

Analisis butir soal latihan penilaian akhir semester ganjil mata pelajaran bahasa Indonesia kelas VIII SMPN I Bambanglipuro Bantul menggunakan program ITEMAN (Analysis of exercise items for odd semester end of semester Indonesian language subjects class VIII SMPN I Bambanglipuro Bantul using the ITEMAN program)

Riswanda Himawan^{1*}, Burhan Nurgiyantoro²

Magister Pendidikan Bahasa dan Sastra Indonesia, Fakultas Bahasa dan Seni, Universitas Negeri Yogyakarta, Karangmalang Caturtunggal, Depok, Kabupaten Sleman, Daerah Istimewa Yogyakarta, Indonesia.

¹riswandahimawan.2021@student.uny.ac.id; burhan@uny.ac.id²

*Corresponding author: riswandahimawan.2021@student.uny.ac.id

Sejarah Artikel

Diterima: 10 Maret 2022

Direvisi: 15 April 2022

Tersedia Daring: 29 April 2022

ABSTRAK

Tes merupakan instrumen yang digunakan untuk melakukan evaluasi pembelajaran kepada peserta didik sesuai dengan materi yang diajarkan. Instrumen tes yang digunakan pun harus diuji kualitasnya. Selaras dengan pernyataan tersebut. Penelitian ini bertujuan untuk mengetahui kualitas instrumen soal latihan Penilaian Akhir Semester Ganjil (Latihan PAS Ganjil) kelas VIII SMPN I Bambanglipuro Bantul tahun 2020/2021 serta mendeskripsikan hasil analisis butir soal berkaitan dengan validitas, reliabilitas, IDB, ITK dan keberfungsian pengecoh. Validitas isi dilakukan dengan memberikan angket kepada sejawat berkaitan dengan kesesuaian soal dengan ksis-kisi soal. Reliabilitas dilakukan dengan melihat skor Alpha yang diperoleh dari program ITEMAN. Analisis Indeks Daya Beda (IDB), Indeks Tingkat Kesulitan (ITK) dan keberfungsian pengecoh dihitung berdasarkan skor yang diperoleh dari program ITEMAN. Terdapat tiga tahapan yang dilakukan dalam penelitian ini. Tahap pertama yaitu memperoleh gambaran umum berkaitan dengan soal yang diujikan kepada peserta didik. Tahap kedua adalah pengumpulan data. Dalam tahapan ini dilakukan pengumpulan data dengan mendokumentasikan soal-soal yang diujikan oleh Guru kepada peserta didik. Tahap ketiga adalah pemeriksaan keabsahan data. Analisis data dalam penelitian ini dilakukan terhadap butir-butir soal Latihan Penilaian Akhir Semester Ganjil (Latihan Akhir Semester Ganjil) dengan menganalisis validitas berdasarkan komentar teman sejawat dilanjutkan dengan menganalisis tingkat kesukaran, indeks daya beda, dan keberfungsian pengecoh. Masing-masing kriteria tersebut dihitung dengan menggunakan bantuan komputer melalui program *Item and Test Analysis* (ITEMAN). Hasil penelitian menunjukkan terdapat 40 soal yang dianalisis validitas dan reliabilitasnya. Hasil validitas menunjukkan seluruh soal yang dianalisis valid. Berkaitan dengan uji reliabilitas diperoleh skor Alpha 0.820 dengan kategori sangat tinggi. Analisis butir soal berkaitan dengan ITK, IDB dan distraktor dilihat dari skor yang diperoleh berdasarkan hasil analisis ITEMAN. Berdasarkan hasil analisis ITEMAN dapat diketahui bahwa analisis IDB menunjukkan 20 soal yang layak untuk diujikan, 4 butir soal dengan kategori sedang, 8 soal dengan kategori cukup, dan 8 soal dengan kategori gugur atau tidak layak untuk diujikan. Hasil analisis soal berkaitan dengan ITK, dapat disimpulkan bahwa terdapat 6 butir soal dengan kategori ITK sangat mudah, 17 butir soal dengan ITK mudah, 8 butir soal dengan Indeks tingkat kesulitan sedang, 9 soal dengan Indeks tingkat kesulitan sukar atau tinggi. Dari hasil analisis keberfungsian pengecoh, dapat disimpulkan bahwa terdapat 31 butir soal dengan keberfungsian pengecoh baik, 9 butir soal dengan memiliki pengecoh yang tidak berfungsi. Berdasarkan hasil analisis butir soal tersebut, soal-soal yang dianggap tidak layak, kemudian direvisi dan diperbaiki. Sehingga memperoleh soal yang berkualitas dan dapat mengukur tingkat ketercapaian tujuan pembelajaran.

Kata Kunci

Analisis Butir Soal, IDB, ITK, Pengecoh

ABSTRACT

The test is an instrument used to evaluate learning to students in accordance with the material being taught. The test instrument used must also be tested for quality. In line with that statement, this study aims to determine the quality of the Odd Semester Final Assessment practice questions (PAS Odd Practice) class VIII SMPN I Bambanglipuro Bantul in 2020/2021 and to describe the item analysis results to validity, reliability, IDB, ITK and the functioning of distractors. Content validity is carried out by giving colleagues a questionnaire regarding the question's suitability with the question grid. Reliability is done by looking at the Alpha score obtained from the ITEMAN program. Analysis of the Distinctive Power Index (IDB), Difficulty Index (ITK) and distractor function were calculated based on the scores obtained from the ITEMAN program. There are three stages carried out in this research. The first stage is to get an overview of the questions tested on students. The second stage is data collection. In this stage, data collection is carried out by documenting the questions tested by the teacher to students. The third stage is checking the validity of the data. Data analysis in this study was carried out on the Odd Semester Final Assessment Exercise (Odd Semester Final Practice) items by analyzing the validity based on the comments of peers, followed by examining the level of difficulty, discriminatory index, and functioning of the distractor. These criteria are calculated using computer assistance through the *Item and Test Analysis* (ITEMAN) program. The results showed that 40 questions were analyzed for validity and reliability. The results of the validity show that all the questions analyzed are valid. Concerning the reliability test, the Alpha score of 0.820 was obtained in the very

high category. Analysis of items related to ITK, IDB, and distractor seen from the scores obtained based on the results of the ITEMAN analysis. Based on the results of the ITEMAN analysis, it can be seen that the IDB analysis shows 20 questions that are eligible to be tested. 4 items in the moderate category, 8 questions in the sufficient category, and 8 questions in the fall category or not eligible to be tested. The results of the analysis of questions related to ITK show that there are 6 questions with a very easy ITK category, 17 questions with an easy ITK, 8 questions with a medium difficulty index, 9 questions with a difficult or high difficulty index. , it can be concluded that there are 31 questions with good distractor function, there are 9 questions with distractors that do not function. Based on the results of the analysis of these items, the questions that were deemed inappropriate were then revised and corrected so that it can be used as a teacher's foothold in presenting evaluation questions.



Copyright@2022, Riswanda Himawan, Burhan Nurgiyantoro
This is an open access article under the [CC-BY-3.0](https://creativecommons.org/licenses/by/3.0/) license



Keywords	Problem Item Analysis, ITEMAN, End-of-Semester Assessment
How to Cite	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. <i>KEMBARA: Jurnal Keilmuan Bahasa, Sastra, dan Pengajarannya (e-journal)</i> , 8(1) 160-180. https://doi.org/10.22219/kembara.v8i1.20530

PENDAHULUAN

Penilaian akhir semester (PAS) dapat didefinisikan sebagai salah satu kegiatan yang dilakukan pada jenjang satuan pendidikan menengah pertama dalam mengevaluasi hasil belajar peserta didik selama satu semester. Hal ini sangat berkaitan dengan kegiatan mengukur kemampuan peserta didik dalam kemampuan kognitif yang telah diajarkan guru selama proses pembelajaran (Setiawan et al., 2020). Berdasarkan pendapat tersebut. Dapat diketahui bahwa penilaian akhir semester merupakan kegiatan yang dilakukan untuk mengevaluasi pembelajaran yang dilakukan guru dan peserta didik selama satu semester. Mengingat pentingnya kegiatan ini, maka perlu dilakukan persiapan yang matang. Persiapan perlu dilakukan oleh guru maupun peserta didik agar hasil penilaian nantinya dapat digunakan sebagai referensi untuk memperbaiki pembelajaran yang akan dilakukan berikutnya.

Penilaian Akhir Semester masuk pada bagian penilaian pembelajaran (Fernanda & Hidayah, 2020), penilaian pembelajaran sendiri dapat diartikan sebagai proses pengumpulan dan pengolahan informasi untuk mengetahui ketercapaian hasil dari proses pembelajaran, penilaian tersebut dilakukan oleh guru dengan memperhatikan berbagai macam aspek, yakni aspek sikap, pengetahuan, dan keterampilan (Setiawati & Arsana, 2018). Penilaian pendidikan dapat dilakukan melalui tiga cara yakni; (1) penilaian oleh pendidik, (2) penilaian oleh satuan pendidikan (sekolah), dan (3) penilaian oleh pemerintah (Sulistiawan, 2016). Penilaian pembelajaran dapat dilakukan dengan menghimpun fakta serta dokumen belajar peserta didik yang dapat dipercaya (Anggraini & Suyata, 2014).

Penilaian pembelajaran atau evaluasi pembelajaran merupakan perihal penting yang harus dilakukan guru dalam setiap proses pembelajaran (Ramadhanti, 2019). Guru merupakan evaluator pertama dan terutama dalam dan menentukan keberhasilan Pendidikan (Suyata et al., 2011). Evaluasi pembelajaran sangat erat kaitannya dengan kurikulum pembelajaran (Mardiana & Suyata, 2017). Kegiatan evaluasi pembelajaran telah diatur dalam Undang-Undang Republik Indonesia No. 20 Tahun 2003 (melalui Kurniawan, 2015) mengenai Sistem Pendidikan Nasional Bab XVI Pasal 58 Ayat 1. Undang-undang tersebut menyatakan bahwa "Evaluasi hasil belajar peserta didik dilakukan oleh guru untuk memantau proses, kemajuan, dan perbaikan hasil belajar peserta didik secara berkesinambungan". Berdasarkan uraian tersebut, dapat disimpulkan bahwa evaluasi pembelajaran dilakukan oleh guru sebagai fasilitator pembelajaran dalam rangka memantau hasil belajar peserta didik.

"Evaluasi merupakan salah satu komponen penting dan tahap yang harus ditempuh oleh guru sebagai pijakan untuk mengetahui keefektifan pembelajaran" (Oktanin, et al 2015). Menurut Arifin (2019:15) tujuan dari penilaian hasil belajar peserta didik dapat dijelaskan sebagai berikut; (a) penilaian



pembelajaran bertujuan untuk mengetahui tingkatan penguasaan materi peserta didik terhadap materi yang telah diberikan dan diajarkan di kelas; (b) penilaian pembelajaran bertujuan untuk mengetahui kecakapan, minat, bakat, motivasi, peserta didik terhadap pembelajaran yang dilakukan; (c) penilaian pembelajaran bertujuan untuk mengetahui kesesuaian hasil belajar yang dihasilkan oleh peserta didik terhadap kompetensi dan indikator ketercapaian materi yang telah ditetapkan; (d) penilaian pembelajaran bertujuan untuk memberikan suatu diagnosis terhadap keunggulan dan kelemahan peserta didik dalam mengikuti proses pembelajaran yang telah dilakukan; (e) penilaian pembelajaran dilakukan untuk menyeleksi berkaitan dengan memilih dan menentukan peserta didik sesuai dengan jenis pendidikan tertentu; (f) penilaian pembelajaran bertujuan untuk menentukan peserta didik dalam kenaikan kelas, kelulusan berkaitan dengan studi jenjang berikutnya; (g) penilaian pembelajaran bertujuan untuk menempatkan peserta didik terhadap bidang atau kompetensi keahlian yang sesuai dengan diri peserta didik itu sendiri. Setiap jenjang sekolah perlu melakukan evaluasi pembelajaran, agar tujuan pembelajaran dapat tercapai (Retnawati et al., 2017).

Evaluasi merupakan komponen yang tidak dapat dipisahkan dari kegiatan belajar mengajar (Pangesti et al., 2020). Terdapat dua Teknik, yang dapat digunakan dalam melakukan evaluasi pembelajaran, teknik tersebut adalah tes dan non tes. Berdasarkan pendapat Susanto et al., (2015) tes sangat erat kaitannya dengan pertanyaan yang harus dijawab dan membuktikan kemampuan berpikir seseorang. Tes dilakukan untuk mengetahui penguasaan materi peserta didik dalam memahami materi, yang telah diajarkan guru. Tes yang baik merupakan tes yang memiliki kualitas soal yang baik Fernanda & Hidayah (via Pangesti et al., 2020).

Pendapat lain, dikemukakan oleh Sudijono (melalui Kurniawan, 2015), yang menyatakan bahwa tes dapat didefinisikan sebagai cara dalam mengukur dan menilai pada bidang pendidikan, tes dapat dilakukan dalam bentuk pemberian tugas atau serangkaian tugas yang harus dikerjakan oleh peserta didik, sehingga dapat diketahui nilai prestasi peserta didik. Nilai tersebut kemudian dibandingkan dengan nilai-nilai yang dicapai. Cronbach (2004:26) menyatakan bahwa tes dapat didefinisikan sebagai suatu prosedur atau tahapan yang sistematis untuk mengamati atau mendeskripsikan satu atau lebih karakteristik seseorang ditandai dengan standar numerik. Kualitas tes merupakan salah satu alat evaluasi yang penting untuk dilakukan dan diperhatikan. Kualitas tes dapat dijadikan pedoman dalam melakukan penilaian, dan menetapkan nilai kepada peserta didik (Anas Sudijono via Susanto et al., 2015). Thoha (melalui Guntur, 2017) berpendapat bahwa tes merupakan pertanyaan yang harus dijawab dan atau perintah-perintah yang harus dijalankan. Berdasarkan pendapat tersebut, maka dapat disimpulkan bahwa tes merupakan sebuah pertanyaan yang harus dijawab oleh penjawab suatu tes.

Arikunto (2016) menyatakan bahwa, tes dikatakan baik apabila memenuhi syarat validitas, reliabilitas, objektivitas, praktikabilitas, dan ekonomis. Instrumen tes harus mampu memberikan pengalaman kepada peserta didik untuk berpikir kritis. Hal ini selaras dengan pendapat yang menyatakan bahwa *critical thinking is one of the 21st century skill components. Critical thinking is one of the abilities of a person to interpret and analyze a thing carefully* (Safi'i et al., 2021). Pendapat tersebut memberikan pengertian bahwa, keterampilan untuk berpikir kritis merupakan keterampilan yang dituntut dalam kecakapan abad 21, seperti yang terjadi saat ini. Guru dapat memberikan pengalaman kepada peserta didik, menuju hal itu melalui soal evaluasi. *The evaluation instrument on determining the credibility of a source directs students' critical thinking attitudes so that in receiving information it is not easy to believe, but must be questioned and proven about its correctness from several points of view* by (Chang, Zhang, & Gwizdka via Safi'i et al., 2021). Pendapat tersebut, memberikan penguatan bahwa instrumen evaluasi harus memberikan pengalaman kepada peserta didik, berkaitan dengan berpikir kritis.

Pentingnya penilaian dalam evaluasi pembelajaran menuntut para guru memiliki kemampuan dalam melakukan penilaian (Pujiastuti & Kulup, 2021). Guru dapat menggunakan berbagai inovasi dalam pembelajaran, terlebih dalam evaluasi pengajaran (Hardiana et al., 2018). Salah satu upaya untuk menuju kepada hal tersebut adalah memperoleh soal atau instrumen tes yang bermutu dengan menganalisis butir soal. Selain bertujuan untuk memperoleh soal yang bermutu, analisis butir soal juga dapat membantu

meningkatkan tes melalui revisi atau membuang soal yang tidak efektif, serta untuk mengetahui informasi apakah peserta didik telah menguasai materi yang diajarkan oleh guru. Analisis butir soal dapat dilakukan secara kualitatif yang berkaitan dengan isi dan bentuk soal maupun kuantitatif yang berkaitan dengan ciri-ciri statistiknya. Menurut Zainal Arifin (dalam [Amalia & Widayati, 2012](#)), instrumen butir soal yang baik memiliki delapan karakteristik yaitu valid, reliabel, relevan, representatif, praktis, diskriminatif, spesifik, dan proporsional. Dalam mengembangkan instrumen tes, terdapat beberapa tahapan yang harus dilalui, yaitu; (1) merancang tes, (2) melakukan uji coba tes (3) validitas, (4) reliabilitas, (5) menginterpretasikan skor tes ([Mardapi & Kartowagiran, 2011](#)).

Sumadi Suryabrata (via [Hanifah, 2014](#)) berpendapat mengenai enam persyaratan tes yang baik yaitu (a) tes yang baik harus reliabel, (b) tes yang baik harus valid, (c) tes yang baik harus objektif, (d) tes yang baik harus diskriminatif, (e) tes yang baik harus komprehensif, (f) tes yang baik harus mudah digunakan. Instrumen tes dikatakan baik untuk digunakan sebagai alat evaluasi pembelajaran apabila memiliki validitas dan reliabilitas yang tinggi ([Purniasari et al., 2021](#)).

Berdasarkan pendapat tersebut, maka dapat diketahui bahwa soal tes yang baik, adalah soal yang memiliki tingkat validitas dan reliabilitas tinggi. Validitas, berasal dari kata "*validity*" yang memiliki arti ketepatan dan kecermatan. Dalam bahasa Arab, kata Sahih berarti tepat, benar, sehingga validitas dapat diartikan sebagai ketepatan, kesahihan, atau kebenaran ([Syarifudin, 2020](#)). Validitas juga dapat diartikan sebagai suatu indeks yang memberikan pernyataan bahwa alat ukur itu benar-benar dapat mengukur apa yang akan diukur. Sesuai dengan tujuannya, validitas merupakan suatu ukuran yang menunjukkan tingkat kevalidan atau kesahihan suatu instrumen. Prinsip validitas adalah pengukuran atau pengamatan yang berarti mengutamakan prinsip keandalan instrumen dalam mengumpulkan data. Instrumen harus dapat mengukur dengan baik apa yang seharusnya diukur. Jadi, validitas lebih menekankan pada alat pengukuran atau pengamatannya.

Sudjana (via [Guntur, 2017](#)) menyatakan bahwa validitas merupakan perihal yang berkaitan dengan kebenaran alat penilaian terhadap apa yang akan dinilai, sehingga benar-benar menilai apa yang sudah menjadi keharusan untuk dinilai. Dari pendapat tersebut, dapat kita ketahui bahwa validitas merupakan kebenaran yang sangat berkaitan dengan kebenaran alat penilaian itu sendiri.

Validitas menunjukkan perkiraan terbaik yang ada terhadap kebenaran atau kesalahan proposisi, termasuk proposisi mengenai sebab. Konsep validitas menggunakan kata "kira-kira" (approximately) atau "secara tentatif" oleh karena siapa pun tidak pernah tahu apa yang benar ([Hyman, 1982](#)). Apabila suatu instrumen memiliki kevalidan yang tinggi jika dapat menjalankan fungsi sebagai alat ukurannya, namun jika instrumen memiliki tingkat validitas yang rendah, maka akan menghasilkan data yang tidak selaras dengan tujuan pengukuran ([Arifin & Retnawati, 2017](#)).

Validitas berkaitan dengan kelayakan penafsiran penggunaan khusus skor hasil tes, sedangkan validasi merupakan pengumpulan bukti-bukti konkret untuk menunjukkan dasar saintifik penafsiran skor seperti yang telah direncanakan. Terdapat tiga pendekatan validasi yang banyak dipakai, yaitu (a) bukti berdasarkan isi, (b) bukti berdasarkan kriteria dan, (c) bukti berdasarkan konstruk Gronlund (via [Nurgiyantoro 2016:172](#)) dan Popham (melalui [Nurgiyantoro 2016:172](#)).

Reliabilitas tes menunjuk pada pengertian apakah suatu tes dapat mengukur secara konsisten sesuatu yang akan diukur dari waktu ke waktu. Hasil pengukuran dapat dipercaya hanya apabila dalam beberapa kali pelaksanaan pengukuran terhadap kelompok subjek yang sama diperoleh hasil yang relatif sama, selama aspek yang diukur dalam diri subjek memang belum berubah. Pengertian reliabilitas alat ukur dan reliabilitas hasil ukur biasanya dianggap sama. Joni (via [Hanifah, 2014](#)) mengemukakan bahwa, reliabilitas tes memberi nilai yang sama kepada individu pada tiap hasil tes.

Fernandes (melalui [Nurgiyantoro, 2016](#)) membedakan jenis reliabilitas tes ke dalam tiga hal, yaitu; (a) jenis konsistensi internal yang terdiri dari; (i) koefisien Alpha, (ii) Kuder-Rechardson 20, (iii) Kuder-Richardson 21, (iii) belah-dua spearman-bold. Berikutnya (b) stabilitas atau teknik ulang uji dan (c) equivalensi atau bentuk paralel. Uji reliabilitas menggunakan rumus Alpha Cronbach, klasifikasi



reliabilitas: 0,00-0,20 (sangat rendah); 0,21-0,40 (rendah); 0,41-0,70 (sedang); 0,71-0,90 (tinggi); 0,91-1,00 (sangat tinggi) (Arikunto, 2016).

Analisis butir soal adalah suatu prosedur sistematis, yang akan memberikan informasi-informasi yang sangat khusus terhadap butir tes yang kita susun. Analisis butir soal tes dilakukan untuk mendapatkan informasi penting yang berguna untuk evaluasi hasil pembelajaran siswa (Daryanto, via Himawan & Suyata, 2021). Sementara itu, analisis butir soal atau analisis item juga diartikan sebagai pengkajian pertanyaan-pertanyaan tes agar diperoleh perangkat pertanyaan yang memiliki kualitas yang memadai (Sudjana melalui Himawan & Suyata, 2022). Dari pengertian analisis butir soal menurut beberapa para ahli disimpulkan bahwa analisis butir soal adalah pengkajian pernyataan-pernyataan tes, sehingga diperoleh informasi-informasi yang khusus atau terperinci dari butir tes yang kita susun. Tujuan dari analisis butir soal adalah untuk mengetahui informasi tiap butir soal untuk nantinya menjadi evaluasi hasil pembelajaran.

Menurut Fitriyanawati (via Mania et al., 2020) manfaat analisis butir soal yaitu memahami soal yang layak dan tidak layak. Analisis butir soal dapat dilakukan dengan mudah dan benar melalui sebuah program komputer. Langkah-langkah analisis butir soal, menurut (Nurgiyantoro, 2016) adalah sebagai berikut; (1) mengurutkan skor pada lembar jawaban peserta didik dari skor yang tertinggi sampai yang terbawah; (2) mengambil sebanyak 27,5% dari jumlah peserta didik, antara skor yang tertinggi dan 27,5 dari skor terendah; (c) menganalisis jawaban benar atau salah tiap butir soal.

Berkaitan dengan analisis butir soal, soal evaluasi dapat dikatakan layak apabila indeks tingkat kesulitan butir soal (ITK) dapat ditoleransi dan indeks daya beda (IDB) memenuhi persyaratan. Indeks tingkat kesulitan butir soal (ITK) adalah indeks yang menunjukkan seberapa mudah atau seberapa sulit sebuah butir soal bagi peserta tes yang diuji, sedangkan Indeks daya beda (IDB) adalah suatu pernyataan tentang seberapa besar daripada sebuah butir soal dapat membedakan kemampuan antara peserta kelompok tinggi dan kelompok rendah (Nurgiyantoro, 2016: 218-221). Menurut Charles D. Hopkins dan Richard L. Antes (via Hanifah, 2014) menyatakan bahwa indeks daya beda (IDB) mengukur soal dalam membedakan tingkat kemampuan peserta didik. Indeks tingkat kesulitan butir soal (ITK) dapat ditoleransi jika berkisar 0,20-0,80, yaitu dengan angka indeks 0,20-0,40 kategori sulit, 0,41-0,60 kategori sedang, dan 0,61-0,80 kategori mudah. Indeks daya beda (IDB) dapat dinyatakan memenuhi syarat layak apabila indeks $\geq 0,20$.

Telaah mengenai Indeks Daya Beda (IDB), Indeks Tingkat Kesukaran (ITK) dan Keberfungsian Pengecoh atau Distraktor. Sangat penting dilakukan. Hal tersebut merupakan perihal penting yang harus diketahui guru, dalam rangka menganalisis butir soal. Analisis butir soal, dapat dikategorikan sebagai identifikasi jawaban benar dan salah, tiap butir soal yang diujikan kepada peserta didik (Nurgiyantoro, 2016).

Dalam penelitian ini, analisis butir soal dilakukan menggunakan program ITEMAN. ITEMAN merupakan program komputer yang digunakan untuk menganalisis butir soal secara klasik. ITEMAN merupakan perkembangan teknologi dalam lingkup penilaian. Hal ini selaras dengan pernyataan yang menyatakan bahwa teknologi digital dapat digunakan sebagai sarana untuk menunjang aktivitas pembelajaran (Yot-Domínguez & Marcelo, 2017). Program ini dapat digunakan untuk menganalisis data file jawaban butir soal yang dihasilkan melalui manual *entry* data, menskor dan menganalisis data soal pilihan ganda dan skala Likert untuk 30.000 siswa dan 250 butir soal, menganalisis sebuah tes yang terdiri dari 10 skala dan memberikan informasi tentang validitas setiap butir, standar *error of measurement*, *mean*, *variance*, standar deviasi, *skew*, kurtosis untuk jumlah skor pada jawaban benar, skor minimum dan maksimum, skor median, dan frekuensi distribusi skor. ITEMAN merupakan program komputer yang digunakan untuk menganalisis butir soal secara klasik. Program ini termasuk satu paket program dalam MicroCATn yang dikembangkan oleh *Assessment Systems Corporation* pada tahun 1982 (Nuryanti et al., 2018).

Nurgiyantoro (2016: 233-234) menjelaskan hasil analisis ITEMAN terdiri dari *Item Statistics* dan *Alternative Statistics* sebagai berikut. *Item Statistics* (statistik untuk butir) terdiri dari; *Seq No.*

(nomor urut) sesuai dengan urutan pemasukan data, *Scale Item* yaitu nomor urut butir soal, *Pop. Correct* (proporsi jawaban betul/tingkat kesulitan) berisi indeks proporsi jawaban betul per butir soal yang menunjukkan indeks kesulitan butir soal. *Biser* yaitu korelasi biserial antara jawaban benar per butir dengan skor jawaban betul. *Point Biser* yaitu korelasi point biserial antara jawaban benar per butir dengan total skor. Koefisien korelasi inilah yang dinyatakan sebagai indeks daya beda (IDB) (Nurgiyantoro, 2016). *Alternative Statistics* (statistik untuk opsi) meliputi; *Alt* yaitu opsi atau pilihan jawaban. *Prop. Endorsing* yaitu proporsi pilihan jawaban untuk tiap opsi (distribusi jawaban). *Biser* yaitu korelasi biserial untuk tiap opsi (validitas alternative jawaban). *Point Biser* yaitu korelasi point biserial untuk tiap opsi (validitas alternatif jawaban). *Key* yaitu kunci jawaban benar dengan diberi kode asterisk *. Pada bagian akhir ITEMAN memberikan data-data statistik (*Scale Statistics*) yaitu *N Of Items* (jumlah butir yang dianalisis), *N Of Examinees* (jumlah peserta ujian), *Mean* (rata-rata hitung), *Variance* (varians), *Std. Dev* (simpangan baku), *SEM* (kesalahan baku).

Monica (2013) berpendapat bahwa terdapat beberapa tahapan untuk melakukan analisis butir soal, menggunakan program komputer ITEMAN. Pertama, tahapan input data dapat dilakukan dengan cara; *Pertama*, Buka (klik) start, pilih program, pilih Accesories dan klik Notepad, Berikutnya, menetik data fail dengan memerhatikan langkah berikut: a) Baris ke-Pertama: diisi jumlah butir soal, kode Omitted (O atau 0) untuk butir soal yang tidak terjawab, kode populasi (N), dan ketik angka untuk identitas peserta; b) Baris kedua: diisi kunci jawaban yang dapat diisi dengan pilihan opsi jawaban, seperti (A, B, C, D, E); c) Baris ketiga: diisi jumlah opsi jawaban; d) Baris keempat: berisi permintaan analisis untuk setiap butir soal Y untuk dianalisis dan N yang tidak dianalisis; e) Baris kelima diisi data identitas dan hasil jawaban siswa. Berikutnya, jangan lupa untuk menyimpan hasil input data ke dalam satu folder dengan program ITEMAN.

Langkah berikutnya, pada bagian menganalisis dapat dilakukan melalui tahapan; Pertama, buka kembali program ITEMAN dengan cara mengklik sebanyak dua kali, berikutnya; Isikan fail data yang akan dianalisis, kemudian tekan enter pada komputer; Berikutnya ketik fail hasil analisis misal kemudian tekan enter; Berikutnya, Ketik 'Y' untuk memulai kegiatan analisis, lalu enter, terakhir silakan keluar dari program analisis, dan data sudah selesai.

Berkaitan dengan penelitian mengenai Analisis Butir Soal menggunakan Program ITEMAN. Penelitian yang relevan dan lebih dahulu dilakukan oleh (Purniasari et al., 2021) dengan penelitiannya yang berjudul *Analisis butir soal ujian sekolah mata pelajaran Kimia SMA N I Kutowinangun tahun pelajaran 2019/2020 menggunakan model ITEMAN dan RASCH*. Persamaan penelitian yang dilakukan oleh Purniasari, et al (2021) dengan penelitian yang dilakukan dalam penelitian ini, yaitu sama-sama menganalisis butir soal, dengan menggunakan program ITEMAN. Perbedaannya, jika dalam penelitian Purniasari, et al (2021) menganalisis mengenai soal ujian sekolah, pada mata pelajaran Kimia, di SMAN I Kutowinangun, penelitian ini menganalisis butir soal latihan Penilaian Akhir Semester Ganjil (Latihan PAS Ganjil) di SMPN I Bambanglipuro, Bantul.

Kedua, penelitian yang relevan dan lebih dahulu dilakukan oleh Rotama et al (2020) dengan penelitiannya yang berjudul *Analisis butir soal penilaian tengah semester mata pelajaran PPKN kelas VII di SMP Muhammadiyah 7 Yogyakarta*. Persamaan penelitian ini, yaitu sama-sama menganalisis butir soal dengan teori klasik menggunakan program ITEMAN. Perbedaannya, jika dalam penelitian yang dilakukan oleh Rotama et al (2020) mengaji mengenai soal dalam mata pelajaran PPKN. Penelitian ini mengaji, butir soal dalam pembelajaran bahasa Indonesia.

Ketiga, penelitian yang relevan dan lebih dahulu dilakukan oleh (Pasi & Yusrizal, 2018) dengan penelitiannya yang berjudul *Analisis butir soal ujian bahasa Indonesia buatan guru MTsN di Kabupaten Aceh Besar*. Persamaan, penelitian yaitu sama-sama menganalisis butir soal, evaluasi dalam pembelajaran bahasa Indonesia tataran SMP/MTs. Perbedaannya, jika dalam penelitian ini mengaji soal buatan guru di MTsN Kabupaten Aceh Besar, jika penelitian yang dilakukan dalam penelitian ini, mengaji soal buatan Guru di SMPN I Bambanglipuro Bantul.



Berdasarkan uraian di atas, maka analisis butir soal perlu dilakukan. Pada dasarnya, penelitian mengenai analisis butir soal menggunakan program komputer, saat ini masih belum banyak dilakukan. Hal ini selaras dengan pendapat (Arvianto, 2016) yang menyatakan bahwa penelitian mengenai analisis butir soal, dengan menggunakan program komputer masih kurang, sehingga perlu dilakukan penelitian untuk memenuhi kebutuhan referensi, serta memunculkan penelitian yang relevan. Selaras dengan pendapat tersebut. Penelitian ini bertujuan untuk memenuhi kebutuhan referensi berkaitan dengan penelitian yang akan dilakukan selanjutnya. Kebaruan yang ditampilkan dalam penelitian ini, yaitu menyajikan sebuah analisis berkaitan dengan analisis butir soal, menggunakan teknologi, sesuai dengan perkembangan zaman. Teknologi yang dimaksud adalah program komputer adalah ITEMAN.

Penelitian ini akan membahas mengenai hasil analisis butir soal, yang dibuat oleh Guru dalam rangka melakukan latihan Penilaian Akhir Semester Ganjil (Latihan PAS) tahun Pelajaran 2020/2021 di Kelas VIII, SMPN I Bambanglipuro, Kabupaten, Bantul. Berdasarkan hasil analisis yang dilakukan berkaitan dengan guru sebagai pembuat soal evaluasi di SMPN I Bambanglipuro, selama ini, guru belum pernah melakukan analisis butir soal terhadap soal yang diujikan kepada peserta didik. Guru di SMPN I Bambanglipuro hanya melalui langkah membuat dan mengujikannya kepada peserta didik. Belum ada langkah lanjutan mengenai analisis butir secara klasik menggunakan program komputer ITEMAN. Hal tersebut menjadi dasar dilakukannya penelitian di SMPN I Bambanglipuro. Secara keseluruhan penelitian ini, akan mendeskripsikan hasil analisis butir soal, dengan teori klasik menggunakan program ITEMAN.

Berkaitan dengan uraian tersebut, maka penelitian ini memiliki kontribusi. Dari lingkup keilmuan bidang Pengajaran Bahasa dan Sastra Indonesia, hasil kajian dalam penelitian ini dapat dijadikan acuan dalam penelitian lanjutan, memberikan referensi terhadap penelitian yang relevan dan akan dilakukan mengenai topik evaluasi pembelajaran bahasa Indonesia, khususnya analisis butir soal, yang berupa telaah validitas soal, telaah reliabilitas soal, telaah Indeks Tingkat Kesulitan (ITK), Indeks Daya Beda (IDB), dan analisis butir pengecoh (distraktor). Dalam penelitian ini, akan disampaikan hasil analisis berkaitan dengan soal yang layak digunakan dan tidak layak digunakan. Sehingga, memunculkan banyak penelitian mengenai analisis butir soal dengan menggunakan program komputer seperti ITEMAN yang relevan dan dapat memenuhi kebutuhan referensi bagi kaum akademisi.

METODE

Penelitian ini merupakan penelitian analisis dokumen dengan pendekatan deskriptif kuantitatif. Penelitian deskriptif dapat didefinisikan sebagai penelitian yang menyelidiki permasalahan penelitian, dengan cara dengan menggambarkan keadaan subjek ataupun objek penelitian Hadari Nawawi (melalui Rotama et al., 2020). Analisis deskriptif kuantitatif digunakan untuk mendeskripsikan hasil perolehan skor pada program ITEMAN. Sumber data yang digunakan dalam penelitian ini yaitu kisi-kisi soal yang dibuat oleh guru, berkaitan dengan soal evaluasi yang dibuat, lembar soal latihan Penilaian Akhir Semester Ganjil mata Pelajaran Bahasa Indonesia, Kelas VIII SMPN I Bambanglipuro Bantul tahun ajaran 2020/2021, yang berupa respon jawaban peserta didik. Dalam penelitian ini dilakukan analisis terhadap dokumen berupa soal Penilaian Akhir Semester Ganjil, tahun ajaran 2020/2021 mata pelajaran Bahasa Indonesia, kelas VIII SMPN I Bambanglipuro Bantul. Dengan menganalisis butir soal dilihat dan ditinjau melalui validitas, reliabilitas, tingkat kesukaran, daya pembeda dan efektivitas pengecoh dengan teori klasik dengan bantuan program komputer ITEMAN.

Terdapat beberapa tahapan, yang dilakukan dalam penelitian ini. Tahapan, pertama yaitu melakukan uji validitas, berkaitan dengan soal yang dibuat. Validitas dilakukan dengan cara menguji kesesuaian kisi-kisi dan soal evaluasi, dengan lembar telaah butir soal pilihan ganda. Validasi dilakukan oleh penulis soal kepada teman sejawat, sesama guru bidang bahasa Indonesia di MGMP Bahasa Indonesia SMP Kabupaten Bantul. Tahapan berikutnya, yaitu pengumpulan data. Pengumpulan data dilakukan dengan cara dokumentasi, yaitu mendokumentasikan hasil jawaban peserta didik, berkaitan dengan soal. Tahap ketiga adalah pemeriksaan keabsahan data. Pada tahap ini, dilakukan pengecekan dengan membandingkan data berupa dokumen yang berkaitan dengan topik penelitian. Analisis data yang

dilakukan terhadap 40 butir soal latihan Penilaian Akhir Semester Ganjil mata Pelajaran Bahasa Indonesia, Kelas VIII SMPN I Bambanglipuro Bantul tahun ajaran 2020/2021, dengan menganalisis validitas dan reliabilitas. Hasil validitas diperoleh dengan cara menganalisis lembar penilaian yang dilakukan kepada teman sejawat. Reliabilitas soal diperoleh dari hasil *output* analisis menggunakan program ITEMAN pada statistik perangkat tes bagian Alpha. Indeks Daya Beda (IDB) butir soal dilihat melalui koefisien korelasi *point biserial*, yaitu *biser* dan *point biser* yang dihitung dengan program ITEMAN. Indeks Tingkat Kesukaran (ITK) butir soal dapat diketahui melalui *proportional correct* yang dihitung dengan program ITEMAN, dan keberfungsian *distractor* (pengecoh) dapat diketahui melalui nilai *proportional endorsing* dalam program ITEMAN.

HASIL DAN PEMBAHASAN

Hasil analisis, berkaitan dengan Soal Penilaian Tengah Semester Ganjil (PTS Ganjil) tahun ajaran 2020/2021, Kelas VIII, SMPN I Bambanglipuro Bantul, adalah sebagai berikut.

Validitas

Validitas yang digunakan untuk menganalisis kualitas butir soal yang diujikan kepada peserta didik dalam soal latihan Penilaian Akhir Semester (Latihan PAS Ganjil) tahun ajaran 2020/2021, di Kelas VIII SMPN I Bambanglipuro adalah validitas isi. Validitas tersebut, menguji kesesuaian soal evaluasi yang dikembangkan dengan indikator yang telah dituliskan dalam lembar telaah soal bentuk pilihan ganda. Telaah tersebut, berisi meliputi aspek; (1) materi, (2) konstruksi dan (3) bahasa. Aspek materi berisi mengenai hal-hal yang berkaitan dengan; (a) kesesuaian butir soal dengan indikator, yang ada dalam kisi-kisi soal; (b) Kesesuaian isi materi dengan keilmuan; (c) Kunci jawaban; (d) keberfungsian opsi pengecoh dalam soal. Aspek Konstruksi, berisi mengenai; (a) kejelasan perumusan pokok-pokok soal; (b) kejelasan pilihan jawaban dalam soal; (c) pilihan jawaban homogen; (d) kepastian tidak ada bentuk negatif ganda; (e) pemilihan panjang jawaban tiap butir soal; (f) Tidak ada kebergantungan antar butir soal; (g) pengurutan pilihan jawaban yang berbentuk angka dan waktu. Aspek bahasa, meliputi; (a) bahasa soal yang komunikatif; (b) kalimat gramatikal; (c) kalimat tidak bermakna ganda; (d) pemilihan kosakata.

Aspek-aspek validitas tersebut, sesuai dengan aspek-aspek validitas isi dalam soal evaluasi yang dikemukakan oleh (Nurgiyantoro, 2016). Validitas dalam penelitian ini, dilakukan oleh guru kepada rekan sejawat, sesama guru di MGMP Bahasa Indonesia SMP Kabupaten Bantul. Hasil validasi terhadap 40 butir soal pilihan ganda yang dilakukan oleh guru, kepada rekan sejawat didapatkan hasil sebagai berikut.



Tabel I
 Tabel. Hasil Validitas Soal ditinjau Dari Validitas Isi No. I- 10

Aspek	Jenis Persyaratan	Nomor butir soal									
		1	2	3	4	5	6	7	8	9	10
Materi	1. Butir soal sesuai dengan indikator	√	√	√	√	√	√	√	√	√	√
	2. Isi materi benar secara keilmuan	√	√	√	√	√	√	√	√	√	√
	3. Hanya ada satu kunci jawaban benar	√	√	√	√	√	√	√	√	√	√
	4. Isi materi sesuai dengan kelas/jenjang pendidikan	√	√	√	√	√	√	√	√	√	√
	5. Butir pengecoh berfungsi dengan baik	√	√	√	√	√	√	√	√	√	√
Konstruksi	6. Pokok soal dirumuskan dengan jelas	√	√	√	√	√	√	√	√	√	√
	7. Pokok soal tidak mengarah ke jawaban benar	√	√	√	√	√	√	√	√	√	√
	8. Pilihan jawaban dirumuskan dengan jelas	√	√	√	√	√	√	√	√	√	√
	9. Pilihan jawaban homogen	√	√	√	√	√	√	√	√	√	√
	10. Tidak ada bentuk negatif ganda	√	√	√	√	√	√	√	√	√	√
	11. Panjang pilihan jawaban kurang lebih sama	√	√	√	√	√	√	√	√	√	√
	12. Antarbutir soal tidak bergantung satu sama lain	√	√	√	√	√	√	√	√	√	√
	13. Pilihan dalam bentuk angka/waktu diurutkan	√	√	√	√	√	√	√	√	√	√
Bahasa	14. Bahasa Komunikatif	√	√	√	√	√	√	√	√	√	√
	15. Kalimat gramatikal	√	√	√	√	√	√	√	√	√	√
	16. Kalimat tidak bermakna ganda	√	√	√	√	√	√	√	√	√	√
	17. Kosakata baku/umum/netral	√	√	√	√	√	√	√	√	√	√

Tabel 2
 Tabel. Hasil Validitas Soal ditinjau Dari Validitas Isi No. II-20

Aspek	Jenis Persyaratan	Nomor butir soal									
		11	12	13	14	15	16	17	18	19	20
Materi	1. Butir soal sesuai dengan indikator	√	√	√	√	√	√	√	√	√	√
	2. Isi materi benar secara keilmuan	√	√	√	√	√	√	√	√	√	√
	3. Hanya ada satu kunci jawaban benar	√	√	√	√	√	√	√	√	√	√
	4. Isi materi sesuai dengan kelas/jenjang pendidikan	√	√	√	√	√	√	√	√	√	√
	5. Butir pengecoh berfungsi dengan baik	√	√	√	√	√	√	√	√	√	√
Konstruksi	6. Pokok soal dirumuskan dengan jelas	√	√	√	√	√	√	√	√	√	√
	7. Pokok soal tidak mengarah ke jawaban benar	√	√	√	√	√	√	√	√	√	√
	8. Pilihan jawaban dirumuskan dengan jelas	√	√	√	√	√	√	√	√	√	√
	9. Pilihan jawaban homogen	√	√	√	√	√	√	√	√	√	√
	10. Tidak ada bentuk negatif ganda	√	√	√	√	√	√	√	√	√	√
	11. Panjang pilihan jawaban kurang lebih sama	√	√	√	√	√	√	√	√	√	√
	12. Antarbutir soal tidak bergantung satu sama lain	√	√	√	√	√	√	√	√	√	√
	13. Pilihan dalam bentuk angka/waktu diurutkan	√	√	√	√	√	√	√	√	√	√
Bahasa	14. Bahasa Komunikatif	√	√	√	√	√	√	√	√	√	√
	15. Kalimat gramatikal	√	√	√	√	√	√	√	√	√	√
	16. Kalimat tidak bermakna ganda	√	√	√	√	√	√	√	√	√	√
	17. Kosakata baku/umum/netral	√	√	√	√	√	√	√	√	√	√

Tabel 3
Tabel. Hasil Validitas Soal ditinjau Dari Validitas Isi No. 21-30

Aspek	Jenis Persyaratan	Nomor butir soal									
		21	22	23	24	25	26	27	28	29	30
Materi	1. Butir soal sesuai dengan indikator	√	√	√	√	√	√	√	√	√	√
	2. Isi materi benar secara keilmuan	√	√	√	√	√	√	√	√	√	√
	3. Hanya ada satu kunci jawaban benar	√	√	√	√	√	√	√	√	√	√
	4. Isi materi sesuai dengan kelas/jenjang pendidikan	√	√	√	√	√	√	√	√	√	√
	5. Butir pengecoh berfungsi dengan baik	√	√	√	√	√	√	√	√	√	√
Konstruksi	6. Pokok soal dirumuskan dengan jelas	√	√	√	√	√	√	√	√	√	√
	7. Pokok soal tidak mengarah ke jawaban benar	√	√	√	√	√	√	√	√	√	√
	8. Pilihan jawaban dirumuskan dengan jelas	√	√	√	√	√	√	√	√	√	√
	9. Pilihan jawaban homogen	√	√	√	√	√	√	√	√	√	√
	10. Tidak ada bentuk negatif ganda	√	√	√	√	√	√	√	√	√	√
	11. Panjang pilihan jawaban kurang lebih sama	√	√	√	√	√	√	√	√	√	√
	12. Antarbutir soal tidak bergantung satu sama lain	√	√	√	√	√	√	√	√	√	√
	13. Pilihan dalam bentuk angka/waktu diurutkan	√	√	√	√	√	√	√	√	√	√
Bahasa	14. Bahasa Komunikatif	√	√	√	√	√	√	√	√	√	√
	15. Kalimat gramatikal	√	√	√	√	√	√	√	√	√	√
	16. Kalimat tidak bermakna ganda	√	√	√	√	√	√	√	√	√	√
	17. Kosakata baku/umum/netral	√	√	√	√	√	√	√	√	√	√

Tabel 4
Tabel. Hasil Validitas Soal ditinjau Dari Validitas Isi No. 31-40

Aspek	Jenis Persyaratan	Nomor butir soal									
		31	32	33	34	35	36	37	38	39	40
Materi	1. Butir soal sesuai dengan indikator	√	√	√	√	√	√	√	√	√	√
	2. Isi materi benar secara keilmuan	√	√	√	√	√	√	√	√	√	√
	3. Hanya ada satu kunci jawaban benar	√	√	√	√	√	√	√	√	√	√
	4. Isi materi sesuai dengan kelas/jenjang pendidikan	√	√	√	√	√	√	√	√	√	√
	5. Butir pengecoh berfungsi dengan baik	√	√	√	√	√	√	√	√	√	√
Konstruksi	6. Pokok soal dirumuskan dengan jelas	√	√	√	√	√	√	√	√	√	√
	7. Pokok soal tidak mengarah ke jawaban benar	√	√	√	√	√	√	√	√	√	√
	8. Pilihan jawaban dirumuskan dengan jelas	√	√	√	√	√	√	√	√	√	√
	9. Pilihan jawaban homogen	√	√	√	√	√	√	√	√	√	√
	10. Tidak ada bentuk negatif ganda	√	√	√	√	√	√	√	√	√	√
	11. Panjang pilihan jawaban kurang lebih sama	√	√	√	√	√	√	√	√	√	√
	12. Antarbutir soal tidak bergantung satu sama lain	√	√	√	√	√	√	√	√	√	√
	13. Pilihan dalam bentuk angka/waktu diurutkan	√	√	√	√	√	√	√	√	√	√
Bahasa	14. Bahasa Komunikatif	√	√	√	√	√	√	√	√	√	√
	15. Kalimat gramatikal	√	√	√	√	√	√	√	√	√	√
	16. Kalimat tidak bermakna ganda	√	√	√	√	√	√	√	√	√	√
	17. Kosakata baku/umum/netral	√	√	√	√	√	√	√	√	√	√

Berdasarkan hasil pencermatan terhadap 40 butir soal, dapat disimpulkan bahwa seluruh soal dikategorikan sebagai soal yang valid, karena antara indikator dengan soal dapat dikatakan sesuai. Berdasarkan tabel hasil pencermatan di atas, terhadap soal yang dikembangkan. Didapatkan hasil, bahwa soal tidak valid berjumlah 0% sedangkan soal yang dianggap valid sejumlah 100%. Jadi dapat disebutkan bahwa keseluruhan butir soal, dalam Penilaian Akhir Semester (PAS Ganjil) tahun ajaran 2020/2021, di Kelas VIII SMPN 1 Bambanglipuro memiliki kadar validitas tinggi. Hal tersebut selaras dengan pendapat Thoha (via Utomo, 2019) yang menyatakan bahwa, dalam menguji validitas isi berkaitan dengan soal evaluasi yang dikembangkan, Guru dapat melakukan perbandingan antara soal dengan kisi-kisi soal atau kurikulum yang telah diajarkan, perbandingan tersebut dapat dilakukan guru kepada teman sejawat.



Reliabilitas

Reliabilitas tes merupakan suatu alat ukur yang digunakan untuk mengetahui konsistensi pengukuran tes yang hasilnya menunjukkan keajegan. Seorang dikatakan dapat dipercaya apabila orang tersebut berbicara ajeg, tidak berubah-ubah pembicaraannya dari waktu ke waktu [Azwar \(2019:7\)](#). Dalam sebuah tes pentingnya diamati keajegan dan kepastian tes tersebut dilihat dari hasil tes yang didapat. Instrumen dapat dikatakan realibel, jika instrumen tersebut, memberikan hasil yang tepat atau benar walaupun dilakukan oleh siapa saja dan kapan saja, sehingga instrumen dapat dipercaya atau dapat diandalkan. Pada tahapan pengukuran realibilitas penulis menggunakan ITEMAN dengan melihat nilai Alpha. Hasil analisis realibilitas terhadap soal latihan Penilaian Akhir Semester (Latihan PAS Ganjil) tahun ajaran 2020/2021, di Kelas VIII SMPN I Bambanglipuro dapat dilihat melalui Tabel 5.

Tabel 5
Hasil Analisis Reliabilitas melalui Program ITEMAN

Karakteristik	Nilai
Scale	0
	40
N of Items	25
N of Examinees	24.160
Mean	39.494
Variance	6.284
Std. Dev.	-1.542
Skew	2.009
Kurtosis	6.000
Minimum	32.000
Maximum	25.000
Alpha	0,820
SEM	2.663
Mean P	0.604
Mean Item- Tot	0.372
Mean Biserial	0.502

Berdasarkan Tabel 5, dapat dilihat bahwa nilai Alpha sebesar **0.820** hal ini menunjukkan bahwa realibilitas butir soal pilihan ganda soal Pra-Penilaian Akhir Semester Ganjil, kelas VIII SMPN I Bambanglipuro Bantul, tahun ajaran 2020/2021 dapat dikategorikan memiliki reliabilitas sangat tinggi. Berdasarkan Tabel 5 menunjukkan bahwa butir soal pilihan ganda soal evaluasi pra-Penilaian Akhir Semester (PAS) kelas VIII B SMPN I Bambanglipuro Bantul, sebesar 0.820. Realibilitas pada suatu soal menunjukkan koefisien $0,80 < r \leq 1,00$ dengan kategori sangat tinggi. Hal tersebut, sesuai dengan pendapat [Nuryanti et al, \(2018\)](#) yang menyatakan bahwa uji reliabilitas dalam program ITEMAN dapat dilihat melalui perolehan skor Alpha dalam program ITEMAN dengan klasifikasi reliabilitas: 0,00-0,20 (sangat rendah); 0,21-0,40 (rendah); 0,41-0,70 (sedang); 0,71-0,90 (tinggi); 0,91-1,00 (sangat tinggi).

Indeks Daya Beda (IDB)

Daya beda butir soal (*item discrimination*) dapat didefinisikan sebagai suatu pernyataan tentang seberapa besar daya sebuah butir soal dapat membedakan kemampuan antara peserta kelompok tinggi dan kelompok rendah. Indeks daya beda (IDB) adalah indeks yang menunjukkan seberapa besar daya sebuah butir soal kemampuan antara peserta kelompok tinggi dan kelompok rendah ([Nurgiantoro, 2017:221](#)). Butir soal yang baik daya bedanya harus memiliki koefisien yang positif dan tinggi pada kunci jawaban ([Azwar, 1996:151](#)). Hasil analisis mengenai Indeks Daya Beda (IDB) terhadap soal latihan Penilaian Akhir Semester dalam Pembelajaran Bahasa Indonesia, di SMPN I Bambanglipuro Bantul dapat dijelaskan melalui Tabel 6.

Tabel 6
Hasil Analisis Indeks Daya Beda

Kategori	Butir	Jumlah
Buruk/Gugur	2,3,8,11,15, 28, 35 dan 37	8
Cukup	4, 6, 7, 10, 14, 17, 18 dan 40	8
Sedang	16, 19, 20 dan 30	4
Baik	1, 5, 9, 12, 13, 21, 22, 23, 24, 25, 26, 27, 29, 31, 32, 33, 34, 35, 36, dan 39	20
Baik Sekali	-	-

Berdasarkan hasil analisis, dapat disimpulkan bahwa kategori butir-butir soal memiliki daya beda yang baik adalah di atas 0.40 sampai 1.00. Berdasarkan tabel hasil ITEMAN daya beda butir soal di atas, terdapat 20 butir soal yang memiliki kategori baik, yakni butir nomor 1, 5, 9, 12, 13, 21, 22, 23, 24, 25, 26, 27, 29, 31, 32, 33, 34, 35, 36, dan 39. Kategori sedang memiliki batas antara 0.30-0.39, terdapat 4 butir soal yang termasuk kategori sedang yaitu 16, 19, 20, dan 30. Soal yang memiliki daya beda cukup berada di antara 0.20-0.29, terdapat 8 butir soal yang termasuk kategori tersebut, yaitu 4, 6, 7, 10, 14, 17, 18, dan 40. Berikutnya, terdapat 8 soal dengan daya beda yang buruk atau gugur berkisar antara negatif-0.19. Hal tersebut, selaras dengan pendapat (Monica, 2013) yang menyatakan bahwa Indeks Daya Beda berkisar antara 0,00 sampai 0,20 dianggap memiliki daya beda jelek, daya beda yang berkisar antara 0,21 sampai 0,40 dianggap cukup, daya pembeda berkisar antara 0,41 sampai 0,70 dapat dikatakan baik, daya pembeda yang berkisar antara 0,71 sampai 1,00 dikatakan baik sekali dan hasil negatif memiliki daya pembeda tidak baik.

Indeks Tingkat Kesulitan (ITK)

Tingkat kesulitan (*item difficulty*) adalah indeks yang menunjukkan seberapa mudah atau sulit butir soal bagi peserta tes yang diuji (Nurgiyantoro, 2017:218). Rentang indeks kesulitan ini berkisar antara 0,00 sampai 1,00. Jika indeks kesulitan menunjukkan angka 1,0 maka itu soal tidak sulit dikerjakan. Sebaliknya jika indeks kesukarannya 0,00 maka soal tersebut sangat sulit dikerjakan. Pendapat lain, disampaikan oleh (Monica, 2013) yang menyatakan bahwa Tingkat kesukaran soal 0,00 sampai 0,30 merupakan soal sukar, tingkat kesukaran 0,31 sampai 0,70 soal sedang, dan tingkat kesukaran 0,71 sampai 1,00 soal mudah. Hasil Analisis Indeks Tingkat Kesulitan (ITK) berkaitan dengan soal yang dianalisis dapat dilihat melalui Tabel 7.

Tabel 7
Tabel Hasil Analisis Indeks Tingkat Kesulitan (ITK)

Kategori	Rentang	Butir Soal	Jumlah
Sangat Mudah	0.81-1.00	1, 6, 12, 16, 22, 39	6
Mudah	0.61-0.80	5, 8, 9, 10, 14, 20, 21, 23, 24, 25, 26, 29, 31, 32, 33, 34, 36	17
Sedang	0.41-0.60	3, 7, 13, 18, 19, 27, 28, 38	8
Sukar	0.21-0.40	2, 4, 11, 15, 17, 30, 35, 37, 40	9
Sangat Sukar	-	-	-

Berdasarkan Tabel 7, dapat disimpulkan bahwa terdapat 6 butir soal dengan kategori ITK sangat mudah yaitu soal dengan nomor 1, 6, 12, 16, 22, dan 39. Terdapat, 17 butir soal dengan ITK mudah, yaitu soal dengan nomor 5, 8, 9, 10, 14, 20, 21, 23, 24, 25, 26, 29, 31, 32, 33, 34, 36. 8 butir soal dengan Indeks tingkat kesulitan sedang, yaitu butir nomor 3, 7, 13, 18, 19, 27, 28, 38. Berikutnya, 9 soal dengan Indeks tingkat kesulitan sukar atau tinggi yaitu butir nomor 2, 4, 11, 15, 17, 30, 35, 37, 40.

Keberfungsian Pengecoh

Pengecoh dapat dikatakan berfungsi apabila semua pengecoh ada pemilihnya, dan dikatakan tidak berfungsi dengan baik jika ada pengecoh yang tidak ada pemilihnya (Azwar, 2015:143). Analisis butir pengecoh dalam soal evaluasi yang dikembangkan, dapat dilihat melalui Tabel 8.



Tabel 8
 Hasil Analisis Keberfungsian Distraktor (Butir Pengecoh)

Butir Soal	Pengecoh	Hasil	Kategori
I	A. 0.920*	92%	Kunci Jawaban
	B. 0.080	8%	Pengecoh berfungsi dengan baik
	C. 0.000	0%	Pengecoh tidak berfungsi
	D. 0.000	0%	Pengecoh tidak berfungsi
2	A. 0.000	0%	Pengecoh berfungsi dengan baik
	B. 0.600	60%	Pengecoh berfungsi dengan baik
	C. 0.000	0%	Pengecoh tidak berfungsi
	D. 0.400*	40%	Kunci Jawaban
3	A. 0.120	12%	Pengecoh berfungsi dengan baik
	B. 0.440*	44%	Kunci Jawaban
	C. 0.400	40%	Pengecoh berfungsi dengan baik
	D. 0.040	4%	Pengecoh berfungsi dengan baik
4	A. 0.120	12 %	Pengecoh berfungsi dengan baik
	B. 0.240*	24%	Kunci Jawaban
	C. 0.600	60%	Pengecoh berfungsi dengan baik
	D. 0.040	4%	Pengecoh berfungsi dengan baik
5	A. 0.160	16%	Pengecoh berfungsi dengan baik
	B. 0.160	16%	Pengecoh berfungsi dengan baik
	C. 0.04	4%	Pengecoh berfungsi dengan baik
	D. 0.640*	64%	Kunci Jawaban
6	A. 0.880*	88%	Kunci Jawaban
	B. 0.000	0%	Pengecoh tidak berfungsi
	C. 0.120	12%	Pengecoh berfungsi dengan baik
	D. 0.000	0%	Pengecoh tidak berfungsi
7	A. 0.040	4%	Pengecoh berfungsi dengan baik
	B. 0.360	36%	Pengecoh berfungsi dengan baik
	C. 0.440*	44%	Kunci Jawaban
	D. 0.160	16%	Pengecoh berfungsi dengan baik
8	A. 0.160	16%	Pengecoh berfungsi dengan baik
	B. 0.120	12%	Pengecoh berfungsi dengan baik
	C. 0.680*	68%	Kunci Jawaban
	D. 0.040	4%	Pengecoh berfungsi dengan baik
9	A. 0.120	12%	Pengecoh berfungsi dengan baik
	B. 0.080	8%	Pengecoh berfungsi dengan baik
	C. 0.200	20%	Pengecoh berfungsi dengan baik
	D. 0.600*	60%	Kunci jawaban
10	A. 0.600*	60%	Kunci jawaban
	B. 0.240	24%	Pengecoh berfungsi dengan baik
	C. 0.040	4%	Kunci Jawaban
	D. 0.120	12%	Pengecoh berfungsi dengan baik
11	A. 0.080	8%	Pengecoh berfungsi dengan baik
	B. 0.320	32%	Pengecoh berfungsi dengan baik
	C. 0.240*	24%	Kunci jawaban
	D. 0.360	36%	Pengecoh berfungsi dengan baik
12	A. 0.960*	96%	Kunci jawaban
	B. 0.000	0%	Pengecoh tidak berfungsi
	C. 0.040	4%	Pengecoh berfungsi dengan baik
	D. 0.000	4%	Pengecoh tidak berfungsi
13	A. 0.000	0%	Pengecoh tidak berfungsi
	B. 0.320	32%	Pengecoh berfungsi dengan baik
	C. 0.120	12%	Pengecoh berfungsi dengan baik
	D. 0.560*	56%	Kunci Jawaban
14	A. 0.800*	80%	Kunci jawaban
	B. 0.120	12%	Pengecoh berfungsi dengan baik

Butir Soal	Pengecoh	Hasil	Kategori
	C. 0.080	8%	Pengecoh berfungsi dengan baik
	D. 0.000	0%	Pengecoh tidak berfungsi
15	A. 0.120	12%	Pengecoh berfungsi dengan baik
	B. 0.520	52%	Pengecoh berfungsi dengan baik
	C. 0.240*	24%	Kunci jawaban
	D. 0.120	12%	Pengecoh berfungsi dengan baik
16	A. 0.880*	88%	Kunci Jawaban
	B. 0.040	4%	Pengecoh berfungsi dengan baik
	C. 0.040	4%	Pengecoh berfungsi dengan baik
	D. 0.040	4%	Pengecoh berfungsi dengan baik
17	A. 0.240	24%	Pengecoh berfungsi dengan baik
	B. 0.400*	40%	Kunci Jawaban
	C. 0.240	24%	Pengecoh berfungsi dengan baik
	D. 0.120	12%	Pengecoh berfungsi dengan baik
18	A. 0.040	4%	Pengecoh berfungsi dengan baik
	B. 0.200	20%	Pengecoh berfungsi dengan baik
	C. 0.200	20%	Pengecoh berfungsi dengan baik
	D. 0.560*	56%	Kunci jawaban
19	A. 0.040	4%	Pengecoh berfungsi dengan baik
	B. 0.600*	60%	Kunci Jawaban
	C. 0.160	16%	Pengecoh berfungsi dengan baik
	D. 0.200	20%	Pengecoh berfungsi dengan baik
20	A. 0.000	0%	Pengecoh tidak berfungsi
	B. 0.240	24%	Pengecoh berfungsi dengan baik
	C. 0.720*	72%	Kunci Jawaban
	D. 0.040	4%	Pengecoh berfungsi dengan baik
21	A. 0.040	4%	Pengecoh berfungsi dengan baik
	B. 0.040	4%	Pengecoh berfungsi dengan baik
	C. 0.120	12%	Pengecoh berfungsi dengan baik
	D. 0.800*	80%	Kunci Jawaban
22	A. 0.880*	88%	Kunci Jawaban
	B. 0.080	8%	Pengecoh berfungsi dengan baik
	C. 0.000	0%	Pengecoh tidak berfungsi
	D. 0.040	4%	Pengecoh berfungsi dengan baik
23	A. 0.200	20%	Pengecoh berfungsi dengan baik
	B. 0.720*	72%	Kunci Jawaban
	C. 0.080	8%	Pengecoh berfungsi dengan baik
	D. 0.000	0%	Pengecoh tidak berfungsi
24	A. 0.040	4%	Pengecoh berfungsi dengan baik
	B. 0.240	24%	Pengecoh berfungsi dengan baik
	C. 0.040	4%	Pengecoh berfungsi dengan baik
	D. 0.680*	68%	Kunci Jawaban
25	A. 0.720*	72%	Kunci Jawaban
	B. 0.040	4%	Pengecoh berfungsi dengan baik
	C. 0.080	8%	Pengecoh berfungsi dengan baik
	D. 0.160	16%	Pengecoh berfungsi dengan baik
26	A. 0.200	20%	Pengecoh berfungsi dengan baik
	B. 0.120	12%	Pengecoh berfungsi dengan baik
	C. 0.680*	68%	Kunci Jawaban
	D. 0.000	0%	Pengecoh tidak berfungsi
27	A. 0.040	4%	Pengecoh berfungsi dengan baik
	B. 0.440*	44%	Kunci Jawaban
	C. 0.440	44%	Pengecoh berfungsi dengan baik
	D. 0.080	8%	Pengecoh berfungsi dengan baik
28	A. 0.120	12%	Pengecoh berfungsi dengan baik
	B. 0.400	40%	Pengecoh berfungsi dengan baik



Butir Soal	Pengecoh	Hasil	Kategori
	C. 0.440*	44%	Kunci Jawaban
	D. 0.040	4%	Pengecoh berfungsi dengan baik
29	A. 0.720*	72%	Kunci Jawaban
	B. 0.040	4%	Pengecoh berfungsi dengan baik
	C. 0.240	24%	Pengecoh berfungsi dengan baik
	D. 0.000	0%	Pengecoh tidak berfungsi
30	A. 0.520	52%	Pengecoh berfungsi dengan baik
	B. 0.120	12%	Pengecoh berfungsi dengan baik
	C. 0.080	8%	Pengecoh berfungsi dengan baik
	D. 0.280*	28%	Kunci Jawaban
31	A. 0.000	0%	Pengecoh berfungsi dengan baik
	B. 0.800*	80%	Kunci Jawaban
	C. 0.080	8%	Pengecoh berfungsi dengan baik
	D. 0.120	12%	Pengecoh berfungsi dengan baik
32	A. 0.800	80%	Pengecoh berfungsi dengan baik
	B. 0.160	16%	Pengecoh berfungsi dengan baik
	C. 0.760*	76%	Kunci Jawaban
	D. 0.000	0%	Pengecoh tidak berfungsi
33	A. 0.680*	68%	Kunci Jawaban
	B. 0.040	4%	Pengecoh berfungsi dengan baik
	C. 0.240	24%	Pengecoh berfungsi dengan baik
	D. 0.040	4%	Pengecoh berfungsi dengan baik
34	A. 0.280	20%	Pengecoh berfungsi dengan baik
	B. 0.600*	60%	Kunci Jawaban
	C. 0.000	0%	Pengecoh tidak berfungsi
	D. 0.120	12%	Pengecoh berfungsi dengan baik
35	A. 0.120	12%	Pengecoh berfungsi dengan baik
	B. 0.440	44%	Pengecoh berfungsi dengan baik
	C. 0.080	8%	Pengecoh berfungsi dengan baik
	D. 0.360*	36%	Kunci Jawaban
36	A. 0.720*	72%	Kunci Jawaban
	B. 0.080	8%	Pengecoh berfungsi dengan baik
	C. 0.000	0%	Pengecoh tidak berfungsi
	D. 0.200	20%	Pengecoh berfungsi dengan baik
37	A. 0.320*	32%	Pengecoh berfungsi dengan baik
	B. 0.200	20%	Pengecoh berfungsi dengan baik
	C. 0.200	20%	Pengecoh berfungsi dengan baik
	D. 0.280*	28%	Pengecoh berfungsi dengan baik
38	A. 0.600*	60%	Kunci Jawaban
	B. 0.120	12%	Pengecoh berfungsi dengan baik
	C. 0.200	20%	Pengecoh berfungsi dengan baik
	D. 0.080	8%	Pengecoh berfungsi dengan baik
39	A. 0.160	16%	Pengecoh berfungsi dengan baik
	B. 0.840*	84%	Kunci Jawaban
	C. 0.000	0%	Pengecoh tidak berfungsi
	D. 0.000	0%	Pengecoh tidak berfungsi
40	A. 0.360*	36%	Kunci Jawaban
	B. 0.280	28%	Pengecoh berfungsi dengan baik
	C. 0.240	24%	Pengecoh berfungsi dengan baik
	D. 0.120	12%	Pengecoh berfungsi dengan baik

Berdasarkan Tabel 8, analisis butir pengecoh terhadap 40 soal Latihan Penilaian Akhir Semester Genap (Latihan PAS) di kelas VIII SMPN I Bambanglipuro Bantul menunjukkan bahwa terdapat 26 butir soal dengan keberfungsian pengecoh baik, yaitu pada nomor 1, 3, 4, 5, 7, 8, 9, 10, 11, 12, 16, 17, 18, 19, 21, 24, 25, 27, 28, 30, 31, 33, 35, 37, 38, 40 dan terdapat 14 butir soal dengan pengecoh yang

tidak berfungsi dengan baik yaitu pada nomor 2, 6, 13, 14, 15, 20, 22, 23, 26, 29, 32, 34, 36, 39. Berdasarkan tabel di atas, dari 40 butir soal yang dianalisis, terdapat pengecoh yang berfungsi dan pengecoh yang tidak berfungsi. Pengecoh butir soal berfungsi jika semua atau 3 pengecohnya berfungsi. Pengecoh tidak berfungsi jika memiliki 1, 2 atau 3 pengecoh yang tidak berfungsi.

Hasil tersebut sesuai dengan hasil penelitian dan pendapat (Rotama et al., 2020) yang menyatakan bahwa keberfungsian pengecoh dapat dilihat melalui nilai *proportional endorsing* dalam program Iteman. (Rotama et al., 2020) menyatakan bahwa suatu butir dapat dikategorikan sebagai soal yang layak dan baik apabila keberfungsian pengecoh berfungsi dengan baik. Pengecoh yang berfungsi dengan baik ditandai dengan dipilih oleh sedikitnya 5% dari peserta tes. Pendapat lain disampaikan oleh (Nurgiyantoro, 2016) yang menyatakan bahwa analisis butir soal dengan menggunakan model klasik, juga sampai pada analisis distraktor.

Hal ini menyatakan bahwa distraktor merupakan hasil analisis jawaban peserta tes terhadap opsi jawaban yang salah. Model tersebut beranggapan bahwa semua opsi pilihan jawaban harus efektif, artinya, walau opsi jawaban itu salah, opsi-opsi jawaban tersebut harus ada beberapa peserta tes yang memilihnya. Opsi salah yang baik, yaitu opsi yang mampu berperan sesuai dengan fungsinya, baik sebagai perusak, penjebak, dan pengecoh terhadap beberapa peserta tes.

Contoh Hasil Iteman soal layak dan tidak layak

Berikut adalah contoh hasil ITEMAN, mengenai soal layak dan tidak layak.

Kota Palangkaraya dan sekitarnya kembali diselubungi oleh kabut asap. Akibatnya, aktivitas berkendara di daerah ini sangat terganggu karena jarak pandang hanya berkisar 50-100 meter. Meski jarak pandang menipis, warga tetap menjalankan aktivitas kesehariannya dengan memakai masker. Kabut asap yang kian pekat ini diduga akibat pembakaran lahan gambut. Masyarakat tidak hanya mengeluhkan kabut asap yang mengganggu aktivitas mereka. Akan tetapi, mereka juga mengeluhkan gangguan pada tubuh berupa pusing dan sesak napas yang diakibatkan kabut asap. Selain itu, proses belajar mengajar pada beberapa sekolah di Palangkaraya juga terganggu. Sejak 2 hari lalu, beberapa sekolah di Palangkaraya telah meliburkan siswa selama 5 hari.

- I. Pokok berita yang dominan dalam kutipan teks berita tersebut adalah
- A. bagaimana
 - B. mengapa
 - C. kapan
 - D. apa

Berdasarkan hasil analisis lembar validitas mengenai validitas soal berkaitan dengan materi, konstruksi, dan bahasa soal nomor I dinyatakan valid. Analisis ITEMAN pada soal nomor I, menunjukkan bahwa soal tersebut dikategorikan dalam soal yang sangat mudah, karena indeks tingkat kesulitan butir soal (ITK) menunjukkan angka 0.920. Besarnya indeks daya beda (IDB) adalah 0.641. Jadi, soal tersebut dapat dikatakan layak. Berkaitan dengan keberfungsian pengecoh yang dilihat pada nilai *proportional endorsing* dalam program ITEMAN. Distraktor atau keberfungsian pengecoh soal nomor I dapat disebutkan sebagai berikut butir A yang merupakan kunci jawaban menunjukkan skor 0.820 atau 82%, artinya 82% dari jumlah keseluruhan peserta didik memilih butir tersebut, butir B menunjukkan skor 0.080 atau 8% dengan keberfungsian pengecoh baik, butir C menunjukkan skor 0.050 atau 5% sehingga dapat diketahui tidak ada yang memilih butir ini berarti dapat diartikan pengecoh tidak berfungsi, sedangkan butir D juga menunjukkan skor 0.050 atau 5% yang menunjukkan tidak ada peserta didik yang memilih butir tersebut berarti dapat diartikan pengecoh tidak berfungsi. Berdasarkan analisis mengenai validitas dan reliabilitas. Soal nomor I dikatakan layak.



I5. Perhatikan kalimat iklan berikut!

- (1) Balsem Kerik dengan *aromatherapy* efektif meredakan masuk angin.
- (2) Top Herbal memiliki kandungan jahe merah, adas, madu, dan *royal jelly*.
- (3) Asaklin lantai benar-benar *cling*, pilihan paling tepat untuk keluarga Indonesia.
- (4) Madu Tirta Ginseng sudah mendapatkan izin dari Dinkes P-IRT No. 2093205011152-23

Kalimat fakta pada kalimat iklan tersebut adalah kalimat nomor . . .

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

Berdasarkan hasil analisis lembar validitas mengenai validitas soal berkaitan dengan materi, konstruksi, dan bahasa soal nomor 15 dinyatakan valid artinya antara keterkaitan antara indikator dan soal dinilai baik, penggunaan bahasa serta penulisan soal dalam hal ini tidak ada kata atau kalimat yang mengalami salah ketik dan sebagainya dinilai valid. Hasil analisis ITEMAN pada soal nomor 15 menunjukkan bahwa soal nomor 15 dikategorikan dalam soal yang sangat sukar, karena indeks tingkat kesulitan butir soal (ITK) menunjukkan skor 0.240. Besarnya indeks daya beda (IDB) adalah -0.193. Jadi, soal tersebut dapat dikatakan tidak layak. Berkaitan dengan keberfungsian pengecoh yang dilihat pada nilai *proportional endorsing* dalam program ITEMAN. Menunjukkan bahwa butir A pengecoh berfungsi dengan baik perolehan skor 0.120 atau 12%, butir B pengecoh berfungsi dengan baik perolehan skor 0.520 atau 52%, butir C merupakan kunci jawaban dengan skor 0.240 atau 24%, sedangkan butir D keberfungsian pengecoh dikatakan dengan baik, karena menunjukkan angka 0.120 atau 12%. Berdasarkan analisis hasil ITEMAN dapat disimpulkan bahwa soal tersebut tergolong pada soal yang sulit. Sehingga dengan mempertimbangkan hasil analisis pada ITEMAN soal tersebut dikatakan tidak layak untuk digunakan.

SIMPULAN

Analisis butir soal, Penilaian Akhir Semester Ganjil, Mata Pelajaran Bahasa Indonesia, di SMPN I Bambanglipuro Bantul, tahun ajaran 2020/2021 dilakukan dengan menggunakan program komputer ITEMAN. Selaras dengan tujuan penelitian, Berdasarkan hasil Validitas, Reliabilitas dan Analisis Butir Soal yang dilakukan diperoleh hasil sebagai berikut; Terdapat 40 butir soal yang dikembangkan, soal berbentuk pilihan ganda. Soal tersebut digunakan sebagai latihan siswa, untuk menghadapi Penilaian Akhir Semester (PAS). Soal diujikan kepada 25 peserta didik kelas VIII B, SMPN I Bambanglipuro Bantul. Uji coba dilakukan secara tatap muka terbatas, mengingat pandemi Covid-19 yang masih terjadi hingga saat ini.

Validitas dilakukan dengan cara menguji kesesuaian materi, konstruksi dan bahasa yang terdapat dalam soal. Validitas dilakukan oleh penulis soal kepada teman sejawat, sesama guru bidang bahasa Indonesia di MGMP Bahasa Indonesia SMP Kabupaten Bantul. Berdasarkan hasil validitas, soal yang dikembangkan dikatakan valid atau sesuai. Reliabilitas instrumen ditinjau dari skor hasil Alpha berdasarkan ITEMAN. Nilai Alpha yang diperoleh berkaitan dengan soal evaluasi yang dikembangkan, sebesar 0.820 hal ini menunjukkan bahwa realibilitas butir soal pilihan ganda soal latihan Penilaian Akhir Semester Ganjil (latihan PAS), kelas VIII B SMPN I Bambanglipuro Bantul, dapat dikategorikan memiliki reliabilitas sangat tinggi.

Hasil analisis butir soal dilakukan dengan menghitung daya beda, berdasarkan analisis IDB, dapat disimpulkan bahwa terdapat 20 soal yang layak untuk diujikan. 4 butir soal dengan kategori sedang, 8 soal dengan kategori cukup, dan 8 soal dengan kategori gugur atau tidak layak untuk diujikan. Hasil analisis soal berkaitan dengan ITK, dapat disimpulkan bahwa terdapat 6 butir soal dengan kategori ITK sangat mudah yaitu soal dengan nomor 1, 6, 12, 16, 22, dan 39. Terdapat, 17 butir soal dengan ITK

mudah, yaitu soal dengan nomor 5, 8, 9, 10, 14, 20, 21, 23, 24, 25, 26, 29, 31, 32, 33, 34, 36. 8 butir soal dengan Indeks tingkat kesulitan sedang, yaitu butir nomor 3, 7, 13, 18, 19, 27, 28, 38. Berikutnya, 9 soal dengan Indeks tingkat kesulitan sukar atau tinggi yaitu butir nomor 2, 4, 11, 15, 17, 30, 35, 37, 40.

Dari hasil analisis keberfungsian pengecoh, dapat disimpulkan bahwa terdapat 26 butir soal dengan keberfungsian pengecoh baik, yaitu pada nomor 1, 3, 4, 5, 7, 8, 9, 10, 11, 12, 16, 17, 18, 19, 21, 24, 25, 27, 28, 30, 31, 33, 35, 37, 38, 40 dan terdapat 14 butir soal dengan pengecoh yang tidak berfungsi dengan baik yaitu pada nomor 2, 6, 13, 14, 15, 20, 22, 23, 26, 29, 32, 34, 36, 39. Berdasarkan analisis butir soal tersebut, soal-soal yang dianggap tidak layak, kemudian direvisi dan diperbaiki sesuai dengan hasil analisis butir soal menggunakan program ITEMAN.

Penelitian ini diharapkan mampu memberikan kontribusi sebagai pijakan dalam melakukan penelitian lanjut, berkaitan dengan analisis butir soal menggunakan program ITEMAN, memberikan referensi berkaitan dengan cara menganalisis butir soal menggunakan program ITEMAN. Secara keseluruhan kegiatan analisis butir soal perlu dilakukan, mengingat hal ini merupakan tahapan terpenting yang harus dilakukan guru sebagai fasilitator dan evaluator pembelajaran. Soal yang baik akan mendukung guru dan peserta didik untuk mencapai tujuan pembelajaran. Mengingat, saat ini referensi berkaitan dengan hal yang selaras dengan penelitian kurang dilakukan.

UCAPAN TERIMA KASIH

Terima kasih, diucapkan kepada Universitas Negeri Yogyakarta (UNY), Program Studi Magister Pendidikan Bahasa dan Sastra Indonesia (S2 PBSI UNY) yang telah memberikan kesempatan kepada penulis, untuk menimba ilmu dan memiliki pengalaman berkaitan dengan mata kuliah “*Pengembangan Assemen Pembelajaran Bahasa dan Sastra Indonesia*” dengan Dosen pengampu Prof. Dr. Burhan Nurgiyantoro, sehingga mampu menghasilkan luaran, berupa artikel relevan dengan mata kuliah yang diajarkan. Terima kasih, juga, kami ucapkan kepada SMPN 1 Bambanglipuro Bantul, MGMP Bahasa Indonesia SMP Kabupaten Bantul yang telah memberikan kesempatan kepada Kami, untuk melakukan analisis kebutuhan, pengembangan, hingga analisis butir soal, berkaitan dengan *soal latihan Penilaian Akhir Semester (Latihan PAS)*, mata pelajaran bahasa Indonesia, tahun ajaran 2020/2021 di SMPN 1 Bambanglipuro Bantul. Secara keseluruhan, penelitian ini diharapkan mampu memberikan dukungan untuk kemajuan ilmu Pendidikan Bahasa dan Sastra Indonesia, khususnya rumpun pengajaran bahasa, dengan topik evaluasi pembelajaran. Semoga dengan adanya hasil penelitian ini dapat menjadi pijakan untuk melakukan penelitian lanjut, yang relevan dan akan dilakukan, berkaitan dengan evaluasi pembelajaran bahasa Indonesia, khususnya topik “Analisis Butir Soal”. Sehingga, guru sebagai fasilitator pembelajaran dapat mencapai indikator yang sesuai, dalam menyajikan soal evaluasi, karena pada dasarnya evaluasi merupakan komponen penting dalam pembelajaran.

DAFTAR PUSTAKA

- Anggraini, D., & Suyata, P. (2014). Karakteristik soal UASBN mata pelajaran Bahasa Indonesia di Daerah Istimewa Yogyakarta pada tahun pelajaran 2008/2009. *Jurnal Prima Edukasia*, 2(1), 57-65. <https://doi.org/10.21831/jpe.v2i1.2644>
- Arifin, Zaenal, & Retnawati, H. (2017). Pengembangan instrumen pengukur higher order thinking skills matematika siswa SMA kelas X. *PYTHAGORAS: Jurnal Pendidikan Matematika*, 12(1), 98-108. <https://doi.org/10.21831/pg.v12i1.14058>
- Arikunto, S. (2016). *Dasar-dasar evaluasi pendidikan (2nd ed.)*. Jakarta: Bumi Aksara.
- Arvianto, I. R. (2016). Pemanfaatan program Iteman 3.0 untuk analisis butir soal lomba cerdas cermat teknologi informasi dan komunikasi tingkat SMA sederajat. *Jurnal Teknologi Informasi*, XI(33), 1-13. <https://doi.org/10.35842/jtir.v11i33.106>
- Azwar, Saifuddin. 2012. *Reliabilitas dan validitas*. Yogyakarta: Pustaka Pelajar.
- Bond, T. G., & Fox, C. M. (2007). Applying the rasch model: Fundamental measurement in the human



- sciences, Second Edition. *Journal of Educational Measurement*, 2(2nd), 360.
<https://doi.org/10.1111/j.1745-3984.2003.tb01103.x>
- Crombach, Lee J. 2004. *Essential of psychologycal testing*. New York: Harper & Row.
- Dhini, M. S., Sunarti, I., & Riadi, B. (2020). Pengembangan instrumen penilaian berbasis HOTS (higher order thinking skill) teks deskripsi. *Jurnal Kata (Bahasa, Sastra, dan Pembelajarannya)*, 8(2), 1-6. Retrieved from <http://jurnal.fkip.unila.ac.id/index.php/BINDOI/article/view/20138/14221>
- Fernanda, J. W., & Hidayah, N. (2020). Analisis kualitas soal ujian statistika menggunakan classical test theory dan rasch model. *Square: Journal of Mathematics and Mathematics Education*, 2(1), 49-60. <https://doi.org/10.21580/square.2020.2.1.5363>
- Guntur, R. (2017). Analisis uji validitas dan reliabilitas tes buatan guru Bahasa Makassar kelas VIII SMPN 2 Binamu Kabupaten Jeneponto (*Doctoral dissertation, FBS*).
- Hanifah, N. (2014). Perbandingan tingkat kesukaran, daya pembeda butir soal dan reliabilitas tes bentuk pilihan ganda biasa dan pilihan ganda asosiasi mata pelajaran Ekonomi. *SOSIO E-KONS*, 6(1), 41-55. <http://dx.doi.org/10.30998/sosioekons.v6i1.1715>
- Hopkins, Charles D dan Antes, Richard L. (2009). *Classroom Measurement and Evaluation*. Illionis: Peacock.
- Hardiana, Suyata (2018). The effectiveness of VAK (Visual, Auditory, Kinesthetic) model in learning of summary writing. *International Journal of Research and Review*, 5(8), 43-49
- Himawan, R, & Suyata, P. (2021). Analisis kebutuhan pengembangan soal HOTS dalam pembelajaran teks pidato persuasif di MGMP SMP Wilayah Kabupaten Bantul. *GHANCARAN: Jurnal Pendidikan Bahasa dan Sastra Indonesia*, 117-128. <https://doi.org/10.19105/ghancaran.vi.5399>
- Himawan, R, Suyata. P (2022). Developing HOTS questions: Evaluating persuasive speech text learning in grade IX of junior high school. *Gramatika: Jurnal Penelitian Pendidikan Bahasa dan Sastra Indonesia*. 8(1), 50-64.
- Hyman, R. (1982). Quasi-experimentation: Design and analysis issues for field settings (Book). *Journal of Personality Assessment*, 46(1), 96-97. https://doi.org/10.1207/s15327752jpa4601_16
- Joni, T. Raka. (2006) *Pengukuran dan penilaian pendidikan*. Surabaya: Karya Anda.
- Kurniawan, T. (2015). Analisis butir soal ulangan akhir semester gasal mata pelajaran IPS sekolah dasar (Analysis of odd semester final test items in elementary school of social studies subjects). *Journal of Elementary Education*, 4(1), 1-6. Retrieved from <https://journal.unnes.ac.id/sju/index.php/jee/article/view/7488>
- Mania, S., Fitriani, F., Majid, A. F., Ichiana, N. N., & Abrar, A. I. P. (2020). Analisis butir soal ujian akhir sekolah. *Al Asma: Journal of Islamic Education*, 2(2), 274-284. <https://doi.org/10.24252/asma.v2i2.16569>
- Mardapi, D., & Kartowagiran, B. (2011). Instruments for student achievements. *Jurnal Penelitian dan Evaluasi Pendidikan*, 2(4), 326-341.
- Mardiana, M., & Suyata, P. (2017). Evaluating the philosophical foundation of 2013 Curriculum. *Jurnal Penelitian dan Evaluasi Pendidikan*, 21(2), 175-188. <https://doi.org/10.21831/pep.v21i2.13336>
- Monica, S., & Sudarman, Y. (2019). Analisis butir soal ujian tengah semester ganjil seni budaya kelas VII di SMPN 29 Sijunjung. *Jurnal Sendratasik*, 8(3), 1-8. <https://doi.org/10.24036/jsu.v7i3.103276>
- Nayla Amalia, A., & Widayati, A. (2012). Analisis butir soal tes kendali mutu kelas XII SMA mata pelajaran ekonomi akuntansi di Kota Yogyakarta. *Jurnal Pendidikan Akuntansi Indonesia*, 10(1), 1-26. <https://doi.org/10.21831/jpai.v10i1.919>
- Nurgiyantoro, Burhan. 2016. *Penilaian pembelajaran bahasa berbasis kompetensi (Edisi Kedua)*. Yogyakarta: BPFEE.

- Nuryanti, S., Masykuri, M., & Susilowati, E. (2018). Analisis iteman dan model rasch pada pengembangan instrumen kemampuan berpikir kritis peserta didik sekolah menengah kejuruan. *Jurnal Inovasi Pendidikan IPA*, 4(2), 224–233. <https://doi.org/10.21831/jipi.v4i2.21442>
- Oktanin, W. S., & Sukirno, S. (2015). Analisis butir soal ujian akhir mata pelajaran ekonomi akuntansi. *Jurnal Pendidikan Akuntansi Indonesia*, 13(1), 35–44. <https://doi.org/10.21831/jpai.v13i1.5183>
- Pangesti, F., Fauzan, F., & Risnawati, R. (2020). Kualitas butir soal *try out* uji pengetahuan dalam memprediksi tingkat kelulusan mahasiswa PPG. *Jurnal Pendidikan Profesi Guru*, 1(2), 91–98. <https://doi.org/10.22219/jppg.v1i2.13503>
- Pasi, S. N., & Yusrizal, Y. (2018). Analisis butir soal ujian bahasa Indonesia buatan guru MTsN di Kabupaten Aceh Besar. *Master Bahasa*, 6(2), 195–202. <https://doi.org/10.24173/mb.v6i2.II666>
- Pujiastuti, R., & Kulup, L. I. (2021). Penyusunan Instrumen penilaian kognitif berbasis HOTS melalui problem based learning dan peer assessment. *Indonesian Language Education and Literature*, 7(1), 88–105. <https://doi.org/10.24235/ileal.v7i1.9058>
- Purniasari, L., Masykuri, M., & Ariani, S. R. D. (2021). Analisis butir soal ujian sekolah mata pelajaran Kimia SMA N I Kutowinangun tahun pelajaran 2019/2022 menggunakan model iteman dan rasch. *Jurnal Pendidikan Kimia*, 10(2), 205–214. <https://doi.org/10.20961/jpkim.v10i2.48244>
- Rahmawati, L. E., & Setyaningsih, V. I. (2021). Kemandirian belajar siswa dalam pembelajaran daring mata pelajaran bahasa Indonesia. *KEMBARA: Jurnal Keilmuan Bahasa, Sastra, dan Pengajarannya (e-journal)*, 7(2), 353–365. <https://doi.org/10.22219/kembara.v7i2.16326>
- Ramadhanti, D. (2019). Development of explanatory text writing evaluation tools based on a process approach (Pengembangan alat evaluasi menulis teks eksplanasi berbasis pendekatan proses). *Jurnal Gramatika: Jurnal Penelitian Pendidikan Bahasa dan Sastra Indonesia*, 5(2), 194–210. <https://doi.org/10.22202/jg.2019.v5i2.3445>
- Retnawati, H., Hadi, S., Nugraha, A. C., Arlinwibowo, J., Sulistyarningsih, E., Djidu, H., Apino, E., & Iryanti, H. D. (2017). Implementing the computer-based national examination in Indonesian schools: The challenges and strategies. *Problems of Education in the 21st Century*, 75(6), 612–633. <https://doi.org/10.33225/pec/17.75.612>
- 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. *Academy of Education Journal*, 11(1), 24–35. <https://doi.org/10.47200/Aoej.V11i01.314>
- Safi'i, I., Tarmini, W., & Wahdini, L. (2021). Critical thinking in evaluation instruments at BSE Indonesian language. *KEMBARA: Jurnal Keilmuan Bahasa, Sastra, dan Pengajarannya (e-journal)*, 7(2), 11–19. <https://doi.org/10.22219/kembara.v7i2.17300>
- Setiawan, M. A., Susongko, P., & Hayati, M. N. (2020). Pendeteksian DIF pada Perangkat tes objektif penilaian akhir semester IPA dengan menggunakan permodelan rasch. *Pancasakti Science Education Journal*, 5(2), 23–29. Retrieved from <https://scienceedujournal.org/index.php/PSEJ/article/view/25>
- Setiawati, G. A. D., & Arsana, A. A. P. (2018). Pengaruh motivasi belajar dan gender terhadap prestasi belajar IPA siswa kelas bilingual SMP (Slub) Saraswati I Denpasar. *Proceeding Biology Education Conference: Biology, Science, Enviromental, And Learning*, 15(1), 173–179.
- Sulistiawan, C. H. (2016). Kualitas soal ujian sekolah Matematika program IPA dan kontribusinya terhadap hasil ujian nasional. *Jurnal Penelitian dan Evaluasi Pendidikan*, 20(1), 1–10. <https://doi.org/10.21831/pep.v20i1.751>
- Susanto, H., Rinalidi, A., & Novalia, N. (2015). Analisis validitas reliabilitas tingkat kesukaran dan daya beda pada butir soal ujian akhir semester ganjil mata pelajaran Matematika kelas XII IPS di SMA Negeri 12 Bandar Lampung tahun ajaran 2014/2015. *Al-Jabar: Jurnal Pendidikan Matematika*,



-
- 6(2), 203-218. <https://doi.org/10.24042/ajpm.v6i2.50>
- Suyata, P., Mardapi, D., Kartowagiran, B., & Retnawati, H. (2011). Model pengembangan bank soal berbasis guru dan mutu pendidikan model of developing test item bank. *Jurnal Kependidikan*, 41(2), 120–128. <https://doi.org/10.21831/jk.v41i2.2218>
- Syaifudin, S. (2020). Validitas dan reliabilitas instrumen penilaian pada mata pelajaran Bahasa Arab. *Jurnal Kajian Perbatasa Antarnegara*, 3(2), 106–118. Retrieved from <http://www.journal.iaisambas.ac.id/index.php/Cross-Border/article/view/553>
- Utomo, B. (2019). Analisis validitas isi butir soal sebagai salah satu upaya peningkatan kualitas pembelajaran di madrasah berbasis nilai-nilai Islam. *Jurnal Pendidikan Matematika (Kudus)*, 1(2), 155-175. <https://doi.org/10.21043/jpm.v1i2.4883>
- Yot-Domínguez, C., & Marcelo, C. (2017). University students' self-regulated learning using digital technologies. *International Journal of Educational Technology in Higher Education*, 14(1), 1–18. <https://doi.org/10.1186/s41239-017-0076-8>