

**ANALISIS TINGKAT KESUKARAN BUTIR SOAL TRY OUT BUSANA PPG  
BAGI GURU TERTENTU TAHAP 1 UNIVERSITAS NEGERI MALANG  
TAHUN 2025**

<sup>1</sup>Nurul Aini, <sup>2</sup>Mochammad Rizal Ramadhan, <sup>3</sup>Ghaly Wijdan  
<sup>1,2,3</sup>Pendidikan Profesi Guru, Universitas Negeri Malang,  
<sup>1</sup>nurul.aini.ft@um.ac.id, <sup>2</sup>mochhammad.ramadhan.fs@um.ac.id,  
<sup>3</sup>ghaly.wijdan.2104136@students.um.ac.id

**ABSTRACT**

*This study aims to analyze the difficulty level of the PPG Clothing for Certain Teachers (BGT) tryout items for Phase 1 at Malang State University in 2025. The study subjects consisted of 68 participants with a total of 35 questions. The analysis results showed that the average difficulty level index (P) was 0.9. This value indicates that most of the questions are relatively easy, because approximately 90% of participants were able to answer correctly. Whereas in classical test theory, good instrument quality is indicated by an average difficulty index approaching 0.5 (moderate category). Thus, this condition reflects that the test instrument is not fully able to optimally differentiate participants with high and low abilities. These findings have important implications for the quality of evaluation instruments in the context of teacher professional education. Questions with a difficulty level that is too easy tend to reduce the item discrimination power, making it difficult to comprehensively measure variations in participant abilities. In addition, the presence of questions that are too easy has the potential to result in erroneous interpretations regarding participant readiness for the actual PPG exam. Therefore, improvements are needed in the preparation of test instruments. Some strategies that can be adopted include balancing the distribution of easy, medium, and difficult questions, developing questions based on Higher Order Thinking Skills (HOTS), improving the quality of distractors, and piloting questions. This will ensure the resulting instrument is more valid, reliable, and representative of the competencies of fashion teachers.*

*Keywords: level of difficulty, try out, PPG*

**ABSTRAK**

Penelitian ini bertujuan untuk menganalisis tingkat kesukaran butir soal try out PPG Busana bagi Guru Tertentu (BGT) Tahap 1 di Universitas Negeri Malang tahun 2025. Subjek penelitian terdiri atas 68 peserta dengan jumlah soal sebanyak 35 butir. Hasil analisis menunjukkan bahwa rata-rata indeks tingkat kesukaran (P) sebesar 0,9. Nilai ini mengindikasikan bahwa sebagian besar soal tergolong mudah, karena sekitar 90% peserta mampu menjawab dengan benar. Sementara itu, dalam teori tes klasik, kualitas instrumen yang baik ditunjukkan oleh rata-rata indeks kesukaran yang mendekati 0,5 (kategori sedang). Dengan demikian, kondisi ini mencerminkan bahwa instrumen tes belum sepenuhnya mampu membedakan secara optimal antara peserta yang berkemampuan tinggi dan rendah. Temuan ini memiliki implikasi penting terhadap kualitas instrumen evaluasi dalam konteks

pendidikan profesi guru. Soal dengan tingkat kesukaran yang terlalu mudah cenderung menurunkan daya pembeda butir, sehingga menyulitkan pengukuran variasi kemampuan peserta secara komprehensif. Selain itu, keberadaan soal yang terlalu mudah berpotensi menimbulkan interpretasi yang keliru terkait kesiapan peserta menghadapi ujian PPG yang sesungguhnya. Oleh karena itu, diperlukan perbaikan dalam penyusunan instrumen tes. Beberapa strategi yang dapat diterapkan antara lain menyeimbangkan distribusi soal mudah, sedang, dan sukar, mengembangkan soal berbasis Higher Order Thinking Skills (HOTS), meningkatkan kualitas distraktor, serta melakukan uji coba (*try out*) soal. Dengan demikian, instrumen yang dihasilkan akan lebih valid, reliabel, dan representatif terhadap kompetensi guru busana.

Kata kunci: tingkat kesukaran; *try out*; PPG

## **A. Pendahuluan**

Pendidikan Profesi Guru (PPG) bertujuan untuk menghasilkan guru profesional yang memiliki kompetensi pedagogik, kepribadian, sosial, dan profesional sesuai dengan standar nasional pendidikan (Mardhatillah, O., dan Surjanti, J., 2023). Salah satu tahapan penting dalam proses PPG adalah pelaksanaan *try out* atau uji coba soal sebelum peserta menghadapi ujian yang sesungguhnya. *Try out* ini bertujuan untuk memberikan pengalaman, memetakan kesiapan peserta, serta mengidentifikasi kelemahan dan kelebihan dalam penguasaan materi (Pangesti, F., Fauzan, F., Risnawati, R., 2020). Dalam bidang studi Busana, yang merupakan bagian dari rumpun kejuruan Tata Busana, penyusunan butir soal *try out* harus mempertimbangkan karakteristik

keilmuan yang tidak hanya bersifat teoritis, tetapi juga aplikatif. Oleh karena itu, kualitas soal menjadi sangat penting agar hasil *try out* benar-benar dapat mencerminkan kemampuan peserta secara komprehensif.

Salah satu aspek fundamental dalam analisis kualitas soal adalah tingkat kesukaran (*difficulty level*). Tingkat kesukaran mengacu pada sejauh mana suatu butir soal dapat dikerjakan dengan benar oleh peserta tes (Saputri, H. A. S., dan Larasati, N. J., 2023).. Soal yang baik adalah soal yang tidak terlalu mudah dan tidak terlalu sulit, sehingga mampu memberikan gambaran yang proporsional mengenai variasi kemampuan peserta. Jika sebagian besar soal terlalu mudah, maka soal tersebut tidak mampu membedakan peserta yang benar-benar menguasai

materi dengan peserta yang hanya memiliki pengetahuan dasar. Sebaliknya, jika soal terlalu sulit, maka akan membuat peserta merasa frustrasi, bahkan dapat menurunkan motivasi belajar serta mengaburkan penilaian objektif atas kompetensi yang dimiliki.

Dalam praktiknya, penyusunan soal untuk *try out* PPG bidang Busana masih menghadapi sejumlah tantangan. Beberapa laporan dan hasil evaluasi menunjukkan bahwa terdapat butir soal yang terlalu mudah dikerjakan karena hanya mengukur hafalan, bukan keterampilan berpikir kritis dan aplikatif. Di sisi lain, terdapat pula soal yang dinilai terlalu sulit karena memuat materi di luar cakupan atau menggunakan konteks yang kurang familiar bagi peserta. Hal ini tentu menimbulkan permasalahan serius, karena hasil *try out* seharusnya menjadi cerminan nyata kesiapan peserta dalam menghadapi ujian PPG.

Universitas Negeri Malang (UM) sebagai salah satu perguruan tinggi penyelenggara PPG di Indonesia memiliki tanggung jawab besar dalam menjaga mutu instrumen evaluasi, termasuk soal *try out*. Tahun 2025 merupakan periode penting

karena di era ini, pemerintah semakin menekankan pentingnya asesmen berbasis kompetensi, termasuk penerapan soal-soal yang mendorong keterampilan berpikir tingkat tinggi (HOTS). Oleh karena itu, butir soal *try out* Busana PPG harus mampu mengukur penguasaan konsep, kemampuan analisis, evaluasi, hingga kreativitas peserta dalam menyelesaikan masalah kontekstual di bidang busana.

Permasalahan yang kerap muncul adalah ketidakseimbangan proporsi tingkat kesukaran butir soal. Dalam beberapa analisis sebelumnya, ditemukan bahwa terdapat dominasi soal dengan kategori mudah atau sebaliknya, terlalu sulit, sehingga distribusi tingkat kesukaran tidak proporsional. Padahal, dalam teori evaluasi pendidikan, idealnya soal yang baik memiliki komposisi seimbang: sebagian kecil mudah, sebagian besar sedang, dan sebagian kecil sulit. Dengan komposisi tersebut, instrumen akan lebih representatif dalam mengukur kemampuan peserta yang beragam. Lebih jauh, tingkat kesukaran butir soal tidak hanya berdampak pada hasil tes secara langsung, tetapi juga berpengaruh

terhadap aspek lain, seperti daya beda dan efektivitas pengecoh.

Menurut Savika, H. I., & Zuhriyah, I. A. (2024) bahwa menganalisis butir soal itu sangat penting. Bagi PPG adalah tuntutan mutu asesmen PPG, kebutuhan instrumen evaluasi yang representatif, menuntut perubahan paradigma asesmen menuju pengukuran keterampilan berpikir tingkat tinggi, dan instrumen dengan tingkat kesukaran seimbang agar hasil tes benar-benar objektif.

Dengan melihat berbagai faktor di atas, penelitian mengenai analisis tingkat kesukaran butir soal *try out* Busana PPG bagi Guru Tertentu Tahap 1 Universitas Negeri Malang Tahun 2025 menjadi sangat relevan. Penelitian ini diharapkan dapat memberikan gambaran objektif mengenai distribusi tingkat kesukaran soal, mengidentifikasi butir-butir yang perlu direvisi, serta memberikan rekomendasi untuk penyusunan soal yang lebih berkualitas di masa mendatang.

## **B. Metode penelitian**

Penelitian ini menggunakan pendekatan kuantitatif dengan metode deskriptif. Alasan pemilihan pendekatan kuantitatif adalah karena

penelitian ini berfokus pada analisis data numerik hasil jawaban peserta *try out* untuk menentukan tingkat kesukaran butir soal. Sementara itu, rancangan deskriptif digunakan untuk memberikan gambaran secara sistematis dan objektif mengenai kualitas soal berdasarkan kategori mudah, sedang, dan sukar. Penelitian ini tidak berupaya memanipulasi variabel, melainkan menganalisis data yang sudah ada (*ex post facto*). Dengan demikian, penelitian ini termasuk dalam jenis penelitian evaluatif (Rini, S. A., Huda, N., Hermina, D., 2024). yang bertujuan mengkaji kelayakan instrumen evaluasi, khususnya pada aspek tingkat kesukaran soal *try out*.

Populasi penelitian adalah seluruh peserta *try out* PPG Guru Tertentu Bidang Busana Tahap 1 Universitas Negeri Malang Tahun 2025 yang berjumlah 68 orang. Populasi ini terdiri dari guru-guru bidang Tata Busana yang mengikuti program PPG dan diwajibkan untuk mengikuti *try out* sebelum ujian kompetensi. Sampel penelitian ditentukan dengan menggunakan sampel jenuh atau seluruh populasi dijadikan sampel (Fajri, C., Amelya, A., Suworo, S., 2022)., yakni seluruh

peserta *try out* dijadikan sampel penelitian. Hal ini dilakukan karena jumlah peserta relatif terbatas serta data hasil tes tersedia secara menyeluruh. Dengan demikian, penelitian ini memiliki tingkat representativitas yang tinggi terhadap kondisi populasi yang diteliti.

Instrumen penelitian berupa paket soal *try out* bidang Busana yang digunakan pada tahap 1 PPG Guru Tertentu Universitas Negeri Malang tahun 2025. Soal berbentuk pilihan ganda dengan lima opsi jawaban (A–E), yang disusun berdasarkan indikator kompetensi yang tercantum dalam kurikulum PPG. Selain itu, instrumen pendukung berupa lembar jawaban peserta (LJK) atau data hasil tes dalam bentuk digital digunakan sebagai sumber data untuk perhitungan tingkat kesukaran.

Pengumpulan data dilakukan dengan cara dokumentasi. Peneliti memperoleh data yang terdiri dari: 1) naskah soal *try out*, untuk dianalisis kesesuaiannya dengan indikator kompetensi, 2) data hasil pengerjaan soal oleh peserta, baik dalam bentuk LJK maupun rekap skor digital, dan 3) data pendukung, seperti jumlah peserta, distribusi skor, dan informasi

teknis penyelenggaraan *try out*. Data tersebut kemudian diolah untuk menghitung tingkat kesukaran tiap butir soal.

Analisis data difokuskan pada perhitungan tingkat kesukaran butir soal. Tingkat kesukaran menunjukkan proporsi peserta yang mampu menjawab benar suatu soal dibandingkan jumlah keseluruhan peserta. Rumus yang digunakan adalah sebagai berikut (Rajagukguk, M. J. T., dan Naibaho, D., 2023):

$$P = \frac{B}{N}$$

Keterangan:

P = indeks tingkat kesukaran

B = Jumlah peserta yang menjawab benar

N = Jumlah seluruh peserta

Interpretasi nilai P adalah sebagai berikut:

$0.00 \leq P < 0.30$  ; soal sukar

$0.30 \leq P \leq 0.70$  : soal sedang

$0.70 \leq P \leq 1.00$  : soal mudah

Hasil analisis kemudian dikategorikan sesuai klasifikasi di atas, sehingga dapat diketahui distribusi soal yang termasuk kategori mudah, sedang, dan sukar. Selain analisis kuantitatif, dilakukan pula analisis kualitatif sederhana terhadap

soal-soal yang termasuk kategori terlalu mudah atau terlalu sukar. Analisis ini mencakup identifikasi penyebab kemungkinan soal menjadi tidak proporsional, misalnya: penggunaan kata yang ambigu, konteks materi yang terlalu sederhana/kompleks, atau indikator soal yang kurang tepat.

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

#### **Hasil Penelitian**

Penelitian ini dilaksanakan untuk menganalisis kualitas instrumen tes *try out* Program Pendidikan Profesi Guru (PPG) Bidang Studi Busana bagi Guru Tertentu (BGT) Tahap 1 di Universitas Negeri Malang tahun 2025. Instrumen tes berupa 35 butir soal pilihan ganda yang dikerjakan oleh 68 peserta. Fokus penelitian adalah menghitung dan menafsirkan indeks tingkat kesukaran butir soal, yang merupakan salah satu indikator kualitas instrumen penilaian.

Dari hasil analisis menggunakan Indeks tingkat kesukaran (*difficulty index*) merupakan ukuran yang menunjukkan sejauh mana suatu butir soal mudah atau sulit bagi peserta tes., diperoleh

rata-rata indeks tingkat kesukaran (P) sebesar 0,9 (Gambar 1. Data Hasil Screening dan Analisis Butir Soal Try Out Busana). Dengan mengacu pada interpretasi nilai P di atas, maka nilai ini menunjukkan bahwa secara umum soal berada pada kategori sangat mudah. Kondisi dapat disimpulkan bahwa;

1. Sekitar 61–62 peserta dari 68 dapat menjawab benar sebagian besar soal,
2. Hanya sekitar 6–7 peserta yang cenderung melakukan kesalahan pada soal tertentu
3. Hampir tidak ada soal yang termasuk kategori sukar.

Secara umum, hal ini mengindikasikan bahwa:

1. Mayoritas peserta mampu mengerjakan soal dengan benar,
2. Instrumen kurang mampu membedakan kemampuan peserta yang beragam,
3. Memerlukan perbaikan kualitas instrumen agar lebih seimbang dari segi tingkat kesukaran.

Dengan demikian, meskipun soal *try out* memberikan gambaran penguasaan awal peserta, namun untuk tujuan seleksi dan evaluasi kompetensi profesional, instrumen ini masih perlu direvisi agar dapat

memberikan hasil yang lebih akurat dan representatif.

Di bawah ini merupakan Gambar Data Hasil Screening dan Analisis Butir Soal *Try Out* Busana.

Tingkat Kesukaran	0.900420168
-------------------	-------------

374



## **Pembahasan**

Salah satu aspek penting dalam evaluasi kualitas butir soal adalah tingkat kesukaran. Tingkat kesukaran menunjukkan sejauh mana sebuah soal dapat dikerjakan oleh peserta, sehingga menjadi indikator apakah soal tersebut tergolong mudah, sedang, atau sulit (Nurhalimah, S., Hidayati, Y., Rosidi, I., & Hadi, W. P., 2022). Hasil analisis terhadap butir soal *try out* Busana PPG Bagi Guru Tertentu (BGT) Tahap 1 Universitas Negeri Malang Tahun 2025 menunjukkan bahwa rata-rata indeks kesukaran (P) adalah 0,9.

Nilai  $P = 0,9$  berarti 90% peserta mampu menjawab soal dengan benar. Berdasarkan kriteria klasifikasi, soal dengan nilai  $P > 0,71$  termasuk kategori mudah. Dengan demikian, dapat disimpulkan bahwa sebagian besar butir soal *try out* ini memiliki tingkat kesukaran yang rendah atau terlalu mudah bagi peserta. Dalam beberapa teori disebutkan bahwa untuk tingkat kesukaran butir soal ada pada kategori sedang. Seperti dinyatakan Erfan, M., dkk, 2020) menyebutkan bahwa dalam teori tes klasik, kualitas instrumen yang baik ditunjukkan oleh

rata-rata indeks kesukaran mendekati 0,5 (kategori sedang).

*Try out* diikuti oleh 68 guru bidang busana dari berbagai sekolah menengah kejuruan (SMK). Jumlah butir soal yang digunakan adalah 35 soal pilihan ganda yang mencakup kompetensi pedagogik dan profesional guru, serta materi spesifik bidang busana seperti teknik menjahit, desain, pola dasar, dan teknologi tekstil. Adapun distribusi tingkat kesukaran nya adalah sebagai berikut ( Rahayu, R., & Djazari, M., 2016):

- Soal mudah ( $P > 0,71$ ) = 30 butir (85,7%)
- Soal sedang ( $0,31 \leq P \leq 0,70$ ) = 5 butir (14,3%)
- Soal sulit ( $P < 0,30$ ) = 0 butir (0%)

Kondisi ini menunjukkan bahwa hampir seluruh soal termasuk kategori mudah, sementara tidak ditemukan butir soal yang tergolong sulit.

Dominasi soal mudah memiliki beberapa dampak, baik positif maupun negatif (Tarmizi, P. dkk., 2020). Dari sisi positif, peserta merasa lebih percaya diri dan bersemangat karena mampu menjawab sebagian besar soal

dengan benar. Hal ini dapat mengurangi kecemasan peserta dalam menghadapi *try out*. Dari sisi negatif, soal yang terlalu mudah tidak mampu memberikan gambaran akurat tentang variasi kemampuan peserta. Nilai rata-rata yang tinggi bisa menimbulkan kesan seolah-olah semua peserta telah menguasai materi dengan baik, padahal belum tentu. Instrumen *try out* dengan dominasi soal mudah juga memiliki kelemahan dalam hal daya beda, karena tidak bisa membedakan peserta yang benar-benar kompeten dengan peserta yang masih lemah.

#### **1. Faktor-faktor yang berkontribusi sehingga butir soal cenderung mudah bagi peserta**

##### ***Konstruksi Soal yang Terlalu Sederhana***

Banyak soal dalam *try out* hanya mengukur pengetahuan faktual (*recall*), misalnya sekadar menanyakan definisi istilah, nama bahan tekstil, atau fungsi alat menjahit. Soal semacam ini relatif mudah dijawab karena hanya menuntut hafalan, bukan pemahaman mendalam, aplikasi, maupun analisis (Christyowati, Y. I., & Ruffi, R. I., 2023). Selain itu, penggunaan

stimulus soal yang kurang variatif membuat soal lebih cepat ditebak jawabannya.

##### ***Distraktor Kurang Efektif***

Dalam soal pilihan ganda, kualitas pengecoh (distraktor) sangat menentukan tingkat kesukaran (Putri, R. A., dan Rosliyah, Y., 2020). Pada *try out* ini, sebagian besar distraktor dinilai kurang menarik atau tidak logis, sehingga jawaban benar tampak menonjol. Peserta dengan pengetahuan minimal pun bisa dengan mudah mengidentifikasi jawaban benar hanya dengan mengeliminasi opsi yang jelas salah. Hal ini membuat soal terasa mudah bagi hampir semua peserta.

##### ***Materi yang Sangat Familiar***

Materi soal yang diujikan sebagian besar bersifat dasar dan sering muncul dalam praktik pembelajaran busana di SMK. Sebagai guru yang sehari-hari mengajar bidang busana, peserta sudah sangat akrab dengan konsep-konsep dasar seperti jenis kain, teknik menjahit sederhana, atau langkah pembuatan pola. Familiaritas materi ini menjadikan soal lebih mudah dijawab dengan benar (Rajagukguk, M. J. T., dan Naibaho, D., 2023)

### ***Pengalaman Peserta yang Tinggi***

Peserta *try out* adalah guru yang telah memiliki pengalaman mengajar, bahkan sebagian sudah lebih dari 10 tahun mengajar mata pelajaran tata busana. Pengalaman panjang tersebut membuat mereka sudah terbiasa dengan permasalahan teknis maupun teoritis di bidang busana, sehingga butir soal berbasis pengetahuan dasar tidak menjadi tantangan berarti.

### ***Strategi Penyusun Soal yang Terlalu Hati-Hati***

Penyusun soal tampaknya memilih pendekatan konservatif dengan membuat soal aman agar peserta tidak mengalami kesulitan berarti. Tujuannya mungkin untuk mengurangi kecemasan peserta dan memberi pengalaman awal yang positif dalam *try out*. Namun strategi ini justru menghasilkan soal yang dominan mudah dan tidak proporsional dalam mengukur variasi kemampuan peserta.

## **2. Strategi mengatasi butir soal yang memiliki tingkat kesukaran yang rendah atau terlalu mudah bagi peserta**

### ***Menyusun Proporsi Tingkat Kesukaran yang Seimbang***

Distribusi tingkat kesukaran soal sebaiknya disusun dengan proporsi yang seimbang, yaitu: 1) 30% soal mudah untuk mengukur penguasaan konsep dasar, 2) 40% soal sedang untuk menguji kemampuan penerapan dan pemahaman mendalam, dan 3) 30% soal sulit untuk menantang kemampuan analisis, evaluasi, dan sintesis peserta (Zuhri, N. Z., Syihabuddin, S., & Tatang, T., 2024)

Dengan proporsi ini, variasi kemampuan peserta dapat teridentifikasi dengan baik, sehingga hasil *try out* lebih objektif.

### ***Mengembangkan Soal Berbasis HOTS (Higher Order Thinking Skills)***

Soal *try out* sebaiknya tidak hanya berfokus pada level kognitif rendah (C1–C2: mengingat dan memahami), tetapi juga mencakup level kognitif lebih tinggi (C3–C6). Misalnya: 1) C3 (menerapkan): Menggunakan konsep dalam situasi nyata, 2) C4 (menganalisis): Menguraikan suatu permasalahan dalam pembelajaran busana, 3) C5 (mengevaluasi): Menilai efektivitas strategi

pembelajaran tertentu, dan 4.) C6 (mencipta): Merancang solusi kreatif dalam konteks produksi busana atau pembelajaran (Markhamah, N., 2021)

### **Memperbaiki Kualitas Distraktor**

Distraktor (jawaban pengecoh) harus dibuat logis, relevan, dan tidak terlalu mudah dieliminasi. Beberapa langkah yang bisa dilakukan: 1) hindari penggunaan jawaban pengecoh yang jelas salah, 2) buat distraktor berdasarkan kesalahan umum siswa, dan 3) menyeimbangkan panjang kalimat antar opsi agar jawaban benar tidak terlihat menonjol.

Dengan distraktor yang berkualitas, peserta harus benar-benar berpikir kritis untuk menemukan jawaban (Prawiki, S. M., & Helendra, H., 2022).

### **Variasi Bentuk Soal**

Meskipun soal pilihan ganda efektif untuk evaluasi massal, bentuk soal dapat divariasikan agar menantang peserta (Ni'mah, U., dan Novikasari, I. (2024). Misalnya: 1) soal berbasis studi kasus, 2) soal analisis gambar pola busana, 3) soal dengan data tabel atau diagram tekstil, dan 4) soal

yang mengintegrasikan materi pedagogik dengan praktik busana.

### **Uji Coba**

Sebelum digunakan secara luas, soal perlu melalui tahap uji coba pada kelompok kecil peserta. Hasil uji coba kemudian dianalisis menggunakan parameter tingkat kesukaran, daya beda, dan efektivitas distraktor. Dengan demikian, butir soal yang tidak sesuai dapat diperbaiki atau diganti sebelum digunakan dalam try out resmi (Oktaviranda, A., dan Asmara, A. 2021).

Dengan implementasi strategi tersebut, diharapkan soal *try out* ke depan memiliki kualitas yang lebih baik, mampu membedakan kemampuan peserta secara objektif, serta relevan dengan tujuan PPG dalam mencetak guru busana profesional dan kompeten.

### **D. Kesimpulan**

Rata-rata indeks kesukaran sebesar 0,9 menunjukkan bahwa mayoritas peserta mampu menjawab dengan benar hampir seluruh soal. Artinya, butir soal yang diujikan tidak menimbulkan tantangan yang berarti bagi peserta, sehingga kesulitan soal tidak mampu mencerminkan variasi

tingkat kemampuan yang sesungguhnya. Nilai indeks kesukaran yang tinggi berimplikasi pada kurangnya variasi skor. Sebagian besar peserta memperoleh nilai yang relatif tinggi, sehingga instrumen tidak cukup efektif untuk membedakan peserta dengan kemampuan rendah, sedang, dan tinggi. Tes semacam ini lebih merepresentasikan “penguasaan umum” ketimbang “kemampuan selektif”. Dalam teori tes klasik, kualitas instrumen yang baik ditunjukkan oleh rata-rata indeks kesukaran mendekati 0,5 (kategori sedang). Hal ini karena soal sedang lebih mampu menyaring peserta dengan kemampuan beragam. Dengan hasil 0,9, dapat disimpulkan bahwa instrumen tes ini masih perlu revisi agar dapat memenuhi prinsip keseimbangan kualitas butir soal.

Berdasarkan hasil penelitian yang diperoleh, maka terdapat beberapa saran yang perlu dilakukan, yaitu :

Diversifikasi tingkat kesukaran, soal sebaiknya tidak didominasi oleh kategori mudah. Penyusunan instrumen perlu mengacu pada komposisi ideal: 1) 30% soal mudah

( $P = 0,71-1,00$ ), 2) 40% soal sedang ( $P = 0,31-0,70$ ), dan 3) 30% soal sukar ( $P = 0,00-0,30$ ). Dengan pola ini, tes akan memiliki tingkat kesukaran rata-rata mendekati sedang, sehingga lebih representatif dalam mengukur variasi kemampuan peserta.

Meningkatkan level kognitif soal, yaitu dengan menyusun soal diarahkan pada ranah berpikir tingkat tinggi (C4–C6). Misalnya: 1) Soal yang menuntut analisis pola desain busan 2) Studi kasus tentang efisiensi pemakaian bahan, dan 3) Evaluasi kesesuaian proses produksi dengan standar mutu industri. Dengan cara ini, soal lebih menantang dan mencerminkan kemampuan *problem solving* peserta.

Penguatan distraktor (pengecoh), dimana pilihan jawaban dalam soal perlu dibuat lebih logis dan dekat dengan jawaban benar. Distraktor yang terlalu lemah membuat soal mudah ditebak. Dengan distraktor yang baik, peserta dipaksa untuk benar-benar berpikir sebelum menentukan jawaban.

Uji coba instrumen, sebaiknya sebelum digunakan dalam *try out* resmi, soal sebaiknya diujicobakan

pada sampel kecil peserta untuk mengetahui distribusi tingkat kesukaran. Dengan demikian, soal yang terlalu mudah dapat direvisi sebelum digunakan secara luas.

#### **DAFTAR PUSTAKA**

- Christyowati, Y. I., & Ruffi'i, R. I. (2023). Analisis Butir Soal Pilihan Ganda Materi Pernapasan Pada Manusia Kelas V SD Santa Maria Surabaya. *Dharmas Education Journal (DE\_Journal)*, 4(2), 874-880.
- Erfan, M., Maulyda, M. A., Hidayati, V. R., Astria, F. P., & Ratu, T. (2020). Analisis kualitas soal kemampuan membedakan rangkaian seri dan paralel melalui teori tes klasik dan model rasch. *Indonesian Journal of Educational Research and Review*, 3(1), 11-19.
- Fajri, C., Amelya, A., & Suworo, S. (2022). Pengaruh Kepuasan Kerja dan Disiplin Kerja terhadap Kinerja Karyawan PT. Indonesia Applicad. *JlIP-Jurnal Ilmiah Ilmu Pendidikan*, 5(1), 369-373.
- Ni'mah, U. S. W. A. T. U. N., & Novikasari, I. (2024). Analisis Butir Soal Ulangan Akhir Semester I Mata Pelajaran Bahasa Indonesia Kelas II Madrasah Ibtidaiyah: Analisis Butir Soal Ulangan Akhir Semester I Mata Pelajaran Bahasa Indonesia Kelas II Madrasah Ibtidaiyah. *J-PGMI: Jurnal Pendidikan Guru MI*, 7(1), 85-100.
- Nurhalimah, S., Hidayati, Y., Rosidi, I., & Hadi, W. P. (2022). Hubungan antara validitas item dengan daya pembeda dan tingkat kesukaran soal pilihan ganda pas. *Natural Science Education Research*, 4(3), 249-257.
- 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.
- Oktaviranda, A., & Asmara, A. (2021). Kemampuan Literasi Melalui Pengembangan Soal pada Materi Statistika. *JKPM (Jurnal Kajian Pendidikan Matematika)*, 7(1), 15-26.
- Putri, R. A., & Rosliyah, Y. (2020). Analisis Distraktor Butir Soal Bunpou Shokyuu Kouhan Semester 2 Tahun 2018/2019 Prodi Pendidikan Bahasa Jepang Unnes. *Kiryoku*, 4(1), 20-28.
- Prawiki, S. M., & Helendra, H. (2022). Analisis Kualitas Butir Soal Ujian Akhir Semester Ganjil Tahun Pelajaran 2020/2021 Mata Pelajaran Biologi Kelas X SMA Negeri 1 Teluk Sebong. *Biodidaktika: Jurnal Biologi dan Pembelajarannya*, 17(2).
- Pangesti, F., Fauzan, F., & Risnawati, R. (2020). Kualitas butir soal try out uji pengetahuan dalam memprediksi tingkat kelulusan mahasiswa PPG. *Jurnal Pendidikan Profesi Guru*, 1(2), 91-98.

- Rahayu, R., & Djazari, M. (2016). Analisis kualitas soal pra ujian nasional mata pelajaran ekonomi akuntansi. *Jurnal Pendidikan Akuntansi Indonesia*, 14(1).
- Rajagukguk, M. J. T., & Naibaho, D. (2023). Mampu Memilih Soal Berdasarkan Tingkat Kesukaran. *Jurnal Pendidikan Sosial Dan Humaniora*, 2(4), 12736-12747.
- Rini, S. A., Huda, N., & Hermina, D. (2024). Model-model metode penelitian evaluasi. *Jurnal Pendidikan Sosial dan Humaniora*, 3(2), 1427-1435.
- Saputri, H. A. S., & Larasati, N. J. (2023). Analisis Instrumen Assesmen: Validitas, Reliabilitas, Tingkat Kesukaran Dan Daya Beda Butir Soal. *Didaktik: Jurnal Ilmiah PGSD STKIP Subang*, 9(5), 2986-2995.
- Savika, H. I., & Zuhriyah, I. A. (2024). Peran Analisis Butir Soal terhadap Kualitas Soal, Kompetensi Guru, dan Prestasi Belajar Peserta Didik di Sekolah Dasar. *PANDU: Jurnal Pendidikan Anak dan Pendidikan Umum*, 2(2), 43-51.
- Tarmizi, P., Setiono, P., Amaliyah, Y., & Agrian, A. (2020). Analisis butir soal pilihan ganda tema sehat itu penting kelas V SD Negeri 04 Kota Bengkulu. *ELSE (Elementary School Education Journal): Jurnal Pendidikan dan Pembelajaran Sekolah Dasar*, 4(2), 124-132.
- Zuhri, N. Z., Syihabuddin, S., & Tatang, T. (2024). Analisis Validitas, Reliabilitas, dan Tingkat Kesukaran Soal Bahasa Arab Tingkat SMP Berbasis Artificial Intelligence (AI) melalui Platform QuestionWell. *Jurnal Pendidikan Dan Pembelajaran Indonesia (JPPI)*, 4(2), 693-704.