



Analisis Butir Soal IPS Kelas V Sekolah Dasar Menggunakan Software ITEMAN

Fitri Alfarisa^{1✉}, Ima Ni'mah Chudari² & Firman Robiansyah³

^{1✉}Universitas Pendidikan Indonesia, alfarisa@upi.edu, Orcid ID: [0000-0002-6041-7698](https://orcid.org/0000-0002-6041-7698)

²Universitas Pendidikan Indonesia, nimahchudari@upi.edu, Orcid ID: [0000-0001-7799-7441](https://orcid.org/0000-0001-7799-7441)

³Universitas Pendidikan Indonesia, firmanrobiansyah@upi.edu, Orcid ID: [0000-0002-4155-8794](https://orcid.org/0000-0002-4155-8794)

Article Info

History Articles

Received:

Agustus 2019

Accepted:

September 2019

Published:

Oktober 2019

Abstract

The assessment process did not end by only making questions. They will arrive at the analysis of the questions quality. Item analysis is performed to determine the quality of the questions items arranged. Good quality items will greatly affect the results of the assessment and properly capturing the results of learning. The purpose of this study is to analyze the Social Sciences question items for fifth grade of elementary school using the classic ITEMAN focusing on the characteristics of the level of difficulty, the discrimination power, and the functioning of distractor of items. This research is a descriptive study with quantitative research. Subjects were 370 elementary school students in Serang City. The instrument analyzed was made in the form of 20 items of multiple choices with four answer option. The instrument was validated by the content expert and estimated the reliability by using the KR-20 method by reliability index of 0.75 obtained. The data analysis technique used ITEMAN software to see the characteristics of items that contain the level of difficulty, discrimination power and distractor's function. The results showed that 80% of the difficulty level were in the medium category, 90% of the discrimination power items were in the good category, and 90% of the distractors had well-functioned. This result indicated that the question created by teacher for fifth grade of elementary school in Serang City concluded as good category.

Keywords:

Items Analysis, ITEMAN, Difficulty Level, Discrimination Power, Distractor

How to cite:

Alfarisa, F., Chudari, I. I., & Robiansyah, F. (2019). Analisis butir soal IPS kelas V sekolah dasar menggunakan software ITEMAN. *EduBasic Journal: Jurnal Pendidikan Dasar*, 1(2), pp.100-106.

PENDAHULUAN

Setiap penyelenggaraan program pembelajaran selalu melibatkan kegiatan penilaian dan evaluasi. Fokus penilaian hasil belajar adalah individu, sedangkan fokus evaluasi hasil belajar adalah kelompok atau program. Pengertian ini memberi implikasi pada program perbaikan, individu atau kelompok/institusi. Penilaian hasil pembelajaran merupakan suatu hal yang penting dan menjadi kewajiban seorang guru. Mardapi (2017) mengemukakan bahwa proses penilaian meliputi pengumpulan bukti-bukti tentang pencapaian belajar peserta didik, dimana bukti-bukti tersebut tidak selalu diperoleh melalui tes saja tetapi juga bisa dikumpulkan melalui pengamatan atau laporan diri.

Permendikbud Nomor 66 Tahun 2013 tentang Standar Penilaian Pendidikan (Kemendikbud, 2013) menjelaskan bahwa penilaian hasil belajar peserta didik mencakup kompetensi sikap, pengetahuan, dan keterampilan yang dilakukan secara berimbang sehingga dapat digunakan untuk menentukan posisi relatif setiap peserta didik terhadap standar yang telah ditetapkan. Cakupan penilaian merujuk pada ruang lingkup materi, kompetensi mata pelajaran kompetensi muatan atau kompetensi program, dan proses.

Penilaian pendidikan sebagai proses pengumpulan dan pengolahan informasi untuk mengukur pencapaian hasil belajar peserta didik mencakup: penilaian autentik, penilaian diri, penilaian berbasis portofolio, ulangan, ulangan harian, ulangan tengah semester, ulangan akhir semester, ujian tingkat kompetensi, ujian mutu tingkat kompetensi, ujian nasional, dan ujian sekolah/madrasah. Pendekatan penilaian yang digunakan adalah penilaian acuan kriteria (PAK). PAK merupakan penilaian pencapaian kompetensi yang didasarkan pada kriteria ketuntasan minimal (KKM). KKM merupakan kriteria ketuntasan belajar minimal yang ditentukan oleh satuan pendidikan dengan mempertimbangkan karakteristik kompetensi dasar yang akan dicapai, daya dukung, dan karakteristik peserta didik.

Anderson (2003) menyatakan bahwa penilaian terkait dengan pengambilan keputusan tentang siswa dan memiliki dampak jangka panjang, maka guru harus merencanakan dan melaksanakan penilaian dengan serius. Reynolds et al., (2010) mengemukakan bahwa sebelum memberikan penilaian, guru harus mengambil langkah yang tepat untuk mempersiapkan siswa. Langkah-langkah tersebut seperti mengumumkan di depan kelas kapan penilaian akan diberikan, mendiskripsikan isi dan keterampilan yang akan dinilai, memberitahu teknik penskorannya, dan memberitahu hasil tes tersebut akan digunakan untuk apa.

Penilaian yang profesional adalah penilaian yang dilakukan dengan menggunakan keahlian sehingga memenuhi persyaratan akademik. Penilaian autentik merupakan penilaian profesional. Penilaian yang profesional dilakukan dengan cermat, objektif, adil, sistematis, beracuan kriteria, dan akuntabel. Butir-butir prinsip penilaian ini tertuang dalam Permendiknas Nomor 20 Tahun 2007 tentang Standar Penilaian yakni sebagai berikut (kemendiknas, 2007).

1. Sahih, berarti penilaian didasarkan pada data yang mencerminkan kemampuan yang diukur.
2. Objektif, berarti penilaian didasarkan pada prosedur dan kriteria yang jelas, tidak dipengaruhi subjektivitas penilai.
3. Adil, berarti penilaian tidak menguntungkan atau merugikan peserta didik karena berkebutuhan khusus serta perbedaan latar belakang agama, suku, budaya, adat istiadat, status sosial ekonomi, dan gender.
4. Terpadu, berarti penilaian oleh pendidik merupakan salah satu komponen yang tidak terpisahkan dari kegiatan pembelajaran.
5. Terbuka, berarti prosedur dan kriteria penilaian disampaikan kepada personil yang relevan.
6. Menyeluruh dan berkesinambungan, berarti penilaian oleh pendidik mencakup semua aspek kompetensi dengan menggunakan berbagai teknik penilaian yang sesuai untuk memantau

perkembangan kemampuan peserta didik.

7. Sistematis, berarti penilaian dilakukan secara berencana dan bertahap dengan mengikuti langkah-langkah yang baku.
8. Beracuan kriteria, berarti penilaian didasarkan pada ukuran pencapaian kompetensi yang ditetapkan.
9. Akuntabel, berarti penilaian dapat dipertanggungjawabkan, baik dari segi teknik, prosedur, maupun hasilnya.

Dalam prakteknya kesembilan prinsip penilaian di atas, hendaknya dapat dilakukan dilakukan secara berkesinambungan, hal ini bertujuan untuk memantau proses dan kemajuan belajar peserta didik serta untuk meningkatkan efektivitas kegiatan pembelajaran. Penilaian hasil belajar oleh satuan pendidikan dilakukan untuk menilai pencapaian kompetensi peserta didik pada semua mata pelajaran. Tidak terkecuali pada pembelajaran IPS di SD. IPS merupakan ilmu pengetahuan sosial, dimana pembelajaran yang diteri siswa adalah berkaitan dalam kehidupan sehari-hari. Hal ini selaras dengan pengimplementasian dalam Kurikulum 2013.

Dalam Kurikulum 2013 terdapat beberapa aspek yang dinilai yakni meliputi sikap, pengetahuan, dan keterampilan. Semua aspek yang dinilai harus berdasarkan keadaan yang sebenarnya dalam dunia nyata, dan peserta didik harus menerapkan konsep dalam pembelajaran di dunia nyata pula. Untuk mengetahui kompetensi siswa secara komprehensif, dibutuhkan penguasaan teknik penyusunan instrument penilaian dengan baik. Tidak berhenti sampai dengan penyusunan instrument, guru juga seharusnya mampu menganalisis butir soal yang telah disusunnya.

Analisis butir soal merupakan salah satu cara untuk melihat kualitas butir soal atau penilaian yang telah disusun oleh pembuat soal, apakah soalnya sudah layak atau belum. Analisis butir soal ini dapat dilakukan secara modern ataupun klasik (Mardapi, 2017). Analisis butir soal secara klasik lebih mudah digunakan oleh guru SD dalam melihat kualitas butir soal. Dalam analisis butir soal klasik, terdapat tiga karakteristik butir soal yang dapat dilihat

yakni tingkat kesulitan butir, daya beda butir, dan keberfungsian pengecoh (Mardapi, 2017).

Tingkat kesukaran butir terbagi menjadi lima kategori yakni sangat sukar, sukar, sedang, mudah, dan sangat mudah. Kategori tersebut dapat diketahui berdasarkan indeks tingkat kesukaran butir. Indeks tingkat kesukaran butir berkisar antara nilai 0 sampai dengan 1, dengan ketentuan kategori **sangat sukar** (indeks < 0,1); **sukar** (indeks 0,1 - 0,3); kategori **sedang** (indeks 0,3-0,7); dan kategori **mudah** (indeks 0,7-0,9); dan kategori **sangat mudah** (indeks > 0,9) (Allen & Yen, 1979; Azwar, 2016; Mardapi, 2017). Terdapat berbagai cara untuk mengetahui indeks tingkat kesukaran butir, tetapi cara yang paling mudah dan paling banyak digunakan adalah dengan menghitung proporsi menjawab benar atau *proportional correct* (p) (Kartowagiran; 2012). Adapun formulasi tingkat kesukaran butir yakni:

$$P_i = \frac{n}{N} \quad (1)$$

Dalam hal ini;

| | |
|-------|---------------------------------------------------------|
| P_i | = tingkat kesukaran butir soal ke i |
| i | = nomor butir soal |
| n | = banyaknya siswa yang menjawab butir soal dengan benar |
| N | = banyaknya siswa yang menjawab butir soal |

Daya beda (db) butir soal berfungsi untuk menentukan dapat tidaknya suatu butir membedakan antara kelompok *testee* yang berkemampuan tinggi dan *testee* yang berkemampuan rendah (Allen & Yen, 1979). Rentang indeks daya beda butir soal antara nilai -1 sampai dengan 1. Makna indeks harga positif menggambarkan *testee* yang berkemampuan tinggi menjawab benar dan yang berkemampuan rendah menjawab salah, demikian sebaliknya bila indeks ini harganya negatif. Butir soal yang baik adalah butir soal yang mempunyai indeks db > 0,3 (Crocker & Algina, 2008). Lebih lanjut Mardapi (2017) dan Azwar (2016) merincikan daya beda butir soal terbagi menjadi tiga kategori yakni kategori **baik** (indeks db > 0,3); kategori **cukup baik** (indeks db 0,1 - 0,3); dan kategori **tidak baik** (indeks db < 0,1).

Terdapat beberapa cara untuk menghitung indeks daya beda butir soal, diantaranya yaitu: (1) indeks diskriminasi, (2) indeks korelasi, dan (3) indeks keselarasan. Pada materi kali ini, kita akan lebih fokus untuk menghitung indeks daya beda butir soal dengan formula **indeks diskriminasi**. Adapun formula indeks diskriminasi adalah sebagai berikut.

$$d = \frac{n_T}{N_T} - \frac{n_R}{N_R} \quad (2)$$

Dalam hal ini;

- d = daya beda butir atau indeks diskriminasi butir
- n_T = jumlah penjawab benar pada kelompok tinggi
- N_T = jumlah peserta tes dari kelompok berkemampuan tinggi
- n_R = jumlah penjawab benar pada kelompok rendah
- N_R = jumlah peserta tes dari kelompok berkemampuan rendah

Selain tingkat kesukaran dan daya beda butir soal, hal penting yang juga harus diperhatikan dalam menganalisis butir soal adalah keberfungsian distraktor atau alternatif jawaban yang disediakan. Dalam soal pilihan ganda terdapat distraktor atau alternative pilihan jawaban yang digunakan untuk mengetahui kemampuan siswa dalam mencari pilihan jawaban yang paling tepat. Distraktor atau pengecoh dapat dikatakan berfungsi apabila mampu menarik perhatian peserta tes yang kurang menguasai konsep materi tes untuk memilihnya. Pengecoh yang dapat berfungsi dengan baik akan menjadikan butir soal menjadi baik. Butir soal yang baik adalah jika besarnya persentase pengecoh lebih besar dari 5% (Mardapi, 2017; Allen & Yen, 1979). Untuk mengetahui efektivitas setiap pengecoh dapat diketahui dengan rumus sebagai berikut.

$$EP = \frac{\text{jumlah peserta yang memilih pengecoh}}{\text{jumlah peserta tes keseluruhan}} \times 100\% \quad (3)$$

Berdasarkan pemaparan di atas, maka peneliti tertarik untuk melakukan analisis butir soal IPS, untuk mengetahui kualitas

butir soal yang telah dibuat oleh guru secara klasik. Analisis butir soal ini berbantu dengan software ITEMAN.

METODOLOGI

Penelitian ini menggunakan metode deskriptif kuantitatif. Deskriptif di sini bertujuan untuk memaparkan karakteristik butir soal IPS SD secara klasik. Subyek penelitian ini adalah 370 siswa SD kelas V di Kota Serang. Subyek penelitian diambil secara random untuk menjadi sampel ujicoba soal.

Teknik pengambilan data dilakukan dengan tes berjenis pilihan ganda dengan empat pilihan jawaban. Soal yang digunakan adalah soal buatan guru sebanyak 20 butir. Sebelum diujicobakan soal yang telah dibuat divalidasi oleh ahli atau *expert judgment* untuk mengetahui kesesuaian butir dengan materi yang diujikan. Hasil dari analisis ahli diketahui bahwa semua butir telah sesuai dan siap diujikan. Selanjutnya soal juga diestimasi reliabilitasnya menggunakan KR-20 dan diketahui instrument memiliki koefisien reliabilitas 0,75. Hal ini menunjukkan instrument ini sudah reliabel atau konstan digunakan sebagai alat untuk melakukan penilaian kemampuan siswa. Pendapat ini senada dengan Linn, (1989) dan Allen & Yen (1979) yang menyatakan bahwa indeks minimum reliabilitas suatu instrumen sebesar 0,70 sudah masuk kategori cukup baik. Instrument yang reliabel hasil pengukurannya dapat dipercaya mampu merepresentasi kemampuan siswa dengan baik.

Analisis butir soal hasil ujicoba selanjutnya dianalisis menggunakan ITEMAN. Soal yang telah dianalisis dengan menggunakan ITEMAN selanjutnya dilihat karakteristik tingkat kesukaran butir, daya beda butir, dan keberfungsian distraktor masing-masing butir. Ketiga karakteristik ini dapat kita ketahui dengan melakukan analisis pada setiap butir berdasarkan hasil respon atau jawaban siswa dalam menjawab soal.

HASIL DAN PEMBAHASAN

Berdasarkan hasil analisis butir soal IPS Kelas V SD dengan ITEMAN yang terdiri dari 20 butir dan diujikan pada 370

siswa diketahui hasil karakteristik butirnya sebagai berikut.

Tingkat Kesukaran Butir

Analisis tingkat kesukaran butir terbagi menjadi tiga kategori yakni kategori mudah, sedang, dan sulit. Sedangkan kategori butir sangat mudah dan sangat sulit tidak masuk dalam hasil analisis yang tersaji pada Tabel 1 berikut ini.

Tabel 1. Hasil Analisis tingkat kesukaran butir soal IPS kelas V SD

| Butir | Kategori | Jumlah | Ket |
|--------------------------------------------------------|--------------|----------------|----------|
| - | Sangat mudah | 0 butir (0%) | Ditolak |
| 1 & 17 | Mudah | 2 butir (10%) | Direvisi |
| 2, 3, 4, 5, 6, 7, 8, 9, 11, 12, 13, 14, 15, 16, 19, 20 | Sedang | 16 butir (80%) | Diterima |
| 10 & 18 | Sulit | 2 butir (10%) | Direvisi |
| - | Sangat sulit | 0 butir (0%) | Ditolak |

Berdasarkan hasil pada Tabel 1 diketahui rerata soal masuk kategori sedang (80%), terdapat dua butir yang masuk kategori mudah (10%) dan dua butir masuk kategori sulit (10%). Hal ini dapat diartikan bahwa sebagian besar soal diterima, sedangkan empat butir yang masuk kategori mudah dan sulit perlu untuk direvisi. Hal ini senada dengan pernyataan Amalia & Widayati (2012) dalam hasil penelitiannya bahwa soal yang baik adalah soal yang masuk kategori tingkat kesulitan sedang, yakni tidak terlalu mudah atau tidak terlalu sulit. Hal yang baik adalah tidak terdapat soal yang masuk kategori sangat sulit ataupun sangat mudah, sehingga tidak terdapat butir soal yang ditolak.

Daya Beda Butir

Indeks daya beda butir soal dapat digunakan sebagai bahan pertimbangan untuk menentukan apakah sebuah butir dikatakan baik atau tidak baik. Baik artinya dapat membedakan siswa yang berkemampuan tinggi dan rendah, begitu juga sebaliknya, tidak baik yang berarti butir tidak dapat

membedakan siswa yang berkemampuan tinggi dan rendah. Selanjutnya hasil identifikasi karakteristik daya beda butir soal terdapat pada Tabel 2 berikut ini.

Tabel 2. Karakteristik daya beda butir soal IPS kelas V SD

| Butir | Kategori | Jumlah | Ket |
|---------------------------------------------------------------|-----------------------|----------------|----------|
| 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 19, 20 | Baik (> 0,3) | 18 butir (90%) | Diterima |
| 17 | Cukup baik (0,1-0,29) | 1 butir (5%) | Direvisi |
| 18 | Tidak baik (< 0,1) | 1 butir (5%) | Ditolak |

Berdasarkan hasil pada Tabel 2 menunjukkan bahwa dari 20 butir dikatakan 90% butir memiliki daya beda yang sudah baik dan dapat diterima. Selanjutnya 5% butir memiliki daya beda yang cukup baik sehingga perlu direvisi. Sedangkan satu butir lagi (5%) butir ditolak karena indeks $db < 0,1$. Hal ini senada dengan pendapat Azwar (2016) yang menyatakan bahwa butir yang memiliki indeks $db < 0,1$ harus digugurkan atau dibuang. Apabila dicermati butir yang harus digugurkan adalah butir 18 yang juga masuk kategori soal sulit. Daya beda adalah penentu soal dikatakan baik atau tidak. Daya beda yang tidak baik akan sulit dipertahankan sehingga satu butir tersebut akan digugurkan.

Keberfungsian Pengecoh

Pengecoh atau distraktor atau pilihan jawaban selain kunci jawaban perlu untuk dicek keberfungsinya. Hasil analisis butir soal IPS Kelas V SD menunjukkan bahwa hampir semua (90%) pengecoh dapat menjalankan fungsinya dengan baik yakni dipilih minimal 5% dari peserta tes. Sedangkan sisanya 10% (2 butir) masih perlu ada perbaikan atau revisi yakni butir no 1 dan 7 dikarenakan terdapat pengecoh yang dipilih oleh peserta tes kurang dari 5%. Hal ini senada dengan Kartowagiran (2012), Mardapi (2017) dan Mahjabeen, et al. (2017) yang

menyatakan bahwa pengecoh dianggap berfungsi apabila dipilih oleh minimal 5% dari jumlah peserta tes.

Dengan diketahuinya karakteristik butir soal yang telah dibuat, ini dapat dijadikan acuan untuk pembuatan bank soal bagi soal-soal yang telah dianalisis. Hal ini senada dengan hasil penelitian Wibawa (2019) yang menyampaikan bahwa butir soal tes yang diterima tanpa revisi dapat dijadikan sebagai bank soal yang terstandar sehingga dapat bermanfaat bagi penilaian pembelajaran. Dengan adanya bank soal yang berkualitas, penilaian yang dilakukanpun dapat merepresentasikan kemampuan siswa yang sebenarnya.

KESIMPULAN

Hasil penelitian ini menunjukkan bahwa kualitas butir soal IPS Kelas V SD buatan guru di kota Serang masuk dalam kategori baik. Hal ini diketahui berdasarkan karakteristik butir soal menunjukkan bahwa 80% tingkat kesukaran butir masuk kategori sedang, 90% daya beda butir masuk kategori baik, dan 90% pengecoh sudah berfungsi dengan baik. Estimasi reliabilitas instrument yang dihasilkan guru juga sudah baik yakni dengan koefisien 0,75.

Berdasarkan hasil penelitian di atas, peneliti memiliki beberapa saran diantaranya: (1) Menjadikan soal yang sudah berkualitas baik masuk dalam bank soal, (2) guru dapat menggunakan ITEMAN untuk menganalisis butir soal secara klasik dan 3) Untuk penelitian selanjutnya dapat menganalisis butir soal secara modern menggunakan Item Respon Theory (IRT) untuk mendapatkan kualitas soal yang lebih baik lagi.

DAFTAR PUSTAKA

- Allen, M. J & Yen, W. M. (1979). *Introduction to Measurement Theory*. Texas: McGraw-Hill.
- Anderson, L. W. (2003). *Classroom Assessment: Enhancing The Quality of Teacher Decision Making*. New Jersey: Lawrence Erlbaum Associates Publishers.
- Amalia, A. N. & Widayati, A. (2012). Analisis butir soal tes kendali mutu kelas XII SMA mata pelajaran ekonomi akuntansi di Kota Yogyakarta tahun 2012. *Jurnal Pendidikan Akuntansi Indonesia*, 10(1), 1-26.
- Azwar, S. (2016). *Tes Prestasi: Fungsi Pengembangan Pengukuran Prestasi Belajar*. Yogyakarta: Pustaka Pelajar.
- Crocker, L. & Algina, J. (2008). *Introduction to Classical and Modern Test Theory*. Mason, Ohio: Nelson Education, Ltd.
- Kartowagiran, B. (Oktober 2012). Penulisan butir soal. *Makalah* disampaikan pada Pelatihan Penulisan dan Analisis Butir Soal bagi Sumber Daya PNS Dik- Rekinpeg, di Hotel Kawanua Aerotel.
- Kemendikbud. (2013). *Peraturan Menteri Pendidikan dan Kebudayaan RI Nomor 66 Tahun 2013 tentang Standar Penilaian Pendidikan*. Jakarta: Kementerian Pendidikan dan Kebudayaan.
- Kemendiknas. (2007). *Permendiknas Nomor 20 Tahun 2007 tentang Standar Penilaian*. Jakarta: Kementerian Pendidikan dan Kebudayaan Nasional.
- Mahjabeen, W., Alam, S., Hassan, U., Zafar, T., Butt, R., Konain, S., & Rizvi, M. (2017). Difficulty index, discrimination index and distractor efficiency in multiple choice questions. *Annals of Pakistan Institute of Medical Sciences*, 13(4), 310-315.
- Mardapi, D. (2017). *Pengukuran Penilaian dan Evaluasi Pendidikan Edisi 2*. Yogyakarta: Nuha Medika.
- Reynolds, C. R., Livingston, R. B., & Willson, V. (2010). *Measurement and Assessment in Education*. New Jersey: Pearson Education International.
- Linn L. R. (1989). *Educational Measurement*. New York: Mc Millan Publishing.

Wibawa, E. A. (2019). Karakteristik butir soal tes ujian akhir semester hukum bisnis. *Jurnal Pendidikan Akuntansi Indonesia*, 17(1), 87-96.