

Analisis kuantitatif butir soal TPA tryout mahasiswa PPG prodi pendidikan jasmani, olahraga dan kesehatan UM

Ari Wibowo Kurniawan^{*1}, Siti Nurrochmah²

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

²Prodi Pendidikan Profesi Guru (PPG), Sekolah Pascasarjana, Universitas Negeri Malang

*Corresponding author: ari.wibowo.fik@um.ac.id

Abstrak

Peningkatan kualitas pendidikan nasional berkaitan dengan pengembangan kompetensi tenaga pengajar di lapangan. Untuk mencapai tujuan tersebut dengan menggunakan Tes Potensi Akademik (TPA) yang seringkali digunakan diberbagai institusi perguruan tinggi maupun dunia kerja. Tujuan TPA ialah untuk menilai kualitas intelektual, bakat serta kemampuan seseorang dalam bidang akademik. Tujuan penelitian ini menganalisis hasil tryout TPA pada mahasiswa PPG program studi PJOK. Teknik analisis data yang digunakan adalah analisis butir soal secara kuantitatif. Tahapan dalam analisis data meliputi 1) analisis tingkat kesukaran butir soal tryout, 2) mengukur daya beda, 3) analisis validitas butir soal tryout, 4) menguji distraktor pilihan jawaban dan 5) analisis reliabilitas menggunakan Cronbach's Alpha. Dari 35 butir soal tryout yang telah diberikan pada 523 mahasiswa PPG program studi PJOK masih harus ditingkatkan dengan hasil tingkat kesukaran mudah sebanyak 20 butir soal ($p > 0,70$), sedangkan tingkat kesukaran sedang sebanyak 14 butir soal ($0,30 \leq p \leq 0,70$) dan 1 butir soal pada tingkat kesukaran sulit ($p < 0,30$). Hasil analisis daya beda terdapat 20 butir soal termasuk dalam kategori cukup dan 15 butir soal termasuk kategori jelek. Hasil analisis validitas butir soal terdapat 11 butir soal yang dianggap tidak valid dikarenakan $r < 0,3$. Namun terdapat data yang menunjukkan butir soal valid tapi beberapa butir soal memiliki korelasi yang rendah dikategorikan sebagai tidak valid dan perlu untuk digugurkan. Hasil analisis distraktor terdapat 7 butir soal yang tidak berfungsi sehingga butir soal tersebut perlu dibuang atau digugurkan. Hasil analisis reliabilitas dihasilkan sebesar 0,823 untuk keseluruhan 35 butir soal atau pada tahap awal sebelum digugurkan atau revisi. Kemudian meningkat menjadi 0,831 untuk 28 butir soal yang tersisa atau setelah dilakukan pengguguran. Keterkaitan atau hubungan dari 5 indikator tersebut saling terhubung satu sama lain sehingga butir soal tes yang telah diuji dengan cermat akan memiliki tes memiliki tingkat reliabilitas yang tinggi.

Kata Kunci: Tes Potensi Akademik; PPG ; Pendidikan Jasmani, Olahraga dan Kesehatan; Tes

Abstract

Improving the quality of national education is related to developing the competence of teaching staff in the field. To achieve this goal, the Academic Potential Test (TPA) is often used in various higher education institutions and in the world of work. The purpose of the TPA is to assess a person's intellectual quality, talent, and abilities in the academic field. The purpose of this study is to analyze the results of the TPA tryout for PPG students in the PJOK study program. The data analysis technique used is quantitative item analysis. The stages of data analysis include 1) analyzing the level of difficulty of the tryout items, 2) measuring the discrimination power, 3) analyzing the validity of the tryout items, 4) testing the answer choice distractors, and 5) analyzing the reliability using Cronbach's Alpha. Of the 35 tryout questions given to 523 PPG students in the PJOK study program, 20 questions were still too easy ($p > 0.70$), while 14 items were at a moderate level of difficulty ($0.30 \leq p \leq 0.70$) and 1 item was at a difficult level of difficulty ($p < 0.30$). The results of the discriminant analysis showed that 20 items were classified as adequate and 15 items were classified as poor. The results of the item validity analysis showed that 11 items were considered invalid because $r < 0.3$. However, there was data showing that some items were valid, but several items had low correlations and were categorized as invalid and needed to be discarded. The distractor analysis results showed that 7 items did not function properly and therefore needed to be discarded or rejected. The reliability analysis results were 0.823 for all 35 items or at the initial stage before discarding or revising. Then it increased to 0.831 for the remaining 28 items or after discarding. The correlation or relationship of the 5 indicators is interconnected so that test items that have been carefully tested will have a high level of reliability.

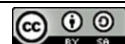
Keywords: Academic Potential Test, PPG 3, physical Education, Sports and Health, Test

Copyright © 2025 Author(s)

Received: 18 08 2025

Revised: 08 09 2025

Accepted: 29 09 2025



Authors' Contribution: A – Conceptualization; B – Methodology; C – Software; D – Validation; E - Formal analysis; F – Investigation; G – Resources; H – Data Curation; I - Writing - Original Draft; J - Writing - Review & Editing; K – Visualization; L – Supervision; M - Project administration; N - Funding acquisition

PENDAHULUAN

Guru memiliki peran sentral dalam sistem pendidikan sebagai kunci penentu utama mutu lulusan (Pratama et al., 2021). Peningkatan kualitas pendidikan nasional berkaitan dengan pengembangan kompetensi tenaga pengajar di lapangan. Guru juga harus memiliki kompetensi sesuai dengan bidangnya untuk mendukung proses pembelajaran. Dalam meningkatkan kompetensi tersebut, guru dapat mengikuti program Pendidikan Profesi Guru (PPG). (Aini et al., 2023) PPG bertujuan untuk menciptakan guru yang tidak hanya ahli di bidangnya, tetapi juga terampil dalam memahami peserta didik serta mampu merancang pembelajaran sesuai dengan kebutuhan mereka. Untuk mencapai tujuan tersebut, dengan dilakukannya penilaian pada tahap evaluasi penting untuk dilakukan agar kualitas proses pembelajaran terjamin.

Tahapan proses pembelajaran khususnya penilaian merupakan komponen vital yang tidak terpisahkan dalam proses pendidikan maupun pengajaran (Mustafa & Masgumelar, 2022). Penilaian dalam program PPG bertujuan untuk mengukur apakah mahasiswa telah mencapai kompetensi yang diharapkan selama dan setelah mengikuti program perkuliahan (Wantoro et al., 2019). Salah satu penerapan penilaian menggunakan Tes Potensi Akademik (TPA) yang seringkali digunakan diberbagai institusi perguruan tinggi maupun dunia kerja.

TPA merupakan tes yang dirancang untuk mengukur pencapaian intelektual seseorang. Tujuan TPA ialah untuk menilai kualitas intelektual, bakat serta kemampuan seseorang dalam bidang akademik (Saputra et al., 2017). Meskipun TPA telah sering digunakan diberbagai institusi namun, instrumen TPA yang berfokus pada mahasiswa PPG khususnya program studi Pendidikan Jasmani, Olahraga dan Kesehatan (PJOK) masih terbatas.

Mahasiswa PPG program studi PJOK dituntut tidak hanya mampu sebagai demonstrator namun juga sebagai perancang, menganalisis serta menjadi pendidik yang berkompeten. Kompetensi yang harus dimiliki oleh guru ahli pertama terdiri dari kompetensi pedagogik, kepribadian, sosial dan profesional (Peraturan Direktur Jenderal GTK, 2023). (Ridwan & Fadli, 2024) memaparkan guru PJOK memiliki tantangan dan peluang baru pada era society 5.0 dimana menekankan pada teknologi digital dan AI dengan nilai-nilai kemanusiaan.

Implementasi pada era society 5.0 meliputi penerapan teknologi secara kreatif dan inovatif dalam proses pembelajaran didukung oleh pendekatan pembelajaran berbasis nilai, serta peningkatan keterampilan interpersonal dan literasi digital. Selaras dengan (Rambe et al.,

2024) mengungkapkan tantangan yang dihadapi guru PJOK ialah mengoptimalkan penggunaan media pembelajaran. Keterbatasan kompetensi tersebut secara langsung menurunkan kualitas pengajaran sehingga berdampak pada hasil belajar peserta didik sehingga perlunya instrumen TPA untuk mengukur kompetensi akademik pada bidang keolahragaan.

Kemampuan TPA tersebut dasarnya untuk mengukur kemampuan kognitif meliputi penalaran, analisis dan pemecahan masalah (Sarea & Ruslan, 2019). Instrumen TPA yang berfokus pada mahasiswa PPG khususnya program studi Pendidikan Jasmani, Olahraga dan Kesehatan (PJOK) masih minim dan terbatas serta bagaimana kualitas dari butir soal *tryout* TPA pada program studi PJOK. Berdasarkan permasalahan tersebut, tujuan penelitian ini adalah menganalisis hasil *tryout* TPA pada mahasiswa PPG program studi PJOK..

METODE

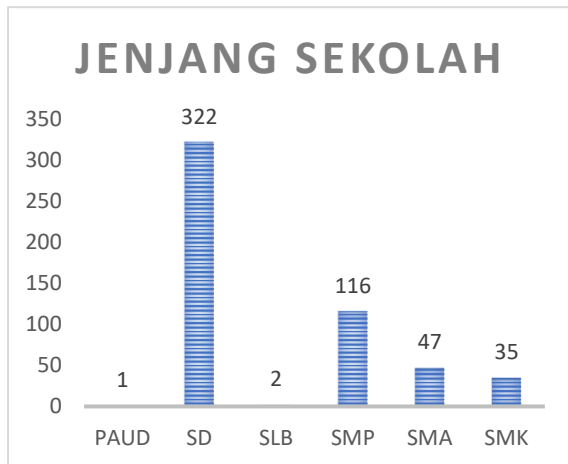
Penelitian tentang analisis TPA pada mahasiswa PPG program studi PJOK menggunakan pendekatan deskriptif kuantitatif untuk menganalisis hasil *tryout* TPA mahasiswa PPG program studi PJOK. Soal *tryout* merupakan hasil pengembangan soal berbasis soal HOTS dengan studi kasus kontekstual. Total soal sebanyak 35 butir soal, terdiri dari 25 pilihan ganda sederhana dan 10 pilihan ganda kompleks. Kompetensi yang di gunakan sebagai dasar penyusunan soal mengacu pada kompetensi Guru Ahli Pertama Level 2 yang terdiri dari kompetensi pedagogik, kepribadian, sosial dan profesional. Subjek penelitian ini berjumlah 523 mahasiswa PPG program studi PJOK. Teknik analisis data yang digunakan adalah analisis butir soal secara kuantitatif.

Tabel 1.1. Jumlah soal pada masing-masing kompetensi lulusan

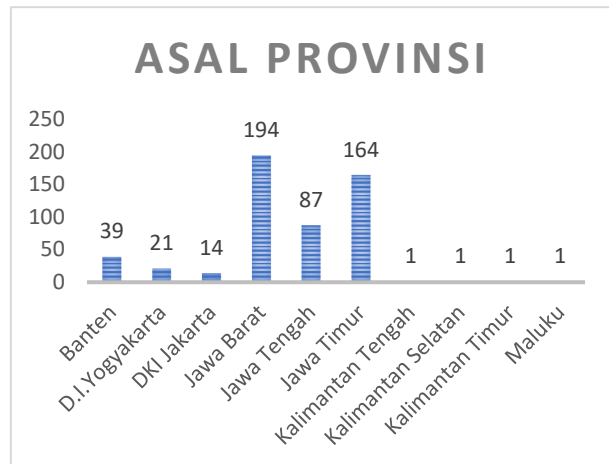
No.	Kompetensi lulusan	Jumlah soal
1.	Pedagogik	11
2.	Kepribadian	6
3.	Sosial	6
4.	Profesional	12

Tahapan dalam analisis data meliputi (1) analisis tingkat kesukaran butir soal *tryout*, (2) mengukur daya beda, (3) analisis validitas butir soal *tryout*, (4) menguji distraktor pilihan jawaban dan 5) analisis reliabilitas menggunakan Cronbach's Alpha. Butir soal *tryout* terdiri dari 35 soal dengan 5 opsi jawaban (A,B,C,D,E) dan 1 kunci jawaban. Setiap pilihan jawaban mendapat skor 1 jika jawaban benar dan 0 jika pilihan jawaban salah.

Grafik 1.1. Persebaran Jenjang Sekolah



Grafik 1.2. Persebaran Asal Provinsi Mahasiswa PPG PJOK



Berdasarkan grafik 1.1 mengenai mahasiswa PPG program studi PJOK didapatkan persebaran jenjang sekolah meliputi jenjang PAUD sebanyak 1 mahasiswa, jenjang SD sebanyak 322 mahasiswa, jenjang SLB sejumlah 2 mahasiswa, jenjang SMP sejumlah 116 mahasiswa, jenjang SMA sebesar 47 mahasiswa dan jenjang SMK sebesar 35 mahasiswa. Sedangkan pada grafik 1.2 merupakan persebaran mahasiswa PPG program studi PJOK berasal berdasarkan asal provinsi. Persebaran mahasiswa berasal dari Provinsi Banten sejumlah 39 mahasiswa, Provinsi D.I.Yogyakarta sejumlah 21 mahasiswa, Provinsi DKI Jakarta sebanyak 14 mahasiswa, Provinsi Jawa Barat sebanyak 194 mahasiswa, Provinsi Jawa Tengah sebanyak 87 mahasiswa, Provinsi Jawa Timur sejumlah 164 mahasiswa dan Provinsi Kalimantan Tengah, Kalimantan Selatan, Kalimantan Timur dan Maluku masing-masing Provinsi terdiri dari 1 mahasiswa.

HASIL

1. Deskripsi Umum Skor Hasil TPA

Tabel 1.1. Hasil Analisis Deskripsi Umum Skor Hasil TPA

No.	Kelompok Nilai	Jumlah Sampel	FR (%)	Skor Min.	Skor Maks.	Rata-rata Skor	SD.S	Rata-rata Waktu	KV (%)
1.	Tinggi	139	26,58	30	35	31,42	1,27	77,12	4,04
2.	Sedang	230	43,98	24	29	26,6	1,68	77	6,30
3.	Rendah	154	29,45	5	23	19	3,80	80	19,82
4.	Jumlah	523	100						

Keterangan:

FR = Frekuensi Relatif

SD.S = Standar Deviasi Sampel

KV = Koefisien Variansi

Diantara ketiga kelompok tersebut, data/skor/nilai dari kelompok rendah menunjukkan lebih menyebar dibandingkan dengan kelompok sedang dan kelompok tinggi. Khusus pada kelompok skor/nilai tinggi hasil analisis penyebaran diperoleh KV 4,04% berarti data atau nilai pada kelompok tinggi menunjukkan lebih memusat, hal tersebut didukung dengan temuan hasil analisis SD.S diperoleh 1,27. Hasil analisis koefisien variansi kelompok nilai sedang ditemukan 6,40% dan KV nilai tinggi ditemukan 4,04%, berarti kelompok sedang, nilai yang diperoleh dari responden lebih menyebar dibandingkan pada kelompok nilai tinggi. Kelompok rendah nilai yang diperoleh lebih menyebar diantara kelompok tinggi dan sedang. Kelompok rendah mengerjakan/merespon soal rata-rata dalam 80 menit sedangkan kelompok nilai sedang dan tinggi rata-rata waktu mengerjakan soal antara 77-77.12.

2. Hasil Analisis Tingkat Kesukaran Butir Soal *Tryout*

Tabel 1.2 Hasil Analisis Tingkat Kesukaran Butir Soal *Tryout*

Tingkat kesukaran butir soal	<i>Proportion Correct</i>	Hasil
Mudah	$P > 0,70$	20
Sedang	$0,30 \leq p \leq 0,70$	14
Sulit	$P < 0,30$	1

Sebanyak 35 butir soal *tryout* yang diujikan kepada 523 mahasiswa PPG program studi PJOK didapatkan hasil bahwa tingkat kesukaran butir soal *tryout* cenderung pada tingkat kesukaran mudah sebanyak 20 butir soal, sedangkan tingkat kesukaran sedang sebanyak 14 butir soal dan 1 butir soal pada tingkat kesukaran sulit. Tingkat kesukaran pada tiap kategori memiliki *proportion correct* berbeda. Tingkat kesukaran mudah ($p > 0,70$), tingkat kesukaran sedang ($0,30 \leq p \leq 0,70$) dan tingkat kesukaran sulit ($p < 0,30$).

Analisis tingkat kesukaran butir soal *tryout* memiliki 3 kategori yaitu sulit, sedang dan mudah. Tingkat kesukaran mudah ditunjukkan pada nomor butir soal 5,6,7,8,9,11,12,13,14,15,16,19,20,22,23,26,29,31,32,33,34 dan 35. Tingkat kesukaran sulit ditunjukkan pada nomor butir soal 4.

3. Hasil Analisis Daya Beda

Tabel 1.3. Hasil Analisis Daya Beda

Daya Beda	Rentang Skor (D)	Hasil
Jelek	$<0,30$	20
Cukup	$0,30 - 0,70$	15
Baik	$0,70 - 1,00$	0

Sebanyak 35 butir soal *tryout* yang diberikan kepada 523 mahasiswa PPG program studi PJOK terdapat 20 butir soal termasuk dalam kategori cukup dan 15 butir soal termasuk kategori jelek. Tingkat analisis daya beda bertujuan untuk menghasilkan instrumen tes yang berkualitas. Butir soal yang tidak memiliki daya beda yang baik dapat digugurkan atau direvisi dikarenakan dapat merusak validitas, tidak efektif serta mendeteksi kualitas butir soal.

4. Hasil Analisis Validitas Butir Soal

Tabel 1.4. Hasil Analisis Validitas Butir Soal

Validitas	R	Hasil
Valid	$R > 0,3$	24
Tidak Valid	$R < 0,3$	11

Sebanyak 35 butir soal *tryout* yang diberikan kepada 523 mahasiswa PPG program studi PJOK terdapat kategori valid dan tidak valid pada tiap butir tes. Validasi butir soal digunakan untuk mengukur apakah instrumen tes mengukur apa yang seharusnya diukur. Butir soal dianggap valid jika nilai korelasi butir-total berada di atas ambang batas (umumnya $r > 0,3$). Terdapat 11 butir soal yang dianggap tidak valid dikarenakan $r < 0,3$ pada butir soal nomor 3,4,10, 19,21,23,27,28,29,31 dan 33. Namun terdapat data yang menunjukkan butir soal valid tapi beberapa butir soal memiliki korelasi yang rendah pada butir soal nomor 4,19,27, 28,29,31 dan 33 dan dikategorikan sebagai tidak valid dan perlu untuk digugurkan.

5. Hasil Analisis Distraktor

Dari 35 butir soal *tryout* yang diberikan kepada 523 mahasiswa PPG program studi PJOK terdapat dua kategori analisis distraktor yaitu berfungsi dan tidak berfungsi. Terdapat 7 butir soal yang tidak berfungsi sehingga butir soal tersebut perlu dibuang atau digugurkan.

Distraktor bertujuan untuk mengevaluasi fungsi dari pilihan jawaban pengecoh pada tiap butir soal. Distraktor dianggap berfungsi jika pilihan opsi jawaban dipilih oleh responden sedangkan distraktornya dianggap tidak berfungsi atau tidak ada yang memilih opsi jawaban tersebut maka dapat di gugurkan atau di revisi.

6. Hasil Analisis Reliabilitas

Tabel 1.5. Perbandingan Hasil Analisis Reliabilitas Butir Tes

Tahap I		Tahap II	
Cronbach's Alpha	0,823	Cronbach's Alpha	0,831
N of Items	35	N of Items	28

Analisis reliabilitas digunakan untuk mengukur konsistensi suatu instrumen tes. Analisis reliabilitas menggunakan Cronbach's Alpha menggunakan SPSS. Dihasilkan sebesar 0,823 untuk keseluruhan 35 butir soal atau pada tahap awal sebelum digugurkan atau revisi. Kemudian meningkat menjadi 0,831 untuk 28 butir soal yang tersisa atau setelah dilakukan pengguguran. Hasil tersebut menunjukkan bahwa instrumen tes memiliki tingkat reliabilitas yang tinggi.

7. Hasil Analisis Waktu Pengerjaan *Tryout*

Tabel 1.6. Perbandingan Waktu Pengerjaan *Tryout* Mahasiswa PPG PJOK

Kategori Mahasiswa	Rata-rata pengerjaan (menit)	waktu Hasil
Tinggi	77,12	139
Sedang	77,43	230
Rendah	80	154
Total	78,18	523

Berdasarkan tabel 1.6 sebanyak 523 mahasiswa PPG program studi PJOK mengerjakan soal *tryout* dengan waktu pengerjaan yang berbeda-beda. Setelah dilakukan analisis data dan dikategorikan menjadi 3 yaitu kategori mahasiswa tinggi, sedang dan rendah. Kategori mahasiswa tinggi memiliki rata-rata waktu pengerjaan selama 77,12 menit dengan jumlah mahasiswa sebanyak 139. Kategori mahasiswa sedang memiliki rata-rata waktu pengerjaan selama 77,43 menit dengan jumlah mahasiswa sejumlah 230. Sedangkan kategori mahasiswa rendah rata-rata waktu pengerjaan selama 80 menit dengan jumlah mahasiswa 154.

Dapat disimpulkan bahwa jumlah mahasiswa yang lebih banyak mengerjakan soal *tryout* tidak selalu berkorelasi dengan waktu pengerjaan yang lebih cepat. Artinya, meskipun pada kategori mahasiswa tinggi dengan rata-rata waktu pengerjaan tercepat

hanya memiliki jumlah mahasiswa yang sedikit. Sebaliknya pada kategori mahasiswa sedang memiliki rata-rata waktu yang sedikit lebih lama memiliki jumlah mahasiswa paling banyak.

PEMBAHASAN

Butir soal tes yang dianalisa digunakan untuk mengukur potensi akademik mahasiswa PPG program studi PJOK. Dalam proses pengambilan keputusan untuk mereduksi butir soal dari 35 butir soal menjadi 28 butir soal, peneliti mempertimbangkan lima kriteria yaitu 1) tingkat kesukaran butir soal, 2) hasil analisis daya beda, 3) hasil analisis validitas butir soal, 4) hasil analisis distraktor dan 5) hasil analisis reliabilitas.

Tingkat kesukaran butir soal berguna menunjukkan seberapa sulit butir soal tersebut dapat di respon oleh responden (Zein et al., 2018). Diperoleh hasil bahwa tingkat kesukaran butir soal *tryout* cenderung pada tingkat kesukaran mudah sebanyak 20 butir soal, sedangkan tingkat kesukaran sedang sebanyak 14 butir soal dan 1 butir soal pada tingkat kesukaran sulit. Hal tersebut menunjukkan bahwa soal terlalu mudah untuk dikerjakan dikarenakan hampir seluruh responden mampu menjawab pada tiap butir soal tersebut (Basuki & Hariyanto, 2016; Saputra et al., 2017). Data tersebut tidak akan memberikan informasi yang diperlukan tentang kemampuan atau potensi responden (Crocker & Algina, 2008). Selaras dengan pernyataan tersebut, (Saputri et al., 2023) menyatakan bahwa jika butir soal terlalu sulit, responden akan merasa putus asa dan tidak termotivasi menyelesaikan soal ujian TPA tersebut.

Hasil analisis daya beda berguna untuk mengukur kemampuan butir soal untuk membedakan antara responden yang memiliki potensi akademik tinggi maupun rendah (Sudjana, 2012). Hasil uji beda menunjukkan terdapat 20 butir soal termasuk dalam kategori cukup dan 15 butir soal termasuk kategori jelek. Butir soal yang tidak memiliki daya beda yang baik dapat digugurkan atau direvisi dikarenakan dapat merusak validitas, tidak efektif serta mendeteksi kualitas butir soal (Garvin & Ebel, 1980). Tes dikatakan tidak memiliki daya pembeda apabila tes tersebut, jika diujikan kepada responden yang memiliki potensi akademik tinggi hasilnya rendah, tetapi jika diujikan kepada responden yang memiliki potensi akademik rendah hasilnya tinggi. Atau diujikan pada responden yang memiliki kategori tersebut hasilnya sama (Sudjana, 2013).

Hasil analisis validitas butir soal ialah ukuran standar yang menunjukkan tingkat ketepatan suatu instrumen (Nurhalimah et al., 2022). Terdapat 11 butir soal yang dianggap tidak valid dikarenakan $r < 0,3$ pada butir soal nomor 3,4,10, 19,21,23,27,28,29,31 dan 33.

Namun terdapat data yang menunjukkan butir soal valid tapi beberapa butir soal memiliki korelasi yang rendah pada butir soal nomor 4,19,27, 28,29,31 dan 33 dan dikategorikan sebagai tidak valid dan perlu untuk digugurkan. Validitas butir soal tidak ditentukan oleh seberapa banyak responden menjawab pertanyaan dengan benar, namun validitas butir diukur berdasarkan hubungan yang kuat antara skor pada butir soal dengan skor total keseluruhan tes (Zein et al., 2018). Meskipun banyak responden menjawab benar, hal tersebut tidak otomatis butir soal tersebut valid.

Hasil analisis distraktor menunjukkan dua kategori analisis distraktor yaitu berfungsi dan tidak berfungsi. Terdapat 7 butir soal yang tidak berfungsi sehingga butir soal tersebut perlu dibuang atau digugurkan. Hasil distraktor yang baik akan dipilih oleh sebagian besar responden dengan potensi akademik rendah dibandingkan dengan responden akademik tinggi. Jika distraktor tidak berfungsi maka tidak ada yang dipilih oleh responden akademik tinggi sehingga butir soal tersebut perlu direvisi direvisi (Aji & Winarno, 2016; Nurhalimah et al., 2022; Saputra et al., 2017).

Hasil analisis reliabilitas berguna untuk mengukur konsistensi instrumen tes. Hasil diperoleh sebesar 0,823 untuk keseluruhan 35 butir soal atau pada tahap awal sebelum digugurkan atau revisi. Kemudian meningkat menjadi 0,831 untuk 28 butir soal yang tersisa atau setelah dilakukan pengguguran. Hasil tersebut menunjukkan bahwa instrumen tes memiliki tingkat reliabilitas yang tinggi. TPA yang reliabel akan memberikan hasil yang konsisten jika diujikan kembali pada responden dan kondisi yang sama (S. Et al., 1961).

Hasil analisis rata-rata waktu pengerjaan soal *tryout* menunjukkan bahwa jumlah mahasiswa yang lebih banyak mengerjakan soal *tryout* tidak selalu berkorelasi dengan waktu pengerjaan yang lebih cepat. Artinya, meskipun pada kategori mahasiswa tinggi dengan rata-rata waktu pengerjaan tercepat hanya memiliki jumlah mahasiswa yang sedikit. Sebaliknya pada kategori mahasiswa sedang memiliki rata-rata waktu yang sedikit lebih lama memiliki jumlah mahasiswa paling banyak.

Keterkaitan atau hubungan dari 5 indikator tersebut saling terhubung satu sama lain. Hal tersebut didukung pada hasil butir tes yang memiliki reliabilitas yang tinggi terhadap hasil TPA. Didukung pernyataan (Maulina & Novirianthy, 2020) bahwa butir soal yang telah diuji secara mendalam serta memiliki kualitas yang baik akan meningkatkan reliabilitas tes secara keseluruhan. Selaras dengan (Augustia et al., 2025) bahwa tes yang baik harus memenuhi kriteria validitas, reliabilitas, daya pembeda, tingkat kesukaran dan distraktor.

KESIMPULAN

Dapat disimpulkan dari 35 butir soal *tryout* yang telah diberikan pada 523 mahasiswa PPG program studi PJOK masih harus ditingkatkan dengan hasil tingkat kesukaran mudah sebanyak 20 butir soal ($p > 0,70$), sedangkan tingkat kesukaran sedang sebanyak 14 butir soal ($0,30 \leq p \leq 0,70$) dan 1 butir soal pada tingkat kesukaran sulit ($p < 0,30$). Hasil analisis daya beda terdapat 20 butir soal termasuk dalam kategori cukup dan 15 butir soal termasuk kategori jelek. Hasil analisis validitas butir soal terdapat 11 butir soal yang dianggap tidak valid dikarenakan $r < 0,3$. Namun terdapat data yang menunjukkan butir soal valid tapi beberapa butir soal memiliki korelasi yang rendah dikategorikan sebagai tidak valid dan perlu untuk digugurkan. Hasil analisis distraktor terdapat 7 butir soal yang tidak berfungsi sehingga butir soal tersebut perlu dibuang atau digugurkan. Hasil analisis reliabilitas dihasilkan sebesar 0,823 untuk keseluruhan 35 butir soal atau pada tahap awal sebelum digugurkan atau revisi. Kemudian meningkat menjadi 0,831 untuk 28 butir soal yang tersisa atau setelah dilakukan pengguguran. Hasil tersebut menunjukkan bahwa instrumen tes memiliki tingkat reliabilitas yang tinggi

UCAPAN TERIMA KASIH

Seiring dengan selesainya penulisan Artikel karya Ilmiah ini, tak lupa penulis ucapkan banyak terima kasih kepada Prodi Pendidikan Profesi Guru Sekolah Pascasarjana Universitas Negeri Malang atas bantuan dana untuk publish artikel pada jurnal yang terakreditasi.

DAFTAR PUSTAKA

- Aini, S. N., Nurjanah, S., & Pratama, A. (2023). Pengaruh Penggunaan Learning Management System (LMS) dan Motivasi Terhadap Kompetensi Mahasiswa PPG Universitas Negeri Jakarta Tahun 2021. *Jurnal Ilmiah Wahana Pendidikan*, 9(6), 558–569. <https://doi.org/10.5281/zenodo.7789237>
- Aji, B. S., & Winarno, M. E. (2016). Pengembangan Instrumen Penilaian Pengetahuan Mata Pelajaran Pendidikan Jasmani Olahraga dan Kesehatan (PJOK) Kelas VIII Semester Gasal. *Jurnal Pendidikan: Teori, Penelitian, Dan Pengembangan*, 1(7), 1449–1463. <https://doi.org/http://dx.doi.org/10.17977/jp.v1i7.6594>
- Augustia, A. D., Agustia, C. N., Azzahra, D., Hakim, L., & Pratiwi, V. (2025). Analisis Validitas dan Reliabilitas Soal Pilihan Ganda dengan Menggunakan Software Anates pada Mata Pelajaran Perpajakan. *AKADEMIK: Jurnal Mahasiswa Ekonomi & Bisnis*, 5(1), 250–265. <https://doi.org/10.37481/jmeb.v5i1.1165>
- Basuki, I., & Hariyanto. (2016). Asesmen Pembelajaran. Remaja Rosdakarya.
- Crocker, L., & Algina, J. (2008). Introduction to classical and modern test theory- Procedures for Estimating Reliability. In Harcourt Brace Jovanovich College.
- Garvin, A. D., & Ebel, R. L. (1980). Essentials of Educational Measurement. *Educational Researcher*, 9(9). <https://doi.org/10.2307/1175572>

- Maulina, N., & Noviriany, R. (2020). Item Analysis And Peer-Review Evaluation Of Specific Health Problems And Applied Research Block Examination. *Jurnal Pendidikan Kedokteran Indonesia: The Indonesian Journal of Medical Education*, 9(2). <https://doi.org/10.22146/jpki.49006>
- Mustafa, P. S., & Masgumelar, N. K. (2022). Kajian Review: Pengembangan Instrumen Penilaian Sikap, Pengetahuan, dan Keterampilan dalam Pendidikan Jasmani dan Olahraga. *Biomatika: Jurnal Ilmiah Fakultas Keguruan Dan Ilmu Pendidikan*, 8(1), 35–36.
- Nurhalimah, S., Hidayati, Y., Rosidi, I., & Hadi, W. P. (2022). The Relationship Between Item Validity With Distinguishing Power And Difficulty Level Of Multiple Choice Questions PAS (in Indonesian). *Natural Science Education Research*, 4(3), 249–257.
- Peraturan Direktur Jenderal GTK. (2023). Tentang Model Kompetensi Guru Nomor 2626/B/HK.04.01/2023. Peraturan Pemerintah, 1–14.
- Pratama, D. C., Sulaiman, S., & Soegiyanto, S. (2021). The Effect of Teaching Certification on the Performance of Physical Education Sports and Health Teacher in Paser Regency , East Kalimantan. *Journal of Physical Education and Sports*, 10(4).
- Rambe, F. A., Siregar, S., Yadaini, S. S., & Ramahsita, S. (2024). Tantangan Guru PJOK Dalam Mengoptimalkan Media Pembelajaran. 2(2), 81–83.
- Ridwan, N. A., & Fadli, M. (2024). Implementasi Etos Kerja Guru Pjok Di Era Society 5.0. *Jurnal Pembelajaran, Bimbingan, Dan Pengelolaan Pendidikan*, 4(8), 17. <https://doi.org/10.17977/um065.v4.i8.2024.17>
- S., R., Thorndike, R. L., & Hagen, E. (1961). Measurement and Evaluation in Psychology and Education. *Journal of the American Statistical Association*, 56(296). <https://doi.org/10.2307/2282039>
- Saputra, M. H. Y., Arthana, I. K. R., & Santyadiputra, G. S. (2017). Simatik : Aplikasi Simulasi Bank Soal Tes Potensi Akademik (TPA) Berbasis Multi Platform. *JST (Jurnal Sains Dan Teknologi)*, 5(2). <https://doi.org/10.23887/jst-undiksha.v5i2.8621>
- Saputri, H. A., Zuhijrah, 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).
- Sarea, M. S., & Ruslan, R. (2019). Karakteristik Butir Soal: Classical Test Theory Vs Item Response Theory? *Didaktika : Jurnal Kependidikan*, 13(1). <https://doi.org/10.30863/didaktika.v13i1.296>
- Sudjana, N. (2013). Penilaian Hasil Proses Belajar Mengajar Cetakan ketujuh belas. In *Penilaian dan Hasil Belajar Mengajar*.
- Wantoro, J., Utama, S., Zuhriah, S., & Hafida, S. H. N. (2019). Pengembangan Instrumen Penilaian Pendidikan Profesi Guru Sekolah Dasar Berbasis Hots. *Profesi Pendidikan Dasar*, 1(1). <https://doi.org/10.23917/ppd.v1i1.8453>
- Zein, A., Fadillah, M., & Novianti, R. (2018). Hubungan Antara Validitas Butir, Reliabilitas, Tingkat Kesukaran Dan Daya Pembeda Soal Ujian Semester Genap Bidang Studi Biologi Kelas Xi Sma / Ma Negeri Di Kota Padang. *Semirata 2013 FMIPA Unila*, 2009, 39–47.