

## Quality Analysis of Indonesian Language Test Questions Class VII SMP Negeri 11 Ruteng Using IteMan Programme Version 4.3

Yunarius Ugi Suharjo<sup>1\*</sup>, Vinsensius Sumardi<sup>2</sup>

Universitas Negeri Yogyakarta

**Corresponding Author:** Yunarius Ugi Suharjo

[yunariusugi.2022@student.uny.ac.id](mailto:yunariusugi.2022@student.uny.ac.id)

---

### ARTICLE INFO

*Keywords:* Question Quality, Indonesian Language, Grade VII

*Received :* 3 November

*Revised :* 17 December

*Accepted:* 17 January

©2024 Suharjo, Sumardi: This is an open-access article distributed under the terms of the [Creative Commons Atribusi 4.0 Internasional](https://creativecommons.org/licenses/by/4.0/).



### ABSTRACT

The purpose of this study was to describe the validity, reliability, difficulty index, differentiation index, and the effectiveness of distractors on the even semester Indonesian language test questions of class VII SMP Negeri 11 Ruteng, Manggarai Regency in the 2022/2023 academic year using the IteMan version 4.3 programme. The method used was descriptive quantitative. The validity results show that the 40 items of multiple choice questions compiled are valid questions. The reliability results of 40 questions tested on 25 students obtained an Alpha Cronbach score of 0.788. Thus, the question is declared reliable because the amount of the reliability index is more than 0.60. The average index of difficulty (ITK) of all items is 0.433, with 9 questions at the easy level, 10 questions at the medium level, 12 questions at the difficult level, and 9 other questions at the very difficult level. The average index of discriminating power (IDB) is 0.246 with 13 questions declared feasible, 5 questions declared moderate (feasible and do not need to be revised), 7 questions declared good enough (feasible and need to be revised), 15 questions declared unfit. The results of the effectiveness of the distractors showed that 39 of the questions were functioning and 1 question option was not functioning

## Analisis Kualitas Soal Ulangan Bahasa Indonesia Kelas VII SMP Negeri 11 Ruteng Menggunakan Program IteMan Versi 4.3

Yunarius Ugi Suharjo<sup>1\*</sup>, Vinsensius Sumardi<sup>2</sup>

Universitas Negeri Yogyakarta

**Corresponding Author:** Yunarius Ugi Suharjo

[yunariusugi.2022@student.uny.ac.id](mailto:yunariusugi.2022@student.uny.ac.id)

---

### ARTICLE INFO

*Kata Kunci:* Kualitas Soal, Bahasa Indonesia, Kelas VII

*Received :* 3 November

*Revised :* 17 December

*Accepted:* 17 January

©2023 Suharjo, Sumardi: This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International](https://creativecommons.org/licenses/by/4.0/).



### ABSTRAK

Tujuan penelitian ini yakni mendeskripsikan validitas, reliabilitas, indeks tingkat kesulitan, indeks daya beda, dan efektivitas distraktor soal ulangan mata pelajaran Bahasa Indonesia semester genap kelas VII SMP Negeri 11 Ruteng Kabupaten Manggarai tahun pelajaran 2022/2023 menggunakan program IteMan versi 4.3. Metode yang digunakan adalah deskriptif kuantitatif. Hasil validitas menunjukkan 40 butir soal pilihan ganda yang disusun merupakan soal yang valid. Hasil reabilitas 40 soal yang diujikan kepada 25 peserta didik diperoleh skor Alpha Cronbach sebesar 0,788. Dengan demikian, soal dinyatakan reliabel karena besaran indeks reliabilitasnya lebih dari 0,60. Rata-rata indeks tingkat kesulitan (ITK) seluruh butir soal sebesar 0,433, dengan 9 soal berada pada level mudah, 10 soal berada pada level sedang, 12 soal berada pada level sulit, dan 9 soal lainnya berada pada level sangat sulit. Rata-rata indeks daya beda (IDB) sebesar 0,246 dengan 13 soal dinyatakan layak, 5 soal dinyatakan sedang (layak dan tidak perlu direvisi), 7 soal dinyatakan cukup baik (layak dan perlu direvisi), 15 soal dinyatakan tidak layak. Hasil efektivitas distraktor menunjukkan 39 soal pengecohnya berfungsi dan 1 opsi soal pengecohnya tidak berfungsi

## PENDAHULUAN

Salah satu kebijakan dalam dunia pendidikan yaitu ditetapkannya Bahasa Indonesia sebagai mata pelajaran di setiap satuan pendidikan. Hal tersebut sejalan dengan Peraturan Pemerintah Republik Indonesia No 57 Tahun 2014 Tentang Pembinaan Bahasa Indonesia, pasal 18 ayat 1 yang menyatakan bahwa setiap satuan pendidikan, baik pendidikan dasar, pendidikan menengah, pendidikan tinggi, maupun program pendidikan kesetaraan, wajib menyelenggarakan pembelajaran Bahasa Indonesia. Selanjutnya, pembelajaran Bahasa Indonesia tersebut mengacu pada standar kemahiran berbahasa Indonesia yang termuat dalam standar isi dan standar kompetensi lulusan di setiap jenjang pendidikan. Oleh karena itu, agar terwujudnya tujuan pendidikan maka kebijakan yang berkaitan dengan bahasa mesti dirumuskan secara komprehensif agar pelaksanaannya optimal.

Lebih lanjut, untuk melengkapi kebijakan implementasi peran Bahasa Indonesia sebagai salah satu mata pelajaran dalam dunia pendidikan, maka perlu dilakukan sebuah penilaian guna melihat pencapaian kompetensi peserta didik dalam proses pembelajaran. Sarker & Ullah (2023) menyatakan bahwa mutu pendidikan dapat dilihat melalui dimensi proses yang mencakup pembelajaran, isi mata pelajaran, pengetahuan guru, keakuratan isi kurikulum, media pembelajaran, dan penilaian terhadap siswa. Oleh karenanya, penilaian itu sendiri penting untuk dilakukan guna mengetahui pencapaian peserta didik. Penilaian menurut Gronlund (via Nurgiyantoro, 2017) merupakan serangkaian proses yang sistematis dilakukan dalam pengumpulan, analisis, dan penafsiran informasi guna menentukan seberapa jauh kemampuan peserta didik dalam mencapai tujuan yang diinginkan. Pengertian tersebut senada dengan pendapat Kurniawati (2020) yang menyatakan bahwa penilaian diadakan tidak hanya sekadar mengukur sejauh mana tujuan yang diinginkan tercapai, melainkan hasil penilaian nantinya akan dijadikan sebagai acuan untuk pengambilan suatu keputusan.

Dengan demikian, penilaian merupakan salah satu aspek terpenting dalam menyimpulkan hasil capaian sebuah pembelajaran. Gladushyna & Strietholt (2023) mengemukakan bahwa pendekatan yang paling umum untuk menilai kualitas pendidikan adalah dengan mengukur pengetahuan dan keterampilan yang dimiliki peserta didik melalui tes standar. Di pihak lain, Nurgiyantoro (2011) mengemukakan bahwa penilaian dapat memberikan dampak positif bagi keberhasilan kompetensi peserta didik, oleh karenanya penilaian itu sendiri merupakan bagian dari kebijakan yang mesti dijadikan salah satu tolok ukur dalam rangkaian proses pembelajaran. Penilaian yang dimaksud harus sesuai dengan rancangan kompetensi dasar dan semua materi yang sudah dipelajari sebelumnya. Hal tersebut dapat dilakukan melalui pengukuran dengan menggunakan instrumen berupa tes.

Penilaian hasil belajar oleh satuan pendidikan dapat dilakukan pada akhir semester dalam bentuk tes ulangan. Pengumpulan informasi lewat teknik tes menurut Nurgiyantoro (2017) biasanya dilakukan dengan memberikan tes berupa tugas, latihan, ulangan atau sejumlah pertanyaan yang harus dikerjakan oleh peserta tes. Selanjutnya hasil pekerjaan peserta tes tersebut dapat

dimanfaatkan sebagai acuan bahan evaluasi. López-Hernández, et al (2023) melihat perlunya menghubungkan penilaian dan umpan balik terhadap hasil belajar siswa, daripada menggunakannya hanya sebagai alat untuk memeriksa standar di akhir pembelajaran. Oleh karena itu, guna mencapai tujuan yang diinginkan dalam proses penilaian hasil belajar, penting untuk mengetahui kualitas alat tes yang akan dikerjakan oleh peserta didik. Kualitas soal atau alat tes tersebut dapat dilakukan melalui analisis butir soal. Kegunaan analisis butir soal menurut Arikunto (2018) yakni untuk menemukan butir soal yang berkualitas baik, cukup baik, dan kurang baik serta agar mendapatkan informasi guna memperbaiki atau menyempurnakan soal terkait. Dengan demikian, penelitian ini berusaha menganalisis kualitas soal menggunakan program komputer IteMan versi 4.3. Program tersebut merupakan salah satu program komputer yang dapat dimanfaatkan untuk mengukur kualitas butir soal sebagai penunjang kegiatan pembelajaran.

Fokus penelitian ini yakni menganalisis kualitas soal ulangan Bahasa Indonesia kelas VII menggunakan program IteMan versi 4.3. Adapun penelitian terdahulu yang berkaitan dengan penelitian ini telah dilakukan oleh Agustiana, et al (2018), Himawan & Nurgiyantoro (2022), dan Pradani & Efendi (2023). Pertama, penelitian yang dilakukan Agustiana, et al (2018) memiliki persamaan dengan penelitian ini yaitu menganalisis butir soal Bahasa Indonesia. Sementara itu, perbedaannya yaitu penelitian Agustiana, et al (2018) menggunakan program SPSS versi 22 dan Anates versi 4 dalam menganalisis butir soal pilihan ganda dan soal uraian, sedangkan penelitian ini menganalisis soal pilihan ganda menggunakan program IteMan versi 4.3. Kedua, penelitian Himawan & Nurgiyantoro (2022) memiliki persamaan dengan penelitian ini yaitu menganalisis butir soal Bahasa Indonesia menggunakan program IteMan. Sementara itu, perbedaannya adalah penelitian Himawan & Nurgiyantoro (2022) menganalisis butir soal latihan kelas VIII SMPN 1 Bambanglipuro Bantul, sedangkan penelitian ini menganalisis soal ulangan kelas VII SMPN 11 Ruteng Kabupaten Manggarai. Ketiga, penelitian yang dilakukan oleh Pradani & Efendi (2023) memiliki persamaan dengan penelitian ini yaitu menganalisis butir soal Bahasa Indonesia menggunakan program IteMan. Sementara itu, perbedaannya yaitu penelitian Pradani & Efendi (2023) menganalisis butir soal Ujian Sekolah SMP di Kabupaten Rembang tahun pelajaran 2021/2022, sedangkan penelitian ini menganalisis butir soal ulangan kelas VII SMPN 11 Ruteng tahun pelajaran 2022/2023.

Berdasarkan uraian di atas, diketahui bahwa program IteMan dapat dimanfaatkan untuk mengukur kualitas soal yang disusun. Penelitian mengenai pemanfaatan program komputer seperti IteMan saat ini masih jarang dilakukan. Hal tersebut selaras dengan pendapat Arviyanto (2017) yang menyatakan bahwa penelitian analisis kualitas butir soal menggunakan program komputer belum banyak dilakukan, sehingga perlu adanya penelitian lanjutan guna melengkapi kebutuhan rujukan/referensi dan menghadirkan penelitian relevan. Sejalan dengan hal tersebut, maka penelitian ini sengaja dihadirkan guna memberikan sumbangsih rujukan untuk penelitian selanjutnya.

Oleh karena itu, kebaruan yang dihadirkan dalam penelitian ini yakni menampilkan analisis butir soal dengan memanfaatkan teknologi berupa program komputer Iteman versi 4.3, guna menyesuaikan tuntutan perkembangan zaman. Soal-soal yang dianalisis pada penelitian ini berjumlah 40 butir dan diujikan kepada 25 peserta didik kelas VII SMP Negeri 11 Ruteng sebagai soal penilaian ulangan harian semester genap tahun pelajaran 2022/2023. Dengan menganalisis validitas isi, reliabilitas, indeks tingkat kesulitan soal, indeks daya beda soal, dan efektivitas distraktor, penyusun soal dapat mengetahui dan menjadikan temuan penelitian sebagai bahan evaluasi, dan guru mata pelajaran dapat memanfaatkannya untuk meningkatkan kualitas setiap butir soal tes.

### **TINJAUAN PUSTAKA**

Pengertian penilaian menurut Gronlund merupakan serangkaian proses yang sistematis dilakukan dalam pengumpulan, analisis, dan penafsiran informasi guna mengetahui sejauh mana kemampuan peserta tes/peserta didik dalam mencapai tujuan yang diinginkan (via Nurgiyantoro, 2017). Penilaian yang diterapkan dalam mengukur pencapaian kompetensi peserta didik dapat menggunakan alat penilaian berupa tes. Alat berupa tes tersebut haruslah memperhatikan beberapa ketentuan. Dalam penentuan tes, Nurgiyantoro (2017) mengemukakan bahwa terdapat empat aspek yang harus dilakukan oleh pembuat tes tersebut, yakni harus menentukan kompetensi dasar, membuat deskripsi bahan uji yang akan diujikan, pembuatan atau penyusunan kisi-kisi soal, dan penentuan atau penyusunan soal serta durasi waktu tes.

Penyusunan butir soal yang akan dikerjakan oleh peserta didik harus memenuhi syarat atau kaidah penulisan. Syarat atau kaidah penyusunan soal pilihan ganda dalam Depdiknas (via Utami & Nurgiyantoro, 2016), harus memperhatikan tiga aspek yang terdiri dari aspek materi, konstruksi, dan bahasa. Pertama, dalam aspek materi penting diperhatikan kesesuaian antara soal dengan indikator yang akan dicapai, artinya soal tersebut harus menanyakan atau mempertimbangkan perilaku dan materi yang akan diukur berdasarkan rumusan indikator dalam kisi-kisi. Selain itu, opsi pengecohnya harus berfungsi, dan setiap soal mesti memiliki satu jawaban yang benar, artinya di dalam satu soal hanya memiliki satu kunci jawaban. Kedua, aspek konstruksi yakni pokok soal yang disusun harus dirumuskan secara tegas dan jelas. Artinya, materi yang akan diukur dalam soal harus jelas dan tidak menimbulkan penafsiran yang berbeda. Berikutnya, setiap butir soal yang disusun hanya mengandung satu persoalan. Ketiga, dalam aspek bahasa setiap butir soal harus menggunakan bahasa yang sesuai dengan kaidah penulisan bahasa Indonesia.

Selain itu, dalam proses penilaian perlu diperhatikan kualitas butir soal yang disusun. Dalam hal ini, Nurgiyantoro (2017) menjelaskan beberapa hal penting dalam menganalisis butir soal, yakni melakukan penghitungan tingkat kesulitan butir soal, penghitungan daya beda soal, dan menganalisis efektivitas distraktor (butir pengecoh) dengan program komputer. Tingkat kesulitan suatu butir soal dapat dinyatakan dengan sebuah indeks yang berkisar antara 0,00 - 1,00. Indeks 0,00 berarti butir soal yang disusun sangat sulit, karena tidak

seorang peserta didik pun yang dapat menjawabnya dengan benar. Sebaliknya, indeks 1,00 berarti butir soal yang disusun sangat mudah, karena semua peserta didik dapat menjawabnya dengan benar. Selanjutnya, Oller (via Nurgiyantoro, 2017) mengemukakan bahwa semua butir soal dinyatakan layak jika indeks tingkat kesulitannya berkisar antara 0,15 - 0,85.

## METODOLOGI

Metode yang digunakan dalam penelitian ini yaitu metode deskriptif kuantitatif. Penelitian deskriptif menurut Nawawi (via Rotama, et al, 2020) adalah suatu penelitian yang berusaha menyelidiki atau mencari tahu permasalahan penelitian, baik dengan upaya mendeskripsikan keadaan subjek maupun objek dalam penelitian. Deskriptif kuantitatif dalam penelitian ini digunakan untuk menguraikan hasil pemerolehan skor pada program Iteman versi 4.3. Sumber data yang digunakan yakni kisi-kisi, lembar soal, kunci jawaban, dan lembar jawaban peserta didik. Selanjutnya, dilakukan analisis terhadap dokumen-dokumen tersebut untuk dilihat dan ditinjau validitas isi, reliabilitas, tingkat kesulitan dan daya beda, serta efektivitas distraktor (butir pengecoh) menggunakan program Iteman versi 4.3.

Analisis data dilakukan terhadap 40 butir soal ulangan mata pelajaran Bahasa Indonesia kelas VII SMP Negeri 11 Ruteng Kabupaten Manggarai tahun ajaran 2022/2023. Beberapa tahapan yang dilakukan dalam penelitian ini yakni pertama, peneliti memberikan angket kepada guru mata pelajaran Bahasa Indonesia di sekolah guna menelaah kesesuaian kisi-kisi soal dengan lembar soal yang telah disusun oleh peneliti. Setelah ditelaah, dilakukan perbaikan berdasarkan masukan penelaah, lalu soal dibagikan kepada peserta didik untuk dikerjakan. Tahapan berikutnya yaitu mengumpulkan lembar jawaban peserta didik. Lembar jawaban tersebut selanjutnya akan dijadikan sebagai bahan analisis. Setelah proses di atas, dilakukan analisis reliabilitas soal yang diperoleh dari hasil output analisis program Iteman versi 4.3 pada bagian Alpha Cronbach. Program Iteman ini memberikan data-data statistik yang diperoleh dari data yang dimasukkan. Kemudian, Indeks Tingkat Kesulitan (ITK) soal didapati melalui bagian *proportional correct* sebagai hasil dari perhitungan program Iteman versi 4.3. Indeks Daya Beda (IDB) soal diketahui melalui bagian *point biserial*, yaitu bagian *biser* dan *point biser* sebagai hasil perhitungan program Iteman versi 4.3, dan efektivitas distraktor (butir pengecoh) dapat dilihat pada skor *proportional endorsing* dalam program Iteman.

Setelah semua data-data dalam penelitian ini sudah dimasukkan ke dalam excel, seperti hasil jawaban siswa dan kunci jawaban soal, adapun tahapan yang dilakukan dalam menganalisis butir soal menggunakan program tersebut yakni tahapan input data yang dilakukan dengan klik program Iteman 4.3 yang sudah diunduh. Lalu pada windows *Files*, pilih *Data matrix file* yang kemudian diisi dengan data hasil jawaban siswa yang telah diubah formatnya menjadi *Text (Tab delimited)*. Selanjutnya pada bagian *Item control file* diisi dengan kunci jawaban soal terkait, sedangkan pada bagian *Output file* ditulis nama file, yang akan muncul dalam bentuk excel dan word. Selanjutnya pada bagian *Run title* juga ditulis dengan nama file. Berikutnya pada windows *Input Format (Fixed Width*

Data) bagian *Number of examinee ID columns* diisi dengan jumlah ID yakni 6. Kemudian pada bagian *Examinee IDs begin in column* diisi angka 1 (artinya dimulai dari kolom ke-1). Selanjutnya pada bagian *Item responses begin in column* diisi angka 7 (artinya respon dimulai dari kolom ke-7). Selanjutnya berpindah ke bagian *Delimited Data* dengan klik centang pada bagian *The data matrix file is delimited by a*, lalu klik *Tab*. Kemudian berpindah ke windows *Scoring Options* dan klik pada bagian *Total Score, Each domain separately*, dan *Standardized, with a mean of* yang secara otomatis akan muncul tanda centang. Lalu tahap terakhir, pada bagian *Output Options* klik *Run* lalu pilih *yes* dan data selesai dianalisis serta secara otomatis tersimpan di folder tersendiri.

## HASIL PENELITIAN DAN PEMBAHASAN

### Validitas Kisi-Kisi Soal dan Butir Soal

Proses telaah kisi-kisi dan butir soal merupakan salah satu langkah yang mencerminkan kevalidan butir soal yang akan diujikan kepada peserta didik. Dalam mencapai tujuan pengukuran yang dikehendaki, perlu dilakukan validitas terhadap alat ukur. Adapun jenis validitas yang digunakan dalam analisis ini yaitu validitas isi.

Secara garis besar, telaah tersebut mesti memenuhi kaidah penulisan soal pilihan ganda yang terdiri dari tiga aspek yaitu materi, konstruksi, dan bahasa yang di dalamnya berisi pernyataan-pernyataan. Berikut hasil validitas isi terhadap kisi-kisi dan soal yang ditelaah oleh Oktavianus Agung, S.Pd selaku guru mata pelajaran Bahasa Indonesia di SMP Negeri 11 Ruteng.

Tabel 1. Hasil Validitas Soal Nomor 1-40 Ditinjau dari Validitas Isi

Aspek	Jenis Pernyataan	Nomor Butir			
		1-10	11-20	21-30	31-40
<b>Materi</b>	Kesesuaian butir soal dengan indikator	√	√	√	√
	Isi materinya benar secara keilmuan	√	√	√	√
	Hanya terdapat satu kunci jawaban benar	√	√	√	√
	Kesesuaian isi materi dengan jenjang pendidikan/kelas	√	√	√	√
	Butir pengecohnya berfungsi	√	√	√	√
<b>Konstruksi</b>	Pokok soalnya dirumuskan dengan jelas	√	√	√	√
	Pokok soalnya tidak mengarah ke jawaban benar	√	√	√	√
	Pilihan jawabannya dirumuskan dengan jelas	√	√	√	√
	Pilihan jawabannya homogen	√	√	√	√
	Soal tidak ada bentuk negatif ganda	√	√	√	√
	Panjang pilihan jawabannya kurang lebih sama	√	√	√	√

	Antarbutir soalnya tidak bergantung satu sama lain	√	√	√	√
	Pilihan berbentuk angka/waktu diurutkan	√	√	√	√
<b>Bahasa</b>	Bahasanya komunikatif	√	√	√	√
	Kalimatnya ramatikal	√	√	√	√
	Kalimat yang digunakan tidak bermakna ganda	√	√	√	√
	Soal menggunakan kosa kata yang baku/umum/netral	√	√	√	√

Berdasarkan hasil validitas isi, ke-40 butir soal berupa pilihan ganda yang disusun merupakan soal yang valid. Hal tersebut terlihat pada terpenuhinya ketiga aspek, yakni aspek materi, konstruksi dan bahasa pada setiap soal yang disusun. Hal ini sejalan dengan pendapat Tukman (via Nurgiyantoro, 2017) yang mengemukakan bahwa validitas isi menunjuk pada pemahaman tentang instrumen tes yang sesuai dengan tujuan dan deskripsi bahan materi pelajaran yang telah dipelajari. Di pihak lain, guna mengukur keberhasilan peserta didik, jenis validitas isi harus terpenuhi dalam instrumen tes. Artinya, suatu tes dapat dikatakan valid apabila butir soal secara jelas disusun untuk mengukur indikator dan kompetensi dasar tertentu yang mewakili keseluruhan bahan ajar yang sudah dipelajari. Hal tersebut senada dengan temuan Pradani & Efendi (2023) bahwa seluruh soal dinyatakan valid apabila indikator soal sesuai dengan soal yang disusun. Jadi, dapat disimpulkan bahwa seluruh butir soal ulangan yang diujikan kepada 25 peserta didik kelas VII di SMPN 11 Ruteng tahun pelajaran 2022/2023 mempunyai kadar validitas yang tinggi.

### Indeks Reliabilitas

Sejumlah prosedur dalam memperkirakan tingkat reliabilitas tes sudah banyak digunakan. Jenis reliabilitas yang digunakan pada penelitian ini adalah reliabilitas Alpha Cronbach lewat program komputer IteMan versi 4.3. Program IteMan ini memberikan data-data statistik yang diperoleh dari data yang dimasukkan. Data-data statistik tersebut akan menghasilkan informasi terkait besaran koefisien reliabilitas pada Alpha Cronbach, di antaranya rata-rata skor (Mean), skor minimal (Min Score), skor maksimal (Max Score), rata-rata ITK (Mean P) untuk seluruh butir soal dan rata-rata IDB (Mean R) untuk semua butir soal. Program yang digunakan dalam analisis kualitas butir soal ini yakni IteMan versi 4.3. Di bawah ini disajikan data-data statistik yang diperoleh dari program IteMan.

Tabel 2. Hasil Analisis Reliabilitas Berdasarkan Program IteMan Versi 4.3

Score	Items	Mean	SD	Min Score	Max Score	Mean P	Mean Rpbis	
Scored Items	40	17.320	5.991	5	31	0.433	0.246	
Score	Alpha	SEM	Split-Half (Random)	Split-Half (First-Last)	Split-Half (Odd-Even)	S-B Random	S-B First-Last	S-B Odd-Even
Scored items	0.788	2.760	0.571	0.645	0.716	0.727	0.784	0.835

Berdasarkan data statistik, jumlah butir soal 40 yang diujikan kepada 25 peserta didik diperoleh koefisien Alpha Cronbach sebesar 0,788, rata-rata 17,320, rata-rata indeks tingkat kesulitan (ITK) seluruh butir soal sebesar 0,433, dan rata-rata indeks daya beda (IDB) sebesar 0,246. Nurgiyantoro (2017) mengemukakan bahwa sebuah tes dinyatakan reliabel jika besaran koefisiennya 0,60. Oleh karena itu, alat tes atau butir-butir soal yang telah diujikan kepada 25 peserta didik kelas VII SMP Negeri 11 Ruteng dapat dikatakan reliabilitas tinggi karena besaran indeks reliabilitasnya lebih dari 0,60 yakni sebesar 0,788. Besaran reliabilitas berdasarkan perhitungan program Iteman dalam penelitian ini berbeda dengan tingkat reliabilitas penelitian terdahulu yang dilakukan oleh Himawan & Nurgiyantoro (2022) dan Pradani & Efendi (2023) yang masing-masing memiliki reliabilitas sebesar 0,820 dan 0,768.

### **Indeks Tingkat Kesulitan (ITK) dan Indeks Daya Beda (IDB)**

Seperti yang sudah dijelaskan di atas bahwa analisis kualitas soal dalam penelitian ini menggunakan program Iteman versi 4.3. Dari ke-40 butir soal yang dianalisis akan diperoleh data Indeks Tingkat Kesulitan (ITK) dan Indeks Daya Beda (IDB). ITK merupakan indeks yang menunjukkan seberapa sulit atau mudahnya butir soal yang diujikan. Nurgiyantoro (2017) mengemukakan bahwa tingkat kesulitan butir soal dinyatakan dengan sebuah indeks yang berkisar antara 0,00 - 1,00. Indeks 0,00 berarti butir soal yang disusun sangat sulit, karena tidak seorang peserta didik pun yang dapat menjawabnya dengan benar. Sebaliknya, indeks 1,00 berarti butir soal yang disusun sangat mudah, karena semua peserta didik dapat menjawabnya dengan benar.

Di pihak lain, hasil IDB dapat menunjukkan besarnya daya suatu soal dalam membedakan kemampuan antara peserta tes grup tinggi (peserta tes pandai) dengan peserta tes grup rendah (peserta tes tidak pandai). Untuk keperluan pembelajaran, IDB sebesar 0,20 merupakan ketentuan indeks yang dapat dinyatakan layak. Di bawah ini disajikan hasil perhitungan ITK dan IDB tiap butir soal berdasarkan program Iteman versi 4.3.

Tabel 3. Hasil ITK dan ITB Berdasarkan Program Iteman Versi 4.3

No. Soal	ITK	IDB	Keterangan	
			ITK	IDB
1.	0,560	0,133	Sedang	Soal Tidak layak
2.	0,640	0,501	Mudah	Soal Layak
3.	0,640	0,075	Mudah	Soal Tidak layak
4.	0,120	-0,452	Sangat sulit	Soal Tidak layak
5.	0,120	-0,119	Sangat sulit	Soal Tidak layak
6.	0,400	0,410	Sulit	Soal Layak
7.	0,400	0,349	Sulit	Sedang (layak dan tidak perlu direvisi)
8.	0,320	-0,030	Sulit	Soal Tidak layak
9.	0,680	0,371	Mudah	Sedang (layak dan tidak perlu direvisi)
10.	0,240	-0,057	Sulit	Soal Tidak layak
11.	0,120	-0,098	Sangat sulit	Soal Tidak layak
12.	0,280	0,137	Sulit	Soal Tidak layak

13.	0,680	0,002	Mudah	Soal Tidak layak
14.	0,360	0,273	Sulit	Cukup baik (layak dan perlu direvisi)
15.	0,400	-0,003	Sulit	Soal Tidak layak
16.	0,600	0,280	Sedang	Cukup baik (layak dan perlu direvisi)
17.	0,480	0,555	Sedang	Soal Layak
18.	0,440	0,534	Sedang	Soal Layak
19.	0,720	0,453	Mudah	Soal Layak
20.	0,520	0,459	Sedang	Soal Layak
21.	0,720	0,520	Mudah	Soal Layak
22.	0,200	0,568	Sangat sulit	Soal Layak
23.	0,400	0,690	Sulit	Soal Layak
24.	0,280	0,137	Sulit	Soal Tidak layak
25.	0,760	0,341	Mudah	Sedang (layak dan tidak perlu direvisi)
26.	0,160	0,083	Sangat sulit	Soal Tidak layak
27.	0,640	0,269	Mudah	Cukup baik (layak dan perlu direvisi)
28.	0,560	0,264	Sedang	Cukup baik (layak dan perlu direvisi)
29.	0,400	0,564	Sulit	Soal Layak
30.	0,800	0,479	Mudah	Soal Layak
31.	0,200	0,255	Sangat sulit	Cukup baik (layak dan perlu direvisi)
32.	0,160	-0,125	Sangat sulit	Soal Tidak layak
33.	0,560	-0,093	Sedang	Soal Tidak layak
34.	0,400	0,200	Sulit	Cukup baik (layak dan perlu direvisi)
35.	0,440	0,442	Sedang	Soal Layak
36.	0,240	0,224	Sulit	Cukup baik (layak dan perlu direvisi)
37.	0,360	0,138	Sulit	Soal Tidak layak
38.	0,280	0,314	Sulit	Sedang (layak dan tidak perlu direvisi)
39.	0,440	0,427	Sedang	Soal Layak
40.	0,600	0,385	Sedang	Sedang (layak dan tidak perlu direvisi)

Berdasarkan hasil analisis di atas, terdapat 9 soal yang indeks tingkat kesulitannya berada pada level mudah yang ditunjukkan oleh nomor 2, 3, 9, 13, 19, 21, 25, 27, dan 30. Terdapat 10 soal berada pada level sedang yang ditunjukkan oleh nomor 1, 16, 17, 18, 20, 28, 33, 35, 39, dan 40. Terdapat 12 soal berada pada level sulit yang ditunjukkan oleh nomor 6, 7, 8, 10, 12, 14, 15, 23, 24, 29, 34, 36, 37, dan 38. Selanjutnya, terdapat 9 soal berada pada level sangat sulit yang ditunjukkan oleh nomor 4, 5, 11, 22, 26, 31, dan 32. Selain itu, hasil IDB ke-40 butir soal yang telah diujikan kepada peserta tes, terdapat 25 soal yang layak ditunjukkan oleh nomor 2, 6, 7, 9, 14, 16, 17, 18, 19, 20, 21, 22, 23, 25, 27, 28, 29, 30, 31, 34, 35, 36, 38, 39, dan 40. Dari 25 soal yang dinyatakan layak, terdapat 7 butir soal yang layak dan perlu direvisi (14, 16, 27, 28, 31, 34, 36) serta terdapat 5 butir soal yang dinyatakan layak dan tidak perlu direvisi (7, 9, 25, 38, 40). Sementara itu, 15 soal lainnya dinyatakan tidak layak yang ditunjukkan oleh nomor 1, 3, 4, 5, 8, 10, 11, 12, 13, 15, 24, 26, 32, 33, dan 37.

### Efektivitas Distraktor (Butir Pengecoh)

Dalam analisis butir soal lewat program komputer, Nurgiyantoro (2017) mengemukakan bahwa terdapat beberapa kriteria untuk menentukan efektivitas distraktor (butir pengecoh), yaitu pertama, semua distraktor (opsi-salah) harus ada yang memilih; kedua, jumlah pemilih opsi-salah dari grup tinggi (peserta tes pandai) tidak boleh lebih banyak daripada grup rendah (peserta tes tidak pandai); dan ketiga, jika pemilih opsi-salah hanya satu, ia termasuk grup rendah. Hal tersebut senada dengan pendapat Azwar (2019) yang menyatakan bahwa pengecoh akan berfungsi jika semua opsi dipilih oleh peserta tes. Maka dari itu, agar memperoleh opsi yang efektif, penyusunan opsi-opsi salah tersebut harus dilakukan dengan baik sehingga tidak terlihat sebagai opsi yang salah. Oleh karenanya, opsi-salah yang efektif ialah opsi yang “serupa tetapi tidak sama” dengan opsi-benar, sehingga memiliki peluang dipilih oleh peserta didik. Di bawah ini disajikan hasil analisis butir pengecoh yang dimaksud.

Tabel 4. Hasil Analisis Efektivitas Distraktor (Butir Pengecoh)

No. Soal	Opsi	Jumlah Siswa yang Menjawab	Proportional	Keterangan
1.	A	14	0,560	Kunci Jawaban
	B	4	0,160	Pengecohnya berfungsi
	C	2	0,080	Pengecohnya berfungsi
	D	5	0,200	Pengecohnya berfungsi
2.	A	5	0,200	Pengecohnya berfungsi
	B	3	0,120	Pengecohnya berfungsi
	C	1	0,040	Pengecohnya berfungsi
	D	16	0,640	Kunci jawaban
3.	A	5	0,200	Pengecohnya berfungsi
	B	1	0,040	Pengecohnya berfungsi
	C	16	0,640	Kunci jawaban
	D	3	0,120	Pengecohnya berfungsi
4.	A	3	0,120	Pengecohnya berfungsi
	B	3	0,120	Kunci Jawaban
	C	7	0,280	Pengecohnya berfungsi
	D	12	0,480	Pengecohnya berfungsi
5.	A	15	0,600	Pengecohnya berfungsi
	B	3	0,120	Pengecohnya berfungsi
	C	3	0,120	Kunci jawaban
	D	4	0,160	Pengecohnya berfungsi
6.	A	7	0,280	Pengecohnya berfungsi
	B	7	0,280	Pengecohnya berfungsi
	C	1	0,040	Pengecohnya berfungsi
	D	10	0,400	Kunci jawaban
7.	A	10	0,400	Kunci jawaban
	B	7	0,280	Pengecohnya berfungsi
	C	3	0,120	Pengecohnya berfungsi

	D	5	0,200	Pengecohnya berfungsi
8.	A	8	0,320	Kunci jawaban
	B	5	0,200	Pengecohnya berfungsi
	C	10	0,400	Pengecohnya berfungsi
	D	2	0,080	Pengecohnya berfungsi
9.	A	3	0,120	Pengecohnya berfungsi
	B	2	0,080	Pengecohnya berfungsi
	C	17	0,680	Kunci Jawaban
	D	3	0,120	Pengecohnya berfungsi
10.	A	5	0,200	Pengecohnya berfungsi
	B	6	0,240	Kunci jawaban
	C	10	0,400	Pengecohnya berfungsi
	D	4	0,160	Pengecohnya berfungsi
11.	A	5	0,200	Pengecohnya berfungsi
	B	2	0,080	Pengecohnya berfungsi
	C	15	0,600	Pengecohnya berfungsi
	D	3	0,120	Kunci jawaban
12.	A	11	0,440	Pengecohnya berfungsi
	B	7	0,280	Kunci jawaban
	C	2	0,080	Pengecohnya berfungsi
	D	5	0,200	Pengecohnya berfungsi
13.	A	3	0,120	Pengecohnya berfungsi
	B	4	0,160	Pengecohnya berfungsi
	C	17	0,680	Kunci jawaban
	D	1	0,040	Pengecohnya berfungsi
14.	A	9	0,360	Kunci jawaban
	B	9	0,360	Pengecohnya berfungsi
	C	1	0,040	Pengecohnya berfungsi
	D	6	0,240	Pengecohnya berfungsi
15.	A	6	0,240	Pengecohnya berfungsi
	B	10	0,400	Kunci jawaban
	C	4	0,160	Pengecohnya berfungsi
	D	5	0,200	Pengecohnya berfungsi
16.	A	4	0,160	Pengecohnya berfungsi
	B	3	0,120	Pengecohnya berfungsi
	C	3	0,120	Pengecohnya berfungsi
	D	15	0,600	Kunci jawaban
17.	A	4	0,160	Pengecohnya berfungsi
	B	12	0,480	Kunci jawaban
	C	4	0,160	Pengecohnya berfungsi
	D	5	0,200	Pengecohnya berfungsi
18.	A	6	0,240	Pengecohnya berfungsi
	B	6	0,240	Pengecohnya berfungsi
	C	2	0,080	Pengecohnya berfungsi
	D	11	0,440	Kunci jawaban

19.	A	3	0,120	Pengecohnya berfungsi
	B	2	0,080	Pengecohnya berfungsi
	C	18	0,720	Kunci Jawaban
	D	2	0,080	Pengecohnya berfungsi
20.	A	2	0,080	Pengecohnya berfungsi
	B	13	0,520	Kunci jawaban
	C	7	0,280	Pengecohnya berfungsi
	D	3	0,120	Pengecohnya berfungsi
21.	A	18	0,720	Kunci jawaban
	B	4	0,160	Pengecohnya berfungsi
	C	2	0,080	Pengecohnya berfungsi
	D	1	0,040	Pengecohnya berfungsi
22.	A	9	0,360	Pengecohnya berfungsi
	B	6	0,240	Pengecohnya berfungsi
	C	5	0,200	Pengecohnya berfungsi
	D	5	0,200	Kunci jawaban
23.	A	7	0,280	Pengecohnya berfungsi
	B	10	0,400	Kunci jawaban
	C	2	0,080	Pengecohnya berfungsi
	D	6	0,240	Pengecohnya berfungsi
24.	A	9	0,360	Pengecohnya berfungsi
	B	4	0,160	Pengecohnya berfungsi
	C	5	0,200	Pengecohnya berfungsi
	D	7	0,280	Kunci jawaban
25.	A	1	0,040	Pengecohnya berfungsi
	B	1	0,040	Pengecohnya berfungsi
	C	19	0,760	Kunci jawaban
	D	4	0,160	Pengecohnya berfungsi
26.	A	5	0,200	Pengecohnya berfungsi
	B	7	0,280	Pengecohnya berfungsi
	C	9	0,360	Pengecohnya berfungsi
	D	4	0,160	Kunci jawaban
27.	A	3	0,120	Pengecohnya berfungsi
	B	4	0,160	Pengecohnya berfungsi
	C	2	0,080	Pengecohnya berfungsi
	D	16	0,640	Kuncinya jawaban
28.	A	14	0,560	Kuncinya jawaban
	B	4	0,160	Pengecohnya berfungsi
	C	4	0,160	Pengecohnya berfungsi
	D	3	0,120	Pengecohnya berfungsi
29.	A	8	0,320	Pengecohnya berfungsi
	B	4	0,160	Pengecohnya berfungsi
	C	3	0,120	Pengecohnya berfungsi
	D	10	0,400	Kunci jawaban

30.	A	3	0,120	Pengecohnya berfungsi
	B	0	0,000	Pengecohnya tidak berfungsi
	C	2	0,080	Pengecohnya berfungsi
	D	20	0,800	Kunci jawaban
31.	A	11	0,440	Pengecohnya berfungsi
	B	4	0,160	Pengecohnya berfungsi
	C	5	0,200	Pengecohnya berfungsi
	D	5	0,200	Kunci jawaban
32.	A	4	0,160	Kunci jawaban
	B	10	0,400	Pengecohnya berfungsi
	C	8	0,320	Pengecohnya berfungsi
	D	3	0,120	Pengecohnya berfungsi
33.	A	4	0,160	Pengecohnya berfungsi
	B	14	0,560	Kunci jawaban
	C	4	0,160	Pengecohnya berfungsi
	D	3	0,120	Pengecohnya berfungsi
34.	A	5	0,200	Pengecohnya berfungsi
	B	10	0,400	Kunci jawaban
	C	6	0,240	Pengecohnya berfungsi
	D	4	0,160	Pengecohnya berfungsi
35.	A	11	0,440	Kunci jawaban
	B	5	0,200	Pengecohnya berfungsi
	C	7	0,280	Pengecohnya berfungsi
	D	2	0,080	Pengecohnya berfungsi
36.	A	10	0,400	Pengecohnya berfungsi
	B	4	0,160	Pengecohnya berfungsi
	C	6	0,240	Kunci jawaban
	D	5	0,200	Pengecohnya berfungsi
37.	A	8	0,320	Pengecohnya berfungsi
	B	9	0,360	Kunci jawaban
	C	5	0,200	Pengecohnya berfungsi
	D	3	0,120	Pengecohnya berfungsi
38.	A	7	0,280	Kunci jawaban
	B	7	0,280	Pengecohnya berfungsi
	C	6	0,240	Pengecohnya berfungsi
	D	5	0,200	Pengecohnya berfungsi
39.	A	3	0,120	Pengecohnya berfungsi
	B	4	0,160	Pengecohnya berfungsi
	C	7	0,280	Pengecohnya berfungsi
	D	11	0,440	Kunci jawaban
40.	A	15	0,600	Kunci jawaban
	B	1	0,040	Pengecohnya berfungsi
	C	6	0,240	Pengecohnya berfungsi
	D	3	0,120	Pengecohnya berfungsi

Berdasarkan tabel di atas, analisis butir pengecoh ke-40 soal ulangan mata pelajaran Bahasa Indonesia kelas VII SMP Negeri 11 Ruteng Kabupaten Manggarai tahun ajaran 2022/2023 menunjukkan terdapat 39 butir soal pengecohnya berfungsi yang ditunjukkan oleh nomor 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 31, 32, 33, 34, 35, 36, 37, 38, 39, dan 40. Selain itu, terdapat 1 opsi soal pengecohnya tidak berfungsi. Butir soal yang pengecohnya tidak berfungsi yaitu soal nomor 30 opsi B. Hasil tersebut sejalan dengan temuan dan pendapat Himawan & Nurgiyantoro (2022) bahwa pengecoh dikatakan berfungsi apabila semua atau tiga pengecohnya berfungsi dengan baik, sedangkan pengecoh dikatakan tidak berfungsi apabila salah satunya tidak berfungsi. Keberfungsian pengecoh dalam penelitian ini dapat dilihat pada nilai yang ditunjukkan pada *proportional endorsing*. Butir soal yang pengecohnya berfungsi menurut Rotama, et al (2020) akan ditandai dengan dipilih setidaknya 5% oleh peserta tes/peserta didik.

Di bawah ini disajikan contoh soal dan hasil perhitungan program Iteman versi 4.3.

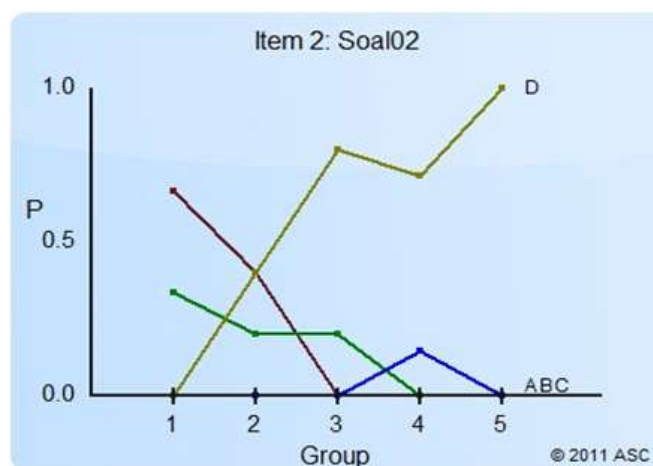
## 02. Bacalah teks-teks di bawah ini!

- |   |   |
|---|---|
| <p>(1) Jika kamu ke pantai,<br/>Jangan lupa membawa baju;<br/>Jika kamu ingin pandai,<br/>Ranjin-rajinlah membaca<br/>buku.</p> | <p>(3) Janganlah engkau berbuat maksiat<br/>Janganlah engkau berbuat jahat<br/>Segeralah engkau bertobat<br/>Agar selamat dunia akhirat</p> |
| <p>(2) Siapa menggemari silang<br/>sengketa,<br/>Kelaknya pasti berduka cita.</p>   | <p>(4) Barang siapa meninggalkan zakat,<br/>Tiadalah hartanya yang beroleh<br/>berkat.</p>  |

Puisi yang termasuk gurindam ditandai oleh nomor ....

- A. (1) dan (2)      B. (1) dan (3)      C. (2) dan (3)      D. (2) dan (4)

Dari soal di atas, diperoleh skor ITK, ITB, dan efektivitas distraktor (butir pengecoh). Berikut grafik hasil perhitungan soal yang dimaksud.



Gambar 1. Hasil Perhitungan Soal

Tabel 5. Information Item

Seq.	ID	Key	Scored	Num Options	Domain	Flags
2	Soal02	D	Yes	4	1	

Tabel 6. Statistics Item

N	P	Total Rpbis	Total Rbis	Alpha w/o
25	0.640	0.501	0.643	0.774

Tabel 7. Option Statistics

Option	N	Prop.	Rpbis	Rbis	Mean	SD	Warna	
A	5	0.200	-0.425	-0.607	12.000	5.745	Merah tua	
B	3	0.120	-0.269	-0.437	12.667	4.726	Hijau	
C	1	0.040	0.086	0.196	19.000	0.000	Biru	
D	16	0.640	0.501	0.643	19.750	5.145	Kuning	**KEY**
Omit	0							
Not Admin	0							

Tabel 8. Quantile Plot Data

Option	N	0-20%	20-40%	40-60%	60-80%	80-100%	Warna	
A	5	0.667	0.400	0.000	0.143	0.000	Merah tua	
B	3	0.333	0.200	0.200	0.000	0.000	Hijau	
C	1	0.000	0.000	0.000	0.143	0.000	Biru	
D	16	0.000	0.400	0.800	0.714	1.000	Kuning	**KEY**

Berdasarkan hasil perhitungan program Iteman versi 4.3 di atas, pada tabel *item statistics* dapat diketahui bahwa Indeks Tingkat Kesulitan soal tersebut sebesar 0,640 atau soal nomor 02 tersebut berkategori mudah. Sementara itu, Indeks Daya Bedanya sebesar 0,501 atau soal tersebut dikategorikan sebagai soal yang layak. Di pihak lain, pada tabel *option statistics* diketahui bahwa semua distraktor dipilih oleh peserta didik. Sebanyak 5 peserta didik yang memilih opsi A, sebanyak 3 peserta didik yang memilih opsi B, sebanyak 1 orang yang memilih opsi C, dan sebanyak 16 orang yang memilih opsi D. Hasil tersebut sejalan dengan salah satu kriteria untuk menetapkan efektivitas distraktor menurut Nurgiyantoro (2017), yaitu semua distraktor (opsi-salah) harus ada yang memilih. Dengan demikian, semua opsi pilihan ganda dalam soal tersebut dapat dikatakan pengecohnya berfungsi.

## KESIMPULAN DAN REKOMENDASI

Berdasarkan penjelasan di atas, kualitas ke-40 butir soal yang diujikan kepada 25 peserta tes/peserta didik merupakan soal yang valid. Selanjutnya, berdasarkan data yang diperoleh melalui program IteMan versi 4.3, jumlah 40 soal yang diujikan kepada 25 peserta didik diperoleh reliabilitas Alpha Cronbach yakni sebesar 0,788. Nurgiyantoro (2017) mengemukakan bahwa sebuah tes dinyatakan reliabel jika memiliki koefisien sebesar 0,60. Dengan demikian, soal yang diujikan kepada peserta didik dinyatakan reliabel karena besaran indeks reliabilitasnya lebih dari 0,60. Selain itu, diperoleh rata-rata indeks tingkat kesulitan (ITK) seluruh butir soal sebesar 0,433, dan rata-rata indeks daya beda (IDB) sebesar 0,246. Selain itu, efektivitas distraktor menunjukkan bahwa terdapat 39 soal yang pengecohnya berfungsi dan 1 opsi soal pengecohnya tidak berfungsi.

Di pihak lain, rekomendasi kepada guru dan lembaga satuan pendidikan agar lebih memperhatikan kualitas butir soal yang akan diujikan kepada peserta didiknya di sekolah. Hal tersebut dapat dilakukan melalui telaah kesesuaian antara kisi-kisi soal dengan lembar soal. Selanjutnya, perlu dicek reliabilitasnya, tingkat kesulitan soal yang disusun, daya beda soalnya, dan efektivitas distraktor (butir pengecoh) agar bisa dijadikan sebagai bahan evaluasi untuk perubahan yang lebih baik.

## PENELITIAN LANJUTAN

Penelitian ini tentu masih memiliki banyak kekurangan, baik dalam penyusunan maupun dalam proses penelitian. Oleh karena itu, perlu dilakukan penelitian lanjutan dalam hal menganalisis butir soal guna menghasilkan soal-soal yang berkualitas dan dapat mencapai tujuan dalam hal penilaian hasil belajar peserta didik.

## UCAPAN TERIMA KASIH

Begitu banyak pihak yang telah membantu peneliti, baik dalam penyusunan maupun dalam proses penelitian. Maka dari itu, peneliti mengucapkan terima kasih kepada semua pihak yang telah membantu, khususnya kepada segenap keluarga besar SMP Negeri 11 Ruteng yang telah mengizinkan peneliti untuk melaksanakan penelitian tentang analisis kualitas soal Bahasa Indonesia.

## DAFTAR PUSTAKA

- Agustiana, M., Mayrita, H., & Muchti, A. (2018). Analisis Butir Soal Ulangan Akhir Semester Mata Pelajaran Bahasa Indonesia Kelas XI. *Jurnal Ilmiah Bina Edukasi*, 11(01), 26–35. <https://doi.org/10.33557/jedukasi.v11i01.203>
- Arikunto, S. (2018). *Dasar-Dasar Evaluasi Pendidikan, Edisi 3*. Jakarta: Bumi Aksara.
- Arvianto, I. R. (2017). Pemanfaatan Program IteMan 3.0 untuk Analisis Butir Soal Lomba Cerdas Cermat Teknologi Informasi dan Komunikasi Tingkat SMA Sederajat. *Jurnal Teknologi Informasi, 11(33)*. <https://doi.org/10.35842/jtir.v11i33.106>
- Azwar, S. (2019). *Reliabilitas dan Validitas*. Yogyakarta: Pustaka Belajar.

- Gladushyna, O., & Strietholt, R. (2023). Measuring education: Do we need a plethora of assessment studies or just a single score? *International Journal of Educational Research Open*, 5(February), 100281. <https://doi.org/10.1016/j.ijedro.2023.100281>
- Himawan, R., & Nurgiyantoro, B. (2022). Analisis Butir Soal Latihan Penilaian Akhir Semester Ganjil Mata Pelajaran Bahasa Indonesia Kelas VIII SMPN 1 Bambanglipuro Bantul Menggunakan program ITEMAN (Analysis of exercise items for odd semester end of semester Indonesian language subjects class. *KEMBARA: Jurnal Keilmuan Bahasa, Sastra, Dan Pengajarannya*, 8(1), 160–180.
- Kurniawati, E. (2020). Analisis Soal Ujian Nasional Bahasa Indonesia pada Tingkat SMP Tahun Ajaran 2018/2019 dalam Perspektif Kebijakan Pendidikan. *Jurnal Pendidikan Bahasa*, 9(2), 201. <https://doi.org/10.31571/bahasa.v9i2.1895>
- López-Hernández, A., Buckingham, L. R., & Strotmann, B. (2023). Enhancing learning-oriented assessment through co-teaching in higher education. *Studies in Educational Evaluation*, 79(October), 0–2. <https://doi.org/10.1016/j.stueduc.2023.101307>
- Nurgiyantoro, B. (2011). Model Penilaian Otentik dalam Pembelajaran Bahasa. *LITERA: Jurnal Penelitian Bahasa, Sastra, dan Pengajarannya*, 10(2), 114–125. <https://doi.org/https://doi.org/10.21831/ltr.v10i2.1157>
- Nurgiyantoro, B. (2017). *Penilaian Pembelajaran Bahasa Berbasis Kompetensi (Edisi Kedua)*. Yogyakarta: BPFE-Yogyakarta.
- Peraturan Pemerintah Republik Indonesia No 57 tahun 2014 tentang Pengembangan, Pembinaan dan Perlindungan Bahasa dan Sastra, serta Peningkatan Fungsi Bahasa Indonesia. (2014). <https://www.regulasip.id/book/354/read>
- Pradani, R. A., & Efendi, A. (2023). Analisis Butir Soal Ujian Sekolah Menggunakan Program Iteman (Analysis of School Exam Questions Using the Iteman Program). *Indonesian Language Education and Literature*, 8(2), 275. <https://doi.org/10.24235/ileal.v8i2.11002>
- Rotama, A. D., Budiutomo, T. W., & Bowo, A. N. A. (2020). Analisis Butir Soal Penilaian Tengah Semester Mata Pelajaran PPKn Kelas VII di SMP Muhammadiyah 7 Yogyakarta. *AoEJ: Academy of Education Journal*, 11(01), 24–35. <https://doi.org/10.47200/aoej.v11i01.314>
- Sarker, M. F., & Ullah, M. S. (2023). A review of quality assessment criteria in secondary education with the impact of the COVID-19 pandemic. *Social Sciences and Humanities Open*, 8(1), 100740. <https://doi.org/10.1016/j.ssaho.2023.100740>
- Utami, S. A., & Nurgiyantoro, B. (2016). Kualitas Soal dan Daya Serap Tes Pendalaman Materi UN Bahasa Indonesia SMP di Gunungkidul. *Diksi*, 24(1), 52–62. <https://doi.org/10.21831/diksi.v24i1.11499>