

KUALITAS SOAL PENILAIAN TENGAH SEMESTER MATA PELAJARAN BAHASA INDONESIA JENJANG SEKOLAH MENENGAH PERTAMA

Ajeng Setia Budi¹, Ayu Wulandari², Mashud Syahroni³

^{1,2,3} Pendidikan Bahasa dan Sastra Indonesia, Universitas Tidar

E-mail: syahronifkip@untidar.ac.id

Abstrak

Tujuan penelitian ini untuk mengetahui kualitas soal Penilaian Tengah Semester Bahasa Indonesia Jenjang Sekolah Menengah Pertama ditinjau dari aspek validitas, reliabilitas, indeks tingkat kesukaran, indeks daya beda, dan efektivitas butir pengecoh. Soal dapat menjadi alat ukur penilaian sebagai bahan evaluasi pembelajaran yang menjadi tugas pokok guru. Subjek penelitian adalah Soal Penilaian Tengah Semester Bahasa Indonesia pada Kelas IX SMP Negeri 4 Magelang Tahun Ajaran 2022/2023. Pengumpulan data menggunakan teknik studi dokumen. Sampel penelitian 63 lembar jawab siswa yang dipilih sesuai purposive sampling. Soal merupakan soal pilihan ganda berjumlah 30 butir. Teknik analisis data menggunakan pendekatan kuantitatif dengan statistik software SPSS Versi 25 dan Microsoft Excel. Hasil penelitian menunjukkan bahwa; 1) Pada aspek validitas terdapat 57% (17 butir) soal valid, sisanya tidak valid; 2) Aspek reliabilitas koefisien alpha cronbach sebesar 0,551 berarti tingkat reliabel soal sedang; 3) Indeks Tingkat Kesukaran, terdapat 7% (2 soal) sulit, 40% (12 soal) sedang, dan 53% (16 soal) mudah; 4) Indeks Daya Beda menghasilkan 17% (5 soal) tidak baik, 40% (12 soal) jelek, 40% (12 soal) cukup, 3% (1 soal) baik; 5) Efektivitas pengecoh 13% (4 soal) sangat baik, 3% (1 soal) baik, 57% (17 soal) kurang baik, dan 27% (8 soal) jelek. Ditinjau dari lima aspek, soal pilihan ganda pada Penilaian Tengah Semester Mata Pelajaran Bahasa Indonesia Kelas IX SMP Tahun Ajaran 2022/2023 memiliki kualitas yang kurang baik dan belum mampu mengukur hasil belajar siswa.

Kata kunci : evaluasi pembelajaran, analisis butir soal, hasil belajar

Abstract

The aim of this research was to determine the quality of the Mid-Semester Indonesian Language Assessment questions at Junior High School Level in terms of validity, reliability, difficulty level index, discrimination index and effectiveness of distractor items. Questions can be used as a measuring tool for evaluating learning, which is the teacher's main task. The research subject is the Indonesian Language Mid-Semester Assessment Questions in Class IX of SMP Negeri 4 Magelang for the 2022/2023 Academic Year. Data collection uses document study techniques. The research sample was 63 student answer sheets selected according to purposive sampling. The questions are multiple choice questions totaling 30 items. The data analysis technique uses a quantitative approach with statistical software SPSS Version 25 and Microsoft Excel. The research results show that; 1) In the validity aspect, 57% (17 items) of the questions are valid, the rest are invalid. 2) The reliability aspect of the Cronbach's alpha coefficient is 0.551, meaning the level of reliability of the questions is medium. 3) The Difficulty Level Index for the questions is 7% (2 questions) difficult, 40% (12 questions) moderate, and 53% (16 questions) easy. 4) The Different Power Index produces 17% (5 questions) not good, 40% (12 questions) bad, 40% (12 questions) fair, 3% (1 question) good, and. 5) The effectiveness of the distractor was 13% (4 questions) very good, 3% (1 question) good, 57% (17 questions) poor, and 27% (8 questions) poor. The results of the five aspects show that the Indonesian Language Mid-Semester Assessment questions for Class IX SMP at the 2022/2023 Academic Year are of poor quality and are not yet able to measure students' learning output.

Keywords: analysis of item, learning evaluation, learning output

PENDAHULUAN

Dalam proses pembelajaran guru memiliki tugas merencanakan, melaksanakan, dan menilai hasil belajar sampai tahap evaluasi pembelajaran. Evaluasi pembelajaran digunakan untuk mengukur kemampuan dan penguasaan hasil belajar peserta didik. Evaluasi dilakukan sebagai tahap akhir dalam proses pembelajaran. Pelaksanaan evaluasi dilaksanakan sesuai dengan standar nasional pendidikan Indonesia. Salah satu bentuk evaluasi adalah tes. Guru selain sebagai seorang pengajar juga berperan sebagai evaluator. Pada proses evaluasi pendidikan sangat dibutuhkan kemampuan untuk menganalisis soal, sehingga soal yang digunakan dapat mencerminkan kemampuan siswa. (Nasir, 2015)

Tes merupakan cara pengukuran yang berisi pertanyaan, pernyataan, atau serangkaian tugas lain yang harus diselesaikan atau dijawab peserta didik. Melalui tes, baik guru maupun siswa dapat hasil berupa umpan balik yang berguna untuk meningkatkan minat dan motivasi memperoleh hasil belajar lebih baik. Tes berfungsi sebagai informasi yang akurat mengenai hasil belajar peserta didik selama proses pembelajaran berlangsung. Selain itu tes juga dapat membantu guru dalam melakukan tindak lanjut terhadap peserta didik, terutama bagi peserta didik yang memperoleh hasil belajar kurang baik. (Rahmasari & Ismiyati, 2016). Salah satu tes yang sering digunakan dalam mengukur hasil pembelajaran adalah Penilaian Tengah Semester (PTS). Tes yang sering digunakan dalam PTS berbentuk pilihan ganda dan uraian.

Kualitas soal akan menentukan kualitas tes yang dibuat. Sehingga soal tes harus disusun sesuai dengan prinsip dan prosedur penyusunannya. (Arifin, 2019). Salah satu cara mengetahui kualitas soal tes dapat dilakukan dengan menganalisis butir soal yang disusun. Kompetensi seorang guru tidak hanya menyusun alat evaluasi guna mengetahui pencapaian hasil belajar peserta didik akan tetapi juga dapat mengevaluasi apakah evaluasi yang telah disusun sudah dapat menjalankan fungsinya sebagai alat pengukur hasil belajar. (Fitrianawati, 2017).

Analisis butir soal atau analisis item sangat penting, guna memperoleh kualitas soal yang baik, sehingga dapat memperoleh gambaran

tentang prestasi siswa yang sebenarnya. Ada beberapa cara untuk melakukan analisis butir soal, yakni analisis validitas, analisis reliabilitas, analisis tingkat kesukaran, dan analisis daya pembeda. (Nurjanah & Marlianingsih, 2015). Soal dikatakan dapat menjalankan fungsinya sebagai alat pengukur hasil belajar apabila mampu memenuhi seluruh karakteristik penilaian butir soal tersebut.

Analisis butir soal dilakukan dengan tujuan memperoleh soal yang bermutu sebelum disebar ke peserta didik. Selain itu analisis butir soal juga dapat meningkatkan mutu tes melalui revisi atau menghilangkan yang tidak efektif digunakan. Soal yang telah lolos analisis butir dapat digunakan untuk mendapatkan informasi diagnostik peserta didik terkait pemahaman terhadap materi. (Aiken, 1994). Selain itu analisis butir soal bermanfaat antara lain, (1) menentukan soal-soal yang cacat atau tidak berfungsi dengan baik; (2) meningkatkan butir soal melalui tiga komponen analisis yaitu tingkat kesukaran, daya pembeda, dan pengecoh soal; (3) meningkatkan validitas soal dan reliabilitas; (4) merevisi soal yang tidak relevan dengan materi yang diajarkan. (Sari, Utomo, & Sumarwati, 2022)

Akan tetapi, pada kenyataan di lapangan tidak semua guru mampu melaksanakan analisis terhadap soal penilaian yang dibuat. Kesibukan akademik dan administrasi menjadikan guru tidak memiliki kesempatan untuk menganalisis butir soal yang dibuat. Salah satu kasus yang terjadi di SMP Negeri 4 Magelang. Fakta menunjukkan hasil PTS dari peserta didik kelas IX berjumlah 170 hanya 51 anak yang mencapai nilai ketuntasan minimal atau passing grade. Sisanya 119 anak tidak mencapai batasa ketuntasan dengan berbagai alasan yang dikemukakan. Salah satu indikasi penyebab banyaknya siswa yang tidak lolos batas ketuntasan minimal karena kualitas soal tidak baik.

Berbagai penelitian terdahulu menunjukkan pentingnya analisis butir soal dalam pembuatan soal ujian kepada siswa. Analisis kualitas butir soal ulangan semester ganjil kelas VIII SMP dari aspek kebahasaan menunjukkan dari dua puluh soal ditemukan 5 butir soal yang dianggap baik dan 15 butir soal yang dianggap tidak baik. (Nurjanah &

Marlianingsih, 2015). Analisis kualitas soal yang lain dilakukan untuk soal UAS bentuk pilihan ganda dan uraian secara kuantitatif menghasilkan reliabilitas soal pilihan ganda sebesar 0,68 artinya soal tidak reliabel. (Rahmasari & Ismiyati, 2016). Analisis butir soal sangat diperlukan untuk memastikan soal yang dibuat benar benar layak untuk dilaksanakan. Sebagaimana soal UAS gasal mata pelajaran IPS kelas III SD yang diujikan tanpa melalui tahapan analisis butir soal. Hasilnya menunjukkan bahwa soal ditinjau dari aspek materi, konstruksi, dan bahasa berkategori sangat tinggi tetapi dari aspek reliabilitasnya memiliki kriteria rendah. (Kurniawan, 2015)

Oleh karena itu tujuan penelitian adalah untuk menganalisis butir soal penilaian tengah semester Bahasa Indonesia kelas IX SMP Negeri 4 Magelang tahun ajaran 2022/2023. Analisis butir soal diperlukan untuk mengetahui kualitas butir soal yang disusun oleh guru. Terlebih soal ditujukan untuk jenjang SMP kelas IX yang membutuhkan pengukuran hasil belajar yang tepat dan terukur. Analisis butir soal ditinjau dari aspek validitas, reliabilitas, indeks tingkat kesukaran, indeks daya beda, dan efektivitas butir pengecoh.

METODE PENELITIAN

Penelitian ini menggunakan pendekatan penelitian deskriptif kuantitatif. Subjek penelitian adalah naskah soal penilaian tengah semester pada mata pelajaran Bahasa Indonesia kelas IX di SMP Negeri 4 Magelang tahun ajaran 2022/2023. Sampel dalam penelitian ini jawaban peserta didik berjumlah 63 yang dipilih untuk mewakili populasi sejumlah 170 orang. Sampel ditentukan dengan teknik simple random sampling. Teknik pengumpulan data menggunakan teknik studi dokumentasi. Data berupa pilihan jawaban pada setiap nomor soal serta kunci jawaban yang telah dibuat guru. Analisis data menggunakan teknik analisis data kuantitatif. Jawabn yang telah dikumpulkan kemudian dimasukkan dalam tabel yang kemudian diuji berdasarkan aspek pengujian. Uji validitas, reliabilitas, tingkat kesukaran, dan indeks daya beda dihitung dengan bantuan program SPSS Ver. 25, sedangkan efektivitas butir pengecoh dihitung dengan Microsoft Office Exel.

HASIL DAN PEMBAHASAN

Hasil analisis butir soal terhadap soal penilaian tengah semester Bahasa Indonesia kelas IX SMP Negeri 4 Magelang tahun ajaran 2022/2023 pada bagian pilihan ganda berdasarkan aspek validitas, reliabilitas, indeks tingkat kesukaran, indeks daya beda dan efektivitas butir pengecoh adalah sebagai berikut.

Aspek Validitas

Uji validitas menggunakan rumus *point biserial* dengan bantuan program *SPSS Ver 25*. Hasil perhitungan uji validitas dibandingkan dengan r_{tabel} pada taraf signifikansi 5%. Sampel dalam penelitian ini berjumlah 63, maka r_{tabel} dari N 63 dengan tingkat signifikan 5% adalah 0,244. Butir soal dapat dikatakan valid apabila r_{hitung} lebih besar dari r_{tabel} (0,244), sebaliknya jika r_{hitung} lebih kecil dari r_{tabel} maka soal dikatakan tidak valid. Berikut adalah hasil uji validitas butir soal. Berikut hasil uji validitas butir soal.

Tabel 1. Hasil uji validitas

Nomor Soal	r_{tabel}	r_{hitung}	Kategori
1	0,244	,337	Valid
2	0,244	,259	Valid
3	0,244	,411	Valid
4	0,244	0,218	Tidak valid
5	0,244	-0,046	Tidak valid
6	0,244	,415	Valid
7	0,244	,318	Valid
8	0,244	,466	Valid
9	0,244	,357	Valid
10	0,244	0,054	Tidak valid
11	0,244	0,145	Tidak valid
12	0,244	,324	Valid
13	0,244	0,219	Tidak valid
14	0,244	,500	Valid
15	0,244	0,207	Tidak valid
16	0,244	,429	Valid
17	0,244	,346	Valid
18	0,244	,459	Valid
19	0,244	,563	Valid
20	0,244	,252	Valid
21	0,244	0,167	Tidak valid
22	0,244	0,119	Tidak valid
23	0,244	0,170	Tidak valid
24	0,244	,431	Valid
25	0,244	0,214	Tidak valid
26	0,244	-0,103	Tidak valid
27	0,244	-0,095	Tidak valid
28	0,244	,413	Valid

Nomor Soal	r_{tabel}	r_{hitung}	Kategori
29	0,244	0,105	Tidak valid
30	0,244	410	Valid

Berdasarkan tabel tersebut menunjukkan sebanyak 17 butir soal atau 57% dinyatakan valid karena r_{hitung} lebih besar dari r_{tabel} . Sedangkan sebanyak 13 butir soal atau 43% dinyatakan tidak valid karena r_{hitung} kurang dari r_{tabel} . Sehingga dapat disimpulkan bahwa pada aspek validitas, butir soal penilaian tengah semester Bahasa Indonesia kelas IX SMP Negeri 4 Magelang tahun ajaran 2022/2023 memiliki tingkat validitas sedang karena perbandingan jumlah butir soal valid dengan butir soal tidak valid hampir seimbang.

Sebuah instrumen penilaian berupa soal dapat dikatakan valid apabila mampu mengukur apa yang diukur dan mampu mendapatkan data dari variabel yang diteliti dengan tepat. Sehingga diperlukan uji validitas untuk mengetahui seberapa tepat dan cermat suatu instrumen penilaian dalam melakukan fungsi ukurnya. Validitas yang baik menunjukkan kesahihan atau ketepatan suatu instrumen penilaian. (Arifin, 2019). Soal penilaian tengah semester yang dibuat perlu direvisi untuk ditingkatkan validitas sehingga tujuan tes digunakan dapat tercapai. Diperlukan peningkatan validitas soal agar tes yang digunakan mampu mendapatkan data kemampuan peserta didik dalam memahami materi pembelajaran.

Semakin banyak butir item yang dapat dijawab dengan betul oleh testee, maka skor- skor total hasil tes akan semakin tinggi. Sebaliknya, semakin sedikit butir- butir item yang dapat dijawab dengan betul oleh testee, maka skor- skor total hasil tes itu akan semakin rendah atau menurun. (Nurjanah & Marlianingsih, 2015). Dalam soal penilaian tengah semester matpel bahasa indonesia tersebut masih banyak soal yang dijawab oleh siswa salah. Melihat masih adanya siswa yang benar kemungkinan besar yang menjadi sumber masalah adalah jenis soal.

Aspek Reliabilitas

Uji reliabilitas soal pilihan ganda dapat dilakukan dengan teknik *Cronbach's Alpha* atau Koefisien *Alpha*. Berikut hasil hitung uji reliabilitas butir soal.

Cronbach's Alpha	N of Items
0,551	30

Sukiman (2012) menyebutkan interpretasi terhadap hasil uji reliabilitas dibagi menjadi lima kategori yaitu sangat tinggi, tinggi, sedang, rendah, dan sangat rendah. Berdasarkan standar tersebut maka disimpulkan bahwa butir soal pilihan ganda penilaian tengah semester Bahasa Indonesia kelas IX SMP Negeri 4 Magelang tahun ajaran 2022/2023 masuk kategori sedang. Artinya butir soal dinilai belum mampu secara maksimal dalam memberikan hasil yang sama apabila butir soal diujikan pada peserta didik yang sama dalam waktu atau kesempatan yang berbeda.

Peningkatan reliabilitas perlu dilakukan guru. Reliabilitas merupakan tingkat konsistensi atau keajegan dari suatu instrumen penilain. Tes dapat dikatakan reliabel apabila dapat memberikan hasil yang sama terhadap kelompok yang sama meskipun diujikan pada waktu yang berbeda. (Arifin, 2016). Atau dengan kata lain tes dikatakan reliable jika memberikan hasil yang tepat apabila diteskan berkali-kali. Susunan tes dikatakan objektif apabila dalam melaksanakan tes itu tidak ada faktor subjektif yang mempengaruhi. (Nurjanah & Marlianingsih, 2015). Tes yang reliabel tentu akan mempermudah guru dalam melaksanakan tes ulang dengan soal yang sama di kemudian hari. Sehingga peningkatan reliabilitas perlu dilaksanakan.

Pada soal PTS Bahasa Indonesia tingkat reliabilitas sedang sehingga belum mampu jika diterapkan berulang pada siswa yang lain. Terdapat dua faktor yang mempengaruhi tingkat reliabilitas butir soal yaitu panjang tes dan tingkat kesukaran. Panjang tes berarti juga banyaknya jumlah soal tes, hal ini berpengaruh karena semakin banyak soal dalam suatu tes, maka semakin banyak juga sampel yang harus diukur dan jumlah jawaban benar akan semakin banyak, sehingga mempengaruhi faktor tebakan yaitu faktor tebakan semakin rendah. Faktor tingkat kesukaran juga turut mempengaruhi tingkat reliabilitas. Semakin mudah tingkat kesukaran pada butir soal akan menghasilkan reliabilitas yang tinggi.

Indeks Tingkat Kesukaran

Analisis tingkat kesukaran dilakukan untuk mengukur derajat kesulitan suatu butir soal. Arifin (2019) mengemukakan bahwa soal dapat dikatakan baik apabila memiliki tingkat kesukaran seimbang atau sedang, artinya soal disusun tidak terlalu mudah dan tidak terlalu sukar. Hasil perhitungan indeks tingkat kesukaran diinterpretasikan dalam kategori sukar, sedang, dan mudah. Berikut adalah hasil analisis butir soal berdasarkan tingkat kesukaran. Berikut hasil analisis tingkat kesukaran butir soal.

Tabel 3. Indeks Tingkat kesukaran

Indeks Tingkat Kesukaran	Kategori	Jumlah	%
$P < 0,30$	Sukar	2	7%
$0,30 \leq p \leq 0,70$	Sedang	12	40%
$P > 0,70$	Mudah	16	53%

Dapat disimpulkan berdasarkan hasil analisis, bahwa soal penilaian tengah semester Bahasa Indonesia kelas IX SMP Negeri 4 Magelang tahun ajaran 2022/2023 adalah soal yang tidak baik karena memiliki tingkat kesukarannya dominan yang mudah dan sedang. Akibatnya indeks tingkat kesukaran pada butir soal tidak berfungsi dengan baik sehingga tidak mampu membedakan peserta didik kelompok tinggi dengan peserta didik kelompok rendah.

Faktor tidak berfungsinya indeks tingkat kesukaran adalah banyaknya butir soal yang tergolong mudah daripada kategori yang lain yaitu sebanyak 16 butir soal berada pada interval $> 0,70$. Semakin tinggi indeks tingkat kesukaran, maka soal tersebut semakin mudah. Sebaliknya, semakin rendah indeks tingkat kesukaran, maka soal semakin sulit. Butir soal yang masuk dalam kategori sukar dan mudah menunjukkan bahwa dari segi materi, butir soal tersebut belum mewakili materi yang telah diajarkan. Selain itu, soal yang terlalu mudah tidak akan mendorong peserta didik untuk memecahkan soal tersebut, sedangkan soal yang terlalu sukar menyebabkan peserta didik putus asa dan tidak mau mencoba karena di luar batas kemampuannya. (Dias Rasmawati, 2016). Terdapat 2 butir soal berkategori sukar karena berada pada interval $< 0,30$. Adanya soal sukar ini menyebabkan peserta didik dari kelompok tinggi dan kelompok rendah kesulitan dalam menjawab butir soal. Agar soal dikatakan baik, maka indeks tingkat kesukaran

harus berada pada interval $0,30 - 0,70$. Maka perlu dilakukan revisi perbaikan soal.

Indeks Daya Beda

Analisis daya beda suatu butir soal berfungsi untuk membedakan kemampuan peserta didik dalam menguasai kompetensi yang sudah ditentukan. Soal dapat dikatakan baik apabila memiliki tingkat koefisien yang tinggi, sehingga soal tersebut mampu membedakan peserta didik yang menguasai kompetensi dengan peserta didik yang kurang menguasai kompetensi. Indeks daya pembeda diartikan sebagai selisih hasil jawaban benar kelompok atas dengan hasil jawaban benar kelompok bawah. Kelompok atas adalah kelompok peserta didik yang mampu atau pandai, sedangkan kelompok bawah adalah kelompok peserta didik yang kurang mampu atau kurang pandai. Angka hasil perhitungan daya beda disebut indeks diskriminasi atau D dengan angka yang berkisar antara $0,00 - 1,00$. Adapun penggunaan tanda negatif (-) pada indeks diskriminasi ini menunjukkan bahwa terbalikny kualitas peserta didik tes. Hal ini berarti peserta didik tes berkemampuan rendah mampu menjawab soal dengan benar, sedangkan peserta didik berkemampuan tinggi menjawab dengan salah.

Daryanto (2001) menyebutkan hasil perhitungan indeks daya beda diinterpretasikan menjadi lima kategori, yaitu jelek, cukup, baik, baik sekali, dan tidak baik atau negatif. Berikut hasil analisis indeks daya beda terhadap butir soal.

Tabel 4. Indeks Daya Beda

Indeks Daya Beda	Kategori	Jumlah	Persentase
Negatif	Tidak baik	5	17%
$0,00 - 0,20$	Jelek	12	40%
$0,21 - 0,40$	Cukup	12	40%
$0,41 - 0,70$	Baik	1	3%
$0,71 - 1,00$	Sangat baik	-	0%

Berdasarkan tabel tersebut distribusi hasil analisis indeks daya daya terdapat 12 butir soal atau 40% berkategori jelek yaitu soal nomor 1, 2, 4, 11, 13, 15, 17, 20, 21, 22, 23, dan 25. Sebanyak 12 butir soal atau 40% berkategori cukup yaitu soal nomor 3, 6, 7, 8, 9, 12, 14, 16, 18, 24, 28, dan 30. Butir soal berkategori baik hanya 1 soal atau 3% yaitu soal nomor 19. Butir soal berkategori tidak baik sejumlah 5 butir soal atau 17% yaitu soal nomor 5, 10, 26, 27, dan 29.

Dapat disimpulkan bahwa soal penilaian tengah semester Bahasa Indonesia kelas IX SMP Negeri 4 Magelang tahun ajaran 2022/2023 memiliki indeks daya beda cukup sampai jelek. Tujuan pokok mencari daya pembeda ialah untuk menentukan apakah butir soal tersebut memiliki kemampuan membedakan kelompok dalam aspek yang diukur, sesuai dengan perbedaan yang ada pada kelompok tersebut. (Nasir, 2015). Soal tersebut belum mampu secara maksimal dalam membedakan kemampuan peserta didik kelas tinggi dengan peserta didik kelas bawah dalam menguasai materi pembelajaran. Hal ini diketahui dari hasil analisis yang menunjukkan bahwa butir soal yang memiliki daya beda berkategori tidak baik dan berkategori cukup lebih banyak dari butir soal berkategori lain.

Efektivitas Butir Pengecoh

Butir pengecoh merupakan alternatif jawaban pada soal pilihan ganda yang harus dipilih peserta didik. Butir pengecoh berfungsi untuk mengecoh peserta didik dalam menentukan jawaban soal. Semakin banyak peserta didik yang memilih terkecoh, maka pengecoh tersebut berfungsi dengan baik. Sebaliknya, jika pengecoh tersebut tidak dipilih oleh peserta didik, maka pengecoh tidak berfungsi dengan baik. Butir pengecoh pada tiap soal dihitung lalu diinterpretasikan menjadi lima kategori, yaitu sangat baik, baik, kurang baik, jelek, dan sangat jelek. Hasil analisis butir pengecoh tiap butir soal selanjutnya dianalisis dengan kategori sebagai berikut:

- Apabila tiga jawaban pengecoh dapat berfungsi maka soal tersebut berkategori sangat baik.
- Apabila dua jawaban pengecoh dapat berfungsi maka soal tersebut berkategori baik sebut berkategori kurang baik.
- Apabila semua jawaban pengecoh tidak berfungsi maka soal tersebut berkategori jelek.

Tabel 5. Efektivitas Butir Pengecoh

Kategori	Jumlah	Persentase
Sangat baik	4	13%
Baik	1	3%
Kurang baik	17	57%
Jelek	8	27%

Hal ini selaras dengan hasil analisis tingkat kesukaran butir soal yang berkategori mudah. Berdasarkan hasil analisis butir soal penilaian tengah semester Bahasa Indonesia kelas

IX SMP Negeri 4 Magelang tahun ajaran 2022/2023 memiliki efektivitas belum mampu mengecoh peserta didik dalam memilih jawaban sehingga belum bisa mengukur kemampuan peserta didik dengan baik. Pengecoh adalah pilihan jawaban bukan kunci jawaban yang berfungsi untuk mengecoh peserta tes yang tidak menguasai persoalan/materi yang diujikan. Semua pengecoh harus mempunyai daya tarik untuk dipilih peserta tes sebagai jawaban terhadap butir soal. (Sari, Utomo, & Sumarwati, 2022). Pada soal tersebut pengecoh kurang dibuat secara baik. Sehingga mampu dijawab oleh semua peserta didik.

Kualitas Soal

Hasil analisis butir soal penilaian tengah semester Bahasa Indonesia kelas IX SMP Negeri 4 Magelang tahun ajaran 2022/2023 selanjutnya dideskripsikan tiap-tiap butir soalnya untuk mengetahui kualitas butir soal dan tindak lanjut yang perlu dilakukan. Hasil analisis diinterpretasi kualitas butir soal yang diadaptasi dari skala *Likert* sebagai berikut.

- Butir soal yang memenuhi empat kriteria soal yang baik, yaitu validitas, tingkat kesukaran, daya beda, dan efektivitas pengecoh maka soal tersebut memiliki kualitas sangat baik dan soal dapat disimpan di bank soal.
- Butir soal yang memenuhi tiga dari empat kriteria soal yang baik, yaitu validitas, tingkat kesukaran, daya beda, dan efektivitas pengecoh maka soal tersebut memiliki kualitas baik, sehingga tidak dapat disimpan di bank soal dan soal perlu dilakukan revisi agar dapat memenuhi empat kriteria.
- Butir soal yang memenuhi dua dari empat kriteria soal yang baik, yaitu validitas, tingkat kesukaran, daya beda, dan efektivitas pengecoh maka soal tersebut memiliki kualitas sedang, sehingga tidak dapat disimpan di bank soal dan soal perlu dilakukan revisi agar dapat memenuhi empat kriteria.
- Butir soal yang memenuhi satu dari empat kriteria soal yang baik, yaitu validitas, tingkat kesukaran, daya beda, dan efektivitas pengecoh maka soal tersebut memiliki kualitas tidak baik, sehingga tidak dapat disimpan di bank soal dan soal perlu dibuang.
- Butir soal yang tidak memenuhi empat kriteria soal yang baik, yaitu validitas, tingkat kesukaran, daya beda, dan efektivitas

pengecoh maka soal tersebut memiliki kualitas sangat tidak baik, sehingga tidak dapat disimpan di bank soal dan soal perlu dibuang.

- f. Selain syarat berbasis butir, maka tes secara keseluruhan harus reliabel dengan ketentuan yang sudah dijelaskan sebelumnya.

Tabel 6. Kualitas Soal

Kriteria Kualitas Soal	Jumlah	Persentase
Sangat baik	1	4%
Baik	-	0%
Sedang	10	33%
Tidak baik	10	33%
Sangat tidak baik	9	30%

Kualitas butir soal penilaian tengah semester Bahasa Indonesia kelas IX SMP Negeri 4 Magelang tahun ajaran 2022/2023 berada pada tingkat sedang dan tidak baik. Dapat disimpulkan butir soal belum memiliki kualitas yang sangat baik ataupun baik, karena hasil analisis menunjukkan kualitas berada pada tingkat sedang dan tidak baik. Hasil analisis butir soal menentukan tindak lanjut yang perlu dilakukan pada soal-soal tersebut. Berikut hasil analisis tindak lanjut terhadap butir soal.

Tabel 7. Tindak Lanjut Soal

Bentuk Tindak Lanjut	Jumlah	Persentase
Dihapus	19	64%
Direvisi	10	33%
Disimpan dalam bank soal	1	3%

Soal penilaian tengah semester Bahasa Indonesia kelas IX SMP Negeri 4 Magelang tahun ajaran 2022/2023 belum bisa menjalankan fungsinya atau gagal sebagai instrumen penilaian dengan baik. Kegagalan butir soal dapat diketahui melalui aspek validitas, indeks tingkat kesukaran, indeks daya beda, dan efektivitas butir pengecoh.

Tabel 8. Penyebab Kegagalan Butir Soal

Penyebab Kegagalan	Jumlah	Persentase
Validitas (tidak valid)	13	18%
Indeks Tingkat Kesukaran (Sukar dan mudah)	18	25%
Indeks Daya Beda (tidak baik)	17	23%
Efektivitas Butir Pengecoh (Kurang baik)	25	34%

Penyebab kegagalan yang paling utama adalah efektivitas butir pengecoh. Hal ini menunjukkan bahwa butir pengecoh belum

mampu menjalankan fungsinya dengan baik. Penyebab kegagalan butir soal yang kedua adalah tingkat kesukaran soal yang mudah. Penyebab yang ketiga yaitu daya beda yang jelek dan negatif, sehingga butir soal tidak mampu membedakan antara peserta didik kelompok tinggi dengan peserta didik kelompok rendah. Penyebab yang keempat yaitu terdapat butir soal yang tidak valid. Selain keempat aspek tersebut, butir soal juga memiliki tingkat reliabel yang sedang, sehingga tingkat keajegkan atau konsistensi soal yaitu cukup.

Soal penilaian tengah semester Bahasa Indonesia kelas IX SMP Negeri 4 Magelang tahun ajaran 2022/2023 belum memiliki kualitas yang baik, karena jumlah soal yang dibuang lebih banyak, yaitu sebanyak 19 butir soal atau 64%. Jumlah ini merupakan jumlah yang lebih banyak dari soal yang direvisi yaitu 10 soal atau 33% dan soal yang dimasukkan ke dalam bank soal hanya 1 soal atau 3%. Soal tes yang layak untuk mengukur kemampuan kognitif mahasiswa jika soal tes kategori valid, indeks reliabilitas soal pada kriteria reliabilitas tinggi, indeks kesukaran dengan kategori butir soal baik/layak, dan daya pembeda butir soal termasuk klasifikasi soal yang baik. (Saputra et al, 2022). Sehingga soal yang tersedia berdasarkan analisis butir soal yang dilakukan belum mampu mengukur tingkat kemampuan kognitif siswa. Soal masih perlu diperbaiki dari kelima aspek yang dianalisis.

SIMPULAN

Berdasarkan hasil analisis butir soal penilaian tengah semester Bahasa Indonesia kelas IX SMP Negeri 4 Magelang tahun ajaran 2022/2023 yang ditinjau dari segi validitas, reliabilitas, indeks tingkat kesukaran, indeks daya beda, dan efektivitas butir pengecoh dapat disimpulkan bahwa (1) butir soal memiliki validitas yang sedang karena perbandingan butir soal yang valid dengan butir soal tidak valid jumlahnya hampir seimbang, (2) butir soal memiliki tingkat reliabel sedang dengan hasil 0,551, (3) butir soal memiliki tingkat kesukaran mudah, (4) butir soal memiliki indeks daya beda berkategori jelek sampai cukup, dan (5) butir soal memiliki efektivitas butir pengecoh yang kurang baik. Berdasarkan keseluruhan aspek analisis butir soal, soal penilaian tengah semester Bahasa Indonesia kelas IX SMP Negeri 4 Magelang tahun

ajaran 2022/2023 adalah butir soal yang kualitasnya belum baik. Hal ini dapat diketahui dari jumlah butir soal yang dibuang atau dihilangkan adalah jumlah paling banyak yaitu 19 butir soal atau 64%, sedangkan butir soal yang direvisi berjumlah 10 soal atau 33%, dan hanya ada 1 atau 3% soal yang dapat dimasukkan dalam bank soal. Adapun penyebab kegagalan butir soal yang paling besar adalah dari aspek efektivitas butir pengecoh (kurang baik dan jelek), lalu aspek indeks tingkat kesukaran (sukar dan mudah) ketiga oleh indeks daya beda (jelek dan negatif), dan validitas (tidak valid) butir soal yang berjumlah. Selain keempat aspek tersebut, butir soal juga memiliki tingkat reliabel yang sedang, sehingga tingkat keajegkan atau konsistensi soal yaitu cukup.

DAFTAR PUSTAKA

- Arifin, Z. (2019). *Evaluasi Pembelajaran*. Bandung: PT Remaja Rosdakarya.
- Daryanto. (2001). *Evaluasi Pendidikan*. Jakarta: PT Rineka Cipta.
- Fitriawan, M. (2017). Peran Analisis Butir Soal Guna Meningkatkan Kualitas Butir Soal, Kompetensi Guru Dan Hasil Belajar Peserta Didik. *Prosiding Seminar Nasional dan Call for Papers Pendidikan 2017 (PGSD UMS & HDPGSDI Wilayah Jawa)*. Diakses dari <https://publikasiilmiah.ums.ac.id/xmlui/handle/11617/11576>
- Kurniawan, T. (2015). ANALISIS BUTIR SOAL ULANGAN AKHIR SEMESTER GASAL MATA PELAJARAN IPS SEKOLAH DASAR. *Journal of Elementary Education*. 4(1). Retrieved from <http://journal.unnes.ac.id/sju/index.php/je>
- Nasir, M. (2015). ANALISIS EMPIRIK PROGRAM ANALISIS BUTIR SOAL DALAM RANGKA MENGHASILKAN SOAL YANG BAIK DAN BERMUTU SEBAGAI ALAT EVALUASI PEMBELAJARAN FISIKA. *Prosiding Semirata 2015 bidang MIPA BKS-PTN Barat Universitas Tanjungpura Pontianak*. Retrieved from <https://jurnal.untan.ac.id/index.php/semirata2015/article/view/13534/0>
- Nurjanah, & Marlianingsih, N. (2015). ANALISIS BUTIR SOAL PILIHAN GANDA DARI ASPEK KEBAHASAAN. *Faktor : Jurnal Ilmiah Kependidikan*. 2(1). <http://dx.doi.org/10.30998/fjik.v2i1.377>
- Rahmasari, D., & Ismiyati, I. (2016). ANALISIS BUTIR SOAL MATA PELAJARAN PENGANTAR ADMINISTRASI PERKANTORAN. *Economic Education Analysis Journal*. 5(1). Retrieved from <https://journal.unnes.ac.id/sju/index.php/eaj/article/view/10007>
- Saputra, H.D., Fernandez, D., Purwanto, W., Setiawan, D., & Putra, R. (2022). Hasil Belajar Mahasiswa: Analisis Butir Soal Tes. *Edukasi Jurnal Pendidikan*. 20(1). <https://doi.org/10.31571/edukasi.v20i1.3432>
- Sari, V., Utomo, A., & Sumarwati, S. (2022). Kualitas Soal Bahasa Indonesia di SMP Muhammadiyah 1 Pontianak: Analisis Butir Soal. *Jurnal Pendidikan Bahasa Dan Sastra Indonesia*, 11(2), 112-119. <https://doi.org/10.15294/jpbsi.v11i2.58091>
- Sukiman. (2012). *Pengembangan Sistem Evaluasi*. Yogyakarta: Insan Madani.