

**PENGEMBANGAN INSTRUMEN TRYOUT PPG BERBASIS HIGHER ORDER  
THINKING SKILLS DAN PEDAGOGICAL CONTENT KNOWLEDGE UNTUK  
GURU IPS**

Neni Wahyuningtyas<sup>1</sup>, Fatiya Rosyida<sup>2</sup>  
<sup>1, 2</sup>PPG, Sekolah Pascasarjana, Universitas Negeri Malang  
[1neni.wahyuningtyas.fis@um.ac.id](mailto:neni.wahyuningtyas.fis@um.ac.id), [2fatiya.rosyida.fis@um.ac.id](mailto:fatiya.rosyida.fis@um.ac.id)

**ABSTRACT**

*21st century education requires professional teachers who are able to foster 6C skills, while Indonesia's PISA 2022 achievements are still low. This study aims to develop a tryout instrument for students of Teacher Professional Education (PPG) in the field of social studies at Malang State University to improve their readiness for the Computer-Based Written Test (UTBK). The method used was Research and Development (R&D) with a 4D model. The instrument consisted of 35 questions based on Higher Order Thinking Skills (HOTS) and Pedagogical Content Knowledge (PCK), which were validated by experts, revised, and tested on 950 students. The results showed that 82.86% of the questions were valid, with 100% reliability, good to very good discrimination, a balanced level of difficulty (40% easy, 42.86% medium, 17.14% difficult), and a distractor effectiveness of 63.44%. These findings confirm that the tryout questions are suitable for use as an adaptive assessment instrument to measure the professional competence of social studies teachers.*

**Keywords:** *Teacher Professional Education (PPG), Tryout Test Development, HOTS, PCK*

**ABSTRAK**

Pendidikan abad ke-21 menuntut guru profesional yang mampu menumbuhkan keterampilan 6C, sementara capaian PISA 2022 Indonesia masih rendah. Penelitian ini bertujuan mengembangkan instrumen soal *tryout* bagi mahasiswa Pendidikan Profesi Guru (PPG) bidang studi IPS di Universitas Negeri Malang untuk meningkatkan kesiapan menghadapi Ujian Tertulis Berbasis Komputer (UTBK). Metode yang digunakan adalah *Research and Development* (R&D) dengan model 4D. Instrumen berupa 35 soal berbasis HOTS dan *Pedagogical Content Knowledge* (PCK), divalidasi ahli, direvisi, serta diuji pada 950 mahasiswa. Hasil menunjukkan 82,86% soal valid, reliabilitas 100%, daya beda baik–sangat baik, tingkat kesulitan seimbang (40% mudah, 42,86% sedang, 17,14% sulit), dan efektivitas distraktor 63,44%. Temuan ini menegaskan soal *tryout* layak digunakan sebagai instrumen asesmen adaptif untuk mengukur kompetensi profesional guru IPS.

**Kata Kunci :** Pendidikan Profesi Guru (PPG), Pengembangan Soal *Tryout*, HOTS, PCK

## **A. Pendahuluan**

Pendidikan abad ke-21 menuntut guru untuk menjadi agen transformasi yang mampu menyiapkan peserta didik menghadapi dinamika global yang semakin dinamis. Tugas seorang guru tidak hanya sebagai penyampai pengetahuan, tetapi juga fasilitator yang membimbing peserta didik untuk mengembangkan potensi diri (Munirah et al., 2025). Guna mendukung tujuan tersebut, keterampilan abad 21 berupa *critical thinking* (berfikir kritis), *communication* (komunikatif), *collaboration* (kolaboratif), *creativity* (kreatifitas), *citizenship* (kewarganegaraan), dan *character* (karakter), yang dikenal dengan 6C menjadi sangat penting untuk ditanamkan sejak dini.

Enam keterampilan abad 21 menjadi fondasi utama dalam membekali peserta didik untuk menghadapi tantangan kompleks di era globalisasi. *Critical thinking* membantu mengasah keterampilan peserta didik untuk melakukan analisis permasalahan secara mendalam; *communication* menstimulasi keterampilan peserta didik dalam menyampaikan ide atau gagasan; *collaboration* berusaha

membangun keterampilan dalam bekerja sama; *creativity* membantu menumbuhkan keterampilan kreatif dan inovatif; *citizenship* mendorong solidaritas kemanusiaan; sedangkan *character* menuntun peserta didik untuk selalu memegang teguh nilai moral dan etika (Anggraeni et al., 2023).

Dalam sistem pendidikan Indonesia, penguasaan keterampilan 6C masih menjadi tantangan besar. Pernyataan ini dibuktikan dengan adanya hasil survei *Programme for International Student Assessment* (PISA) 2022 yang menunjukkan bahwa Indonesia hanya meraih skor rata-rata 383 dalam literasi sains, jauh tertinggal dari rata-rata OECD yaitu 485 (OECD, 2023). Laporan resmi PISA 2022 Volume I menunjukkan bahwa peserta didik Indonesia masih lemah dalam penalaran ilmiah dan pemecahan masalah berbasis sains.

Rendahnya capaian ini bukan sekadar angka statistik, tetapi cerminan lemahnya kemampuan peserta didik dalam menghubungkan konsep akademik dengan realitas sehari-hari. Padahal, kemampuan tersebut berkaitan erat dengan komponen utama 6C, terutama *critical thinking* dan *creativity*. Tanpa

penguasaan dua aspek ini, peserta didik akan kesulitan beradaptasi dengan kebutuhan abad ke-21 yang berbasis inovasi dan pemecahan masalah.

PISA 2022 juga menilai domain baru, yaitu *creative thinking*. Hasilnya memperlihatkan hanya 31% peserta didik Indonesia yang mampu mencapai level dasar (Level 3) dalam kreativitas berpikir, jauh di bawah rata-rata OECD (OECD, 2024). Temuan ini semakin menegaskan bahwa sistem pendidikan Indonesia memerlukan transformasi besar pada pola belajar berbasis hafalan menuju pendekatan yang mampu memberikan dorongan peserta didik untuk mengembangkan berbagai ide baru, menghasilkan inovasi, dan mengeksplorasi masalah dengan kreatif.

Data-data di atas menunjukkan urgensi peningkatan kualitas guru demi tercapainya pendidikan yang adaptif dan selaras dengan perkembangan zaman. Guru profesional tidak hanya mampu menguasai teori, tetapi juga kompetensi pedagogik yang mampu mengembangkan keterampilan 6C (Ramadhanti, 2024). Dengan demikian, Program Pendidikan Profesi Guru (PPG) memiliki peran yang

sangat penting dalam mempersiapkan guru profesional yang visioner dan adaptif terhadap perkembangan zaman.

Dalam hal ini, Universitas Negeri Malang (UM) sebagai salah satu Lembaga Pendidikan Tenaga Kependidikan (LPTK) terkemuka di Indonesia memiliki komitmen penuh untuk melahirkan guru yang tidak sekadar mahir dalam penguasaan materi, tetapi juga mampu merancang dan mengelola yang efektif, adaptif dan sesuai karakteristik kebutuhan peserta didik. Bentuk komitmen ini diwujudkan melalui berbagai inovasi pembelajaran, salah satunya melalui pengembangan soal *tryout* yang diperuntukkan untuk mahasiswa PPG (Pendidikan Profesi Guru) Bagi Guru Tertentu.

Soal *tryout* yang dikembangkan tidak hanya bertujuan untuk mengukur sejauh mana kompetensi mahasiswa PPG Bagi Guru Tertentu, tetapi juga memiliki makna strategis sebagai sarana pembelajaran. Soal *tryout* membantu melatih mahasiswa PPG Bagi Guru Tertentu untuk mengembangkan kemampuan analitis, kreatif dalam menyelesaikan masalah kompleks, serta melatih mereka untuk mampu

menghubungkan konsep akademik dengan realitas sosial budaya yang berada di sekitarnya.

Lebih lanjut, pemilihan desain soal yang selaras dan adaptif dengan kebutuhan zaman juga membantu mahasiswa PPG Bagi Guru Tertentu untuk dapat membiasakan diri mengenal instrumen asesmen yang tidak hanya menguji aspek kognitif, tetapi juga afektif dan psikomotorik. Cara ini juga berperan sebagai sarana pembekalan guru agar mampu menyeimbangkan kompetensi akademik dengan penguatan karakter dan kepedulian sosial, sesuai prinsip keterampilan 6C. Hasil dari penilaian soal ini tidak hanya melatih mahasiswa PPG Bagi Guru Tertentu, tetapi juga sebagai sarana refleksi kemampuan guru dalam menghadirkan pembelajaran yang kontekstual dan transformatif.

Penelitian ini bertujuan untuk menganalisis hasil pengembangan soal *tryout* untuk mahasiswa PPG Bagi Guru Tertentu. Hasil analisis diharapkan mampu memberikan gambaran secara kompleks tentang kualitas soal dan relevansinya dengan perkembangan zaman. Selain itu, hasil pengembangan juga diharapkan dapat menjadi sarana efektif bagi

Universitas Negeri Malang dalam mencetak guru profesional abad 21, sehingga lulusan PPG Bagi Guru Tertentu mampu hadir sebagai guru yang inovatif, berkarakter, dan adaptif dengan dinamika global.

## **B. Metode Penelitian**

Penelitian ini menggunakan pendekatan penelitian dan pengembangan R&D (*Research and Development*). Model R&D yang diterapkan adalah model 4D. Menurut Thiagarajan et al. (1974), model 4D terdiri dari empat tahap:

### **1. Tahap *Define* (Pendefinisian)**

Tahap ini berfokus pada penentuan produk yang akan dikembangkan, termasuk spesifikasi dan karakteristiknya. Kegiatan yang dilakukan pada tahap ini meliputi analisis masalah, analisis mahasiswa, analisis tugas, dan eksplorasi solusi alternatif. Proses pendefinisian dilakukan melalui studi pustaka, dengan memanfaatkan buku, jurnal, dan bahan referensi lainnya.

### **2. Tahap *Design* (Perancangan)**

Pada tahap ini, dilakukan perancangan kerangka instrumen soal *try out*. Langkah ini dilakukan setelah menganalisis materi

pembelajaran dan menetapkan tujuan pembelajaran. Tahap ini meliputi penentuan indikator pertanyaan untuk setiap tujuan dan pemilihan format serta level tes. Hasil dari tahap ini adalah draf awal instrumen evaluasi.

3. Tahap *Develop* (Pengembangan)

Tahap pengembangan instrumen evaluasi dilakukan berdasarkan kisi-kisi pertanyaan yang telah disiapkan. Instrumen yang dihasilkan kemudian divalidasi oleh ahli dan direvisi berdasarkan rekomendasi yang didapatkan. Apabila sudah dinyatakan layak oleh validator, maka dapat dilanjutkan uji coba produk.

4. Tahap *Disseminate* (Penyebaran)

Tahap ini meliputi penerapan atau penggunaan instrumen soal yang telah dikembangkan dalam pembelajaran PPG Bagi Guru Tertentu.

Penelitian ini dilaksanakan di Universitas Negeri Malang yang berlokasi di Jalan Semarang No. 5, Kota Malang, Jawa Timur. Penelitian dilakukan pada bulan Juni hingga Juli 2025. Subjek penelitian adalah seluruh mahasiswa PPG Bagi Guru Tertentu Universitas Negeri Malang

pada Bidang Studi IPS yang dipilih menggunakan teknik *purposive sampling*. Penelitian ini menggunakan kuesioner terbuka dan tertutup untuk memperoleh masukan dari para validator dan responden uji coba. Jawaban dari kuesioner terbuka kemudian dianalisis untuk menentukan modus. Sementara itu, dalam kuesioner tertutup, responden memiliki kesempatan untuk mempertimbangkan jawaban sesuai opsi yang telah diberikan. Kuesioner ini menggunakan skala Likert dengan skor penilaian, yaitu Sangat Setuju (SS) dengan nilai = 4, Setuju (S) dengan nilai = 3, Tidak Setuju (TS) dengan nilai = 2, Sangat Tidak Setuju (STS) dengan nilai = 1.

Teknik analisis data dilakukan apabila peneliti telah memperoleh seluruh data yang dibutuhkan dalam penelitian. Data tersebut meliputi hasil perolehan data dari uji validasi dan uji coba produk, yang kemudian dihitung dengan menggunakan rumus sebagai berikut:

$$P = \frac{\Sigma x}{\Sigma xi} \times 100 \%$$

Gambar 1. Penghitungan Persentase

Keterangan :

P = Persentase Validitas

X = Skor penilaian dalam satu item

Xi = Skor penilaian ideal dalam satu item

100% = Konstanta

Setelah menganalisis data hasil validasi, peneliti kemudian menafsirkan data dan menyimpulkannya dengan mengacu pada standar penilaian Arikunto (2018) sebagai berikut:

**Tabel 1 Kriteria Validitas Instrumen**

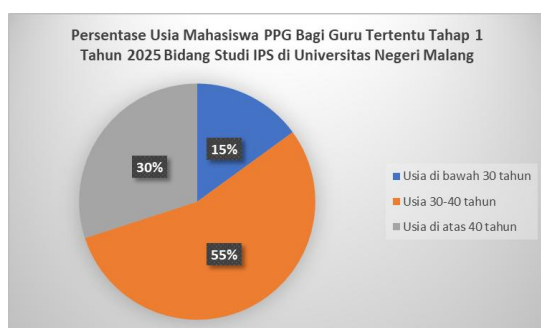
Presentase	Kriteria	Kategori
81-100%	Sangat Valid	Tidak Revisi
61-80%	Valid	Revisi Sebagian
41-60%	Cukup Valid	Revisi Sebagian (Uji Coba Ulang)
21-40%	Kurang Valid	Revisi Total

### **C. Hasil Penelitian dan Pembahasan**

Soal *try out* dilakukan dengan model 4D dilakukan selama 2 bulan (Juni-Juli 2025). Tahap *Define* dalam model pengembangan 4-D Thiagarajan et al. (1974) merupakan fase penting dalam mengidentifikasi

dan merumuskan permasalahan, kebutuhan, dan tujuan spesifik sebelum kegiatan merancang soal. Tahap ini menjamin relevansi dan efektivitas solusi yang dikembangkan dalam pembuatan soal. Pada tahap ini mencakup empat komponen utama: analisis masalah, analisis mahasiswa, analisis tugas, dan eksplorasi solusi alternatif. Analisis masalah yang telah dilakukan menunjukkan bahwa terdapat permasalahan penting yang dihadapi oleh mahasiswa Program Pendidikan Profesi Guru (PPG) Bagi Guru Tertentu Tahap 1 Tahun 2025, terutama terkait tantangan dalam mempersiapkan Ujian Tertulis Berbasis Komputer (UTBK). Temuan ini diperkuat bahwa soal-soal *tryout* terstruktur saat ini belum tersedia dan berfungsi sebagai sumber belajar mandiri, sehingga membuat mahasiswa tidak siap (Putri et al., 2024; Sudarsono et al., 2024). Analisis karakteristik mahasiswa dilakukan untuk menunjang kebutuhan dalam menyusun soal *try out*. Hasil dari analisis ini menggambarkan karakteristik demografis dan akademik mahasiswa. Mahasiswa dalam program PPG Bagi Guru Tertentu menunjukkan keragaman yang signifikan terkait asal

geografis, usia, dan pengalaman mengajar, sehingga membutuhkan fleksibilitas dan adaptabilitas dalam desain produk. Hal ini sejalan dengan prinsip-prinsip pembelajaran orang dewasa yang menekankan pentingnya sumber belajar yang relevan dan dapat diakses secara mandiri (Saputra et al., 2024; Suharyani et al., 2025).



**Grafik 1. Persentase Usia Mahasiswa PPG Bagi Guru Tertentu Tahap 1 Tahun 2025 Bidang Studi IPS di Universitas Negeri Malang**



**Grafik 2. Persentase Asal Provinsi Mahasiswa PPG Bagi Guru Tertentu Bidang Studi IPS di Universitas Negeri Malang**

Berdasarkan grafik 1 data menunjukkan bahwa distribusi usia mahasiswa sebagian besar terkonsentrasi pada kelompok usia produktif. Lima puluh lima persen peserta didik berada dalam rentang usia 30 hingga 40 tahun. Kelompok usia di bawah 30 tahun dan di atas 40 tahun memiliki persentase yang lebih kecil, masing-masing sebesar 15% dan 30%. Proporsi ini menunjukkan bahwa mahasiswa sebagian besar adalah guru berpengalaman, sejalan dengan tujuan program Pendidikan Profesi Guru (PPG) untuk meningkatkan kompetensi pendidik saat ini. Secara geografis, sebaran asal mahasiswa tersebar luas di seluruh Indonesia, dengan konsentrasi yang signifikan di beberapa provinsi. Berdasarkan grafik 2, Sekitar 20% mahasiswa berasal dari Provinsi Jawa Timur, diikuti oleh 18% dari Jawa Tengah, dan 12% dari Jawa Barat. Provinsi lain, seperti Sumatera Selatan dan Sumatera Utara, masing-masing menyumbang 8% dan 7% dari total populasi mahasiswa. Data ini menegaskan jangkauan program secara nasional, meskipun terdapat aglomerasi mahasiswa di daerah padat penduduk, seperti Jawa.

Selanjutnya, analisis tugas dilakukan dengan meninjau Sistem Manajemen Pembelajaran (LMS) yang digunakan. Hal tersebut menunjukkan bahwa tidak adanya sumber daya yang secara eksplisit ditujukan untuk persiapan UTBK. Analisis ini menegaskan bahwa salah satu tugas utama yang harus dikuasai mahasiswa, yaitu menjawab soal UTBK, tidak didukung oleh modul atau bank soal yang memadai, sehingga menciptakan kebutuhan mendesak untuk intervensi. Berdasarkan temuan dari ketiga analisis ini, eksplorasi solusi alternatif menghasilkan pengembangan soal *tryout* sebagai solusi yang paling praktis dan efisien.

Tahap selanjutnya adalah *Design* (membuat rancangan), Pada tahap ini dilakukan untuk menentukan validitas, reliabilitas, dan kualitas dari soal *tryout*. Soal *tryout* IPS dirancang dengan menggunakan studi kasus kontekstual untuk menekankan pendekatan Keterampilan Berpikir Tingkat Tinggi (HOTS) (Adi et al., 2025; Tasrif, 2023) Hal ini dilakukan agar kemampuan analisis, evaluasi, dan pemecahan masalah yang merupakan kompetensi profesional guru dapat diukur. Soal dirancang dengan pendekatan *Pedagogical*

*Content Knowledge* (PCK) sebanyak 35 soal. Kombinasi dari PCK dengan konten studi kasus kontekstual menjamin bahwa soal-soal tersebut tidak hanya menilai penerapan materi dalam skenario dunia nyata, tetapi juga penguasaan materi tersebut (Kind & Chan, 2019; Sutamrin et al., 2022).

Soal *tryout* ini dirancang dengan menggunakan 2 jenis untuk memberikan variasi dalam penilaian. Jenis soal dirancang dengan Pilihan Ganda Sederhana dan Pilihan Ganda Kompleks. Pilihan ganda kompleks dirancang sebanyak 10 soal. Jenis soal ini meliputi, multi jawaban, pernyataan benar-salah, fakta-opini, atau menjodohkan (Hajaroh et al., 2025). Hal ini bertujuan untuk menilai kemampuan kognitif tingkat tinggi. Soal pilihan ganda sederhana dirancang sebanyak 25 soal. Soal ini digunakan untuk mengukur pemahaman dasar hingga menengah. Kombinasi jenis soal dapat meningkatkan validitas instrumen (Fossett et al., 2004; Moreno-Llamas et al., 2020; Widiyawati et al., 2019). Tingkat kesukaran soal menggunakan prinsip kurva normal, yang menyatakan bahwa distribusi soal mudah, sedang, dan sulit secara



proporsional. Distribusi ini menjamin bahwa instrumen memiliki daya pembeda terbaik agar mampu membedakan tingkat kompetensi peserta (Saputri et al., 2023; Shen et al., 2022).

Soal dirancang berdasarkan standar kompetensi yang tercantum dalam Perdirjen 2626/B/HK/04.01/2023 untuk Guru Ahli Pertama Level 2. Kompetensi ini meliputi pedagogik, sosial, kepribadian, dan profesional. Konten materi pada soal disusun pada konsep bidang studi IPS yang wajib dikuasai oleh seorang Guru Ahli Pertama Level 2. Hal ini dilakukan untuk memastikan bahwa isi soal relevan dan sejalan dengan persyaratan profesi guru. Untuk menghasilkan instrumen yang memiliki akuntabilitas profesional dan relevansi substantif, soal harus disusun menggunakan kerangka kompetensi ini. Untuk menjamin keseragaman dan kemudahan administrasi, setiap soal disusun menggunakan templat terstruktur yang mencakup soal, kunci jawaban, dan pedoman penskoran (Simalango, 2021).

Soal akan diunggah pada *Learning Management System* (LMS)

sehingga perlu mempertimbangkan faktor teknis untuk memudahkan administrasi dan implementasi. Stimulus dicantumkan pada setiap soal yang digunakan untuk maksimal tiga pertanyaan. Meskipun repetitif, teknik ini penting untuk menjamin bahwa dalam bank soal dapat diacak secara independen tanpa kehilangan konteks. Metode ini mendukung integritas proses penilaian dengan mengurangi kesalahan unggah administrator dan menjamin fleksibilitas dalam membuat set pertanyaan untuk setiap mahasiswa (Adiarta & Divayana, 2019; Kasidi et al., 2022). Agar pelaksanaan tryout dalam skala besar dapat dilaksanakan dengan lancar, standarisasi ini merupakan langkah awal yang penting.

Tahap ketiga dari model 4D adalah *Develop* (Pengembangan) meliputi pengembangan soal, validasi, uji coba produk. Pengembang soal yang menyeluruh sebagai dasar pengembangan butir soal tes diperlukan untuk menghasilkan alat evaluasi yang valid dan relevan. Pengembangan soal dilakukan dengan mempertimbangkan bahan kajian soal yang didasarkan pada Perdirjen 2626/B/HK/04.01/2023

untuk Guru Ahli Pertama Level 2, yang mencakup unsur pedagogis, evaluatif, dan teoritis. Fokus utama adalah pada strategi lingkungan pembelajaran yang aman dan nyaman, yang merupakan bagian penting untuk mendorong terlaksana pembelajaran yang menarik (Antoniuk et al., 2023; Sholeh, 2023). Soal-soal ini juga bertujuan untuk menilai kemampuan guru dalam menerapkan pembelajaran yang berpusat pada peserta didik, yang mencakup pemilihan sumber belajar yang relevan dan penyusunan pembelajaran yang relevan dengan konteks lokal. Dengan integrasi elemen-elemen ini, soal tryout dapat menilai kompetensi guru secara lebih komprehensif, bukan hanya penguasaan materi. Dengan pertanyaan yang disusun dengan materi pembelajaran tertentu, seperti strategi pembelajaran IPS pada mata pelajaran seperti masyarakat multikultural, konflik sosial, dan ketimpangan sosial ekonomi, keterkaitan antara materi pembelajaran dan konten IPS juga krusial. Metode ini memungkinkan penilaian kompetensi guru dalam menggunakan strategi pengajaran

yang sesuai dengan dinamika sosial dan situasi permasalahan aktual.

Selain itu, bahan kajian soal difokuskan pada pemahaman guru tentang penelitian tindakan kelas (PTK), penilaian, dan instrumen kurikulum. Soal-soal tersebut menilai pemahaman guru tentang elemen-elemen kurikulum dan cara penerapannya, serta kapasitas mereka dalam menilai hasil belajar siswa. Pemanfaatan PTK sebagai bahan kajian soal menitikberatkan pada pentingnya guru sebagai peneliti yang dapat menerapkan teknik-teknik penelitian untuk meningkatkan standar pembelajaran di kelas (Oestar & Marzo, 2022; Risan, 2025). Soal-soal ini akan menguji praktik sekaligus memberikan landasan teori yang kokoh karena penekanannya pada kajian teoritis seperti hakikat evaluasi, teori pembelajaran, dan model evaluasi pembelajaran serta teori-teori psikologis seperti teori kognitif dan sosial. Bahan kajian soal ini menciptakan kerangka kerja menyeluruh yang memungkinkan terciptanya soal-soal yang andal, valid, dan relevan dengan persyaratan kompetensi profesional guru.

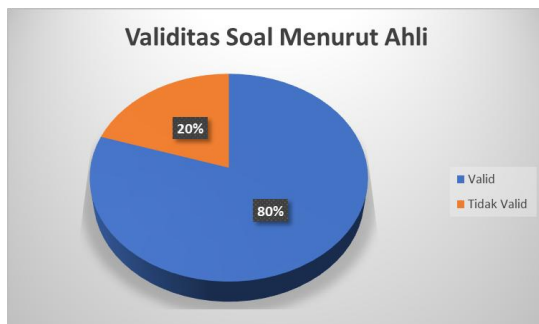
Pengembangan soal ini menggabungkan materi pelajaran yang relevan dengan konteks pedagogis dalam pendekatan *Pedagogical Content Knowledge* (PCK) dalam bidang studi IPS. Soal yang dikembangkan tidak hanya menguji pengetahuan faktual, tetapi untuk menilai kemampuan guru dalam mengajar mata pelajaran penting IPS. Topik-topik materi pelajaran seperti perdagangan internasional, masyarakat multikultural, konflik dan integrasi sosial, keragaman sosiokultural, ketimpangan sosioekonomi, pembangunan daerah, dan tata ruang digunakan untuk mengukur seberapa baik guru dapat mengimplementasikan materi dalam strategi pembelajaran yang efektif. Strategi ini sejalan dengan konsep PCK, yang menekankan pemahaman guru tentang cara menyampaikan materi pembelajaran dengan cara yang mudah dipahami siswa (Sutamrin et al., 2022).

Integrasi topik penting materi pembelajaran IPS dilakukan agar dapat t mengevaluasi seberapa baik guru mampu menyusun rencana pembelajaran yang mengatasi permasalahan sosial saat ini. Misalnya, soal tentang ketimpangan

sosial ekonomi di Indonesia menilai pengetahuan siswa tentang definisi serta kemampuan mereka untuk mengajarkan konsep tersebut secara sensitif dan efektif menggunakan teknik seperti analisis data dan studi kasus lokal. Selain itu,, pertanyaan tentang konflik dan integrasi sosial dirancang untuk menilai kapasitas guru dalam memfasilitasi diskusi produktif dan menyampaikan teknik pemecahan masalah sosial. Hasilnya, materi pembelajaran dalam bidang studi ini memastikan bahwa pertanyaan-pertanyaan tersebut memiliki validitas tinggi untuk mengevaluasi PCK guru secara menyeluruh dengan menjembatani kesenjangan antara konten studi sosial dan praktik pembelajaran kreatif.

Setelah dilakukan pengembangn soal, langkah selanjutnya adalah validasi soal yang dilakukan oleh ahli. Hasil validasi soal adalah sebagian besar dari 35 pertanyaan yang direview dinyatakan sah dan layak digunakan. Melalui proses review ini, terbukti bahwa Soal yang disusun telah memenuhi persyaratan penting, termasuk kesesuaian dengan kompetensi yang diuji dan keselarasan antara deskripsi masalah

dan pilihan jawaban. Secara keseluruhan, soal tryout berkualitas baik dan dapat digunakan, meskipun beberapa pertanyaan perlu diperbaiki. Tujuh pertanyaan perlu diperbaiki atau dinyatakan tidak valid, sementara dua puluh delapan pertanyaan dinyatakan valid. Sebagian besar koreksi mengatasi kesalahan dalam pilihan jawaban, seperti beberapa jawaban yang benar atau pilihan jawaban yang tidak terlalu relevan dengan permasalahan yang dibahas.



Grafik 3. Validitas Soal Menurut Ahli

Proses review menghasilkan sejumlah temuan penting sebagai dasar revisi soal sebagai berikut:

1. Hasil review menunjukkan bahwa terdapat beberapa pilihan jawaban dengan jawaban yang benar dalam soal Pilihan Ganda Kompleks. Hal ini menunjukkan bahwa untuk memastikan tidak ada ambiguitas, validitas setiap pilihan jawaban perlu ditinjau.
2. Mayoritas soal mendapatkan peringkat "selaras" dan "sangat tepat" antara pilihan jawaban dan deskripsi masalah (stem). Ketidakkonsistenan ini terlihat pada sejumlah soal yang dianggap tidak valid, karena pilihan jawaban yang diberikan tidak cukup menjawab permasalahan.
3. Secara khusus, disarankan agar frasa "model pengajaran" diubah menjadi "model pembelajaran" dan kesalahan ketik juga ditemukan. Hal ini menggambarkan betapa pentingnya menulis pertanyaan dengan teliti.
4. Peninjauan kembali pilihan jawaban untuk sejumlah pertanyaan, terutama yang dianggap kurang valid. Hal ini menunjukkan bahwa kegunaan dan relevansi distraksi, selain keakuratan jawaban utama, menentukan validitas pertanyaan.

Langkah selanjutnya adalah revisi soal berdasarkan review ahli dan pengunggahan pada LMS. Setelah itu dilakukan uji coba pada mahasiswa PPG Bagi Guru Tertentu Tahap 1 Tahun 2025 dalam skala besar dengan jumlah responden 950 mahasiswa, Hasil dari uji coba tersebut secara umum telah

memenuhi syarat. Pada penelitian ini, analisis yang dilakukan adalah Validitas, reliabilitas, daya pembeda, tingkat kesulitan, dan efektivitas distraktor. Berdasarkan temuan, sebagian besar butir soal memenuhi persyaratan, sehingga layak digunakan sebagai alat asesmen. Sebanyak 82,86% dari 35 butir soal yang diuji dinyatakan valid, artinya butir soal tersebut berhasil menilai kompetensi yang dirancang. Hanya 17,14% item yang perlu direvisi karena tidak valid, misalnya pilihan jawaban yang tidak jelas. Selain itu, 100% item dalam instrumen memiliki skor reliabilitas di atas 0,40, yang menunjukkan reliabilitas yang sangat baik. Konsistensi ini menunjukkan bahwa soal *tryout* telah dapat menghasilkan hasil yang konsisten dan dapat dipercaya.

Hasil skor daya diskriminatif menunjukkan lebih besar dari 0,20, analisis daya diskriminatif menunjukkan bahwa semua butir soal (100%) efektif membedakan peserta berkemampuan tinggi dan rendah. Khususnya, 37,14% butir soal memiliki daya diskriminatif baik, 17,14% memiliki daya diskriminatif sangat baik, dan 45,71% memiliki daya diskriminatif cukup. Sebagian

besar butir soal terkonsentrasi pada kategori sedang (42,86%) dan mudah (40%) serta sisanya pada kategori sulit (17,14%), distribusi tingkat kesulitannya juga sempurna. Butir-butir soal dalam distribusi ini tidak terlalu mudah maupun terlalu sulit bagi mayoritas peserta. Hal ini menunjukkan bahwa tingkat kesukaran soal telah konsisten dengan praktik pengembangan tes. Uji efektivitas distraktor menunjukkan bahwa 63,44% di antaranya berhasil menarik peserta yang tidak mahir dalam mata pelajaran tersebut. Meskipun demikian, 36,56% distraktor masih tidak efektif. Hal ini menunjukkan bahwa pilihan jawaban distraktor perlu diperbarui agar lebih andal dan relevan.

**Tabel 2 Hasil Analisis Butir Soal *Tryout***  
**Bidang Studi IPS**

No	Analisis Butir Soal	Hasil
1.	Realibitas	100 % reliabel
2.	Validitas	82,86% Valid 17,14% Tidak Valid
3.	Daya Beda (Diskriminan)	45,71% Cukup 37,14% Baik 17,14% Sangat Baik
4.	Tingkat Kesulitan	40% Mudah 42,86% Sedang 17,14% Sulit

5.	Distraktor	63,44% Berfungsi dengan baik. 36,56% Tidak Berfungsi.
----	------------	--

Produk akhir berupa soal *tryout* bidang studi IPS, diimplementasikan pada sasaran selama tahap diseminasi, yang merupakan bagian penting dari siklus pengembangan. Selain diseminasi, prosedur ini menguji validitas dan reliabilitas instrumen dalam konteks dunia nyata dan skala yang lebih luas. Mahasiswa yang terdaftar dalam Pendidikan Profesi Guru (PPG) Khusus Guru Tahap 2 pada tahun 2025 dan PPG Dalam Jabatan Kementerian Agama untuk mata pelajaran umum pada tahun 2025 merupakan dua kelompok sasaran utama penelitian ini, dan soal tersebut didistribusikan dan diimplementasikan secara menyeluruh kepada mereka. Instrumen soal ini diimplementasikan secara independen dan sistematis dengan menggunakan LMS PPG. Strategi ini memungkinkan penyampaian tes *online* yang efektif kepada semua peserta dan menjangkau seluruh peserta tanpa kendala geografis. Penggunaan LMS dilakukan agar pengumpulan data dilakukan secara otomatis,

administrasi terpusat, dan pengacakan soal dilakukan untuk menjaga integritas pelaksanaan *tryout*. Selain memberikan hasil *tryout* kepada peserta, data yang dihasilkan dari fase diseminasi ini juga menyediakan banyak data empiris untuk analisis lebih lanjut.

#### **D. Kesimpulan**

Instrumen soal *tryout* ini memiliki kualitas yang baik dan layak digunakan sebagai alat asesmen, berdasarkan analisis menyeluruh terhadap soal-soal *tryout* IPS yang disusun menggunakan pendekatan kontekstual, PCK (*Pedagogical Content Knowledge*), dan HOTS (*Higher Order Thinking Skills*). Soal ini secara keseluruhan menunjukkan reliabilitas yang sangat tinggi (100%), menjamin konsistensi pengukuran, dan mayoritas soal (82,86%) dinyatakan valid. Selain menguji pengetahuan faktual, pertanyaan-pertanyaan yang disusun menggunakan skenario dunia nyata dan studi kasus kontekstual berhasil menilai penerapan PCK oleh guru. Distribusi tingkat kesulitan butir soal mengikuti aturan kurva normal, dan semua soal memiliki daya pembeda efektif 100%. Efektivitas pengecoh

ditemukan memiliki kekurangan, dengan 36,56% di antaranya tidak berfungsi optimal. Hal ini merupakan area signifikan yang perlu dikaji lebih lanjut. Saran yang dapat diberikan agar butir soal dengan distraksi non-fungsional direvisi secara menyeluruh untuk mengoptimalkan kualitas instrumen. Tujuan utama penyempurnaan ini adalah menciptakan pilihan jawaban distraktor yang lebih meyakinkan, relevan, dan sangat mirip dengan kunci jawaban. Selain itu, terutama untuk butir soal berbasis HOTS, validasi ulang butir soal yang sebelumnya tidak valid diperlukan untuk menjamin tidak ada lagi ambiguitas atau ketidaksesuaian antara stimulus dan pilihan jawaban. Setelah modifikasi ini, soal dapat diuji ulang pada sampel yang lebih besar untuk memverifikasi efikasi perubahan dan memastikan butir soal berfungsi sebaik mungkin sebagai alat untuk menilai kompetensi guru.

#### **DAFTAR PUSTAKA**

- Adi, D., Makki, M., & Novitasari, S. (2025). Analisis Kesulitan Siswa dalam Menyelesaikan Soal Berbasis HOTS (Higher Order Thinking Skill) Kelas IV Sekolah Dasar. *Pendas : Jurnal Ilmiah Pendidikan Dasar*, 10(03).
- Adiarta, A., & Divayana, D. G. H. (2019). Pengembangan Soal Test Digital Mata Kuliah Asesmen dan Evaluasi menggunakan Aplikasi Wondershare. *Jurnal Pendidikan Teknologi Dan Kejuruan*, 16(2). <https://doi.org/10.23887/jptk-undiksha.v16i2.19199>
- Anggraeni, P., Sunendar, D., Maftuh, B., & Sopandi, W. (2023). How Do 6Cs Appearance at Elementary Schools Learning? *DWIJA CENDEKIA: Jurnal Riset Pedagogik*, 7(2). <https://doi.org/10.20961/jdc.v7i2.75668>
- Antoniuk, O., Vertel, A., Kolodyazhna, A., Kyrychenko, R., & Mnozhynska, R. (2023). Creating a Comfortable Learning Environment: The Role of Teachers and Education Seekers. *Cadernos de Educação Tecnologia e Sociedade*, 16(1). <https://doi.org/10.14571/brajets.v16.n1.153-162>
- Arikunto, S. (2018). *Dasar-dasar evaluasi pendidikan*. Bumi Aksara.
- Fossett, T. R. D., McNeil, M. R., Doyle, P. J., Rubinsky, H., Nixon, S., Hula, W., & Brady, J. (2004). Assessing the validity of multiple-choice questions for RAPP story comprehension. *Aphasiology*, 18(5–7). <https://doi.org/10.1080/02687030444000066>

- Hajarah, S., Raehanah, R., & Syawahid, M. (2025). Pelatihan Penyusunan Instrumen HOTS, Pilihan Ganda Kompleks, dan Analisis Butir Soal Melalui JASP untuk Meningkatkan Kompetensi Guru. *Participative Journal: Jurnal Pengabdian Pada Masyarakat*, 5(1), 17–27. <https://doi.org/10.55099/participative.v5i1.180>
- Kasidi, Satyarini, M., & Muryati, S. (2022). *The Influence of Try Out Scores on National Examination Scores*. <https://doi.org/10.4108/eai.14-12-2021.2318347>
- Kind, V., & Chan, K. K. H. (2019). Resolving the Amalgam: Connecting Pedagogical Content Knowledge, Content Knowledge and Pedagogical Knowledge. *International Journal of Science Education*, 41(7). <https://doi.org/10.1080/09500693.2019.1584931>
- Moreno-Llamas, A., García-Mayor, J., & De la Cruz-Sánchez, E. (2020). Concurrent and Convergent Validity of a Single, Brief Question for Physical Activity Assessment. *International Journal of Environmental Research and Public Health*, 17(6). <https://doi.org/10.3390/ijerph17061989>
- Munirah, Rahman, M. H., & Bagiya. (2025). Analysis of the Teacher's Role as a Facilitator in Project-Based Learning to Enhance Student Creativity. *The Journal of Academic Science*, 2. <https://thejoas.com/index.php/>
- OECD. (2023). *PISA 2022 Results Factsheets Indonesia PUBE*. <https://oecdch.art/a40de1dbaf/C108>.
- OECD. (2024). *PISA 2022 Results Volume III: Creative Minds, Creative Schools*.
- Oestar, J., & Marzo, C. (2022). Teachers as Researchers: Skills and Challenges in Action Research Making. *International Journal of Theory and Application in Elementary and Secondary School Education*, 4(2). <https://doi.org/10.31098/ijtaese.v4i2.1020>
- Putri, A. Y., Usman, A., Muhimmah, H. A., & Hendratno, H. (2024). Pembuatan Soal Mandiri dan Diskusi Tertulis Sebagai Pendekatan Inklusif untuk Mahasiswa Generasi Z. *Jurnal Pendidikan Dan Pembelajaran Indonesia (JPPI)*, 4(4), 1429–1439. <https://doi.org/10.53299/jppi.v4i4.776>
- Ramadhanti, I. (2024). Teacher Strategies in Facing in the Challenges of 21st Century Education. *Visipena*, 14(2). <https://doi.org/10.46244/visipena.v14i2.2561>
- Risan, M. (2025). Schoolteachers and teacher educators: making sense of 'research' in the context of



- research-based teacher education. *Teachers and Teaching: Theory and Practice*, 31(4), 691–706. <https://doi.org/10.1080/13540602.2024.2383358>
- Saputra, A. S., Fahrezi, D. wulandari, & Afriza, A. (2024). Andragogi: Adaptasi Pembelajaran Orang Dewasa Pada Era Digitalisasi. *AL-MIKRAJ Jurnal Studi Islam Dan Humaniora (E-ISSN 2745-4584)*, 4(02), 501–513. <https://doi.org/10.37680/almikraj.v4i02.4528>
- Saputri, H. A., Zulhijrah, Larasati, N. J., & Shaleh. (2023). Analisis Instrumen Assesmen: Validitas, Reliabilitas, Tingkat Kesukaran, dan Daya Beda Butir Soal. *Jurnal Ilmiah PGSD FKIP Universitas Mandiri*, 9(5).
- Shen, S., Huang, Z., Liu, Q., Su, Y., Wang, S., & Chen, E. (2022). Assessing Student's Dynamic Knowledge State by Exploring the Question Difficulty Effect. *SIGIR 2022 - Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval*. <https://doi.org/10.1145/3477495.3531939>
- Sholeh, M. I. (2023). Strategi Efektif dalam Manajemen Pendidikan untuk Meningkatkan Kualitas Pembelajaran. *Tarbawi Ngabar: Jurnal of Education*, 4(2). <https://doi.org/10.55380/tarbawi.v4i2.462>
- Simalango, H. M. (2021). Pelatihan Pemanfaatan Aplikasi Word dalam Pembuatan Model Soal Ujian pada Guru SMA Maitreyawira Kota Batam. *WIDYABHAKTI Jurnal Ilmiah Populer*, 4(1). <https://doi.org/10.30864/widyabhakti.v4i1.271>
- Sudarsono, B., Listyaningrum, P., Tentama, F., & Ghozali, F. A. (2024). Developing learning and training within industry model to improve work readiness of vocational high school students. *International Journal of Evaluation and Research in Education*, 13(3), 1731–1739. <https://doi.org/10.11591/ijere.v13i3.26175>
- Suharyani, Purwoko, B., & Rosyanafi, R. J. (2025). Rekonstruksi Pendekatan Andragogi dan Pedagogi Dalam Meningkatkan Kualitas Pembelajaran Mahasiswa. *Jurnal Transformasi*, 11.
- Sutamrin, Rosidah, & Zaki, A. (2022). The Pedagogical Content Knowledge (PCK) of Prospective Teachers. *EduLine: Journal of Education and Learning Innovation*, 2(4). <https://doi.org/10.35877/454ri.eduline1291>
- Tasrif. (2023). Higher Order Thinking Skills (HOTS) dalam pembelajaran social studies di

sekolah menengah atas. *Jurnal Pembangunan Pendidikan: Fondasi Dan Aplikasi*, 10(1).  
<https://doi.org/10.21831/jppfa.v10i1.29347>

Thiagarajan, Semmel, & Semmel. (1974). *Instructional Development for Training Teachers of Exceptional Children: A Sourcebook*. Indiana Univ., Bloomington. Center for Innovation in. *Research and Development, Mc*.

Widiyawati, Y., Nurwahidah, I., & Sari, D. S. (2019). Pengembangan Instrumen Integrated Science Test Tipe Pilihan Ganda Beralasan untuk Mengukur HOTS Peserta Didik. *Saintifika*, 21(2).