

## **SOAL PILIHAN GANDA YANG BERKUALITAS PADA MATA PELAJARAN BAHASA JAWA**

**Rosalia Indah, Mujimin, Sungging Widagdo**

Program Studi Pendidikan Bahasa dan Sastra Jawa,

Universitas Negeri Semarang

rosaliaindah0616@students.unnes.ac.id

DOI: <https://doi.org/10.15294/bsb.v1i1.1397>

QRCCN 62-6861-5651-259

### **ABSTRAK**

Penilaian sumatif merupakan proses evaluasi yang dilaksanakan pada akhir periode pembelajaran dengan tujuan mengukur tingkat pencapaian hasil belajar siswa. Dalam praktiknya, soal pilihan ganda menjadi salah satu instrumen penilaian yang paling umum digunakan karena memiliki keunggulan dari segi objektivitas, kemudahan pelaksanaan, dan efisiensi waktu. Meskipun demikian, pada mata pelajaran Bahasa Jawa, kualitas soal pilihan ganda masih menghadapi tantangan yang cukup signifikan. Permasalahan utama terletak pada tingkat kesukaran soal yang kurang proporsional, daya pembeda yang belum optimal, serta efektivitas pengecoh yang kurang mampu meminimalkan jawaban tebak. Kajian ini bertujuan mendeskripsikan kriteria soal pilihan ganda yang berkualitas pada mata pelajaran Bahasa Jawa melalui kombinasi analisis kualitatif dan kuantitatif. Pendekatan kualitatif digunakan untuk menilai aspek kesesuaian dengan tujuan pembelajaran, kejelasan redaksi, relevansi konteks budaya lokal, dan ketepatan struktur teknis. Sementara itu, analisis kuantitatif diterapkan untuk mengukur tingkat kesukaran, daya pembeda, dan efektivitas pengecoh berdasarkan hasil uji coba soal. Hasil penelitian menunjukkan bahwa soal pilihan ganda yang berkualitas harus memenuhi indikator utama meliputi: (1) keterkaitan langsung dengan kompetensi yang diharapkan; (2) penggunaan bahasa yang jelas, lugas, dan

sesuai kaidah Bahasa Jawa; (3) penyusunan konteks soal yang berakar pada budaya lokal; (4) penerapan kaidah teknis penyusunan butir soal; dan (5) distribusi proporsional terhadap keterampilan menyimak, membaca, dan menulis. Berdasarkan temuan tersebut, kajian ini merekomendasikan adanya pelatihan berkelanjutan bagi guru dalam penyusunan soal, pelaksanaan evaluasi sistematis terhadap instrumen penilaian, serta pengembangan bank soal Bahasa Jawa yang terstandar dan berbasis konteks. Upaya tersebut diharapkan mampu meningkatkan kualitas asesmen sumatif dan mendorong pembelajaran Bahasa Jawa yang lebih efektif.

**Kata Kunci:** soal pilihan ganda, penilaian sumatif, Bahasa Jawa, evaluasi pembelajaran, kualitas soal.

## PENDAHULUAN

Penilaian pembelajaran merupakan bagian integral dalam sistem pendidikan, berfungsi untuk mengevaluasi sejauh mana pencapaian kompetensi peserta didik sesuai tujuan pembelajaran (Brookhart, 2010). Selain sebagai alat ukur hasil belajar, penilaian juga berperan penting dalam menganalisis efektivitas strategi pengajaran serta merumuskan tindak lanjut dalam perbaikan proses belajar mengajar (Azizan, 2023). Berdasarkan fungsinya, penilaian terbagi menjadi tiga tipe: diagnostic-untuk mengidentifikasi kekuatan dan kelemahan siswa sebelum pembelajaran dimulai; formatif-untuk memberikan umpan balik selama pembelajaran; dan sumatif-untuk mengukur pencapaian kompetensi di akhir proses pembelajaran (Brookhart, 2010; Azizan, 2023).

Pada praktiknya, soal pilihan ganda merupakan instrumen penilaian sumatif yang paling banyak digunakan di berbagai jenjang pendidikan. Hal ini dikarenakan soal pilihan ganda bersifat objektif, praktis, serta efisien, dan cakupan materinya bisa luas dalam waktu singkat. Selain itu, hasil soal pilihan ganda mudah dianalisis secara kuantitatif-dari segi validitas, reliabilitas, tingkat kesukaran, hingga daya pembeda (Riinawati, 2021). Sebagai salah satu pendekatan penilaian

terkini, Asesmen Kompetensi Minimum (AKM) menitikberatkan pada literasi membaca dan numerasi. Astuti (2021) dan Riinawati (2021) menyatakan bahwa kemampuan kognitif siswa dapat tercermin melalui soal yang disusun berdasar AKM. Sementara itu, Wulandari dan Suryaman (2022) menekankan pentingnya memasukkan konteks literasi budaya lokal dalam penyusunan soal, agar lebih relevan dengan kehidupan siswa.

Akan tetapi, kualitas soal pilihan ganda masih menjadi permasalahan di berbagai mata pelajaran. Suryani dan Pratami (2022) menemukan bahwa soal Ujian Sekolah Bahasa Indonesia untuk kelas XII belum seimbang dalam distribusi level kognitif menurut taksonomi Bloom. Zamzami dan Fitriani (2020) juga mencatat bahwa banyak guru kesulitan menyusun soal berdasarkan keterampilan berpikir tingkat tinggi (HOTS) karena mereka belum terbiasa merumuskan indikator kognitif dan stimulus soal yang menantang secara berpikir.

Permasalahan serupa muncul pada mata pelajaran Bahasa Jawa. Sebagai muatan local-diajar di jenjang SD sampai SMA di wilayah Jawa Tengah, Yogyakarta, dan Jawa Timur-Bahasa Jawa memiliki karakteristik linguistik dan budaya yang unik. Penyusunan soal yang tidak mempertimbangkan aspek tersebut dapat menurunkan efektivitas penilaian. Mujimin (2010) menunjukkan perlunya peningkatan kompetensi guru dalam memahami teori evaluasi dan kaidah penyusunan soal Bahasa Jawa. Wahyuningrum (2014) juga menemukan bahwa soal ujian akhir semester Bahasa Jawa memiliki tingkat kesukaran dan daya pembeda yang rendah, sehingga pengukuran kemampuan siswa menjadi kurang akurat. Konteks pembelajaran abad ke-21 menuntut instrumen yang tidak hanya mengukur menghafal, tetapi juga berpikir kritis, analitis, dan kreatif.

Di sisi lain, implementasi HOTS dalam soal Bahasa Jawa mulai mendapatkan perhatian. Fitriyani Astuti (2021) menegaskan bahwa soal HOTS—baik pilihan ganda maupun uraian—dapat melatih siswa mengaitkan, menginterpretasikan, menerapkan, dan mengintegrasikan pengetahuan dalam konteks lokal serta kontekstual (termasuk tipe teks, kemampuan berpikir, dan stimulus) Penelitian lain menggunakan parameter PISA untuk mengevaluasi soal baca

Bahasa Jawa di SMP DIY; hasilnya menunjukkan hanya sekitar 52 % ciri-ciri soal PISA yang muncul dalam soal uji Bahasa Jawa—menunjukkan bahwa kemampuan berpikir tingkat tinggi masih belum optimal diwakili dalam soal Bahasa Jawa saat ini.

Selain itu, pengembangan instrumen yang responsif budaya penting untuk Bahasa Jawa. Pudjastawa et al. (2021) berhasil mengembangkan soal ujian berbasis web dengan pendekatan *culturally responsive assessment* menggunakan bahasa Jawa dialek Timur—dan terbukti meningkatkan skor siswa secara signifikan serta mendapatkan respons positif dari guru dan siswa. Di ranah soal pilihan ganda berbasis HOTS, Purwati et al. (2021) menunjukkan bahwa mayoritas soal memiliki kesukaran mudah, daya pembeda baik, pengecoh efektif, valid, dan reliabel, meskipun penelitian ini dilakukan di SD untuk soal umum (bukan Bahasa Jawa). Ini menunjukkan bahwa analisis butir soal berbasis aplikasi seperti Anates dapat memberikan gambaran yang mendalam terhadap kualitas soal.

Dengan demikian, dapat diketahui bahwa soal Bahasa Jawa selama ini belum sepenuhnya mencakup aspek struktur teknis, konteks budaya, serta distribusi ranah kognitif tingkat tinggi secara proporsional. Hal ini membatasi kemampuan penilaian sumatif untuk mengukur keseluruhan kompetensi siswa, terutama yang mencakup kemampuan berpikir kritis dan kreatif dalam konteks budaya lokal dan Kurikulum Merdeka.

Penelitian ini bertujuan untuk mendeskripsikan kriteria soal pilihan ganda yang berkualitas pada mata pelajaran Bahasa Jawa, dengan mempertimbangkan: (1) kesesuaian dengan tujuan pembelajaran dan kurikulum, (2) kejelasan bahasa/redaksi, (3) relevansi konteks budaya lokal, (4) struktur teknis butir soal (validitas, reliabilitas, tingkat kesukaran, daya pembeda, dan efektivitas pengecoh), serta (5) representasi keterampilan berpikir tingkat tinggi (HOTS). Harapannya, melalui kajian ini, instrumen penilaian Bahasa Jawa bisa lebih bermakna, representatif, dan relevan, serta memberikan acuan yang jelas bagi guru dan pengembang instrumen soal.

## **METODE**

Penilaian pembelajaran merupakan komponen esensial dalam sistem pendidikan yang berfungsi mengevaluasi pencapaian kompetensi peserta didik sesuai tujuan pembelajaran (Lombu et al., 2024). Penilaian tidak sekadar mengukur hasil belajar, tetapi juga menjadi sarana refleksi untuk memperbaiki proses pembelajaran. Berdasarkan fungsinya, penilaian dibedakan menjadi diagnostik, formatif, dan sumatif. Penilaian diagnostik digunakan untuk memetakan kekuatan dan kelemahan siswa sebelum pembelajaran, formatif dilaksanakan selama pembelajaran untuk memberikan umpan balik, sedangkan sumatif dilakukan di akhir pembelajaran untuk mengukur pencapaian akhir. Prinsip utama penilaian meliputi validitas, reliabilitas, objektivitas, kepraktisan, dan keterpaduan. Dalam praktiknya, soal pilihan ganda menjadi instrumen sumatif yang paling sering digunakan karena sifatnya yang objektif, efisien, dan mudah dianalisis, baik secara kualitatif maupun kuantitatif (Riinawati & Pd, 2021). Pendekatan mutakhir seperti Asesmen Kompetensi Minimum (AKM) menekankan literasi membaca dan numerasi, yang dapat diintegrasikan dalam penyusunan soal untuk mengukur keterampilan berpikir tingkat tinggi (Firdaini Aulia, 2024).

Taksonomi Bloom yang direvisi oleh Anderson & Krathwohl (2001) mengklasifikasikan keterampilan kognitif menjadi enam tingkatan: mengingat, memahami, menerapkan, menganalisis, mengevaluasi, dan mencipta. Proporsi ideal penyusunan soal mencakup LOTS 30%, MOTS 40%, dan HOTS 30% (Badan Litbang dan Perbukuan, 2021). Penerapan taksonomi ini pada soal Bahasa Jawa dapat memastikan evaluasi tidak hanya berorientasi pada hafalan, tetapi juga pengembangan berpikir kritis dan kreatif. Mata pelajaran Bahasa Jawa memiliki peran strategis dalam pelestarian bahasa, sastra, dan budaya daerah (Permendikbud No. 79 Tahun 2014). Tujuan pembelajarannya mencakup penguasaan kaidah bahasa, apresiasi sastra, serta pemahaman nilai budaya. Evaluasi melalui soal pilihan ganda harus dirancang kontekstual, mengintegrasikan aspek bahasa, sastra, dan budaya, serta mengukur keterampilan menyimak, berbicara, membaca, dan menulis. Hal ini memastikan penilaian tidak

hanya bersifat akademis, tetapi juga membentuk karakter dan identitas budaya peserta didik.

## **PEMBAHASAN**

Dalam hasil kajian dokumen soal Bahasa Jawa SMP, ditemukan bahwa sebagian besar soal masih berfokus pada aspek ingatan (recall) dan belum mendorong siswa untuk berpikir kritis. Hal ini menunjukkan bahwa karakteristik soal berkualitas belum sepenuhnya terwujud. Menurut (Nafs et.al 2023) guru seharusnya melakukan analisis butir soal yang dilakukan secara berkala guna mengetahui kualitas soal. Analisis yang bisa dilakukan meliputi analisis kualitatif dan analisis kuantitatif.

### **Analisis Kualitatif**

Analisis kualitatif terhadap soal pilihan ganda dalam mata pelajaran *Bahasa Jawa* didasarkan pada kaidah penulisan soal. Dalam konteks kaidah penulisan soal harus memperhatikan beberapa aspek penyusunan. Aspek-Aspek penting dalam penyusunan Soal Pilihan Ganda Bahasa Jawa yang perlu diperhatikan sebagai berikut:

#### **Aspek Isi**

Soal yang baik harus mengacu pada materi ajar yang bersumber dari Silabus dan Kompetensi Dasar yang tercantum dalam Kurikulum yang sedang berlaku. Kesesuaian ini penting agar benar-benar mengukur kompetensi yang menjadi tujuan pembelajaran (Nitko & Brookhart, 2013). Untuk mempermudah dalam memahami analisis aspek isi soal perlu memperhatikan beberapa poin berikut :

- a. Soal sudah sesuai dengan indikator. (Kesesuaian soal dengan indikator yang terdapat pada kurikulum)
- b. Pertanyaan serta jawaban sudah sesuai harapan. (Pertanyaan yang diajukan dapat dijawab dengan memuaskan)
- c. Materi menyesuaikan kompetensi. (Materi yang disajikan dirancang sedemikian rupa menyesuaikan tingkat pengetahuan, keterampilan dan sikap)
- d. Isi materi disesuaikan dengan jenjang sekolah atau tingkatan. (Materi disesuaikan dengan pemahaman dan kebutuhan siswa)

### Aspek Konstruksi Soal

Konstruksi soal merupakan unsur teknis yang sangat menentukan kualitas dan keakuratan hasil evaluasi pembelajaran. Sebuah soal yang dikonstruksi dengan baik akan mampu membedakan siswa berdasarkan tingkat penguasaan materi dan mengurangi kemungkinan jawaban yang benar akibat tebakan semata (Brookhart, 2010). Menurut Arikunto (2012), struktur dasar soal pilihan ganda terdiri dari dua komponen utama, yaitu:

- a. Pokok Soal (*Stem*): Merupakan pernyataan atau pertanyaan utama yang mengandung inti permasalahan.
- b. Pilihan Jawaban (*Options*): Biasanya terdiri atas satu jawaban benar (*kunci*) dan beberapa pengecoh (*distraktor*), yang dirancang untuk menyerupai jawaban benar agar mampu mengeliminasi siswa yang tidak memahami materi secara mendalam.

Agar instrumen ini berkualitas tinggi, perlu diterapkan prinsip-prinsip konstruksi soal yang baik sebagaimana dirumuskan oleh Widoyoko (2009) serta Retno Palupi Miftahul Aniyah (2020), yaitu sebagai berikut:

- a. *Stem* dirumuskan dengan singkat, jelas, dan tidak multitafsir.
- b. Hindari penggunaan pernyataan negatif ganda yang membingungkan siswa.
- c. Semua opsi jawaban harus homogen dalam bentuk dan logis dalam isi.
- d. Panjang pilihan jawaban sebaiknya relatif sama untuk menghindari bias visual.
- e. Posisi jawaban benar perlu divariasikan secara sistematis, tidak selalu berada di pilihan yang sama.
- f. Distraktor (pengecoh) harus berfungsi, artinya dapat menarik respon dari siswa yang belum menguasai materi.

### Aspek Budaya dan Bahasa

Dengan memperhatikan aspek budaya dan bahasa dalam penyusunan soal, guru tidak hanya mentransmisikan pengetahuan linguistik, tetapi juga menanamkan nilai-nilai karakter, identitas etnis, dan kearifan lokal yang menjadi inti pendidikan Bahasa Jawa dalam Kurikulum Nasional. Dalam

konteks aspek budaya dan bahasa dalam butir soal Bahasa Jawa sebagai berikut :

- a. Menggunakan bahasa sesuai dengan *unggah-ungguh basa*. Penerapan *unggah-ungguh basa*, yaitu tata krama dalam berbahasa yang mencerminkan relasi sosial antara penutur dan lawan bicara. Dalam tradisi Jawa, penggunaan tingkat tutur seperti ngoko, madya, dan krama memiliki nilai sosial yang sangat tinggi. Soal yang baik seharusnya tidak hanya menanyakan definisi atau fungsi tingkat tutur, tetapi juga menguji pemahaman siswa terhadap konteks penggunaannya dalam kehidupan nyata (Astuti, 2021).
- b. Menggunakan bahasa yang komunikatif. Bahasa yang digunakan harus jelas, sederhana dan dapat diterima.
- c. Menggunakan bahasa yang berlaku di daerah asal (dialek). Aspek dialek dan variasi regional juga layak diperhitungkan. Meskipun Bahasa Jawa yang diajarkan di sekolah-sekolah didominasi oleh dialek krama inggil standar dari Yogyakarta-Surakarta, realitas sosial menunjukkan keberagaman dialek daerah seperti Jawa Banyumasan, Mataraman, hingga Arekan. Soal pilihan ganda yang kontekstual dapat menampilkan ragam dialek atau kosakata khas daerah sebagai bentuk pelestarian bahasa sekaligus penghargaan terhadap identitas lokal (Mujimin, 2010; Irvine, 2019).
- d. Pilihan jawaban tidak mengulang kata/kelompok kata yang sama, kecuali satu kesatuan pengertian. Untuk menghindari jawaban yang ambigu, setiap pilihan jawaban tidak memiliki frasa atau kata yang sama persis dengan pilihan jawaban lain.

Pembelajaran Bahasa Jawa memiliki dimensi yang unik dibandingkan mata pelajaran bahasa lainnya karena mengandung unsur identitas budaya yang sangat kental. Oleh karena itu, penyusunan soal pilihan ganda tidak cukup hanya memperhatikan struktur bahasa secara linguistik, tetapi juga harus mencerminkan nilai-nilai budaya lokal, norma sosial, dan filosofi masyarakat Jawa. Sebagai contoh, soal dapat menampilkan teks percakapan dari masyarakat pesisir atau pegunungan yang menggunakan istilah khas daerah, kemudian meminta siswa untuk menginterpretasikan maknanya dalam

Bahasa Jawa baku. Pendekatan ini sejalan dengan pembelajaran berbasis *multikulturalisme* dan *sociocultural linguistic competence*.

### Analisis Kuantitatif

Analisis kuantitatif terhadap butir soal merupakan langkah penting dalam mengevaluasi kualitas instrumen penilaian secara objektif dan terukur. Pendekatan ini berdasarkan teori tes klasik mencakup aspek tingkat kesukaran, daya pembeda, efektivitas pengecoh, dan reabilitas

### Tingkat Kesukaran

Tingkat kesukaran adalah ukuran yang menunjukkan proporsi peserta tes yang mampu menjawab suatu butir soal dengan benar. Soal dengan tingkat kesukaran tinggi berarti sebagian besar peserta dapat menjawab dengan benar, sedangkan soal dengan tingkat kesukaran rendah menunjukkan bahwa hanya sedikit peserta yang menjawab benar (Arikunto, 2012). Indeks ini berfungsi untuk mengidentifikasi apakah suatu soal termasuk kategori sangat mudah, mudah, sedang, sukar, atau sangat sukar. Dengan adanya variasi tingkat kesukaran, guru atau peneliti dapat menyusun perangkat tes yang seimbang dan mencerminkan kemampuan peserta secara lebih adil (Widoyoko, 2018). Untuk mempermudah dalam menentukan proporsi ideal soal pilihan ganda (Badan Litbang dan Perbukuan, 2021) merumuskan:

LOTS- C1 & C2: 30% MOTS- C3 & C4: 40% HOTS- C5 & C6: 30% Dengan taksonomi ini, pembelajaran tidak hanya berfokus pada hafalan, tetapi juga pemahaman mendalam, keterampilan berpikir kritis, dan kreativitas siswa.

Tabel 3 Kriteria kesukaran butir soal

Rumus	Nilai koefisien korelasi
$P = \frac{\sum b}{N}$	0,0 - 0,20 = kategori sangat sukar, dibuang
	0,2 - 0,30 = kategori sukar, kurang baik, di revisi
	0,31 - 0,70 = kategori sedang, dan cukup baik
	0,71 - 0,90 = kategori mudah, kurang baik, direvisi
	0,91 - 1,00 = kategori sangat mudah, butir tes dibuang

Rumus di atas digunakan untuk menghitung tingkat kesukaran suatu soal, yaitu seberapa banyak peserta tes yang mampu menjawabnya dengan benar.

### **Daya Pembeda (*Discrimination Index*)**

Daya pembeda merupakan indikator yang menunjukkan kemampuan suatu butir soal dalam membedakan antara peserta tes yang memiliki tingkat kemampuan tinggi dan rendah. Soal yang memiliki daya pembeda tinggi akan dijawab benar oleh sebagian besar siswa berkemampuan tinggi, dan dijawab salah oleh sebagian besar siswa berkemampuan rendah. Oleh karena itu, daya pembeda berfungsi untuk menilai apakah suatu soal mampu memilah peserta berdasarkan tingkat penguasaan materi, sehingga dapat meningkatkan validitas instrumen evaluasi yang digunakan (Qomariyah, 2022).

Tabel 4 Menunjukkan rumus dan kriteria daya pembeda butir soal

<b>Rumus</b>	<b>Kriteria besarnya koefisien daya beda</b>
$DP = \frac{B_a - B_b}{\frac{1}{2}N}$	0,51 – 1,00 = kategori sangat baik, dapat digunakan
	0,41 - 0,50 = kategori baik, dapat digunakan
	0,31 – 0,40 = kategori cukup baik, direvisi
	0,21 – 0,30 = kategori kurang baik, perlu pembahasan dan revisi
	0,00 - 0,20 = kategori tidak baik, dibuang

(Widoyoko, 2018)

### **Efektivitas Pengecoh**

Efektivitas pengecoh merupakan indikator penting dalam mengevaluasi kualitas butir soal pilihan ganda, terutama dalam menilai sejauh mana opsi jawaban yang salah (*distraktor*) mampu menarik siswa yang belum memahami materi dengan baik. Secara definisional, efektivitas pengecoh merujuk pada sejauh mana pilihan jawaban salah mampu mengalihkan perhatian siswa yang tidak mengetahui jawaban benar. Dalam konteks evaluasi pendidikan, pengecoh

dikatakan efektif apabila dipilih oleh siswa dari kelompok bawah—yakni mereka yang penguasaannya terhadap materi masih rendah sementara siswa dari kelompok atas cenderung memilih jawaban yang benar. Dengan demikian, pengecoh berperan penting dalam mempertajam perbedaan performa antara siswa berkemampuan tinggi dan rendah (Nitko & Brookhart, 2013; Popham, 2008). Dari sisi fungsional, pengecoh berfungsi sebagai alat diagnosis terhadap miskonsepsi maupun kesalahan konseptual siswa. Efektivitas pengecoh dirumuskan :

$ED = (\text{Jumlah siswa yang memilih pengecoh} / \text{Total siswa}) \times 100\%$ , ED (Efektivitas Distraktor) merepresentasikan proporsi peserta tes yang memilih tiap opsi pengecoh. Berdasarkan kriteria yang dikembangkan oleh Arbiatin dan Mulabbiyah (2020), pengecoh dikategorikan sebagai berfungsi efektif jika dipilih oleh  $\geq 5\%$  peserta; perlu direvisi jika  $< 5\%$ ; dan ditolak jika tidak dipilih sama sekali (0%). Sebagai ilustrasi, misalkan terdapat 80 peserta ujian. Jika pengecoh A dipilih oleh 6 siswa (7,5%), pengecoh B oleh 3 siswa (3,75%), dan pengecoh C oleh 0 siswa (0%), maka dapat disimpulkan bahwa pengecoh A termasuk efektif, pengecoh B perlu direvisi, dan pengecoh C harus ditolak atau diganti karena tidak berfungsi.

Dalam konteks soal pilihan ganda Bahasa Jawa, pengecoh harus pula mempertimbangkan karakteristik bahasa dan budaya lokal. Beberapa kesalahan umum siswa yang dapat dimanfaatkan sebagai basis penyusunan pengecoh antara lain: (1) kesalahan dalam penggunaan unggah-ungguh basa seperti tembung krama atau madya yang tidak tepat, (2) kekeliruan dalam interpretasi teks sastra Jawa, serta (3) kesalahan dalam membaca maupun menulis aksara Jawa. Oleh karena itu, pengecoh tidak hanya menjadi distraktor teknis, melainkan juga dapat berfungsi sebagai indikator cerminan kompetensi kebahasaan dan kognitif siswa terhadap substansi pembelajaran Bahasa Jawa secara utuh dan kontekstual.

## **PENUTUP**

Hasil analisis kualitatif dan kuantitatif terhadap soal pilihan ganda pada mata pelajaran Bahasa Jawa menunjukkan beberapa temuan utama. Pertama, keterkaitan antara soal

dengan tujuan pembelajaran yang tertuang dalam kurikulum, kompetensi dasar, dan indikator capaian menjadi prasyarat utama validitas isi. Kesesuaian ini memastikan materi yang diujikan relevan dengan kompetensi yang ingin dicapai. Kedua, konstruksi soal harus jelas, sistematis, dan bebas ambiguitas, dengan pokok soal yang terarah, opsi jawaban homogen serta logis, dan pengecoh yang efektif untuk membedakan kemampuan siswa. Ketiga, butir soal perlu merepresentasikan keterampilan menyimak, membaca, dan menulis secara proporsional, sekaligus mengintegrasikan unsur kebahasaan khas Bahasa Jawa seperti unggah-ungguh basa, paramasastra, dan kosakata lokal. Integrasi ini berfungsi memperkaya muatan soal sekaligus memperkuat nilai-nilai budaya dalam evaluasi pembelajaran. Keempat, kualitas instrumen harus diuji secara objektif melalui analisis tingkat kesukaran, daya pembeda, dan efektivitas pengecoh, sehingga kelayakan soal dapat terukur secara empiris.

Berdasarkan temuan tersebut, rekomendasi strategis yang diajukan meliputi: (1) peningkatan kompetensi guru dalam penyusunan instrumen evaluasi berbasis taksonomi Bloom revisi; (2) pelaksanaan evaluasi berkala terhadap konstruksi soal, distribusi level kognitif, dan kinerja pengecoh; (3) pengayaan materi dan konteks soal agar lebih kontekstual dengan kehidupan nyata dan budaya lokal; serta (4) pemanfaatan perangkat lunak analisis butir soal seperti Anates, ITEMAN, atau pendekatan Rasch untuk mendukung akurasi analisis. Penerapan rekomendasi ini diharapkan dapat menghasilkan soal pilihan ganda Bahasa Jawa yang tidak hanya memenuhi kriteria teknis penilaian, tetapi juga mendukung pembelajaran yang bermakna, berpijak pada budaya lokal, dan sejalan dengan tuntutan kompetensi abad ke-21.

#### **DAFTAR PUSTAKA**

- Anderson, L. W., & Krathwohl, D. R. (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. Longman.
- Arbiatin, I., & Mulabbiyah, M. (2020). *Analisis kualitas*

- distraktor pada butir soal pilihan ganda*. Jurnal Evaluasi Pendidikan, 11(1), 34–42.
- Arifin, Z. (2022). *Evaluasi pembelajaran: Prinsip, teknik, dan prosedur*. Remaja Rosdakarya.
- Arikunto, S. (2012). *Dasar-dasar evaluasi pendidikan* (Edisi Revisi). Bumi Aksara.
- Astuti, S. (2021). *Pengembangan instrumen penilaian Bahasa Jawa berbasis kearifan lokal*. Jurnal Pendidikan Bahasa dan Sastra Daerah, 3(2), 45–53.
- Azizan, M. (2023). *Strategi evaluasi pembelajaran berbasis kompetensi*. Jurnal Inovasi Pendidikan, 10(1), 15–24.
- Badan Litbang dan Perbukuan. (2021). *Panduan penyusunan soal AKM*. Kementerian Pendidikan dan Kebudayaan.
- Brookhart, S. M. (2010). *How to assess higher-order thinking skills in your classroom*. ASCD.
- Chasanah, U., Widodo, A., & Suprpto, E. (2022). *Evaluasi pembelajaran berbasis digital dan klasik*. UPT Penerbitan dan Percetakan UNISMA.
- Dwi Lactona, & Cahyono, A. (2024). *Relevansi taksonomi Bloom dalam perumusan soal HOTS*. Jurnal Pendidikan Inovatif, 8(1), 21–35.
- Firdaini, A. (2024). *Implementasi Asesmen Kompetensi Minimum dalam pembelajaran Bahasa Jawa*. Prosiding Seminar Nasional Evaluasi Pendidikan, 2(1), 66–75.
- Gronlund, N. E., & Linn, R. L. (2009). *Measurement and assessment in teaching* (10th ed.). Pearson Education.
- Gunawan, D., & Prabowo, A. (2023). *Pengembangan soal berbasis budaya lokal dengan pendekatan digital di mata pelajaran Bahasa Jawa*. Jurnal Evaluasi Pendidikan, 8(2), 115–130.
- Irvine, J. J. (2019). *Culturally responsive pedagogy and practice*. Harvard Educational Review, 89(1), 1–25.
- Kurniawan, T. (2023). *Integrasi HOTS dalam soal berbasis kearifan lokal: Studi pada Kurikulum Merdeka*. Jurnal Pendidikan Daerah, 7(1), 66–75.
- Lombu, A., Nurjanah, N., & Siregar, R. (2024). *Jenis dan fungsi penilaian dalam pembelajaran*. Jurnal Evaluasi Pendidikan, 9(2), 105–115.
- Mujimin. (2010). *Peningkatan kompetensi guru dalam menyusun soal Bahasa Jawa*. Jurnal Pendidikan Daerah,

- 12(1), 17–26.
- Nafs, A., Hartati, S., & Prasetyo, H. (2023). *Pentingnya analisis butir soal dalam evaluasi pembelajaran*. *Jurnal Pendidikan dan Evaluasi*, 5(2), 98–108.
- Nitko, A. J., & Brookhart, S. M. (2013). *Educational assessment of students* (6th ed.). Pearson.
- Otaya, G. N. (2014). *Teori tes klasik dan aplikasi dalam evaluasi pembelajaran*. UNY Press.
- Popham, W. J. (2008). *Classroom assessment: What teachers need to know* (5th ed.). Pearson Merrill Prentice Hall.
- Qomariyah, A. (2022). *Analisis daya pembeda dan validitas soal pilihan ganda Bahasa Indonesia*. *Jurnal Evaluasi dan Penilaian Pendidikan*, 4(1), 33–42.
- Retno Palupi Miftahul Aniyah. (2020). *Konstruksi soal pilihan ganda yang baik*. *Jurnal Pendidikan*, 11(1), 56–68.
- Riinawati, E. (2021). *Penilaian pembelajaran dalam era Kurikulum Merdeka*. *Jurnal Inovasi Pendidikan*, 6(2), 87–95.
- Riinawati, E., & Pd, M. (2021). *Evaluasi pembelajaran berbasis AKM*. *Jurnal Evaluasi Pendidikan*, 3(2), 112–120.
- Sudijono, A. (2018). *Pengantar evaluasi pendidikan*. RajaGrafindo Persada.
- Suryani, N., & Pratami, F. (2022). *Analisis kualitas soal Bahasa Indonesia kelas XII berdasarkan taksonomi Bloom*. *Jurnal Evaluasi Pendidikan*, 7(1), 50–60.
- Sutrisno. (2016). *Pembelajaran Bahasa Jawa di era global*. Pustaka Adi.
- Taufik, M., Hidayah, N., & Lestari, R. (2022). *Evaluasi pendidikan: Teori dan implementasi*. Deepublish.
- Wahyuningrum, R. (2014). *Analisis tingkat kesukaran dan daya pembeda soal Bahasa Jawa semester ganjil SMP Negeri 1 Karanganyar*. *Jurnal Evaluasi Pendidikan*, 4(2), 101–112.
- Widoyoko, E. P. (2018). *Teknik penyusunan instrumen penelitian*. Pustaka Pelajar.
- Wulandari, F., & Suryaman, A. (2022). *Asesmen kompetensi minimum dan relevansi budaya lokal dalam penyusunan soal literasi*. *Jurnal Evaluasi Kurikulum*, 6(3), 98–110.
- Zamzami, Z., & Fitriani, H. (2020). *Kendala guru dalam menyusun soal HOTS dan alternatif solusinya*. *Jurnal Pendidikan Inovatif*, 4(2), 55–62.