang terlalu tinggi cenderung memberikan hasil yang kurang optimal, karena dapat menyebabkan peserta didik kehilangan motivasi atau merasa putus asa untuk mencoba menyelesaikannya kembali. Di sisi lain, soal yang terlalu mudah juga dianggap kurang baik, karena tidak mampu mengukur pemahaman peserta didik dengan tepat.

Dapat diketahui analisis tingkat kesulitan melalui program *Winstep*, soal ASAS mata pelajaran Pendidikan Pancasila kelas VIII memiliki kesulitan sangat sulit 1(2%) butir soal, sulit 11 (22%) butir soal, sedang 25 (50%) butir soal, mudah 7 (14%) butir soal, sangat mudah 6 (12%) butir soal. Butir soal yang dianggap terlalu sulit dan terlalu mudah harus dibuang atau dilakukan revisi. Hasil temuan penelitian ini menunjukkan bahwa terdapat 7 butir soal yang kurang baik yaitu soal nomor 2, 4, 7, 12, 28, 33, dan 35 perlu direvisi atau dibuang karena nilai tingkat kesulitan tidak berada pada interval $-2 < b < 2$. Pada butir soal nomor 2, 4, 12, 28, 33, dan 35 terlihat nilai tingkat kesulitan yang cukup rendah karena nilai *logit* kurang dari $-0,2$ yang menunjukkan bahwa butir-butir soal terlalu mudah, sedangkan pada soal nomor 7 terdapat nilai tingkat kesulitan yang cukup tinggi karena nilai *logit* $< 2,0$ yang menunjukkan soal terlalu sulit. Menurut Hambleton & Swaminathan sebagaimana dikutip Jannah, (2024, p.336), butir soal dianggap memiliki kategori tingkat kesulitan baik jika tingkat kesulitan (nilai *logit*) berada di antara rentang $-2,0$ sampai dengan 2 . Soal dalam rentang tersebut dianggap ideal karena tingkat kesulitannya seimbang, tidak terlalu mudah maupun terlalu sulit, sehingga mampu mengukur kemampuan peserta tes dengan optimal. Dengan demikian, jika terdapat soal yang memiliki tingkat kesulitan sangat sulit atau sangat mudah harus dilakukan perbaikan. Adanya perbaikan pada butir soal dilakukan untuk menjamin efektivitas evaluasi pembelajaran.

Adanya persebaran indeks tingkat kesulitan dari hasil analisis menunjukkan bahwa instrumen penelitian memiliki variasi yang cukup baik, dengan sebagian besar soal berada pada tingkat kesulitan yang sesuai. Hasil analisis dalam penelitian ini sejalan dengan pendapat Arikunto (2018, p.232), yang menyatakan soal yang baik harus memiliki tingkat kesulitan seimbang, sehingga tidak terlalu mudah ataupun terlalu sulit, agar dapat secara optimal mengukur kemampuan peserta didik. Meskipun terdapat sedikit perbedaan dalam proporsi hasil analisis, secara keseluruhan soal ASAS Pendidikan Pancasila tahun ajaran 2024/2025 memiliki tingkat kesulitan yang memadai. Hal ini menunjukkan bahwa instrumen tersebut mampu menyajikan informasi yang tepat mengenai kemampuan peserta tes. Hal tersebut menunjukkan bahwa instrumen tersebut mampu menyajikan informasi yang tepat mengenai kemampuan peserta tes. Berdasarkan hasil distribusi tingkat kesulitan, soal ASAS menunjukkan keseimbangan yang baik. Hasil analisis menunjukkan bahwa proporsi peserta didik yang menjawab benar dan salah mencerminkan soal dengan tingkat kesulitan sedang dan sulit dibandingkan soal yang tergolong mudah. Dengan demikian, disimpulkan bahwa variasi dalam tingkat kesulitan soal ASAS berperan positif dalam proses evaluasi, karena memberikan tantangan yang sesuai bagi peserta didik.

Simpulan

Dapat ditarik kesimpulan bahwa soal ASAS memiliki kualitas baik, Hal ini terlihat dari validitas yang memadai, reliabilitas yang tinggi, tingkat kesesuaian butir memenuhi kriteria fit, serta distribusi

tingkat kesulitan yang seimbang. Selain memenuhi standar pengukuran yang ditetapkan, soal ini juga mampu menggambarkan kemampuan peserta didik secara akurat. Sehingga, instrumen ini dapat dimanfaatkan secara optimal dalam proses evaluasi untuk menilai pencapaian belajar secara objektif.

Dari hasil uraian diatas dapat disimpulkan bahwa soal assesmen sumatif akhir semester (ASAS) mata pelajaran Pendidikan Pancasila memiliki kualitas butir soal yang *fit* karena telah memenuhi kriteria minimal satu dari ketiga kriteria yang telah ditentukan. Di sisi lain pada tingkat kesulitan soal, masih terdapat tujuh butir soal yang perlu mendapat perhatian lebih oleh guru karena tergolong terlalu mudah dan terlalu sulit. Dapat disimpulkan bahwa butir soal pilihan ganda ASAS Pendidikan Pancasila kelas VIII memiliki tingkat kesulitan yaitu sebagai berikut, terdapat 1 (2%) soal tergolong sangat sulit, 11 (22%) soal tergolong sulit, 25 (50%) soal tergolong sedang, 7 (14%) tergolong mudah, dan 6 (12%) soal tergolong sangat mudah. Adanya persebaran indeks tingkat kesulitan dari hasil analisis menunjukkan bahwa instrumen penelitian memiliki variasi cukup baik, dengan sebagian besar soal berada pada tingkat kesulitan yang sesuai. Secara keseluruhan, soal ASAS mata pelajaran Pendidikan Pancasila tahun ajaran 2024/2025 memiliki tingkat kesulitan yang proporsional. Soal yang memenuhi standar kesulitan yang proporsional ini layak disimpan dalam bank soal agar dapat digunakan kembali dalam penilaian pembelajaran di semester yang akan datang.

Berdasarkan beberapa saran dan kesimpulan yang ada diajukan, diantaranya: (1) Bagi guru Pendidikan Pancasila, hasil analisis butir soal ASAS ini dapat dijadikan bank soal. Selain itu, Jika ditemui soal yang terlalu mudah atau terlalu sulit, guru disarankan untuk melakukan perbaikan untuk menyesuaikan tingkat kesulitan tersebut. (2) Bagi kepala sekolah, pentingnya penilaian yang baik menjadikan perlu adanya pelatihan analisis butir soal bagi guru. Langkah ini bertujuan untuk memberikan guru keterampilan dalam menganalisis perangkat tes secara empiris, sehingga dapat diterapkan dalam proses pembelajaran pada mata pelajaran yang diampu.

Referensi

- Abdullah, N., Jahja, M., & Setiawan, D. G. E. (2022). Analisis kualitas butir soal pada mata pelajaran fisika di jurusan fisika fakultas MIPA Universitas Negeri Gorontalo tahun ajaran 2021/2022. *Jurnal Sains Dan Pendidikan Fisika*, 18(1), 44–52. <https://doi.org/10.35580/jspf.v18i1.32358>
- Aeni, N., & Suyato, S. (2020). Analisis butir soal mata pelajaran PPKN kelas IX SMP 1 Karangjambu Purbalingga Tahun 2019/2020. *Agora*, 9(5), 490–501. <https://journal.student.uny.ac.id/index.php/civics/article/view/17284>.
- Afriyanti, A. (2024). Rasch Model dalam Analisis Persepsi Mahasiswa Agroteknologi Terhadap Mata Kuliah Statistika. *Jurnal Ilmiah Profesi Pendidikan*, 9(2), 1009–1015. <https://doi.org/10.29303/jipp.v9i2.2086>
- Anggarani, N., & Hidayat, R. A. (2022). Analisis Kesalahan Siswa dalam Mengerjakan Soal pada Penjumlahan dan Pengurangan Bilangan Bulat Berbasis HOTS pada Siswa Kelas V. *Indonesian Journal of Islamic Elementary*, 2(1), 49-58. <https://doi.org/https://doi.org/10.28918/ijjee.v2i1.5154> Education
- Arifin, Z. (2017). *Evaluasi pembelajaran: prinsip, teknik, dan prosedur*. PT Remaja Rosdakarya.
- Arikunto, S. (2018). *Dasar-dasar evaluasi pendidikan edisi 3*. Bumi Aksara.
- Asrul, A., Saragih, A. H. & Mukhtar, M. (2022). *Evaluasi pembelajaran*. Perdana Publishing.
- Ayub, M. R. S. S. N., Istiyono, E., Munadi, S., Permadi, C., Pattiserlihun, A., & Sudjito, D. N. (2020). Analisa penilaian soal fisika menggunakan model rasch dengan program R. *Jurnal Sains Dan Edukasi Sains*, 3(2), 46-52. <https://doi.org/10.24246/juses.v3i2p46-52>
- Aqida, D. S. (2024). Penggunaan Rasch Model dalam Menganalisis butir soal matematika kelas V Sekolah Dasar. 4, 91-101. <https://doi.org/https://doi.org/10.28918/ijjee.v4i1.6947>

- DeMars, C. (2018). *Classical test theory and item response theory*. The Wiley Handbook of Psychometric Testing.
- Farida & Musyarofah, A. (2021). Validitas dan reliabilitas dalam analisis butir soal. *Al-Muarrib Journal of Arabic Education*, 1(1), 34–44. <https://doi.org/https://doi.org/10.32923/al-muarrib.v1i1.2100>
- Hambelton, R. K., & Swaminathan, H., & Rogers, D. J. (1991). *Fundamentals of item respon theory*. Sage Publications. <https://doi.org/10.5860/choice.29-4185>
- Hanna, W. F. & Retnawati, H. (2022). Analisis kualitas butir soal matematika menggunakan model rasch dengan bantuan software quest. *AKSIOMA: Jurnal Program Studi Pendidikan Matematika*, 11(4), 3695. <https://doi.org/10.24127/ajpm.v11i4.5908>
- Harmurni, L. (2019). *Instrumen penilaian & validasinya*. Uwais Inspirasi Indonesia.
- Ida, F. F., & Musyarofah, A. (2021). Validitas dan Reliabilitas dalam Analisis Butir Soal. *Al Mu'arrib: Jurnal Pendidikan Bahasa Arab*, 1(1), 34-44. <https://doi.org/https://doi.org/10.32923/al-muarrib.v1i1.2100>
- Ismail, M. I. (2020). *Asesmen dan evaluasi pembelajaran*. Cendikia Publisher.
- Jannah, R. (2024). Analisis kualitas butir soal dan abilitas siswa pada asesmen Madrasah Ibtidaiyah dengan Rasch Model. *Jurnal Didaktika Pendidikan Dasar*, 8(1), 331–342. <https://doi.org/10.26811/didaktika.v8i1.1234>
- Jumini, S., Madnasri, S., Cahyono, E., & Parmin, P. (2023). Analisis kualitas butir soal pengukuran literasi sains melalui teori tes klasik dan rasch model. In *Prosiding Seminar Nasional Pascasarjana*, 6(1), 758-765. <https://proceeding.unnes.ac.id/snpasca/article/view/2215>
- Nudin, I., & Hidayatullah, R. S. (2023). Analisis butir soal penilaian tengah semester menggunakan model Rasch di SMK Negeri 5 Surabaya. *Jptm*, 12(2), 132–139. <https://ejournal.unesa.ac.id/index.php/jurnal-pendidikan-teknik-mesin/article/view/52258>
- Prijowuntato, S. W. (2020). *Evaluasi pembelajaran*. Sanata Dharma Universitas Press.
- Pudjiati, I & Madani, F. (2023). Asesmen Autentik Analisis Butir Soal Dengan Rasch Model Di Sekolah Dasar: Literature Review. *Mutiara: Jurnal Penelitian Dan Karya Ilmiah*, 1(4), 01–12. <https://doi.org/10.59059/mutiara.v1i4.325>
- Purnama, D. N., & Alfarisa, F. (2020). Karakteristik Butir Soal Try Out Teori Kejuruan Akuntansi Smk Berdasarkan Teori Tes Klasik Dan Teori Respons Butir. *Jurnal Pendidikan Akuntansi Indonesia*, 18(1), 36–46. <https://doi.org/10.21831/jpai.v18i1.31457>
- Retnawati, H. (2014). *Teori respon butir dan penerapannya untuk peneliti, praktisi pengukuran dan pengujian, mahasiswa pascasarjana*. Nuha Medika.
- Sugiyono. (2019). *Metodelogi penelitian kuantitatif dan kualitatif Dan R&D*. Alfabeta.
- Sumintono, B. & Widhiarso, W. (2014). *Aplikasi model rasch untuk penelitian ilmu-ilmu sosial*. Trim Komunikata.
- Sumintono, B. dan Widhiarso. W. (2015). *Aplikasi pemodelan rasch pada assessment pendidikan*. Trim Komunikata.
- Tarigan, E. F., Nilmarito, S., Islamiyah, K., Darmana, A., & Suyanti, R. D. (2022). Analisis instrumen tes menggunakan Rasch Model dan software SPSS 22.0. *Jurnal Inovasi Pendidikan Kimia*, 16(2), 92–96. <https://doi.org/10.15294/jipk.v16i2.30530>

Utomo, B. (2022). Analisis validitas isi butir soal sebagai salah satu upaya peningkatan kualitas pembelajaran di madrasah berbasis nilai-nilai islam. *Jurnal Pendidikan Matematika (Kudus)*, 1(2). <https://doi.org/http://dx.doi.org/10.21043/jmtk.v1i2.4868>