

ANALISIS KELAYAKAN INSTRUMEN *TRY OUT* UJI KOMPETENSI PESERTA PENDIDIKAN PROFESI GURU BIDANG STUDI LAYANAN KESEHATAN

Zihan Novita Sari^{1*}, Hendra Susanto¹, Paundra Wangsa Fajar Kusuma¹

¹ Pendidikan Profesi Guru (PPG), Sekolah Pascasarjana, Universitas Negeri Malang

Email: zihan.novita.fik@um.ac.id

Abstrak

Penelitian ini bertujuan untuk mengembangkan instrumen try out bidang studi Layanan Kesehatan yang valid dan reliabel sebagai sarana menyiapkan mahasiswa menghadapi UKPPPG. Metode yang digunakan adalah pengembangan tes melalui penyusunan spesifikasi, penulisan dan penelaahan butir soal, uji coba lapangan, serta analisis hasil. Subjek penelitian terdiri atas 85 mahasiswa PPG Layanan Kesehatan Universitas Negeri Malang. Data diolah dengan analisis klasik untuk menilai tingkat kesulitan, daya beda, distraktor, validitas, dan reliabilitas butir soal. Dari 35 butir yang diujikan, 25 butir memenuhi kriteria validitas ($r \geq 0,30$) dan 10 butir berada di bawah kriteria, namun tiga di antaranya dipertimbangkan untuk direvisi karena memiliki validitas mendekati 0,30. Sisanya, direvisi secara total. Distribusi tingkat kesulitan menunjukkan 29 butir soal tergolong mudah, 4 sedang, dan 2 sulit. Sedangkan analisis daya beda memperlihatkan 9 butir ideal dan 26 rendah. Analisis distraktor juga mengindikasikan sebagian besar opsi jawaban belum berfungsi optimal. Setelah dilakukan revisi redaksional dan perbaikan distraktor, seluruh 35 butir soal dapat dipertahankan dengan koefisien reliabilitas instrumen sebesar 0,842 yang termasuk kategori tinggi. Instrumen akhir ini dinyatakan layak digunakan sebagai sarana evaluasi awal kesiapan mahasiswa PPG Layanan Kesehatan menghadapi UKPPPG. Rekomendasi penelitian selanjutnya adalah memperluas uji coba pada bidang studi lainnya serta mengoptimalkan kualitas daya beda dan distraktor.

Kata Kunci: Uji Kompetensi, Pendidikan Profesi Guru, Layanan Kesehatan

Abstract

This study aims to develop a valid and reliable try-out instrument in the field of Health Education as a means to map students' readiness in facing the Teacher Professional Education Competency Examination (UKPPPG). The research was conducted using a test development method consisting of test specification, item writing and review, field trials, and analysis of the test results. The subjects were 85 students of the Teacher Professional Education (PPG) Program in the field of Health Education at Universitas Negeri Malang. Data were analyzed using classical test theory to examine item difficulty, discrimination power, distractors, validity, and reliability. Of the 35 items tested, 25 met the validity criteria ($r \geq 0.30$), while 10 items fell below the threshold. However, three of these items were considered for revision due to their validity values being close to 0.30. The difficulty analysis showed that 29 items were classified as easy, 4 as moderate, and 2 as difficult, while the discrimination analysis indicated 9 items with ideal discrimination and 26 with low discrimination. Distractor analysis also revealed that most answer options did not function optimally. After editorial revisions and improvement of distractors, all 35 items were retained, resulting in an instrument reliability coefficient of 0.842, which is classified as high. The final instrument is therefore considered feasible to be used as an initial evaluation tool for assessing PPG students' readiness in Health Education to face the UKPPPG. Future research is recommended to expand trials to a more diverse population and to optimize item discrimination and distractor quality in order to enhance measurement accuracy

Keywords: Competency Examination, Teacher Professional Education, Health Education

PENDAHULUAN

Pendidikan Profesi Guru (PPG) merupakan salah satu tahap paling strategis dalam sistem pendidikan nasional karena menjadi instrumen utama dalam mempersiapkan tenaga pendidik yang profesional dan berdaya saing global. Pendidikan abad 21 menuntut guru untuk menguasai pengetahuan bidang studi sekaligus berperan sebagai fasilitator pembelajaran yang mendorong peserta didik mengembangkan keterampilan berpikir kritis, kreativitas, kolaborasi, dan komunikasi. Perubahan paradigma pendidikan berimplikasi pada peningkatan kapasitas guru profesional agar mampu mendesain pengalaman belajar yang bermakna dan berorientasi pada pengembangan kompetensi, bukan sekadar transfer materi. Hal ini menegaskan bahwa kualitas guru menjadi faktor kunci dalam keberhasilan pendidikan. Di lain sisi, pembelajaran abad ke-21 hanya akan tercapai apabila guru mampu mengintegrasikan teknologi, kolaborasi, dan pendekatan pembelajaran inovatif (Muhali, 2019). Dengan demikian, PPG menjadi fondasi utama untuk menjawab tantangan kompleks yang dihadapi dunia pendidikan saat ini.

Di Indonesia, PPG menjadi salah satu kebijakan strategis untuk menghasilkan pendidik yang memenuhi standar kompetensi nasional. PPG merupakan program profesional untuk mengintegrasikan kompetensi pedagogik, profesional, sosial, dan kepribadian sehingga calon guru dapat menjalankan peran secara utuh dan berimbang (Abbas et al., 2023). Penyelenggaraan PPG tidak hanya berorientasi pada penguasaan akademik, tetapi juga pada penguatan nilai, sikap, serta kemampuan praktik dalam mengajar. Hal ini sejalan dengan semangat Undang-Undang Nomor 14 Tahun 2005 tentang Guru dan Dosen yang menempatkan sertifikasi guru sebagai penanda profesionalisme. Oleh sebab itu, PPG menjadi instrumen vital dalam membangun mutu pendidikan, memastikan setiap calon guru tidak hanya memiliki ijazah akademik, tetapi juga kompetensi profesional yang diakui secara formal melalui sertifikat pendidik (serdik).

Salah satu bidang studi yang diselenggarakan dalam PPG adalah Layanan Kesehatan. Bidang studi ini memiliki karakteristik khas karena menggabungkan dimensi pengetahuan teoritis dengan keterampilan praktik serta sikap profesional. Bidang studi Layanan Kesehatan dalam PPG berfokus pada pembekalan calon guru profesional agar mampu mengajarkan materi yang terkait dengan kesehatan, gizi, kebugaran jasmani, dan praktik layanan kesehatan dasar di sekolah. Selain itu, calon guru profesional Layanan Kesehatan juga diharapkan memiliki sensitivitas terhadap isu-isu etis dan keselamatan, karena materi yang diajarkan berkaitan erat dengan kehidupan dan kesejahteraan peserta didik.

Calon guru profesional dituntut untuk memiliki kemampuan menyeluruh dalam mengelola kelas, menyusun perangkat pembelajaran, menerapkan pendekatan saintifik, hingga menciptakan suasana belajar yang partisipatif (Febriani et al., 2024). Tugas ini membutuhkan keterampilan berpikir tingkat tinggi, seperti pemecahan masalah dan pengambilan keputusan cepat, serta keterampilan komunikasi efektif yang sangat penting dalam pembelajaran kesehatan. Schleicher (2018) menekankan bahwa guru profesional adalah mereka yang mampu mengintegrasikan teori dengan praktik pembelajaran nyata di kelas. Tidak hanya itu, calon guru Layanan Kesehatan juga ditantang untuk mampu menghubungkan pengetahuan kesehatan dengan praktik sehari-hari peserta didik, sehingga

pembelajaran bersifat aplikatif dan relevan. Tantangan ini menegaskan pentingnya instrumen evaluasi yang mampu mengukur kompetensi calon guru secara komprehensif.

Guna menjawab kebutuhan tersebut, pemerintah melalui Kementerian Pendidikan Dasar dan Menengah (Kemdikdasmen) merancang Uji Kompetensi Peserta Pendidikan Profesi Guru (UKPPPG) sebagai instrumen utama untuk menilai kesiapan calon guru sebelum memperoleh serdik. UKPPPG memiliki fungsi strategis sebagai tolok ukur mutu nasional serta memastikan bahwa guru yang dinyatakan lulus benar-benar memiliki kompetensi yang sesuai dengan standar profesionalisme (Direktur Jenderal Guru dan Tenaga Kependidikan et al., 2024). Serdik yang diperoleh melalui kelulusan UKPPPG bukan hanya simbol administratif, melainkan legitimasi formal bahwa seorang guru telah diakui secara profesional dan layak mengemban tugas pendidikan.

UKPPPG terdiri atas dua komponen utama, yaitu Uji Tertulis Berbasis Komputer (UTBK) dan Uji Kinerja (UKin) (Direktur Jenderal Guru dan Tenaga Kependidikan et al., 2024). UTBK digunakan untuk mengukur penguasaan pengetahuan pedagogik, keilmuan bidang studi, serta kemampuan berpikir kritis dan reflektif. Sementara itu, UKin menilai keterampilan praktik calon guru dalam melaksanakan pembelajaran nyata di kelas, termasuk perencanaan, pelaksanaan, interaksi dengan siswa, penggunaan media, serta evaluasi pembelajaran (Prodi PPG - Universitas Negeri Malang, 2025). Keduanya saling melengkapi untuk menilai kompetensi guru secara komprehensif, baik dari sisi teori maupun praktik.

Materi yang diujikan dalam UKPPPG mengacu pada empat kompetensi utama guru, yaitu pedagogik, profesional, sosial, dan kepribadian. Pada bidang studi Layanan Kesehatan, materi ujian mencakup kemampuan mengajarkan konsep kesehatan, praktik layanan kesehatan dasar, integrasi teknologi pembelajaran, hingga penguatan etika dan sikap profesional. Peserta ujian dituntut untuk mampu merancang perangkat pembelajaran, melaksanakan pembelajaran sesuai rencana, mengelola kelas, serta melakukan evaluasi berbasis asesmen autentik. Hal ini menunjukkan bahwa UKPPPG bukan sekadar ujian teoritis, tetapi instrumen evaluasi komprehensif yang mensimulasikan tugas nyata seorang guru.

Pelaksanaan UKPPPG tidak lepas dari tantangan. Kompleksitas materi ujian, standar penilaian yang tinggi, serta keterbatasan kesiapan calon guru seringkali menjadi faktor penghambat. Penelitian Setiawan (2022) menunjukkan bahwa banyak peserta mengalami kesulitan karena lingkup materi yang luas dan tuntutan penguasaan teknologi yang belum merata. Hal ini berdampak pada tingkat kelulusan yang bervariasi antarperiode, sehingga memerlukan strategi persiapan lebih sistematis bagi mahasiswa PPG. Oleh karena itu, keberadaan program pendukung seperti *try out* menjadi semakin penting sebagai sarana latihan dan pemetaan kesiapan.

Sebagai bentuk respon terhadap kebutuhan tersebut, Lembaga Pendidikan Tenaga Kependidikan (LPTK), termasuk Universitas Negeri Malang (UM), mengembangkan program *try out* UKPPPG. Program ini dirancang untuk memberikan pengalaman simulasi ujian, membantu mahasiswa mengenali format soal, melatih keterampilan manajemen waktu, serta memahami mekanisme penilaian. *Try out* juga dapat menjadi sarana bagi peserta untuk mengevaluasi kelemahan mereka sebelum menghadapi ujian sesungguhnya. Dengan demikian, *try out* berfungsi sebagai jembatan yang menghubungkan hasil pembelajaran dengan standar evaluasi nasional.

Keberadaan *try out* menjadi semakin strategis dalam bidang studi Layanan Kesehatan, mengingat karakteristik kompetensi yang harus dicapai berfokus pada penguasaan kompetensi profesional dalam praktik pedagogik. Melalui *try out*, calon guru dapat berlatih untuk mengintegrasikan berbagai aspek kompetensi tersebut, sekaligus membangun rasa percaya diri menghadapi UKPPPG. Selain itu, *try out* juga menyediakan data awal yang dapat dianalisis untuk mengetahui sejauh mana soal-soal yang digunakan mampu mengukur kompetensi sesuai standar. Hal ini memberi peluang bagi pengembangan instrumen evaluasi yang lebih valid, reliabel, dan sesuai dengan kebutuhan mahasiswa.

Pengembangan instrumen *try out* UKPPPG pada bidang studi Layanan Kesehatan menjadi urgen untuk dilaksanakan. Instrumen yang dikembangkan diharapkan tidak hanya berfungsi sebagai alat ukur capaian pembelajaran, tetapi juga mampu memberikan gambaran komprehensif mengenai kesiapan dan potensi calon guru dalam menghadapi UKPPPG. Selama ini, sebagian besar kajian evaluasi lebih berfokus pada aspek kognitif, padahal konteks pendidikan profesi guru menuntut penilaian yang juga mencakup keterampilan pedagogik dan sikap profesional.

Penelitian ini bertujuan untuk mengembangkan instrumen *try out* UKPPPG bidang studi Layanan Kesehatan yang valid dan reliabel, sehingga mampu mengungkap kompetensi secara menyeluruh sekaligus memprediksi keberhasilan calon guru dalam mengikuti UKPPPG.

METODE

Instrumen dalam penelitian ini dikembangkan dengan mengadaptasi prosedur penyusunan alat ukur sebagaimana dikemukakan oleh Suryabrata (2005), melalui tahapan berikut: (1) Pengembangan spesifikasi alat ukur, (2) Penulisan pernyataan-pernyataan/soal berdasar kisi-kisi, (3) Penelaahan pernyataan/soal, (4) Perakitan instrumen, (5) Uji coba instrumen, (6) Analisis hasil uji coba, (7) Seleksi dan perakitan alat ukur.

Analisis data uji coba dilakukan dengan menggunakan analisis klasik untuk menganalisis tingkat kesulitan, daya beda, distraktor, validitas butir dan reliabilitas instrumen. Validitas butir diperiksa melalui korelasi skor butir dengan skor total, sedangkan reliabilitas instrumen ditentukan menggunakan koefisien reliabilitas internal. Analisis ini bertujuan memastikan bahwa instrumen yang dikembangkan memenuhi persyaratan validitas dan reliabilitas, sehingga dapat digunakan secara sah untuk mengukur kesiapan mahasiswa PPG dalam menghadapi UKPPPG.

Subjek uji coba penelitian ini adalah mahasiswa PPG bidang studi Layanan Kesehatan di UM. Total responden berjumlah 85 mahasiswa yang seluruhnya mengikuti *try out* UKPPPG secara daring. Data yang terkumpul meliputi jawaban tes, waktu pengerjaan, serta skor akhir. Selanjutnya, data tersebut dianalisis untuk memperoleh informasi empiris mengenai karakteristik butir soal.

Tabel 1. Kisi-Kisi Soal

No Soal	Materi	Indikator Kemampuan yang Diharapkan
---------	--------	-------------------------------------

1	Manajemen kelas dalam pembelajaran layanan kesehatan	Menganalisis tindakan guru dalam menciptakan suasana belajar aman, nyaman, dan inklusif
2	Pembelajaran berdiferensiasi	Mengidentifikasi penerapan pembelajaran berpusat pada peserta didik sesuai gaya belajar
3	Layanan kesehatan berbasis klien	Menentukan tujuan penciptaan rasa aman dan nyaman dalam proses layanan/pembelajaran
4	Perencanaan pembelajaran	Mengidentifikasi desain pembelajaran yang terstruktur dan sistematis
5	Pembelajaran kontekstual & keberlanjutan	Menilai relevansi desain pembelajaran dengan lingkungan dan isu kesehatan berkelanjutan
6	Sumber belajar	Menentukan pemilihan sumber belajar yang sesuai dengan tujuan pembelajaran
7	Model pembelajaran berbasis masalah	Menganalisis desain pembelajaran yang mendorong keaktifan dan berpikir kritis
8	Integrasi TIK dalam pembelajaran	Mengidentifikasi penggunaan TIK secara adaptif dalam pembelajaran
9	Asesmen pembelajaran	Menganalisis asesmen berpusat pada peserta didik yang mendorong kreativitas dan refleksi
10	Umpan balik pembelajaran	Mengidentifikasi praktik pemberian umpan balik yang konstruktif
11	Sistem rujukan layanan kesehatan	Menjelaskan fungsi koordinasi antarjenjang fasilitas layanan kesehatan
12	Etika profesi guru	Mengidentifikasi penerapan kode etik guru dalam pembelajaran
13	Ketimpangan teknologi kesehatan	Menganalisis dampak ketimpangan teknologi terhadap pemerataan layanan kesehatan
14	Pengembangan profesional guru	Menilai sikap adaptif guru dalam meningkatkan pembelajaran berpusat pada siswa
15	Prinsip 3R dalam layanan kesehatan	Menentukan penerapan prinsip <i>Reduce</i> dalam pengelolaan limbah medis
16	Healthpreneurship	Mengidentifikasi karakteristik utama profil healthpreneur
17	Etika layanan kesehatan	Menganalisis sikap menghargai hak klien secara adil dan proporsional
18	K3 layanan kesehatan	Mengidentifikasi tindakan guru dalam menjamin keselamatan dan keamanan peserta didik
19	Sistem saraf dan ekskresi	Menjelaskan hubungan sistem saraf dengan pengendalian sistem ekskresi
20	Pengelolaan SDM layanan kesehatan	Mengidentifikasi kemampuan kerja tim dan koordinasi dalam pengelolaan SDM
21	Komunikasi efektif dalam layanan kesehatan	Menganalisis strategi komunikasi tenaga kesehatan yang empatik, jelas, dan beretika dalam menghadapi klien
22	Pembelajaran berpusat pada peserta didik	Mengidentifikasi penerapan prinsip pembelajaran yang menghargai perbedaan kemampuan dan kebutuhan belajar siswa
23	Keselamatan pasien dan etika layanan	Menganalisis pengambilan keputusan layanan kesehatan yang mengutamakan keselamatan dan kepentingan klien
24	Pengelolaan fasilitas layanan kesehatan	Menentukan langkah pengelolaan fasilitas layanan kesehatan yang efisien dan sesuai standar mutu layanan
25	Profesionalisme tenaga kesehatan	Mengidentifikasi sikap profesional tenaga kesehatan dalam menjalankan tugas sesuai kode etik dan tanggung jawab kerja
26	Manajemen SDM kesehatan (PG kompleks)	Menganalisis strategi pengelolaan SDM berorientasi mutu layanan dan kesejahteraan
27	Teknologi kesehatan & <i>patient-centered care</i>	Menentukan dampak positif AI dan IoMT terhadap layanan kesehatan berbasis pasien
28	K3 dan kesehatan kerja	Menganalisis strategi holistik dalam menurunkan stres dan meningkatkan kualitas layanan
29	Manajemen klien	Mengidentifikasi langkah profesional dalam identifikasi kebutuhan klien
30	Layanan kesehatan primer	Menganalisis strategi layanan komunitas yang berkelanjutan
31	Keselamatan klien	Menentukan strategi layanan adaptif dan berorientasi keselamatan klien

32	Isu global kesehatan	Menganalisis pendekatan holistik dalam penanggulangan resistensi antibiotik
33	Etika komunikasi kesehatan	Mengidentifikasi strategi komunikasi etis dan empatik kepada pasien
34	Peluang usaha layanan kesehatan	Menganalisis indikator dan peluang pasar layanan kesehatan modern
35	Profesi baru layanan kesehatan	Mengidentifikasi profesi baru akibat perubahan teknologi dan gaya hidup sehat

HASIL DAN PEMBAHASAN

HASIL

Tingkat Kesulitan

Berdasarkan hasil analisis tingkat kesukaran butir soal pada uji coba instrumen, dari 35 soal yang diujikan ditemukan 29 butir (82,9%) tergolong mudah, 4 butir (11,4%) tergolong sedang atau ideal, dan hanya 2 butir (5,7%) tergolong sulit. Kategori kesukaran ini ditentukan melalui proporsi jawaban benar peserta, yakni soal mudah apabila dijawab benar oleh lebih dari 70% peserta, sedang apabila berada pada rentang 30–69,9%, dan sulit apabila kurang dari 30% (0 s.d. 29,9%) peserta mampu menjawab benar (Atmoko, 2020; Cronbach, 1990). Dominasi soal mudah menunjukkan bahwa komposisi tes masih belum seimbang, sehingga kemampuan peserta yang berada pada tingkat menengah dan tinggi kurang terakomodasi secara optimal. Idealnya, distribusi soal harus lebih proporsional agar instrumen dapat mengukur rentang kemampuan peserta secara lebih komprehensif dan akurat (Dorfman & Kalugin, 2020; Kim, 2020).

Perbandingan antara target jumlah butir soal dan realisasi hasil analisis tingkat kesukaran disajikan pada Tabel 2. Berdasarkan hasil tersebut, diketahui bahwa dari 35 butir soal yang diujikan, sebagian besar termasuk dalam kategori mudah, sementara jumlah soal kategori sedang dan sulit relatif terbatas.

Tabel 2. Target Banyak Butir Soal Berdasarkan Kesukaran

Jumlah dirakit	Banyak Soal berdasarkan Kesukaran					
	Sulit		Sedang		Mudah	
	Target	Realisasi	Target	Realisasi	Target	Realisasi
35	2	2	4	4	29	29

Daya Beda

Hasil analisis terhadap daya beda menunjukkan bahwa dari 35 butir soal yang diujikan, terdapat 26 butir (74,3%) yang termasuk kategori rendah (*jelek*) dan hanya 9 butir (25,7%) yang berkategori sedang atau ideal (*cukup*), sementara tidak ada satupun soal yang mencapai kategori tinggi. Secara prinsip, daya beda ditentukan melalui perbedaan proporsi jawaban benar antara kelompok atas dan kelompok bawah; apabila selisih proporsinya kurang dari 30% (0 s.d. 29,9%) maka dikategorikan rendah, sedangkan selisih 30%–69,9% dikategorikan ideal, dan lebih dari 70% tergolong tinggi (Atmoko, 2020; Cronbach, 1990). Temuan ini mengindikasikan bahwa sebagian besar soal belum optimal dalam membedakan kemampuan peserta dengan capaian tinggi dan rendah, sehingga fungsi diagnostik instrumen menjadi terbatas. Oleh karena itu, revisi butir soal perlu diarahkan pada peningkatan daya

beda agar tes dapat lebih efektif dalam mengidentifikasi variasi kemampuan peserta.

Hasil analisis daya beda menunjukkan variasi kualitas butir soal dalam membedakan kemampuan peserta. Perbandingan antara target dan realisasi jumlah butir soal berdasarkan kategori daya beda disajikan pada Tabel 3. Data tersebut memberikan gambaran mengenai distribusi daya beda butir soal dalam instrumen yang diuji cobakan.

Tabel 3. Target Banyak Butir Soal Berdasarkan Daya Beda

Jumlah dirakit	Banyak Soal berdasarkan Daya Beda					
	Tinggi		Ideal		Sedang	
	Target	Realisasi	Target	Realisasi	Target	Realisasi
35	0	0	9	9	26	26

Validitas Butir

Berdasarkan hasil analisis validitas butir soal, dari total 35 butir yang diujicobakan terdapat 25 butir yang memenuhi kriteria validitas dengan koefisien korelasi $\geq 0,30$, sedangkan 10 butir tergolong tidak valid karena koefisien validitasnya berada di bawah ambang batas 0,30. Perhitungan validitas dilakukan dengan mengkorelasikan skor setiap butir terhadap total skor keseluruhan responden. Skor butir merepresentasikan pencapaian indikator spesifik, sedangkan total skor merefleksikan konstruk yang diukur secara menyeluruh. Mengacu pada Cronbach (1990), batasan 0,30 digunakan sebagai syarat minimum agar suatu butir dapat dianggap memberikan kontribusi yang cukup terhadap pengukuran konstruk. Oleh karena itu, butir yang memiliki korelasi di bawah 0,30 dikategorikan tidak valid karena hanya memberikan kontribusi kurang dari 30% terhadap pengukuran konstruk yang sedang diukur. Temuan ini menunjukkan bahwa secara umum mayoritas butir soal memiliki validitas yang memadai, meskipun beberapa butir tetap memerlukan revisi agar dapat dipertahankan dalam instrumen final.

Distraktor Pilihan Jawaban

Berdasarkan hasil analisis distraktor pada 35 butir soal yang diujikan, ditemukan bahwa sebanyak 25 butir soal masih memiliki distraktor yang tidak berfungsi karena tidak dipilih sama sekali oleh responden, sehingga butir-butir tersebut memerlukan revisi. Sementara itu, sebanyak 10 butir soal lainnya telah memiliki distraktor yang berfungsi secara ideal karena setiap pilihan jawaban (A, B, C, D, maupun E) dipertimbangkan oleh responden meskipun tidak selalu menjadi kunci jawaban. Secara konseptual, distraktor berfungsi efektif apabila mampu mengecoh sebagian peserta, sebab hal ini menunjukkan bahwa opsi tersebut relevan dan tidak sekadar menjadi hiasan dalam pilihan jawaban. Sebaliknya, jika suatu opsi sama sekali tidak dipilih, maka distraktor tersebut dianggap jelek dan harus diperbaiki (Atmoko, 2020). Temuan ini mengindikasikan bahwa sebagian besar soal masih memerlukan perbaikan kualitas distraktor agar instrumen lebih optimal dalam mengukur kompetensi dan membedakan kemampuan responden secara lebih akurat.

Hasil analisis distraktor yang membandingkan target dan realisasi kualitas distraktor pada butir soal disajikan pada Tabel 4. Berdasarkan hasil tersebut, diketahui bahwa masih

terdapat sejumlah butir soal yang memiliki distraktor tidak berfungsi, sehingga memerlukan perbaikan pada tahap revisi instrumen.

Tabel 4. Target Banyak Butir Soal Berdasarkan Distraktor

Jumlah dirakit	Banyak Soal berdasarkan Distraktor					
	Seluruh opsi berfungsi		Sebagian opsi berfungsi		Seluruh opsi TIDAK berfungsi	
	Target	Realisasi	Target	Realisasi	Target	Realisasi
35	0	0	10	10	25	25

Butir yang gugur, digunakan tanpa revisi dan digunakan setelah revisi

Keputusan untuk menggunakan suatu butir soal tanpa revisi, digunakan setelah direvisi, atau digugurkan didasarkan terutama pada tiga indikator yaitu validitas butir, tingkat kesukaran, dan daya beda. Selain itu, saran mengenai distraktor turut menjadi pertimbangan. Dari 35 butir yang diuji terdapat 10 butir yang direkomendasikan untuk dipakai setelah dilakukan revisi, tidak ada butir yang digugurkan, 25 butir yang langsung dipakai tanpa revisi. Butir-butir yang dipilih untuk direvisi umumnya memiliki validitas yang mendekati ambang kritis (mis. berkisar 0,24–0,29), atau memperlihatkan masalah pada distraktor (opsi jawaban yang tidak efektif), sehingga perbaikan difokuskan pada redaksi soal, revisi pilihan jawaban, dan penyesuaian keterwakilan indikator pada kisi-kisi. Pendekatan ini sesuai dengan praktik pengembangan instrumen yang mengutamakan analisis item-total, tingkat kesukaran, dan daya beda sebagai dasar keputusan konstruksi soal. Revisi diarahkan agar soal mempunyai proporsi jawaban benar yang lebih sesuai, daya beda yang meningkat, serta distraktor yang berfungsi sebagai pengecoh. Setelah perbaikan tersebut, 35 butir akan dirakit kembali untuk versi final instrumen (lihat tabel 5).

Tabel 5. Revisi Butir Soal

No Soal	Nilai Validitas	Permasalahan Utama	Fokus Revisi	Status Setelah Revisi
4	0,120	Validitas rendah	Perbaikan redaksi dan opsi jawaban	Digunakan
5	-0,002	Validitas dan daya beda rendah	Penyelarasan stimulus dan distraktor	Digunakan
8	0,239	Distraktor tidak berfungsi	Penyusunan ulang alternatif jawaban	Digunakan
13	0,227	Distraktor tidak berfungsi	Perbaikan dan pemerataan kualitas distraktor	Digunakan
14	0,141	Daya beda kurang optimal	Penyesuaian redaksi dan konteks soal	Digunakan
16	0,285	Redaksi kurang jelas	Penyederhanaan redaksi dan penajaman konteks	Digunakan
24	0	Soal tidak valid	Penyusunan ulang keseluruhan soal	Digunakan
25	0,123	Redaksi dan distraktor	Klarifikasi redaksi dan opsi	Digunakan

		lemah	jawaban	
33	0,107	Validitas rendah	Penajaman indikator dalam redaksi soal	Digunakan
34	0,173	Distraktor tidak optimal	Revisi alternatif jawaban	Digunakan

Reliabilitas

Hasil uji reliabilitas instrumen menunjukkan bahwa koefisien Cronbach's Alpha sebesar 0,842. Nilai ini menandakan bahwa instrumen tes yang terdiri dari 35 butir soal memiliki tingkat reliabilitas yang tinggi, karena berada di atas ambang batas 0,70 yang secara umum digunakan sebagai kriteria minimum reliabilitas (Cronbach, 1990). Dengan kata lain, konsistensi internal antarbutir soal dalam instrumen ini tergolong baik, sehingga instrumen mampu menghasilkan hasil pengukuran yang stabil dan dapat dipercaya. Dengan demikian, meskipun masih terdapat sejumlah butir yang perlu direvisi atau digugurkan berdasarkan validitas, tingkat kesukaran, daya beda, maupun distraktor, secara keseluruhan instrumen telah memenuhi standar reliabilitas yang memadai untuk digunakan dalam pengukuran kemampuan responden.

PEMBAHASAN

Guna pengambilan keputusan mereduksi jumlah soal dari 35 butir yang diujikan, peneliti menggunakan dasar pertimbangan pada (1) tingkat validitas, (2) tingkat kesulitan, (3) daya beda, dan (4) saran berdasarkan hasil analisis distraktor pada setiap butir (Atmoko, 2020). Berdasarkan keempat dasar pertimbangan tersebut, diputuskan apakah suatu butir soal dipakai tanpa revisi, dipakai dengan revisi, atau digugurkan.

Tingkat validitas butir soal dihitung dengan mengkorelasikan skor setiap butir terhadap total skor yang diperoleh responden. Cronbach (1990) menetapkan kriteria sebesar 0,30 sebagai batas minimal validitas suatu butir. Jika koefisien korelasi lebih kecil dari 0,30, maka butir soal dianggap tidak valid karena kontribusinya terhadap pengukuran konstruk kurang dari 30%. Dari 35 butir soal yang diujikan kepada 85 mahasiswa PPG Layanan Kesehatan, terdapat 25 butir soal yang memenuhi kriteria validitas dan 10 butir soal yang dinyatakan tidak valid karena memiliki korelasi di bawah 0,30. Jumlah butir valid ini relatif mencukupi, namun untuk mewakili seluruh aspek kisi-kisi, peneliti juga mempertimbangkan butir yang memiliki validitas mendekati 0,30 (*borderline*).

Butir soal yang masuk kategori *borderline* meliputi pertanyaan nomor 16 (validitas 0,285), nomor 8 (validitas 0,239), dan nomor 13 (validitas 0,227). Ketiga soal tersebut dipertimbangkan untuk direvisi, bukan langsung digugurkan. Sebagai contoh, butir nomor 16 termasuk soal dengan validitas mendekati 0,30 dan daya beda cukup baik, sehingga revisi diarahkan pada redaksi kalimat agar lebih jelas. Pada soal nomor 8 dan 13, distribusi pilihan jawaban kurang merata karena adanya opsi distraktor yang tidak berfungsi. Revisi difokuskan pada penyusunan ulang alternatif jawaban agar semua pilihan memiliki daya tarik dan dapat meningkatkan kualitas butir. Dengan demikian, keputusan tidak hanya bergantung pada angka validitas semata, tetapi juga pada potensi butir untuk diperbaiki agar tetap representatif terhadap indikator kisi-kisi.

Analisis tingkat kesulitan menunjukkan bahwa dari 35 soal yang dirakit, terdapat 2 butir tergolong sulit, 4 butir tergolong sedang, dan 29 butir tergolong mudah (lihat tabel 1). Komposisi ini menunjukkan dominasi soal kategori mudah yang relatif tinggi, sementara soal kategori sedang dan sulit masih sangat terbatas. Padahal, keseimbangan tingkat kesukaran penting agar instrumen mampu membedakan kemampuan peserta secara proporsional.

Berdasarkan daya beda, hasil analisis memperlihatkan bahwa hanya 9 butir yang memiliki daya beda ideal, sedangkan 26 butir berada pada kategori daya beda rendah (lihat tabel 2). Tidak ditemukan soal dengan daya beda tinggi. Kondisi ini menunjukkan bahwa sebagian besar soal belum cukup sensitif dalam membedakan mahasiswa yang memiliki tingkat penguasaan tinggi dengan yang rendah. Oleh karena itu, butir dengan daya beda rendah perlu direvisi, baik melalui perbaikan redaksi maupun pengaturan opsi jawaban, agar instrumen mampu mengukur secara lebih optimal.

Merujuk pada tabel 5, hasil analisis empiris menunjukkan bahwa dari 35 butir soal yang diujicobakan, terdapat 10 butir dengan koefisien validitas di bawah batas minimum 0,30. Meskipun demikian, seluruh butir tersebut tidak digugurkan, melainkan direvisi, karena masih merepresentasikan indikator penting dalam kisi-kisi serta memiliki potensi untuk diperbaiki secara redaksional maupun teknis. Sebagian butir dengan validitas rendah berada pada kategori *borderline*, yaitu memiliki nilai mendekati ambang batas validitas. Butir soal nomor 16 (validitas 0,285), nomor 8 (validitas 0,239), dan nomor 13 (validitas 0,227) termasuk dalam kategori ini (lihat tabel 5). Ketiga butir tersebut dipertimbangkan untuk direvisi karena secara substantif telah sesuai dengan indikator yang diukur dan, pada beberapa aspek, masih menunjukkan daya beda yang cukup memadai.

Pada butir nomor 8 dan 13, kelemahan utama teridentifikasi pada kualitas distraktor, di mana terdapat opsi jawaban yang tidak berfungsi karena tidak dipilih oleh responden. Revisi dilakukan dengan menyusun ulang alternatif jawaban agar setiap opsi memiliki daya tarik yang seimbang dan berfungsi sebagai pengecoh yang efektif. Sementara itu, pada butir nomor 16, permasalahan utama terletak pada redaksi kalimat soal yang berpotensi menimbulkan ambiguitas. Oleh karena itu, revisi difokuskan pada penyederhanaan struktur kalimat dan penajaman konteks agar stimulus lebih jelas dan mudah dipahami peserta.

Selain butir *borderline*, tujuh butir lainnya juga direvisi karena menunjukkan kombinasi antara validitas rendah, daya beda kurang optimal, serta efektivitas distraktor yang belum memadai. Revisi pada butir-butir tersebut dilakukan tanpa mengubah indikator yang diukur, melainkan difokuskan pada perbaikan redaksi, penyelarasan konteks soal, dan penguatan kualitas opsi jawaban.

Mengacu pada hasil akhir perakitan instrumen, seluruh 35 butir soal tetap dapat digunakan kembali karena jumlah butir yang terealisasi sepenuhnya sesuai dengan target. Hal ini tercapai setelah peneliti melakukan revisi redaksional maupun penyusunan ulang opsi jawaban berdasarkan temuan analisis. Dengan demikian, walaupun kualitas awal sebagian butir masih menunjukkan kelemahan pada daya beda dan distraktor, hasil revisi menjadikan keseluruhan soal layak dipertahankan sebagai instrumen *try out* UKPPPG Layanan Kesehatan.

SIMPULAN

Penelitian ini menganalisis instrumen *try out* UKPPPG bidang studi Layanan Kesehatan yang terdiri atas 35 butir soal dengan koefisien reliabilitas tinggi ($\alpha = 0,842$). Analisis butir menunjukkan bahwa sebagian besar soal tergolong mudah dan memiliki daya beda rendah, sedangkan validitas soal relatif memadai dengan 25 butir memenuhi kriteria $\geq 0,30$ dan 10 butir memerlukan perbaikan. Selain itu, analisis distraktor mengindikasikan perlunya revisi karena masih terdapat opsi jawaban yang tidak berfungsi secara optimal. Melalui proses revisi redaksional dan penyusunan ulang opsi jawaban, seluruh butir dapat dipertahankan sehingga instrumen ini layak digunakan sebagai sarana evaluasi awal kesiapan mahasiswa menghadapi UKPPPG.

SARAN

Penelitian ini merekomendasikan agar pengembangan instrumen selanjutnya memperhatikan keseimbangan distribusi tingkat kesulitan, memperkuat daya beda, serta memperbaiki kualitas distraktor. Selain itu, uji coba pada populasi yang lebih luas dengan melibatkan variasi konteks institusi pendidikan disarankan agar instrumen memiliki generalisasi yang lebih tinggi. Dengan demikian, perangkat tes *try out* tidak hanya berfungsi sebagai simulasi, tetapi juga sebagai instrumen prediktif yang mendukung pencapaian standar kompetensi calon guru profesional di bidang Layanan Kesehatan.

REFERENCE

- Abbas, A., Amaliawati, S., Aulia, N., & Agustiningrum, T. (2023). PPG Students' Perception of PPG Program in Developing Their Teaching Ability. *Global Synthesis in Education Journal*, 1(2), 65–74. <https://doi.org/https://doi.org/10.61667/9y6ksm24>
- Atmoko, A. (2020). Pengembangan Tes Potensi Keberhasilan Akademik pada Mahasiswa Program Studi Bimbingan dan Konseling. *Jurnal Konseling Gusjigang*, 6(2), 72–84. <https://doi.org/10.24176/jkg.v6i2.2972>
- Cronbach, L. J. (1990). *Essentials of Psychological Testing* (Fifth Edition). Harper & Row.
- Direktur Jenderal Guru dan Tenaga Kependidikan, Sekretariat Direktorat Jenderal Guru dan Tenaga Kependidikan, & Plt. Direktur Pendidikan Profesi Guru. (2024). *Petunjuk Teknis Bagi Peserta Uji Kompetensi Peserta Pendidikan Profesi Guru (UKPPPG) Bagi Guru Tertentu Tahun Ajaran 2024/2025*. Direktorat Jenderal Guru dan Tenaga Kependidikan. <https://ppg.kemdikbud.go.id/storage/Jul%202024-Juknis%20Peserta%20UKPPPG%20Bagi%20Guru%20Tertentu.pdf>
- Dorfman, L. Y., & Kalugin, A. Y. (2020). Resources, Potentials and Academic Achievements of Students. Part 2. From Differentiation to Integration of Resources, Potentials and Academic Achievements of Students. *Obrazovanie i Nauka*, 22(5), 90–110. <https://doi.org/10.17853/1994-5639-2020-5-90-110>
- Febriani, E. W., Nugroho, P., Mufidah, N., & Aditya, S. R. (2024). Pengembangan Profesionalisme Guru Melalui PKG, PKB, dan PPG (Dalam Jabatan dan Pra Jabatan). *SIGNIFICANT: Journal Of Research and Multidisciplinary*, 3(2), 87–103. <https://doi.org/10.62668/significant.v3i02.1290>

- Kim, J.-Y. (2020). A longitudinal study of the relation between creative potential and academic achievement at an engineering university in Korea. *JEEE: Journal of Engineering Education*. <https://doi.org/https://doi.org/10.1002/jee.20365>
- Muhali, M. (2019). Pembelajaran Inovatif Abad Ke-21. *Jurnal Penelitian Dan Pengkajian Ilmu Pendidikan: E-Saintika*, 3(2), 25–30. <https://doi.org/10.36312/e-saintika.v3i2.126>
- Prodi PPG - Universitas Negeri Malang. (2025). *Sosialisasi UKPPPG Bagi Guru Tertentu* (pp. 1–95). Prodi PPG - Universitas Negeri Malang.
- Suryabrata, S. (2005). *Pengembangan Alat Ukur Psikologis*. Andi Offset.