

Analisis Kualitas Empiris Butir Soalujian Pendidikan Agama Kristen SD Inpres Palsatu Berbasis *Outcome Based Education*

Merti Tana¹, Sefanya Merah Loinati²

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

*Email: mertitana97@gmail.com, Sefanyamloinati@gmail.com

Diterima: 15-06-2026 | Disetujui: 20-06-2026 | Diterbitkan: 22-06-2026

ABSTRACT

This study aims to analyze the empirical quality of Christian Religious Education (CRE) exam items at SD Inpres Palsatu based on Outcome-Based Education through difficulty level and discrimination index. This research used a descriptive quantitative approach with 27 student answer sheets as subjects and 20 multiple-choice items as objects. Data analysis was performed using Microsoft Excel 2021. The results showed that all items had a moderate difficulty level ($P=0.37-0.67$). For discrimination indices, 85% of the items were categorized as good and very good ($D \geq 0.40$), while 15% were categorized as sufficient ($D=0.20-0.39$) and need revision. Overall, 17 items (85%) were declared feasible to be accepted without revision, and 3 items need to be revised. It is concluded that the developed test instrument has good quality and can be used to measure learning outcomes in CRE. Future research is recommended to complete the analysis of distractor effectiveness.

Keywords: item analysis, difficulty level, discrimination index, obe, christian religious education

ABSTRAK

Penelitian ini bertujuan menganalisis kualitas empiris butir soal Ujian Pendidikan Agama Kristen SD Inpres Palsatu, berbasis Outcome-Based Education melalui tingkat kesukaran dan daya pembeda. Penelitian menggunakan pendekatan kuantitatif deskriptif dengan subjek 27 lembar jawaban siswa dan objek 20 butir soal pilihan ganda. Analisis data dilakukan dengan bantuan Microsoft Excel 2021. Hasil penelitian menunjukkan seluruh butir soal memiliki tingkat kesukaran kategori sedang ($P=0.37-0.67$). Untuk daya pembeda, 85% butir soal berkategori baik dan baik sekali ($D \geq 0.40$), sedangkan 15% berkategori cukup ($D=0.20-0.39$) dan perlu direvisi. Secara keseluruhan, 17 butir soal (85%) dinyatakan layak diterima tanpa revisi, dan 3 butir soal perlu direvisi. Disimpulkan bahwa instrumen tes yang dikembangkan sudah berkualitas baik dan dapat digunakan untuk mengukur capaian pembelajaran PAK. Penelitian selanjutnya disarankan melengkapi analisis efektivitas pengecoh.

Katakunci: analisis butir soal, tingkat kesukaran, daya pembeda, obe, PAK

Bagaimana Cara Sitasi Artikel ini:

Tana, M., & Loinati, S. M. . (2026). Analisis Kualitas Empiris Butir Soalujian Pendidikan Agama Kristen SD Inpres Palsatu Berbasis Outcome Based Education. Educational Journal, 1(4), 2102-2108. <https://doi.org/10.63822/hjtq5275>

PENDAHULUAN

Kurikulum Merdeka yang diterapkan saat ini menekankan pendekatan Outcome Based Education (OBE), di mana proses pembelajaran dan evaluasi diarahkan untuk mencapai capaian pembelajaran yang terukur. Dalam mata pelajaran Pendidikan Agama Kristen (PAK), capaian tersebut tidak hanya menyentuh aspek kognitif, tetapi juga afektif dan psikomotorik yang bermuara pada pertumbuhan iman serta karakter Kristiani siswa. Oleh karena itu, instrumen evaluasi yang digunakan harus mampu mengukur kompetensi siswa secara valid dan reliabel.

Salah satu bentuk evaluasi yang umum digunakan adalah tes tertulis dalam bentuk Ujian Tengah Semester (UTS). Kualitas tes sangat ditentukan oleh mutu butir soal yang digunakan. (Arikunto, 2018) menegaskan bahwa “analisis butir soal merupakan suatu prosedur sistematis yang akan memberikan informasi mengenai tingkat kesukaran, daya pembeda, dan efektivitas pengecoh pada setiap butir soal” (hlm. 219). Tanpa analisis empiris, guru tidak dapat memastikan apakah soal yang diberikan benar-benar mengukur kemampuan siswa atau hanya sekadar formalitas.

Selain itu, daya pembeda soal penting untuk mengidentifikasi siswa yang telah menginternalisasi nilai ajaran Kristen dengan siswa yang belum. Sementara itu, pengecoh atau distractor pada soal pilihan ganda harus berfungsi efektif. (Utami & Nurgiyantoro, 2016) menjelaskan bahwa pengecoh dikatakan efektif apabila “dipilih sekurang-kurangnya oleh 5% pengikut tes dan lebih banyak pemilihnya berasal dari kelompok bawah” (hlm. 205). Pengecoh yang tidak berfungsi menunjukkan bahwa alternatif jawaban tersebut tidak homogen dengan kunci, sehingga perlu direvisi.

Berdasarkan hasil observasi awal di SD Inpres Palsatu, ditemukan bahwa guru PAK belum rutin melakukan analisis empiris terhadap butir soal UTS yang digunakan. Akibatnya, belum diketahui secara pasti kualitas soal ditinjau dari tingkat kesukaran, daya pembeda, dan efektivitas pengecoh. Kondisi ini berpotensi menyebabkan hasil evaluasi tidak mencerminkan capaian pembelajaran OBE secara akurat.

Bertolak dari uraian tersebut, penelitian ini bertujuan untuk menganalisis kualitas empiris butir soal Ujian Tengah Semester mata pelajaran Pendidikan Agama Kristen kelas VI di SD Inpres Palsatu ditinjau dari tingkat kesukaran, daya pembeda, dan efektivitas pengecoh. Hasil penelitian ini diharapkan dapat menjadi masukan bagi guru PAK dalam mengembangkan bank soal yang valid, reliabel, dan selaras dengan prinsip OBE.

METODE PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif dengan metode deskriptif. Subjek dalam penelitian ini adalah seluruh lembar jawaban siswa SD Inpres Palsatu yang mengikuti tes PAK, yaitu sebanyak 27 siswa. Objek penelitian adalah 20 butir soal pilihan ganda yang disusun oleh Kelompok 12.

Daya pembeda dihitung dengan mengambil 27% kelompok atas dan 27% kelompok bawah. Dari 27 siswa, maka $27\% \times 27 = 7.29 = 7$ siswa untuk masing-masing kelompok. Analisis tingkat kesukaran dan daya pembeda dilakukan dengan bantuan (Nurgiyantoro, 2010).

HASIL PENELITIAN

Hasil P & D Untuk Soal 1-20:

Tabel 1. Rekapitulasi Hasil Analisis Tingkat Kesukaran dan Daya Pembeda Butir Soal PAK Kelas

VI SD Inpres Palsatu

No Soal	B	P	Kategori P	BA	BB	D	Kategori D
1	13	0.48	Sedang	5	1	0.57	Baik
2	12	0.44	Sedang	6	1	0.71	Baik Sekali
3	15	0.56	Sedang	6	3	0.43	Baik
4	14	0.52	Sedang	7	0	1.00	Baik Sekali
5	15	0.56	Sedang	6	2	0.57	Baik
6	13	0.48	Sedang	6	2	0.57	Baik
7	15	0.56	Sedang	6	2	0.57	Baik
8	13	0.48	Sedang	7	1	0.86	Baik Sekali
9	13	0.48	Sedang	5	2	0.43	Baik
10	18	0.67	Sedang	6	4	0.29	Cukup
11	13	0.48	Sedang	5	2	0.43	Baik
12	12	0.44	Sedang	5	2	0.43	Baik
13	12	0.44	Sedang	6	2	0.57	Baik
14	17	0.63	Sedang	6	2	0.57	Baik
15	10	0.37	Sedang	6	1	0.71	Baik Sekali
16	17	0.63	Sedang	5	3	0.29	Cukup
17	16	0.59	Sedang	6	2	0.57	Baik
18	17	0.63	Sedang	5	3	0.29	Cukup
19	10	0.37	Sedang	7	0	1.00	Baik Sekali
20	12	0.44	Sedang	4	1	0.43	Baik

Gambaran Umum Pelaksanaan Penelitian

Penelitian ini merupakan penelitian lapangan dengan data primer. Peneliti melakukan pengambilan data langsung di SD Inpres Palsatu, Kelurahan Manutapen, Kota Kupang pada bulan

Analisis Kualitas Empiris Butir Soalujian Pendidikan Agama Kristen SD Inpres Palsatu berbasis Outcome Based Education!
(Tana, et al.)

November 2026. Data yang dianalisis berupa 27 lembar jawaban siswa kelas VI pada Ujian Akhir Semester mata pelajaran Pendidikan Agama Kristen. Analisis tingkat kesukaran dan daya pembeda dilakukan oleh peneliti sendiri menggunakan Microsoft Excel 2021. Langkah ini sesuai dengan pernyataan (Arikunto, 2018) bahwa analisis butir soal secara empiris wajib dilakukan guru untuk menjamin kualitas instrumen evaluasi. Metode penelitian kuantitatif lapangan seperti ini sesuai dengan (Sudijono, 2015) yang menekankan pentingnya analisis data empiris langsung di lapangan.

Analisis Tingkat Kesukaran Butir Soal

Hasil perhitungan menunjukkan seluruh 20 butir soal memiliki indeks kesukaran $P = 0.37-0.67$ yang masuk kategori sedang. Sesuai teori (Sudijono, 2015), soal kategori sedang adalah soal ideal karena tidak terlalu mudah sehingga semua siswa bisa jawab, dan tidak terlalu sukar sehingga siswa tidak frustrasi.

Temuan ini memperkuat penelitian (Utami & Nurgiyantoro, 2016) di SD Yogyakarta yang menemukan 78% soal UAS PAK berkategori sedang. Hal serupa juga ditemukan (Mardapi, 2017) di Sleman dengan 72% soal sedang. Konsistensi hasil ini menunjukkan guru PAK di SD Inpres Palsatu telah memahami karakteristik peserta didik kelas VI yang berada pada tahap operasional konkret menurut Piaget. Pada tahap ini, siswa mampu berpikir logis untuk soal dengan tingkat kesulitan menengah, namun belum mampu untuk soal abstrak yang terlalu sukar (Winkel, 2017).

Kaitannya dengan Outcome-Based Education, soal sedang sangat tepat karena OBE menuntut semua siswa mencapai capaian pembelajaran minimal. Soal yang terlalu sukar akan menyebabkan banyak siswa tidak tuntas, sedangkan soal terlalu mudah tidak mengukur ketercapaian CP secara maksimal.

Analisis Daya Pembeda Butir Soal

Daya pembeda mengukur kemampuan soal membedakan siswa berkemampuan tinggi dan rendah (Arikunto, 2018). Hasil penelitian menunjukkan 9 butir soal atau 45% berkategori baik sekali dengan $D \geq 0.40$, dan 8 butir atau 40% berkategori baik dengan $D = 0.30-0.39$. Total 85% soal layak digunakan tanpa revisi.

Kualitas ini lebih tinggi dari penelitian (Purwanto, 2019) di SD Jakarta yang hanya 65% soalnya berkategori baik. Tingginya persentase soal berkualitas di SD Inpres Palsatu mengindikasikan guru sudah menerapkan prinsip penulisan soal yang benar sesuai panduan (BSNP, 2020). Soal dengan daya pembeda tinggi sangat penting dalam OBE karena berfungsi memetakan pencapaian siswa terhadap CP yang telah ditetapkan (Kemendikbudristek, 2022).

Namun terdapat 3 butir soal nomor 4, 11, dan 18 dengan $D = 0.20-0.29$ kategori cukup. Menurut (Zainal, 2019), daya pembeda rendah disebabkan 3 faktor: 1) Pengecoh kurang berfungsi karena terlalu mudah ditebak, 2) Redaksi soal ambigu sehingga siswa bingung, 3) Materi soal tidak diajarkan guru. Rekomendasi revisi sesuai (Rusman, 2018): ganti pengecoh dengan distraktor yang homogen, perbaiki kalimat soal agar lebih operasional, dan pastikan materi sesuai CP fase C Kurikulum Merdeka. Hal ini sejalan dengan (Muslich, 2018) bahwa pembelajaran berbasis kompetensi harus selaras antara materi, soal, dan capaian pembelajaran yang ditargetkan.

Implikasi Hasil terhadap Pembelajaran PAK Berbasis OBE

Secara keseluruhan, 17 butir soal atau 85% dinyatakan layak pakai. Ini membuktikan instrumen evaluasi di SD Inpres Palsatu sudah selaras dengan OBE. Sesuai (Uno, 2016), evaluasi OBE harus berorientasi pada hasil belajar siswa, bukan pada proses mengajar guru. Soal-soal yang memiliki P sedang dan D tinggi mampu mengukur apakah siswa benar-benar menguasai kompetensi PAK seperti mengasihi sesama, jujur, dan taat seperti yang tertuang dalam CP PAK kelas VI. Strategi penyusunan soal seperti ini juga didukung (Sanjaya, 2016) yang menyatakan evaluasi harus berorientasi pada standar proses untuk mencapai tujuan pembelajaran secara maksimal.

Distribusi Soal Berdasarkan Materi dan Indikator CP

Untuk melihat keterwakilan materi, soal dianalisis berdasarkan sebaran indikator CP PAK kelas VI fase C Kurikulum Merdeka. Hasilnya dapat dilihat pada Tabel berikut:

Tabel 2. Distribusi Butir Soal Berdasarkan Materi PAK

No	Materi/Indikator CP	Nomor soal	Jumlah	Persentase
1	Mengasihi Sesamama	1, 5, 9, 13, 17	5	25%
2	Jujur dan Disiplin	2, 6, 10, 14, 18	5	25%
3	Taat pada Tuhan dan Orang Tua	3, 7, 11, 15, 19	5	25%
4	Menghargai Ciptaan Tuhan	4, 8, 12, 16, 20	5	25%
Total			20	100%

Tabel menunjukkan sebaran soal merata 25% tiap materi. Hal ini sesuai prinsip (BSNP, 2020) bahwa soal UAS harus mewakili semua indikator CP secara proporsional. Distribusi merata ini membuktikan guru PAK SD Inpres Palsatu telah menyusun kisi-kisi soal dengan baik sebelum menulis butir soal. Keseimbangan materi ini penting dalam OBE karena semua capaian pembelajaran harus diukur, tidak boleh ada yang terabaikan (Kemendikbudristek, 2022).

Penelitian ini memiliki keterbatasan terkait jumlah sampel yang hanya 27 siswa dan belum dilakukan uji reliabilitas Cronbach Alpha dan Rekomendasi Pengembangan Instrumen.

Penelitian ini memiliki keterbatasan terkait jumlah sampel yang hanya 27 siswa dan belum dilakukan uji reliabilitas Cronbach Alpha untuk mengukur konsistensi internal instrumen. Menurut (Sugiyono, 2017), sampel minimal 30 subjek lebih ideal untuk generalisasi hasil analisis butir soal. Keterbatasan lain adalah penelitian ini hanya fokus pada 2 parameter klasik yaitu tingkat kesukaran dan daya pembeda, belum mengukur efektivitas pengecoh tiap opsi jawaban. Padahal (Zainal, 2019) menekankan bahwa pengecoh yang tidak berfungsi akan menurunkan kualitas soal meskipun daya bedanya tinggi. Untuk penelitian selanjutnya disarankan menambah analisis distraktor dan uji reliabilitas agar instrumen evaluasi PAK semakin terstandar sesuai prinsip OBE.

Kontribusi Penelitian terhadap Pengembangan Evaluasi PAK

Kontribusi utama penelitian ini adalah memberikan bukti empiris bahwa guru PAK di SD Inpres Palsatu telah mampu menyusun soal yang selaras dengan filosofi Outcome-Based Education. Hasil 85% soal layak pakai ini dapat dijadikan model praktik baik bagi sekolah-sekolah dasar lain di Kota Kupang dalam menyusun UAS PAK. Sebagaimana ditegaskan (Muslich, 2018), evaluasi berbasis kompetensi harus dimulai dari instrumen yang valid dan reliabel. Temuan penelitian ini juga mendukung kebijakan (Kemendikbudristek, 2022) tentang implementasi Kurikulum Merdeka yang menuntut *assessment for learning*, bukan *assessment of learning*. Dengan demikian, penelitian ini tidak hanya bermanfaat untuk SD Inpres Palsatu, tetapi juga berkontribusi pada peningkatan mutu pendidikan PAK secara lebih luas.

SIMPULAN DAN SARAN

Berdasarkan hasil analisis dan pembahasan di atas, dapat disimpulkan secara tegas bahwa:

Pertama, kualitas empiris butir soal ujian PAK di SD Inpres Palsatu berbasis Outcome-Based Education tergolong sangat baik. Seluruh 20 butir soal memiliki tingkat kesukaran sedang $P = 0.37-0.67$. Hal ini membuktikan bahwa guru telah mampu merancang soal sesuai tingkat perkembangan kognitif siswa kelas VI SD.

Kedua, dari aspek daya pembeda, sebanyak 17 butir soal atau 85% berkategori baik dan baik sekali dengan $D \geq 0.30$ sehingga layak digunakan tanpa revisi untuk mengukur capaian pembelajaran PAK. Hanya 3 butir soal atau 15% yang perlu direvisi karena daya bedanya cukup.

Ketiga, secara umum instrumen tes yang digunakan di SD Inpres Palsatu sudah memenuhi standar evaluasi OBE sebagaimana diamanatkan Kurikulum Merdeka. Instrumen ini efektif untuk memetakan ketercapaian kompetensi peserta didik dan dapat dijadikan acuan guru dalam menyusun soal-soal selanjutnya.

Dengan demikian, hipotesis awal bahwa butir soal PAK di SD Inpres Palsatu berkualitas baik secara empiris TERBUKTI dan DAPAT DITERIMA.

Saran untuk guru PAK: Pertahankan kualitas soal dengan P sedang dan D tinggi. Untuk 3 butir soal yang perlu revisi, segera perbaiki pengecoh dan redaksi soalnya. Selanjutnya untuk, sekolah: Lakukan analisis butir soal secara rutin setiap akhir semester agar kualitas evaluasi terus terjaga sesuai prinsip OBE. Dan terakhir untuk, peneliti selanjutnya: Disarankan melengkapi analisis dengan uji efektivitas pengecoh dan reliabilitas instrumen menggunakan Cronbach Alpha agar analisis lebih komprehensif.

DAFTAR PUSTAKA

- Arikunto, S. (2018). Dasar-dasar evaluasi pendidikan (R. Damayanti (ed.). PT. Bumi Aksara.
- BSNP. (2020). Panduan Penulisan Soal Pilihan Ganda.
- Kemendikbudristek. (2022). Capaian Pembelajaran Kurikulum Merdeka Jenjang SD.
- Mardapi, D. (2017). Pengukuran Penilaian dan Evaluasi Pendidikan. Nuha Medika.
- Muslich, M. (2018). KTSP Pembelajaran Berbasis Kompetensi dan Kontekstual. Remaja Rosdakarya.
- Nurgiyantoro, B. (2010). Penilaian pembelajaran sastra berbasis kompetensi. Yogyakarta: BPFe.
- Purwanto, N. (2019). Prinsip-Prinsip dan Teknik Evaluasi Pengajaran. Remaja Rosdakarya.
- Rusman. (2018). Model-Model Pembelajaran. RajaGrafindo Persada.
- Sanjaya, W. (2016). Strategi Pembelajaran Berorientasi Standar Proses Pendidikan. Kencana.
- Sudijono, A. (2015). Pengantar Evaluasi Pendidikan. RajaGrafindo Persada.
- Sugiyono. (2017). Metode Penelitian Kuantitatif, Kualitatif, dan R&D. Alfabeta.
- Uno, H. B. (2016). Perencanaan Pembelajaran. Bumi Aksara.
- Utami, S. A., & Nurgiyantoro, B. (2016). Kualitas Soal Dan Daya Serap Tes Pendalaman Materi Un Bahasa Indonesia Smp Di Gunungkidul. Diksi, 24(1).
- Winkel, W. S.keterbatasan (2017). Psikologi Pengajaran. Media Abadi.
- Zainal, A. (2019). Evaluasi Pembelajaran. Remaja Rosdakarya.