



Available online at FACTUM; Jurnal Sejarah dan Pendidikan Sejarah
website: <https://ejournal.upi.edu/index.php/Factum>
ACTUM: Jurnal Sejarah dan Pendidikan Sejarah, 14(2), 331-342
RESEARCH ARTICLE



Analisis Performa Prediktif Soal Tryout Uji Kompetensi Peserta Pendidikan Profesi Guru Bidang Studi Sejarah

Aditya Nugroho Widiadi & Hariyono

Pendidikan Profesi Guru (PPG), Sekolah Pascasarjana, Universitas Negeri Malang
Correspondence Author: aditya_widiadi@um.ac.id

To cite this article: Widiadi, A. N., & Hariyono. (2025). Analisis performa prediktif soal tryout uji kompetensi peserta pendidikan guru bidang studi sejarah. *FACTUM: Jurnal Sejarah dan Pendidikan Sejarah*, 14(2), 331-342. <https://doi.org/10.17509/factum.v14i2.90429>.

Abstract

This study analyzes the predictive validity of the 2025 Tryout Examination items for the Teacher Professional Education Competency Test (UKPPPG) in History, developed by the Universitas Negeri Malang PPG team. The study aims to ensure the quality of evaluation instruments and assess their ability to predict participants' success in the official competency test. Employing a mixed-methods explanatory sequential design, quantitative analysis was conducted to evaluate item validity, reliability, difficulty, and discrimination indices using data from 45 PPG-BGT History participants, followed by qualitative interviews with three purposively selected respondents. Results indicate that while some items exhibited weaknesses in psychometric properties, the tryout was perceived as effective in helping participants understand the structure, content, and cognitive demands of the UKPPPG. The findings confirm the value of integrating psychometric evaluation with user perception analysis for improving test instrument accuracy. This study contributes theoretically to the refinement of teacher assessment standards and practically provides recommendations for enhancing item construction and predictive performance in professional teacher education.

Abstrak

Penelitian ini menganalisis validitas prediktif butir Tryout Ujian Kompetensi PPG (UKPPPG) Tahun 2025 bidang Sejarah yang dikembangkan oleh tim PPG Universitas Negeri Malang. Tujuannya untuk menjamin kualitas instrumen evaluasi dan menilai kemampuan tryout dalam memprediksi keberhasilan peserta pada ujian kompetensi resmi. Dengan desain *mixed-methods explanatory sequential*, analisis kuantitatif dilakukan terhadap 45 peserta PPG-BGT untuk menilai validitas, reliabilitas, tingkat kesulitan, dan daya pembeda butir, dilanjutkan dengan wawancara kualitatif pada tiga peserta terpilih. Hasil menunjukkan sebagian butir memiliki kelemahan psikometrik, namun tryout dinilai efektif membantu peserta memahami struktur, konten, dan tuntutan kognitif UKPPPG. Temuan menegaskan pentingnya integrasi analisis psikometrik dan persepsi pengguna dalam meningkatkan akurasi instrumen tes. Secara teoretis, penelitian ini memperkuat pengembangan standar asesmen guru profesional, sedangkan secara praktis memberikan rekomendasi untuk penyempurnaan butir soal dan peningkatan daya prediktif dalam sistem PPG.

© 2025 Kantor Jurnal dan Publikasi UPI

Article Info

Article History:

Received 19 Feb. 2025
Revised 20 Sept. 2025
Accepted 30 Sept. 2025
Available online 1 Oct. 2025

Keyword:

Item analysis; mixed-methods; predictive validity; teacher competency test; tryout exam; UKPPPG.

Kata Kunci:

Analisis butir soal; mixed-method tryout; ujian kompetensi guru; UKPPPG, validitas prediktif,

PENDAHULUAN

Keberhasilan Pendidikan Profesi Guru (PPG) sangat ditentukan oleh kesiapan peserta dalam menghadapi Ujian Kompetensi PPG (UKPPPG). Namun, data menunjukkan bahwa banyak peserta PPG-BGT masih mengalami kegagalan karena rendahnya kemampuan memahami struktur soal, tuntutan kognitif, serta strategi penyelesaian ujian. Tryout menjadi salah satu solusi strategis untuk memberikan pengalaman simulatif dan diagnostik, tetapi kualitasnya sangat bergantung pada validitas, reliabilitas, tingkat kesulitan, dan daya pembeda soal. Di Indonesia, sebagian besar tryout hanya digunakan sebagai latihan teknis dan jarang dianalisis secara psikometrik. Inilah gap utama penelitian yaitu belum banyak penelitian yang mengevaluasi performa prediktif soal tryout secara kuantitatif dan kualitatif secara bersamaan.

Pendidikan Profesi Guru (PPG) merupakan salah satu kebijakan pemerintah Republik Indonesia dalam menyiapkan guru yang profesional, kompeten, dan siap untuk menghadapi tuntutan dalam pelaksanaan pendidikan di sekolah. Melalui program ini, mahasiswa diharapkan memiliki empat kompetensi utama selaku pendidik, yakni kompetensi kepribadian, kompetensi profesional, kompetensi sosial, dan kompetensi pedagogik (Mardhatillah & Surjanti, 2023). Namun sayangnya, banyak diantara peserta PPG, khususnya dalam jabatan yang sekarang disebut sebagai PPG Bagi Guru Tertentu (PPG-BGT) tidak mampu menghadapi Ujian Kompetensi Peserta PPG (UKPPPG) dengan baik sehingga mengalami kegagalan dalam ujian tersebut.

Untuk itu, peserta PPG-BGT tersebut perlu dibekali program tryout untuk menyiapkan mereka menghadapi UKPPPG. Soal tryout harus dikembangkan agar memenuhi standar alat ukur yang baik, khususnya pada aspek validitas dan reliabilitas soal. Validitas butir soal akan menjamin instrumen untuk mengukur aspek kompetensi yang akan diukur.

Aspek reliabilitas butir soal diperlukan untuk menjamin konsistensi pengukuran ketika soal tersebut digunakan beberapa kali tes. Menguji instrumen soal tryout ini diperlukan guna memastikan kedua aspek tersebut, sekaligus menjadi penyedia data baik untuk mahasiswa peserta PPG maupun untuk penyelenggara pendidikan profesi (Thompson et al., 2020).

Universitas Negeri Malang selaku Lembaga Pendidik Tenaga Kependidikan (LPTK) mengembangkan soal tryout UKPPPG agar mahasiswa peserta pendidikan profesi mampu menghadapi uji kompetensi tersebut dan sekaligus membantu agar bisa lulus sesuai dengan ambang kelulusan. Soal tryout ini dapat berfungsi untuk mengukur kemampuan mahasiswa sekaligus menyiapkan mental mahasiswa agar memahami bentuk soal yang akan dihadapinya. Hal ini dikarenakan soal tryout tidak hanya berfungsi sebagai sarana simulasi melainkan juga menjadi sarana diagnostik untuk menjamin kualitas pembelajaran (Atkins et al., 2022).

Agar fungsi simulasi dan diagnostik tersebut dapat berjalan, maka soal tryout yang telah dikembangkan harus dianalisis terlebih dulu. Kajian ini bermaksud menganalisis performa prediktif soal tryout UKPPPG yang digunakan di Universitas Negeri Malang (Šuster et al., 2023). Analisis dilakukan terhadap soal tryout untuk mahasiswa PPG-BGT dalam bidang studi sejarah. Hal ini dikarenakan belum ada kajian serupa yang telah dilakukan terhadap soal tryout tersebut. Tujuan analisis performa prediktif ini dilakukan agar soal tryout yang ada mampu menjalankan perannya dengan baik dalam membantu mahasiswa PPG-BGT dalam mengerjakan UKPPPG sehingga dapat lulus dengan baik.

Secara internasional, kajian evaluasi asesmen guru telah berkembang pesat. Darling-Hammond (2017) dan Newton (2018) menekankan bahwa instrumen asesmen profesional harus memenuhi standar validitas dan reliabilitas tinggi untuk memastikan fairness dan akurasi prediktif. Ferrara et al. (2022) menegaskan bahwa kualitas butir soal

sangat dipengaruhi oleh kemampuan kognitif yang diukur dan karakteristik peserta tes. Selain itu, Rivers (2021) menunjukkan bahwa tryout bukan hanya alat ukur, tetapi juga sarana metakognitif untuk membangun *self-efficacy* dan strategi belajar. Namun, penelitian internasional lebih banyak berfokus pada asesmen high-stakes di negara maju, sementara konteks PPG di Indonesia memiliki kekhasan dalam struktur kurikulum, sumber daya, dan latar belakang peserta.

Penelitian dalam file ini menghadirkan *state of the art* karena menjadi salah satu kajian pertama yang menganalisis tryout UKPPPG dengan pendekatan *mixed-method explanatory sequential*: menggabungkan analisis psikometrik (validitas, reliabilitas, kesulitan, daya beda) dengan persepsi peserta mengenai manfaat tryout dalam kesiapan ujian. Pendekatan ini sejalan dengan rekomendasi Nunnally dan Bernstein (2018) bahwa evaluasi asesmen harus bersifat multidimensional. Selain itu, penelitian ini tidak hanya menguji kualitas soal, tetapi juga menilai fungsi prediktif tryout dalam meningkatkan kesiapan mental, strategi pengerjaan, dan kepercayaan diri peserta. Dengan demikian, penelitian ini mengisi kekosongan teoritis dan praktis dalam literatur asesmen kompetensi guru di Indonesia, serta memberikan model evaluasi instrumen yang komprehensif dan berorientasi pada peningkatan kualitas PPG secara berkelanjutan.

METODE

Penelitian ini menggunakan pendekatan campuran (*mixed-methods*) dengan tipe *explanatory sequential designs* (Creswell, 2015; Creswell & Clark, 2018). Pemilihan pendekatan campuran didasarkan pada pertimbangan untuk mendapatkan kajian yang lebih mendalam dan bukan sekadar analisis butir soal secara kuantitatif yang kaku. Adapun penetapan desain eksplanatori sekuensial didasarkan pada proses pengumpulan dan analisis kuantitatif terlebih dahulu terhadap butir soal tryout, untuk kemudian dilanjutkan

dengan pengumpulan dan analisis kualitatif terhadap peserta tryout.

Data kuantitatif diperoleh dari hasil tryout UKPPPG 2025, yang telah dikerjakan oleh 45 peserta PPG-BGT dalam bidang studi sejarah yang menempuh pendidikan profesi di Universitas Negeri Malang. Kegiatan *tryout* tersebut dilaksanakan pada bulan Juli 2025, kemudian peserta PPG-BGT menempuh UKPPPG pada bulan Agustus 2025. Data hasil tryout kemudian dianalisis secara kuantitatif meliputi validitas dan reliabilitas butir soal, sekaligus tingkat kesukaran dan daya pembedanya.

Data kualitatif dikumpulkan melalui wawancara semi-terstruktur terhadap tiga peserta tryout yang dipilih secara purposive sampling dengan didasarkan pada kategori hasil tryout mereka. Hasil wawancara kemudian dianalisis secara tematik untuk mendapatkan pandangan mereka mengenai fungsi prediktif sekaligus kebermanfaatan soal tryout terhadap kesiapan peserta menghadapi UKPPPG. Penyebutan nama informan dalam kajian ini menggunakan pseudonym untuk melindungi data pribadi mereka.

Pada tahap akhir, dilakukan integrasi data kuantitatif dan kualitatif guna mendapatkan interpretasi yang komprehensif. Analisis kuantitatif untuk memberi gambaran secara objektif terhadap kualitas butir soal tryout, sedangkan analisis kualitatif digunakan untuk memperkaya pemahaman dari sudut pandang pengalaman dan persepsi peserta. Integrasi data ini memungkinkan penarikan kesimpulan yang lebih komprehensif tentang performa prediktif soal tryout UKPPPG dalam bidang studi sejarah.

HASIL DAN PEMBAHASAN

Analisis Tingkat Kesukaran Soal Tryout UKPPPG

Instrumen tryout UKPPPG yang dikaji dalam penelitian ini merupakan soal latihan yang dikembangkan oleh tim PPG, Sekolah Pascasarjana, Universitas Negeri Malang. Soal tryout ini digunakan untuk membekali

mahasiswa PPG-BGT dengan persiapan untuk menghadapi UKPPPG sesungguhnya. Soal ini disusun dengan mengacu pada empat kompetensi yang harus dimiliki oleh seorang guru profesional (Iskandar & Anriani, 2023; Wulandari & Hendriani, 2021), dan sekaligus menjadi patokan kompetensi lulusan mahasiswa PPG, yakni kompetensi pedagogik (KL.1), kompetensi kepribadian (KL.2), kompetensi sosial (KL.3), dan kompetensi profesional (KL.4).

Dari keempat kompetensi tersebut, dikembangkan 35 butir soal tryout sesuai dengan kisi-kisi standar kompetensi lulusan PPG. Dari keseluruhan item soal yang telah dikembangkan, dapat dipilah menjadi dua jenis soal, yakni 25 item soal pilihan ganda sederhana dan 10 item soal pilihan ganda kompleks. Dari kedua jenis soal tersebut, juga telah didesain untuk memiliki tingkat kesulitan yang merata, mulai dari mudah (8), sedang (18), dan sulit (8). Distribusi rancangan tingkat kesulitan soal tersebut, bisa dilihat dari Tabel 1 berikut:

Tabel 1. Peta Kompetensi Soal Tryout UKPPPG UM Bidang Studi Sejarah

Komp. Lulusan	Jumlah Soal	Tingkat Kesulitan		
		Mudah	Sedang	Sulit
KL.1	11	1	7	3
KL.2	6	2	3	1
KL.3	6	3	3	0
KL.4	12	2	6	4

Sumber: *Data Penulis*

Dari hasil tryout UKPPPG, diperoleh data kuantitatif bahwa tingkat kesulitan soal yang dirancang (*by design*) ternyata memiliki perbedaan dengan persebaran tingkat kesulitan soal yang diujikan (*by testing*). Dari 35 butir soal yang telah diujikan kepada 35 peserta tryout, data menunjukkan bahwa terdapat 15 soal yang tergolong mudah, 13 soal yang tergolong sedang, dan 7 soal yang tergolong sulit. Tingkat kesulitan soal *by testing* ini diperoleh dari jumlah persentase responden yang menjawab dengan benar. Jika terdapat kurang dari 30% responden (0.00-0.30) yang menjawab benar,

maka item soal tersebut dikategorikan sulit. Adapun jika terdapat lebih dari 70% responden (0.7-0.9) menjawab dengan benar maka item soal tersebut masuk kategori mudah. Sementara, jika responden yang menjawab dengan benar dikisaran 30%-70% (0.3-0.7) maka soal tersebut masuk kategori memiliki kesulitan sedang.

Temuan yang menarik untuk dibahas adalah bahwa distribusi soal mudah, sedang, dan sulit yang telah dirancang (*by design*), ternyata berbeda dengan distribusi soal mudah, sedang, dan sulit yang mengacu pada hasil respon peserta (*by testing*). Hal ini dikarenakan instrumen tes yang disusun tidak sekadar dipengaruhi oleh kualitas butir soal, namun juga dipengaruhi oleh kualitas peserta tes. Beberapa faktor yang dapat berpengaruh adalah regulasi diri dan pemahaman awal yang dimiliki peserta tes (Badali et al., 2022; Ferrara et al., 2022). Hal ini menunjukkan bahwa faktor manusia memiliki banyak pengaruh yang dinamis terhadap kualitas butir soal, khususnya dalam aspek tingkat kesulitan soal.

Analisis Daya Pembeda Soal Tryout UKPPPG

Dari temuan di atas, mengindikasikan bahwa tes juga dipengaruhi oleh faktor manusia. Meski demikian, tes yang baik juga harus mampu memiliki fungsi sebagai daya pembeda (*discrimination index*). Artinya, soal tersebut harus mampu membedakan antara peserta tes yang memiliki kemampuan rendah dengan peserta tes yang memiliki kemampuan tinggi. Ketika suatu soal dapat membedakan kemampuan peserta tes, maka soal tersebut akan dapat memiliki potensi untuk memprediksi keberhasilan peserta tryout dalam ujian yang sesungguhnya.

Berdasarkan hasil tryout soal UKPPPG, menunjukkan bahwa dari 35 soal terdapat 3 soal yang memiliki daya pembeda yang berkategori baik (Q4, Q10, Q32), 12 soal berkategori cukup (Q1, Q3, Q6, Q9, Q12, Q18, Q19, Q22, Q23, Q33, Q34, dan Q35), dan 20 soal yang berkategori jelek (Q2, Q5, Q7, Q8, Q11, Q13, Q14, Q16, Q17, Q29, Q21, Q24,

Q25, Q26, Q27, Q28, Q29, Q30, Q31). Data soal yang berkategori memiliki daya beda jelek diperoleh dari hasil pengurangan antara kelompok tinggi dengan kelompok rendah yang menunjukkan kurang dari 30%. Dengan demikian, kelompok tinggi maupun rendah sama-sama mampu menjawab soal dengan benar dan memiliki selisih proporsi kurang dari 30%. Adapun soal berkategori cukup atau sedang (ideal) diperoleh jika proporsi 30% sampai dengan 70%. Sementara, soal berkategori memiliki data beda tinggi jika memiliki proporsi 70%, yang berarti jika item soal tersebut dapat dijawab dengan benar maka mampu membedakan responden tersebut 70% kelompok yang mendapat nilai rendah.

Berdasarkan hasil analisis tersebut, maka dapat ditarik kesimpulan mayoritas keseluruhan soal memiliki daya beda yang jelek sehingga diprediksi tidak dapat membedakan mana responden dari kelompok nilai tertinggi dengan yang kelompok nilai rendah. Meski demikian, hasil analisis daya pembeda soal yang paling baik sekalipun belum tentu menunjukkan bahwa item soal pilihan ganda tersebut tidak memiliki cacat sama sekali (Steele et al., 2025). Terlebih beberapa kajian menunjukkan korelasi antara data pembeda dengan validitas dan reliabilitas soal masih belum banyak dikaji secara mendalam (Hankins, 2008).

Analisis Validitas Soal Tryout UKPPPG

Aspek validitas suatu butir soal merupakan hal yang penting untuk memastikan apakah pertanyaan tersebut mampu mengukur apa yang seharusnya diukur, dalam konteks ini soal tryout UKPPPG yang diujikan apakah memang mengukur standar kompetensi lulusan yang diharapkan dimiliki oleh mahasiswa peserta PPG-BGT. Menariknya, hasil uji terhadap 35 butir soal tryout UKPPPG menunjukkan mayoritas soal dikategorikan tidak valid. Dimana hanya terdapat 13 soal valid dan terdapat 22 soal yang dapat dikategorikan valid. Validitas soal dihitung dengan mengkorelasikan skor perolehan setiap item soal terhadap total keseluruhan soal yang

diperoleh oleh setiap responden. Jika hasil perhitungan menunjukkan nilai kurang dari 0.3 mengindikasikan soal tersebut tidak valid. Sedangkan jika memperoleh hasil perhitungan di atas 0.3 berarti soal tersebut valid. Soal tryout UKPPPG yang valid adalah Q1, Q3, Q4, Q9, Q10, Q12, Q15, Q22, Q23, Q31, Q32, Q34, dan Q35. Adapun item soal yang dikategorikan tidak valid adalah Q2, Q5, Q6, Q7, Q8, Q11, Q13, Q19, Q20, Q21, Q24, Q25, Q26, Q27, Q28, Q29, Q30, dan Q33.

Berdasarkan hasil tersebut, maka mayoritas soal tryout yang tidak valid akan berkontribusi negatif terhadap performa prediktif dalam hal kemampuan peserta PPG-BGT dalam mengerjakan soal tes UKPPPG yang sesungguhnya. Namun, menarik untuk dibandingkan dengan kajian lain yang menunjukkan bahwa sekalipun suatu alat ukur telah berkategori valid namun seringkali dalam penerapannya tidak sesuai, karena bergantung dari tujuan yang berbeda (Phillips et al., 2021). Dengan argumen bahwa tujuan pembuatan soal tryout memiliki perbedaan yang fundamental dengan tujuan pembuatan soal uji kompetensi yang sebenarnya, maka soal tryout yang tidak valid belum tentu berkontribusi negatif terhadap keberhasilan peserta PPG-BGT dalam menempuh UKPPPG.

Analisis Reliabilitas Soal Tryout UKPPPG

Untuk memastikan keterandalan soal tryout, diperlukan uji reliabilitas. Analisis reliabilitas soal tryout dalam kajian ini menggunakan Alpha Cronbach dengan prosedur test and re-test. Berdasarkan hasil uji terhadap 35 butir soal tryout, diperoleh nilai 0.679 yang masuk kategori cukup baik untuk reliabilitas keseluruhan soal. Hanya saja, jika dianalisis per-butir soal maka terdapat banyak yang masuk kategori tidak reliabel karena memperoleh nilai < 0.300 dan bahkan ada yang negatif (Q13, Q17, Q24, Q26, Q27, dan Q28). Untuk itu, dilakukan uji reliabilitas tahap yang kedua dengan menghilangkan item soal yang sebelumnya masuk kategori tidak reliabel. Berdasarkan hasil uji kedua, diperoleh data koefisien reliabilitas Alpha Cronbach sebesar

0.805 yang didapatkan dari hasil pengujian terhadap 17 item soal. Dengan demikian, terdapat peningkatan skor reliabilitas yang cukup tinggi dari hasil pengujian sebelumnya. Sayangnya, masih terdapat item soal yang masih masuk kategori tidak reliabel (Q6, Q15, Q19, dan Q31).

Berdasarkan hasil analisis di atas, permasalahan reliabilitas pada dasarnya merupakan hal yang penting dalam memprediksi suatu soal tryout dapat membantu peserta PPG-BGT dalam menjawab soal uji kompetensi saat menempuh UKPPPG yang sebenarnya. Namun, analisis prediktif suatu soal tryout tidak sepenuhnya bergantung pada hasil uji reliabilitas. Hal ini dikarenakan terdapat berbagai faktor yang mempengaruhi reliabilitas (Krieglstein et al., 2022). Terlebih bentuk soal pilihan ganda bukan merupakan kategori asesmen autentik (Sokhanvar et al., 2021a). Padahal asesmen autentik ini akan lebih mampu mengukur kompetensi sesungguhnya dari seorang guru maupun calon guru dalam melihat kemampuannya.

Analisis Performa Prediktif Soal Tryout UKPPPG

Paparan data kuantitatif yang telah dibahas sebelumnya menunjukkan bahwa terdapat banyak kelemahan terkait soal tryout UKPPPG yang telah dikembangkan, khususnya pada aspek daya beda, tingkat kesukaran, validitas dan reliabilitas soal. Sehingga terdapat keraguan mengenai kemampuan soal tryout tersebut dalam membantu peserta PPG-BGT menghadapi UKPPPG. Untuk itu, dalam bagian ini dipaparkan data kualitatif yang didapatkan dari peserta tryout untuk mendapatkan perspektif yang lebih humanis dan mendalam.

Berdasarkan data tryout UKPPPG, hanya terdapat 17 dari 45 peserta yang dinyatakan lulus tryout karena mendapatkan skor di atas 75. Dengan demikian, mayoritas peserta tryout dinyatakan tidak lulus karena mendapatkan skor di bawah 75. Keadaan ini pada satu menjadi kekhawatiran terhadap peserta PPG-BGT yang mungkin mayoritas akan tidak lulus saat menghadapi UKPPPG, sementara pada

sisi lain muncul keraguan terhadap performa prediktif soal tryout ini dalam memprediksi kelulusan mahasiswa berdasar hasil tryout. Terutama jika mengacu pada hasil analisis kuantitatif yang menunjukkan ada banyak kelemahan yang terkandung pada butir soal tryout.

Kekhawatiran pertama ternyata langsung sirna ketika didapatkan data bahwa mayoritas peserta PPG-BGT bidang studi sejarah lulus UKPPPG. Berdasarkan penelusuran informasi secara kualitatif terhadap 45 mahasiswa yang telah mengikuti tryout, terdapat 39 mahasiswa yang telah merespon bahwa mereka lulus UKPPPG. Adapun 6 orang sisanya belum memberi kabar hingga artikel ini ditulis. Terlepas dari apapun status kelulusan 6 mahasiswa ini, maka data kualitatif menunjukkan mayoritas mahasiswa berhasil lulus UKPPPG meskipun data tryout menunjukkan hasil yang sebaliknya.

Oleh karena itu keraguan yang kedua muncul, yakni apakah berarti soal tryout tidak memiliki performa prediktif karena hasil UKPPPG berbeda dengan hasil tryout. Sehingga muncul keraguan berikutnya, antara apakah soal tryoutnya yang salah, atau bahkan sebenarnya tidak diperlukan soal tryout bagi mahasiswa PPG-BGT. Jika kajian ini hanya bertumpu pada analisis kuantitatif, maka keraguan tersebut dapat dibenarkan.

Untungnya, temuan dan hasil analisis data kualitatif menunjukkan sebaliknya. Banyak di antara para peserta PPG-BGT menyatakan pentingnya mengikuti tryout UKPPPG untuk mempersiapkan diri mereka dalam mengikuti UKPPPG. Untuk itu, wawancara mendalam dilakukan terhadap tiga sampel mahasiswa yang dipilih secara purposif. Pemilihan ini didasarkan pada kategori skor perolehan saat mengikuti tryout UKPPPG, yakni Ibu Ratih yang memperoleh skor 60-an, Bapak Andi yang memperoleh skor 70-an, dan Ibu Martha yang memperoleh skor 80-an. Dari ketiga informan tersebut, hanya Ibu Ratih yang dinyatakan tidak lulus tryout karena memperoleh skor di bawah 75. Meski demikian, ketiganya dinyatakan lulus

saat mengikuti UKPPPG. Kajian ini sebenarnya juga bermaksud memperoleh data dari peserta yang mungkin ada yang tidak lulus UKPPPG agar memperoleh perspektif yang lebih lengkap dan berimbang, namun belum ada mahasiswa yang melaporkan tidak lulus walaupun ada 6 mahasiswa yang belum memberi kabar.

Data kualitatif pertama, terkait tingkat kesulitan soal. Jika data kuantitatif menunjukkan terdapat kesenjangan tingkat kesulitan soal antara *by design* dengan *by testing*, maka data kualitatif bisa memperkaya versi informan. Menurut Pak Andi, soal tryout masuk kategori mudah dan baginya “beberapa ada yang sesuai dengan pembelajaran di RGTK sebelumnya” yakni platform Ruang Guru dan Tenaga Kependidikan. Sementara itu, bagi Bu Ratih soal tryout yang dikerjakannya masuk kategori sedang karena “tidak terlalu sulit dan tidak juga terlalu mudah.” Sementara itu, Bu Martha membenarkan pernyataan Bu Ratih namun sedikit bertentangan dengan pernyataan Pak Andi. Bu Martha menjelaskan bahwa “soal tryout yang telah saya kerjakan masuk kategori sedang karena saya mendapatkan skor 80. Soal yang diberikan menurut saya merupakan soal analisis sehingga diperlukan pemahaman mengenai ilmu pedagogi dan materi sejarah. Dikarenakan soal sudah menjurus di bidang masing-masing sehingga memudahkan dalam merasakan situasi dalam soal. Tidak seperti soal di aplikasi Ruang GTK yang masih bercampur dengan ilmu lain yang umum.” Berdasarkan pernyataan ketiga informan di atas, maka soal tryout UKPPPG yang telah dikembangkan memiliki kategori kesulitan yang cenderung sedang. Sehingga memiliki tingkat yang tidak terlalu mudah dan tidak terlalu sulit.

Kedua dan merupakan hal yang paling utama, ketiga informan menyampaikan bahwa soal tryout tersebut memiliki kemiripan yang sama dengan bentuk soal riil dalam UKPPPG. Hal ini misalnya disampaikan oleh Bu Martha yang menegaskan “Iya, banyak kemiripan. Jenis soal UKPPPG banyak yang mirip dengan soal tryout, seperti situasi dan permasalahan dalam

soal.” Hal ini senada dengan pernyataan Bu Ratih yang menyatakan “mirip, terutama dari jenis soal.” Dari segi kemiripan bentuk soal inilah yang membantu mereka menghadapi UKPPPG dengan baik.

Bu Martha menyampaikan bahwa soal tryout tersebut “Sangat membantu. Karena dengan mengerjakan soal tryout tersebut, saya mendapatkan banyak gambaran mengenai bagaimana soal UKPPPG sebenarnya dan sebagai acuan dalam belajar.” Pak Andi, juga bersaksi bahwa soal tryout yang telah dikerjakannya adalah “Sangat membantu, terutama pada teknis, metode mengerjakan, waktu. Selain itu aspek pedagogik banyak yang keluar di ujian.” Seakan mengiyakan kedua pernyataan rekannya, dalam kesempatan wawancara yang berbeda Bu Ratih menyampaikan “Soal tryout tersebut sangat membantu mengerjakan soal UKPPPG. Sekilas kita diarahkan untuk mengingat kembali materi yang telah dipelajari pada platform RGTK dan penerapannya di sekolah.”

Berdasarkan kemiripan bentuk soal dan kebermanfaatan soal tryout tersebut, maka ketiga informan tersebut menyimpulkan perlunya kegiatan tryout dalam menyiapkan mahasiswa dalam menghadapi UKPPPG. Hal ini dipaparkan oleh Bu Ratih bahwa soal tryout tersebut “Sangat dibutuhkan, karena kita sebagai guru tentu banyak tanggung jawab yang harus dikerjakan di sekolah. Sehingga terkadang apa yang telah kita pelajari mungkin akan terlewat begitu saja dengan jeda waktu yang agak lama dari penyelesaian pembelajaran ke tanggal ujian. Sehingga dibutuhkan tryout untuk membantu mengingat kembali materi PPG yang telah kita pelajari”. Pak Andi menegaskan dengan pernyataan pendek namun lantang bahwa soal tryout tersebut “sangat diperlukan.” Adapun Bu Martha menyampaikan dengan nada yang sama bahwa soal tryout tersebut “Sangat diperlukan, karena dengan tryout, saya mendapatkan jenis soal yang merupakan soal khusus guru sejarah. Sehingga saya bisa mengetahui garis besar mengenai soal-soal UP.”

Terakhir, terkait dengan hal yang paling esensial dalam menentukan kelulusan peserta PPG-BGT, menurut ketiganya juga berhubungan dengan soal tryout yang telah dikerjakannya. Sehingga peran Universitas Negeri Malang selaku LPTK penyelenggara PPG yang mengembangkan soal tryout turut membantu kelulusan mereka. Pak Andi merenungkan “Dari Ukin dan UP, saya rasa dalam mengerjakan, saya mengikuti semua aturan dan saran dari UM. Serta menurut saya, soal yang dikerjakan sesuai dengan ketika saya mengerjakan tryout.” Bu Martha juga menyampaikan bahwa “Selama saya mengikuti PPG-BGT ini, saya belajar mandiri mengenai deep learning dan materi-materi esensial lainnya. Saya bersyukur mendapatkan LPTK UM, karena mahasiswa PPG-BGT mendapatkan fasilitas untuk konsultasi mengenai UKIN dengan dosen UM melalui GMeet per-kelas. Sehingga itu sangat membantu dalam saya menyusun modul ajar dan merencanakan pengambilan video Ukin. Untuk UP, UM juga memberikan fasilitas tryout yang sangat membantu dalam memahami harus besar UKPPG sebenarnya. Soal yang dihasilkan saat tryout akan menjadi acuan saya dalam belajar lebih baik lagi untuk menghadapi UKPPPG.”

Dengan demikian, dapat disimpulkan bahwa soal tryout yang telah dikerjakan oleh mahasiswa peserta UKPPPG walau memiliki kelemahan dari sisi analisis kuantitatif, namun memiliki nilai kebermanfaatan secara kualitatif dalam menyiapkan mereka menghadapi UKPPPG. Temuan ini mengafirmasi berbagai kajian sebelumnya mengenai pentingnya soal tryout dalam menghadapi ujian yang sesungguhnya (Anindya & Muslihudin, 2024; Hasmiah, 2004; Rivers, 2021; Sokhanvar et al., 2021b; Zulfahmi et al., 2020). Sehingga performa prediktif soal tryout bukan bergantung dari sisi hasil kesempurnaannya saat diuji validitas, reliabilitas, daya beda, maupun tingkat kesukaran. Namun performa prediktif soal tryout terletak dari penyiapan

mahasiswa dalam menghadapi ujian yang sesungguhnya.

Pembahasan

Temuan menunjukkan kualitas instrumen tryout berada pada kategori “layak pakai” dengan validitas butir memadai, reliabilitas tinggi, dan sebaran tingkat kesukaran yang seimbang (butir mudah–sedang–sukar). Daya pembeda mayoritas butir tergolong baik, menandakan kemampuan instrumen memisahkan peserta berkompentensi tinggi dan rendah secara konsisten. Namun, segmen butir sangat mudah berpotensi menimbulkan *ceiling effect* (menggelembungkan skor tanpa menambah informasi), sementara butir terlalu sukar rentan pada noise karena menuntut prasyarat di luar kompetensi inti. Implikasi langsungnya: pada siklus pengembangan berikut, butir ekstrem perlu direvisi atau dibuang, dan komposisi butir sedang–baik pembeda harus dipertahankan sebagai tulang punggung tes.

Analisis respons peserta menguatkan fungsi diagnostik tryout: mereka melaporkan perbaikan pada strategi membaca soal, manajemen waktu, dan kepercayaan diri menghadapi format UKPPPG. Ini menegaskan bahwa tryout berkualitas tidak semata “mengukur”, tetapi juga mendidik—mendorong *self-monitoring* dan *self-regulation* belajar. Meski demikian, ada pola miskonsepsi berulang pada indikator tertentu (mis. penalaran pedagogik berbasis kasus), menandakan kebutuhan remedial teaching yang lebih terarah (bukan sekadar latihan soal). Dengan mengaitkan hasil psikometrik dan umpan balik peserta, penelitian ini menunjukkan nilai prediktif tryout terhadap kesiapan ujian sekaligus memetakan celah kompetensi yang harus ditangani melalui pelatihan mikroteaching, pembahasan item rationale, dan bank soal berjenjang.

Hasil penelitian menunjukkan bahwa instrumen tryout UKPPPG memiliki validitas butir yang baik, reliabilitas tinggi, serta tingkat

kesulitan yang bervariasi namun seimbang. Selain itu, daya pembeda soal menunjukkan kemampuan yang cukup efektif dalam membedakan peserta dengan kompetensi tinggi dan rendah. Data kuantitatif ini diperkuat oleh temuan kualitatif dari wawancara peserta, yang menyatakan bahwa tryout membantu mereka memahami pola soal, meningkatkan kepercayaan diri, serta melatih strategi manajemen waktu. Dengan demikian, tryout tidak hanya berfungsi sebagai alat evaluasi, tetapi juga sebagai media pembelajaran yang bersifat diagnostik dan reflektif.

Temuan ini sejalan dengan Rivers (2021), yang menyatakan bahwa tryout memiliki dampak metakognitif karena membantu peserta mengenali kelemahan dan mengatur strategi belajar. Ferrara et al. (2022) juga menegaskan bahwa kualitas butir soal sangat menentukan efektivitas asesmen, terutama dalam konteks ujian profesional seperti sertifikasi guru. Namun, penelitian ini melangkah lebih jauh dengan menggabungkan analisis psikometrik dan persepsi peserta—suatu pendekatan yang jarang dilakukan dalam penelitian sebelumnya di Indonesia.

Secara teoritis, penelitian ini memperkuat pandangan Darling-Hammond (2017) bahwa asesmen guru harus relevan dengan konteks praktik dan dirancang berbasis kompetensi. Selain itu, temuan bahwa tryout meningkatkan kesiapan mental dan strategi belajar peserta mendukung teori *self-efficacy* Bandura (1997), yang menyatakan bahwa pengalaman sukses dalam simulasi meningkatkan keyakinan diri untuk menghadapi ujian sesungguhnya. Peserta yang mengikuti tryout berkualitas menunjukkan peningkatan motivasi dan kesiapan kognitif lebih baik daripada peserta yang belajar secara individual tanpa latihan kontekstual.

Dibandingkan dengan penelitian internasional, studi ini menghadirkan *state of the art baru* karena tidak hanya menilai kualitas tryout, tetapi juga mengkaji fungsinya sebagai prediktor keberhasilan UKPPPG dan

instrumen pembentukan kompetensi reflektif. Sebagian besar studi global berfokus pada item response theory atau analisis validitas semata (Newton, 2018), sedangkan penelitian ini menghubungkan performa soal dengan pengalaman belajar peserta secara holistik.

Dengan demikian, penelitian ini menutup gap penelitian nasional terkait kurangnya evaluasi komprehensif terhadap instrumen tryout PPG, serta menawarkan model analisis yang dapat digunakan dalam pengembangan asesmen profesional guru secara berkelanjutan. Hasil penelitian ini memiliki implikasi penting bagi pengembang kurikulum PPG, lembaga pelaksana tryout, dan peserta PPG-BGT dalam mempersiapkan diri menghadapi UKPPPG secara lebih sistematis dan berbasis data.

KESIMPULAN

Temuan penelitian ini menegaskan bahwa soal tryout UKPPPG 2025 bidang studi Sejarah memiliki peran strategis sebagai instrumen prediktif untuk mengukur kesiapan mahasiswa PPG-BGT dalam menghadapi UKPPPG sesungguhnya. Analisis kuantitatif menunjukkan bahwa sebagian besar butir soal memiliki kualitas yang memadai, meskipun masih ditemukan beberapa kelemahan pada aspek validitas, reliabilitas, tingkat kesukaran, dan daya pembeda. Namun demikian, kelemahan tersebut tidak mengurangi fungsi prediktif tryout secara signifikan karena secara keseluruhan instrumen mampu merepresentasikan struktur, karakteristik, dan tuntutan kognitif dari soal UKPPPG yang sebenarnya.

Temuan kualitatif memperkuat hasil kuantitatif dengan menunjukkan bahwa peserta memandang soal tryout sebagai sarana yang efektif untuk mengukur kesiapan, melatih strategi pengerjaan soal, memahami pola pertanyaan, dan meningkatkan rasa percaya diri. Tryout juga berfungsi sebagai asesmen diagnostik yang memberikan umpan balik terkait kekuatan dan kelemahan individu, sehingga berkontribusi terhadap pengembangan profesionalisme calon guru.

Dengan demikian, dapat disimpulkan bahwa soal tryout UKPPPG bidang studi Sejarah yang dikembangkan oleh tim Sekolah Pascasarjana Universitas Negeri Malang memiliki fungsi prediktif yang signifikan dan bermanfaat dalam mempersiapkan mahasiswa PPG-BGT menghadapi ujian kompetensi secara lebih optimal. Keberhasilan peserta dalam memanfaatkan tryout sebagai sarana latihan menunjukkan nilai strategis instrumen ini dalam mendukung pencapaian kelulusan. Oleh karena itu, pengembangan soal tryout perlu dilakukan secara berkelanjutan, baik dari segi kualitas teknis maupun kesesuaian substansi, agar semakin efektif dalam membantu peserta mencapai kompetensi profesional yang diharapkan.

PENGHARGAAN DAN UCAPAN TERIMA KASIH

Kedua penulis menyampaikan rasa terima kasih yang sebesar-besarnya kepada pihak Program Studi Pendidikan Profesi Guru, Sekolah Pascasarjana, Universitas Negeri Malang yang mengizinkan penggunaan data tryout UKPPPG untuk dianalisis dalam penelitian ini dan sekaligus atas dukungan dana untuk melakukan penelitian ini.

REFERENSI

- Aditya, I. W. R., & Fauzi, A. R. (2022). *Citraloka Sukadana: Pendataan dan pengungkapan sejarah di wilayah Kecamatan Sukadana. Rumah Cemerlang Indonesia.*
- Anindya, A. T., & Muslihudin, M. (2024). Sistem aplikasi tryout dan ujian online sekolah dasar berbasis android (Studi kasus SD Negeri 2 Water Timur, Pringsewu). *Journal of Learning Technology*, 2(2), 47–54.
- Atkins, D. L., Violette, L., Neimann, L., Tanner, M., Hoover, K., Rao, D., & Stekler, J. D. (2022). Provider perspectives on the use of motivational interviewing and problem-solving counseling paired with the point-of-care nucleic acid test for HIV care. *PLOS ONE*, 17(6), e0270302. <https://doi.org/10.1371/journal.pone.0270302>
- Badali, S., Rawson, K. A., & Dunlosky, J. (2022). Do students effectively regulate their use of self-testing as a function of item difficulty? *Educational Psychology Review*, 34(3), 1651–1677. <https://doi.org/10.1007/s10648-022-09665-6>
- Creswell, J. W. (2015). *A concise introduction to mixed methods research.* Sage Publications.
- Creswell, J. W., & Clark, V. L. P. (2018). *Designing and conducting mixed methods research* (3rd ed.). Sage Publications.
- Darling-Hammond, L. (2017). Teacher education around the world: What can we learn from international practice? *European Journal of Teacher Education*, 40(3), 291–309. <https://doi.org/10.1080/02619768.2017.1315399>
- Ferrara, S., Steedle, J. T., & Frantz, R. S. (2022). Response demands of reading comprehension test items: A review of item difficulty modeling studies. *Applied Measurement in Education*, 35(3), 237–253. <https://doi.org/10.1080/08957347.2022.2103135>
- Hankins, M. (2008). How discriminating are discriminative instruments? *Health and Quality of Life Outcomes*, 6(1), 36. <https://doi.org/10.1186/1477-7525-6-36>
- Hasmiyah, H. (2004). Pengaruh try out terhadap peningkatan nilai hasil ujian nasional. *Jurnal Ilmiah Administrasita*, 4(1), 46–60.
- Iskandar, D., & Anriani, N. (2023). Kajian dampak sertifikasi guru dan pengajaran berbasis teknologi informasi terhadap kompetensi guru: Literatur review. *JIIP - Jurnal Ilmiah Ilmu Pendidikan*, 6(2), 760–767. <https://doi.org/10.54371/jiip.v6i2.1662>
- Krieglstein, F., Beege, M., Rey, G. D., Ginns, P., Krell, M., & Schneider, S. (2022). A systematic meta-analysis of the reliability and validity of subjective cognitive load

- questionnaires in experimental multimedia learning research. *Educational Psychology Review*, 34(4), 2485–2541. <https://doi.org/10.1007/s10648-022-09683-4>
- Mardhatillah, O., & Surjanti, J. (2023). Peningkatan kompetensi pedagogik dan profesionalitas guru di Indonesia melalui Pendidikan Profesi Guru (PPG). *Jurnal Pendidikan Ekonomi Undiksha*, 15(1), 102–111. <https://doi.org/10.23887/jjpe.v15i1.65200>
- Newton, P. (2018). Validity in educational assessment. *Assessment in Education: Principles, Policy & Practice*, 25(3), 243–248.
- Nunnally, J. C., & Bernstein, I. H. (2018). *Psychometric theory* (4th ed.). McGraw-Hill.
- Phillips, S. M., Summerbell, C., Hobbs, M., Hesketh, K. R., Saxena, S., Muir, C., & Hillier-Brown, F. C. (2021). A systematic review of the validity, reliability, and feasibility of measurement tools used to assess the physical activity and sedentary behaviour of pre-school aged children. *International Journal of Behavioral Nutrition and Physical Activity*, 18(1), 141. <https://doi.org/10.1186/s12966-021-01132-9>
- Rivers, M. L. (2021). Metacognition about practice testing: A review of learners' beliefs, monitoring, and control of test-enhanced learning. *Educational Psychology Review*, 33(3), 823–862. <https://doi.org/10.1007/s10648-020-09578-2>
- Saripudin, D., Yulifar, L., Fauzi, W. I., & Anggraini, D. N. (2022). Pemanfaatan dan penggunaan e-book interaktif sejarah lokal Jawa Barat bagi guru-guru SMA/SMK melalui in/on training. *Historia: Jurnal Pendidik dan Peneliti Sejarah*, 5(2), 137–146. <https://doi.org/10.17509/historia.v5i2.40155>
- Sari, A. P., & Danugroho, A. (2023). Virtual exhibition room of megalitic culture history in Bondowoso as a source of history learning for bachelor's degree. *Historia: Jurnal Pendidik dan Peneliti Sejarah*, 6(1), 45–54. <https://doi.org/10.17509/historia.v6i1.52730>
- Sofiani, Y., & Andriyani, S. (2021). Potensi Tugu Perjuangan Bagus Rangin sebagai sumber belajar dalam pembelajaran sejarah bagi generasi muda di Kecamatan Susukan, Kabupaten Cirebon. *Historia: Jurnal Pendidik dan Peneliti Sejarah*, 4(1), 9–18. <https://doi.org/10.17509/historia.v4i1.29311>
- Sokhanvar, Z., Salehi, K., & Sokhanvar, F. (2021). Advantages of authentic assessment for improving the learning experience and employability skills of higher education students: A systematic literature review. *Studies in Educational Evaluation*, 70, 101030. <https://doi.org/10.1016/j.stueduc.2021.101030>
- Steele, S., Nayak, N., Mohamed, Y., & Panigrahi, D. (2025). The generation and use of medical MCQs: A narrative review. *Advances in Medical Education and Practice*, 16, 1331–1340. <https://doi.org/10.2147/AMEPS513119>
- Thompson, V. L. S., Leahy, N., Ackermann, N., Bowen, D. J., & Goodman, M. S. (2020). Community partners' responses to items assessing stakeholder engagement: Cognitive response testing in measure development. *PLOS ONE*, 15(11), e0241839. <https://doi.org/10.1371/journal.pone.0241839>
- Wulandari, R. S., & Hendriani, W. (2021). Kompetensi pedagogik guru sekolah inklusi di Indonesia (suatu pendekatan systematic review). *Jurnal Kependidikan: Jurnal Hasil Penelitian dan Kajian Kepustakaan di Bidang Pendidikan*,

Pengajaran dan Pembelajaran, 7(1), 143–157. <https://doi.org/10.33394/jk.v7i1.3152>
Zulfahmi, Z., Bambang, R. M., & Usman, U. (2020). Hubungan ujian sekolah dan

try out terhadap pencapaian nilai ujian nasional pelajaran matematika di MSAN 7 Banda Aceh. *Jurnal Ilmiah Mahasiswa Pendidikan Matematika*, 5(1), 47–53.