

Analisis butir soal pilihan ganda ulangan akhir semester genap mata pelajaran kearsipan

Yulia Fitri Ambarwati *, Ismiyati Ismiyati

Universitas Negeri Semarang, Kampus Sekaran Gunungpati Semarang, 50229, Indonesia

* Corresponding Author. E-mail: yulianafaw@gmail.com; ismiyati@mail.unnes.ac.id

Received: 22 September 2021; Revised: 27 November 2021; Accepted: 30 December 2021

Abstrak: Tujuan dalam penelitian ini yaitu untuk mengatahui kualitas butir soal ulangan akhir semester genap mata pelajaran kearsipan yang digunakan. Metode yang digunakan dalam penelitian ini yaitu kuantitatif, dan subjek yang digunakan sebanyak 105 siswa. Hasil penellitian menunukan bahwa validitas soal mayoritas berada pada kategori valid yaitu 70% atau sebanyak 28 butir soal dan soal yang tidak valid yaitu 30% atau sebanyak 12 butir soal. Nilai reliabilitas menunjukkan $\text{Alpha}=0,81$ yang artinya reliable. Berdasarkan yang dilihat dari tingkat kesukaran sebanyak 1 (2,5%) butir soal kategori sukar, 19 (47,5%) butir soal kategori sedang, dan sebanyak 20 (50%) butir soal kategori mudah. Daya pembeda terdapat 7 (17,5%) butir soal kategori jelek, 15 (37,5%) butir soal kategori cukup, dan 18 (45%) butir soal kategori baik. Keefektifan fungsi distractor, secara keseluruhan terdapat 59 (36,88%) fungsi pengecoh berfungsi dengan baik dan sebanyak 101 (63,13%) tidak berfungsi.

Kata Kunci: analisis butir soal, kualitas butir soal, kearsipan

The analysis of mutiple choice questions of final exam on archive subject

Abstract: The purpose of this study was to determine the quality of the end semester exam items in archiving subjects used. This research uses quantitative research methods and subjects in the study were 105 question sheets. The results showed that the validity of the majority of the questions was in the valid category, namely 70% or 28 items and 30% or 12 invalid items. The reliability value shows $\text{Alpha} = 0.81$ which means the question is reliable. Based on the difficulty level of the questions, there were 1 (2.5%) item in the difficult category, 19 (47.5%) items in the medium category and as many as 20 (50%) items in the easy category. Based on the discrimination power, there are 7 (17.5%) bad items, 15 (37.5%) enough items, and 18 (45%) good category items. Based on the effectiveness of the distractor function, overall there are 59 (36.88%) distractor functions functioning properly and 101 (63.13%) are not functioning.

Keywords: item analysis, questions item quality, archive

How to Cite: Ambarwati, Y., & Ismiyati, I. (2022). Analisis butir soal pilihan ganda ulangan akhir semester genap mata pelajaran kearsipan. *Measurement In Educational Research (Meter)*, 1(2), 64-75. doi:<http://dx.doi.org/10.33292/meter.v1i2.144>



PENDAHULUAN

Pendidikan dipandang dengan cara yang tepat guna membentuk sumber daya manusia berkualitas. Melalui pendidikan guru menjadi salah satu faktor utama dalam kegiatan proses belajar mengajar mulai dari membuat desain pembelajaran hingga evaluasi pembelajaran. Evaluasi pembelajaran merupakan kegiatan yang dilakukan oleh setiap guru setelah adanya proses pembelajaran, dengan tujuan agar guru mampu mencapai tujuan pembelajaran yang telah ditetapkan dan peserta didik dapat menguasai kompetensi yang telah ditetapkan pula.

Untuk mengetahui sejauh mana kemampuan peserta didik menguasai kompetensi, maka guru perlu melakukan tindakan evaluasi pada pembelajaran sebagai tolak ukur kemampuan peserta didik. Alat yang digunakan untuk mengukur peserta didik yaitu tes. Menurut Sudjana (2010) menyatakan bahwa tes sebagai alat penilaian adalah pertanyaan-pertanyaan yang diberikan kepada siswa untuk mendapat jawaban dari siswa dalam bentuk lisan (tes lisan), dalam bentuk tulisan (tes tertulis), atau dalam bentuk perbuatan (tes tindakan).



Dalam penelitian Pasi dan Yusrizal (2018) menyimpulkan bahwa tes yang baik yaitu tes yang dapat mengukur apa yang hendak diukur yang seharusnya diukur. Dengan begitu dalam proses pengukuran sangat diperlukan kualitas tes. Kualitas tes yang baik dapat dilihat melalui analisis butir soal. Dalam teori klasik analisis butir soal terdapat beberapa aspek yang perlu diperhatikan yaitu mulai dari validitas, reliabilitas, tingkat kesukaran, daya pembeda, dan fungsi pengecoh.

Dalam penelitian Solichin (2017) validitas soal dikatakan valid apabila memiliki dukungan besar terhadap skor total. Skor butir soal menyebabkan skor total menjadi tinggi atau rendah. Artinya dapat dikatakan butir soal memiliki validitas tinggi jika skor pada butir soal memiliki kesejajaran dengan skor total. Soal dikatakan reliabel apabila soal diujikan beberapa kali menunjukkan hasil yang sama. Seperti penelitian yang dilakukan Warju et al. (2020) dalam penelitiannya menyimpulkan bahwa reliabilitas adalah sejauh mana suatu soal mampu menghasilkan hasil yang sama pada uji coba secara berulang. Dan didukung dengan pendapat Arikunto (2016) bahwa reliabilitas tes berhubungan dengan masalah kepercayaan yang mana suatu tes yang memiliki taraf kepercayaan yang tinggi apabila tes tersebut dapat memberikan hasil yang tetap.

Tingkat kesukaran merupakan penentuan soal yang berkategori mudah:sedang:sukar. Menurut Arikunto (2016) bahwa butir soal yang baik adalah soal yang tidak terlalu mudah atau tidak terlalu sukar. Soal yang terlalu mudah tidak akan merangsang siswa untuk mempertinggi memecahkannya. Sebaliknya soal yang terlalu sukar akan menyebabkan siswa menjadi putus asa dan tidak mempunyai semangat untuk mencoba lagi karena diluar jangkauannya”.

Daya pembeda yaitu soal yang membedakan peserta didik yang pandai dan peserta didik yang tidak pandai. Seperti pendapat pendapat Sudjana (2010) bahwa analisis daya pembeda mengkaji setiap butir soal agar mengetahui kesanggupan soal dalam membedakan siswa yang tergolong mampu (tinggi prestasinya) dengan siswa yang tergolong kurang (lemah prestasinya). Artinya, Apabila soal tersebut diberikan kepada anak yang mampu/pandai, hasilnya menunjukkan prestasi yang tinggi, dan bila diberikan kepada siswa yang lemah/ tidak pandai, hasilnya rendah.”

Fungsi pengecoh yaitu beberapa opsi pada butir soal pilihan ganda yang bukan kunci jawaban. Iskandar dan Rizal (2017) menyatakan bahwa pengecoh-pengecoh yang baik merupakan yang serupa akan tetapi tidak sama dengan opsi benar sehingga mempunyai peluang untuk dipilih peserta tes yang tidak berhati-hati. Dan menurut Sudijono (2005) adalah option atau alternative yang berjumlah tiga hingga lima butir dan kemungkinan-kemungkinan jawaban yang terpasang pada tiap item itu salah satu diantaranya adalah jawaban benar (kunci jawaban), kemudian sisanya adalah jawaban salah.

Tujuan dari analisis butir soal menurut Arikunto (2016) yaitu untuk mengadakan identifikasi seluruh soal yang baik, kurang baik, dan soal yang jelek. Dengan menganalisis soal-soal dapat diperoleh informasi tentang kejelekan sebuah soal dan ‘petunjuk’ untuk diadakannya perbaikan. Pendapat lain, tujuan dari analisis butir soal menurut Sudjana (2014:149) adalah analisis butir soal yang bertujuan untuk memperoleh kualitas soal tes yang baik, sehingga dapat memperoleh gambaran tentang prestasi siswa yang sebenarnya. Dari pernyataan ahli diatas maka guru wajib melakukan analisis butir soal untuk mendapatkan butir soal yang berkualitas, seperti penelitian yang dilakukan oleh Supandi dan Farikhah (2016) bahwa seorang guru dituntut untuk mampu melakukan analisis butir soal, sehingga dari analisis akan tersebut mendapatkan seperangkat (alat evaluasi) yang berkualitas.

Saat ini dunia sedang dilanda dengan adanya wabah virus covid-19. Dengan adanya wabah ini seluruh kegiatan dilakukan secara daring termasuk kegiatan belajar mengajar (KBM) hingga Ulangan Akhir Semester (UAS) dilakukan secara daring. Salah satu sekolah yang melaksanakan Ulangan Akhir Semester (UAS) secara daring yaitu SMK Negeri 1 Salatiga. Hal ini sangat berbeda dari sebelumnya, semester 1 yang lalu peserta didik melaksanakan Ulangan Akhir Semester (UAS) gasal secara luring, dan semester 2 peserta didik melaksanakan Ulangan Akhir Semester (UAS) genap secara daring sehingga hal ini pertama kalinya untuk sekolah di SMK Negeri 1 Salatiga melaksanakan Ulangan Akhir Semester (UAS) secara daring.

Berdasarkan hasil observasi yang telah dilakukan bahwa di SMK Negeri 1 Salatiga guru diwajibkan untuk menyusun perangkat pembelajaran mulai dari RPP, kisi-kisi, hingga evaluasi termasuk analisis butir soal. Selanjutnya berdasarkan wawancara yang telah dilakukan dengan guru pengampu mata pelajaran kearsipan, bahwa perangkat tes ulangan akhir semester akhir semester genap dibuat dan disusun oleh masing-masing guru mata pelajaran, untuk mata pelajaran kearsipan belum dilakukan analisis secara

mendalam yang menentukan validitas, reliabilitas, tingkat kesukaran, daya pembeda dan fungsi pengecoh yang dikarenakan keterbatasan waktu yang dimiliki, sehingga guru belum mengetahui kualitas soal tersebut.

Berdasarkan data nilai pada ulangan akhir semester genap mata pelajaran kearsipan kelas X tahun pelajaran 2019/2020, sekitar 63,8% siswa belum memenuhi KKM. Tabel 1 merupakan rata-rata Ulangan Akhir Semester Genap pada mata pelajaran kearsipan kelas X di SMK Negeri 1 Salatiga.

Table 1. Rata-rata Hasil UAS Mata Pelajaran Kearsipan

| Kelas | Jml Siswa | KKM | Tuntas | % | Tidak Tuntas | % |
|-------|-----------|-----|--------|------|--------------|------|
| Otkp1 | 34 | 75 | 9 | 26,5 | 25 | 73,5 |
| Otkp2 | 35 | 75 | 19 | 54,3 | 16 | 45,7 |
| Otkp3 | 36 | 75 | 10 | 27,8 | 26 | 72,2 |

Tabel 1 memperlihatkan bahwa jumlah peserta didik yang tuntas KKM pada mata pelajaran Kearsipan di SMK Negeri 1 Salatiga mulai dari kelas X OTKP 1 sebesar 26,5%, OTKP 2 sebesar 54,3%, dan OTKP 3 sebesar 27,8% yang artinya ketuntasan masih rendah. Seperti pendapat Mulyasa (2014:131) suatu pembelajaran dikatakan berhasil jika peserta didik telah tuntas KKM setidaknya 75% dari seluruh peserta didik dalam kelas, dari besarnya ketidak tuntas peserta didik dalam mata pelajaran Kearsipan maka peneliti tertarik untuk melakukan analisis butir soal.

Analisis butir soal sangat bermanfaat di dunia pendidikan, khususnya sebagai evaluasi pembelajaran. Sehingga kualitas soal perlu diperhatikan, seperti penelitian yang dilakukan oleh D'Sa dan Visbal-Dionaldo (2017) menyatakan bahwa item yang cacat tidak cocok untuk menilai domain yang dimaksudkan dan akibatnya mempengaruhi kualitas ujian. Hal ini juga disampaikan dengan pendapat Arifin dan Retnawati (2017) yang menyatakan bahwa jika seluruh butir soal sudah disusun dengan baik, maka butir soal perlu diuji cobakan terlebih dahulu di lapangan. Tujuannya untuk mengetahui soal-soal mana yang perlu diubah, diperbaiki bahkan dibuang, serta soal-soal mana yang baik untuk digunakan.

Berdasarkan penelitian yang dilakukan oleh Negara dan Kurniawati (2018) menunjukkan bahwa pada tingkat kesukaran mayoritas soal pada kategori sedang, dengan daya pembeda mayoritas pada kriteria cukup dan fungsi pengecoh mayoritas tidak berfungsi dengan baik serta untuk reliabilitas dikatakan reliable. Dari kajian tersebut peneliti ingin menggali lebih dalam terkait analisis butir soal pada mata pelajaran kearsipan di SMK Negeri 1 Salatiga. Karena peneliti menganggap hal ini merupakan hal yang baru, dimana yang biasanya analisis butir soal dilakukan dari peneliti sebelumnya menggunakan data secara offline, namun peneliti menggunakan data butir soal ulangan akhir semester genap mata pelajaran Kearsipan secara online atau ujian yang dilaksanakan secara daring. Menurut Sudjana (2010) menyatakan bahwa analisis butir soal adalah pengkajian pertanyaan-pertanyaan tes agar diperoleh perangkat yang memiliki kualitas yang memadai" sehingga analisis butir soal dilakukan sebagai umpan balik untuk melakukan perbaikan soal agar soal memiliki kualitas yang baik. Berdasarkan hal tersebut, hasil penelitian ini diharapkan dapat menggali kualitas butir soal pilihan ganda ulangan akhir semester genap pada mata pelajaran kearsipan.

METODE

Jenis metode yang digunakan dalam penelitian ini yaitu menggunakan metode kuantitatif. Metode kuantitatif merupakan suatu metode yang data penelitiannya berupa angka-angka dan analisis menggunakan statistik (Sugiyono, 2012). Berdasarkan tujuan penelitian yaitu menganalisis kualitas butir soal, sehingga peneliti bermaksud mencari data yang digunakan untuk menggambarkan atau mendeskripsikan kualitas soal yang meliputi validitas, reliabilitas, tingkat kesukaran, daya pembeda, dan fungsi pengecoh (*distractor*) pada soal Ulangan Akhir Semester Genap Mata Pelajaran Kearsipan SMK Negeri 1 Salatiga Tahun Ajaran 2019/2020.

Subjek yang digunakan dalam penelitian ini sebanyak 105 lembar jawab siswa, objek dalam penelitian ini seluruh butir soal Ulangan Akhir Semester genap mata pelajaran Kearsipan di SMK Negeri 1 Salatiga. Teknik sampling yang digunakan adalah non probability sampling dengan pengambilan sampel yaitu teknik sampling jenuh. Menurut Sugiyono (2010) sampling jenuh adalah teknik penentuan sampel bila semua anggota populasi dijadikan sebagai sampel".

Metode pengumpulan data yang digunakan yaitu data primer dan data sekunder. Data primer yang diperoleh langsung peneliti melalui wawancara. Sedangkan data sekunder diperoleh dari dokumentasi. Wawancara, dalam metode ini, peneliti berinteraksi dan berkomunikasi secara langsung untuk mendapatkan informasi yang diinginkan seperti penyusun soal dan analisis butir soal Ulangan Akhir Semester Genap Kelas X SMK Negeri 1 Salatiga tahun ajaran 2019/2020. Dokumentasi, metode ini digunakan untuk mendapatkan data-data dokumen sekolah berupa, soal ulangan akhir semester genap kelas X, lembar jawab siswa, rekap nilai ujian akhir semester, kunci jawaban ulangan akhir semester, kisi-kisi soal, serta daftar nama peserta didik SMK Negeri 1 Salatiga tahun ajaran 2019/2020.

Teknik analisis data yang digunakan dalam penelitian ini yaitu validitas, reliabilitas, tingkat kesukaran, daya pembeda dan fungsi pengecoh. Data yang digunakan yaitu dikotomus dalam bentuk soal pilihan ganda dengan pendekatan teori klasik dan dianalisis menggunakan software ANATES V4.0.9. Rumus yang digunakan sebagai berikut:

Validitas

Menurut Anderson et al. (1975) menyatakan bahwa sebuah tes dikatakan valid apabila tes tersebut mengukur apa yang hendak diukur. Sedangkan Arikunto (2016) menyatakan bahwa mengukur ketepatan instrument validitas tes dapat menggunakan rumus sebagai berikut :

$$Yr_{pbi} = \frac{Mp - Mt}{SDt} \sqrt{\frac{p}{q}}$$

Keterangan :

Yr_{pbi} : korelasi point biserial

Mp : skor rata-rata dari subjek yang menjawab betul bagi item yang dicari validitasnya

Mt : rerata skor total

SDt : standar deviasi dari skor total

p : proporsi siswa yang menjawab benar

q : proporsi siswa yang menjawab salah

kriteria:

$Yr_{pbi} > r_t$: valid

$Yr_{pbi} < r_t$: tidak valid

Reliabilitas

Arikunto (2016) mengatakan bahwa untuk mengetahui reliabilitas dapat menggunakan rumus Kuder-Richardson 20 sebagai berikut :

$$r_{11} = \left(\frac{n}{n-1} \right) \left(\frac{s^2 \Sigma pq}{s^2} \right) \dots\dots\dots 1)$$

Keterangan:

r_{11} : reliabilitas tes secara keseluruhan

n : banyak item

p : proporsi subjek yang menjawab item dengan benar

q : proporsi subjek yang menjawab item dengan salah

s : standart deviasi dari tes

Koefisien reliabilitas yang mendekati angka 1,0 merupakan indeks reliabilitas tinggi (Sudjana, 2002).

Tingkat kesukaran

Menurut pendapat Kusaeri dan Suprananto (2012) bahwa semakin besar indeks tingkat kesukaran yang diperoleh dari hasil perhitungan maka semakin mudah soal tersebut. Menurut Arikunto (2016) rumus yang digunakan dalam mencari tingkat kesukaran soal sebagai berikut :

$$P = \frac{B}{JS} \dots\dots\dots 2)$$

Keterangan:

P : indeks kesukaran

B : banyaknya siswa yang menjawab soal dengan benar

JS : jumlah seluruh peserta didik

Daya pembeda

Menurut pendapat Kusaeri dan Suprananto (2012) semakin tinggi indeks daya pembeda suatu soal maka semakin tinggi kemampuan soal yang bersangkutan dengan membedakan siswa yang telah memahami materi dengan siswa yang belum memahami materi. Menurut Arikunto (2016) rumus yang dapat digunakan menghitung daya pembeda sebagai berikut:

$$D = \frac{B_A}{J_A} - \frac{B_B}{J_B} = P_A - P_B \dots\dots\dots 3)$$

Keterangan:

- D : daya pembeda
- BA : banyaknya siswa kelompok atas yang menjawab benar
- BB : banyaknya siswa kelompok bawah yang menjawab benar
- JA : banyaknya siswa kelompok atas
- JB : banyaknya siswa kelompok bawah
- PA : proporsi siswa kelompok atas yang menjawab benar
- PB : proporsi siswa kelompok bawah yang menjawab benar

Fungsi pengecoh

Rumus yang digunakan untuk mengukur fungsi pengecoh menurut Arifin (2009:279) sebagai berikut:

$$IP = \frac{p}{(N-B)(n-1)} \times 100\% \dots\dots\dots 4)$$

Keterangan :

- IP : indeks pengecoh
- P : jumlah peserta didik yang memilih pengecoh
- N : jumlah peserta didik yang ikut tes
- B : jumlah peserta didik yang menjawab dengan benar setiap butir soal
- n : jumlah jawaban alternative (opsi)
- 1 : bilangan tetap

Semakin banyak peserta didik yang mengikuti tes memilih pengecoh, maka semakin baik pengecoh menjalankan fungsinya. Menurut Arifin (2014:80) kualitas pengecoh berdasarkan indeks pengecoh sebagai berikut:

Table 2. Indeks Pengecoh

| Indeks Pengecoh | Kualitas |
|---------------------------|--------------|
| Indeks Pengecoh: 76%-125% | Sangat baik |
| 51%-75% atau 126%-150% | Baik |
| 26%-50% atau 151-175% | Kurang baik |
| 0%-25% atau 175%-200% | Jelek |
| Lebih dari 200% | Sangat jelek |

HASI DAN PEMBAHASAN

Analisis butir soal bertujuan untuk mengetahui kualitas butir soal yang digunakan. Butir soal ulangan akhir semester genap merupakan soal bentuk soal objektif sebanyak 40 butir soal, dari butir tersebut akan dianalisis mulai dari validitas soal, reliabilitas, tingkat kesukaran soal, daya pembeda soal dan fungsi pengecoh soal yang menggunakan program ANATES V4.0.9.

Validitas

Validitas soal dihitung dengan rumus korelasi point biserial, kemudian hasil dikonsultasikan dengan rtabel (*product moment*) taraf signifikansi 5%, dengan kriteria apabila $r_{hitung} > r_{tabel}$ maka butir soal valid. Interpretasi koefisien validitas dianggap valid, apabila koefisien validitas diatas 0,304. Hasil analisis dengan menggunakan ANATES V4.0.9 maka diperoleh pada table 3.

Tabel 3. Hasil Analisis Validitas Butir Soal UAS Kearsipan

| Indeks validitas | Butir soal | Jumlah | Presentase (%) |
|------------------|----------------------------------------------------------------------------------------------------------|--------|----------------|
| Valid | 1, 2, 4, 5, 6, 7, 10, 14, 15, 16, 17, 18, 19, 20, 21, 24, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 39, 40 | 28 | 70% |
| Tidak valid | 3, 8, 9, 11, 12, 13, 22, 23, 25, 26, 27, 38 | 12 | 30% |
| jumlah | | 40 | 100% |

Berdasarkan hasil perhitungan pada Tabel 3 menunjukkan bahwa 40 butir soal validitas dalam kriteria valid berjumlah 28 butir soal atau sebesar 70%, selanjutnya kriteria tidak valid berjumlah 12 butir soal atau sebesar 30%, Jadi secara keseluruhan validitas butir soal Ulangan Akhir Semester genap mata pelajaran Kearsipan tahun ajaran 2019/2020 mayoritas kategori valid.

Tes dikatakan baik apabila memiliki validitas yang tinggi (Bahar et al., 2021; Fajaruddin et al., 2021; Matondang, 2009). Berdasarkan hasil perhitungan tersebut menunjukkan bahwa sebanyak 40 butir soal yang dikatakan valid sebanyak 28 butir soal atau 70% yang artinya sebanyak 28 butir soal berkorelasi signifikan dengan masing-masing skor total yang telah diperoleh dan soal tersebut dikatakan baik karena memiliki validitas yang tinggi sehingga soal tersebut mampu mengukur dengan tepa tapa yang hendak ditegur seperti penelitian yang dilakukan oleh Sari dan Surtani (2020) semakin tinggi validitas suatu instrumen, berarti semakin baik pula tingkat kebermaknaan maupun kegunaannya. Untuk soal yang dikatakan tidak valid sebanyak 12 butir soal atau 13% yang artinya soal tersebut tidak mengukur dengan tepat apa yang telah diukur. Hal ini didukung dengan pendapat (Bahar et al., 2021; Fajaruddin et al., 2021; Setiawan et al., 2020) yang menyatakan bahwa suatu tes dikatakan baik apabila tes tersebut dapat mengukur apa yang hendak diukur.

Faktor-faktor yang mempengaruhi tidak validnya suatu butir soal menurut Sukardi (2008) yaitu faktor dari dalam tes, faktor yang berasal dari administrasi dan skor, dan faktor yang berasal dari jawaban peserta didik. Dari faktor ketidakvalidan tersebut, untuk validitas soal Ulangan Akhir Semester genap mata pelajaran Kearsipan mungkin disebabkan salah satunya faktor yang berasal dari jawaban peserta didik yaitu pada masa pandemi saat ini, berbeda pada ulangan akhir semester sebelumnya, dimana sebelumnya ujian dilaksanakan secara *offline* yang artinya peserta didik melaksanakan ujian di sekolah, namun kini ujian dilaksanakan secara online, semua kegiatan dilaksanakan dirumah, sehingga pada saat mengerjakan soal beberapa peserta didik mengeluhkan lingkungan rumah dalam mengerjakan tes terlalu ramai dan peserta didik mengeluhkan jangkauan sinyal yang buruk dengan batas waktu pengerjaan yang terus berjalan sehingga mempengaruhi siswa tidak dapat berkonsentrasi dengan baik. Maka hal ini dapat melibatkan banyaknya peserta didik yang menjawab soal tes dengan pilihan jawaban yang salah, dan juga banyaknya peserta didik yang mendapatkan nilai di bawah KKM.

Reliabilitas

Suatu konstruk atau dapat dikatakan reliable jika memberikan Cronbach Alpha > 0,70 (Nunnally, 1981). Warju et al. (2020) dalam penelitiannya menyimpulkan bahwa reliabilitas merupakan sejauh mana suatu butir soal tersebut mampu menghasilkan hasil yang sama pada uji coba yang secara berulang. Lebih singkatnya, reliabilitas berkaitan stabilitas atau konsistensi skor dari waktu ke waktu lintas penilai.

Berdasarkan hasil perhitungan yang menggunakan program ANATES V4.0.9, koefisien Alpha=0,81. Artinya menunjukkan bahwa perangkat tes ulangan akhir semester genap mata pelajaran kearsipan memiliki tingkat reliabilitas yang tinggi karena tingkat reliabilitas 0,81>0,70. Hal ini sesuai dengan pendapat (Siyoto & Sodik, 2015) bahwa butir soal yang baik merupakan soal yang memiliki koefisien reliabilitas lebih dari sama dengan 0,70. Didukung dengan Departemen Pendidikan Nasional Republik Indonesia (2008) bahwa semakin tinggi koefisien reliabilitas suatu tes, semakin tinggi pula keajegan atau ketepatannya”

Menurut pendapat Kusaeri dan Suprananto (2012) faktor-faktor yang mempengaruhi reliabilitas adalah (1) Semakin banyak jumlah butir soal, maka semakin ajeg suatu tes; (2) Semakin lama waktu tes, semakin ajeg; (3) Semakin sempit range kesukaran butir soal, semakin besar keajegan; (4) Soal-soal yang paling berhubungan akan mengurangi keajegan; (5) Semakin objektif dalam pemberian skor semakin besar keajegan.

Dari penjelasan faktor tersebut yang mempengaruhi tingginya reliabilitas, mungkin disebabkan sempitnya range tingkat kesukaran. Dapat dilihat Tabel. 4 pada tingkat kesukaran terdapat satu soal dalam kategori sukar, yaitu butir soal nomor 22.

Tingkat kesukaran

Hasil tingkat kesukaran ulangan akhir semester genap mata pelajaran kearsipan yang dapat dikategorikan menurut Sudjana (2010) seperti pada Tabel 4.

Tabel 4. Hasil Analisis Tingkat Kesukaan Butir Soal Uas Kearsipan

| Tingkat kesukaran | | Nomor soal | Jumlah | (%) |
|-------------------|----------|--------------------------------------------------------------------------|--------|------|
| Rentang | Kriteria | | | |
| 0.00-0.30 | Sukar | 22 | 1 | 2,5 |
| 0.30-0.70 | Sedang | 5, 6, 8, 11, 12, 13, 17, 21, 23, 24, 26, 27, 28, 31, 32, 34, 38, 39, 40 | 19 | 47,5 |
| 0.71-1.00 | Mudah | 1, 2, 3, 4, 7, 9, 10, 14, 15, 16, 18, 19, 20, 25, 29, 30, 33, 35, 36, 37 | 20 | 50 |
| Jumlah | | | 40 | 100 |

Pada Tabel 4 menunjukkan bahwa tingkat kesukaran terdiri dari tiga kriteria antara lain mudah, sedang, dan sukar. Tingkat kesukaran pada kriteria mudah sebanyak 20 butir soal atau 50%, kriteria sedang sebanyak 19 butir soal atau 47,5% dan kriteria sukar sebanyak 1 butir soal atau 2,5%. Tingkat kesukaran soal untuk menentukan proporsi mudah, sedang dan sukar. Dalam penelitian Yuslita et al. (2016) menyimpulkan bahwa dalam menyusun sebuah tes jangan sampai tes tersebut terlalu sukar atau terlalu mudah, arena soa ujian yang baik adalah soal yang tidak terlalu mudah dan tidak terlalu sedang.

Berdasarkan hasil analisis pada tingkat kesukaran ulangan akhir semester genap mata pelajaran kearsipan secara keseluruhan tingkat kesukaran soal menunjukkan bahwa mayoritas soal berada pada kategori mudah sebanyak 50%. Dan ditinjau dari proporsi soal perbandingan antara mudah:sedang:sukar pada soal ulangan akhir semester genap mata pelajaran kearsipan yaitu 5,0:4,75:0,25. Pendapat Sudjana (2010) perbandingan butir soal mudah:sedang:sedang dapat dibuat 1:1:1, 3:4:3 atau 3:5:2. Artinya soal kearsipan pada penelitian ini belum seimbang. Butir soal tingkat kesukaran ulangan akhir semester genap mata pelajaran kearsipan dapat menggunakan perbandingan 3:5:2 yang berarti 40 soal dapat dibagi menjadi 12 soal kategori mudah, 20 soal kategori sedang, dan 8 soal kategori sukar, juga dapat menggunakan perbandingan 3:4:3 yang berarti 40 butir soal dapat dibagi 12 butir soal mudah, 16 butir soal sedang dan 12 butir soal sukar. Atau mungkin dibagi rata perbandingan 1:1:1. Sehingga soal yang tidak seimbang dengan proporsinya, apabila soal terlalu banyak dalam kategori mudah tidak merangsang berfikir siswa. Soal yang baik merupakan soal yang tidak terlalu mudah dan tidak terlalu sukar (Halik et al., 2019; Selvia et al., 2019). Soal yang terlalu mudah tidak akan merangsang peserta didik untuk mempertinggi usahanya dalam memecahkan soal. Sebaliknya apabila soal terlalu sukar akan menyebabkan peserta didik merasa putus asa dan tidak memiliki semangat untuk mencoba lagi karena di luar jangkauannya.

Soal yang berkategori mudah mungkin dapat disebabkan fungsi pengecoh tidak berfungsi. Dilihat dari hasil *output* tabel 5 yang menunjukkan bahwa terdapat tujuh butir soal yang secara keseluruhan fungsi pengecohnya tidak berfungsi yaitu nomor 5, 10, 19, 26, 32, 38, dan 40. Seperti yang dikemukakan oleh Departemen Pendidikan Nasional Republik Indonesia (2008) bahwa suatu soal dalam kategori mudah disebabkan salah satunya pengecoh butir soal yang tidak berfungsi.

Daya Pembeda

Hasil analisis daya pembeda pada butir soal ulangan akhir semester genap mata pelajaran kearsipan kemudian dikategorikan menurut Arikunto (2013) sebagai Tabel 5.

Tabel 5 tersebut menunjukkan bahwa terdiri dari lima kriteria antara lain kriteria jelek, cukup, baik, sangat baik dan tidak baik. Daya pembeda pada kriteria jelek sebanyak 7 butir soal atau 17,5%, kemudian daya pembeda pada kriteria cukup sebanyak 15 butir soal atau 37,5%, dan daya pembeda pada kriteria baik sebanyak 18 butir soal atau 45%. Sehingga daya pembeda mayoritas pada kategori baik sebanyak 18 butir soal atau 45%, hal ini bermakna bahwa peserta didik yang memiliki skor tinggi telah menjawab soal dengan benar dan peserta didik yang memiliki skor rendah menjawab dengan salah. Sehingga artinya soal tersebut mampu membedakan antara peserta didik yang sudah menguasai materi dengan yang belum menguasai materi.

Tabel 5. Analisis Daya Pembeda

| Daya pembeda | | Nomor soal | Jumlah | (%) |
|--------------|-------------------|---------------------------------------------------------------------|--------|------|
| Rentang | Kriteria | | | |
| 0,00-0,20 | Sangat Tidak Baik | 9, 11, 12, 18, 23, 25, 26 | 7 | 17,5 |
| 0,21-0,40 | Cukup | 1, 3, 4, 7, 8, 10, 13, 19, 20, 22, 27, 30, 33, 37, 38 | 15 | 37,5 |
| 0,41-0,70 | Baik | 2, 5, 6, 14, 15, 16, 17, 21, 24, 28, 29, 31, 32, 34, 35, 36, 39, 40 | 18 | 45 |
| 0,71-1,00 | Sangat baik | | | |
| Negatif (-) | Tidak baik | | | |
| Jumlah | | | 40 | 100 |

Sedangkan butir soal yang memiliki daya pembeda dalam kategori jelek sebanyak 7 butir soal atau sebesar 17,5%, faktor yang mempengaruhi daya pembeda menurut Departemen Pendidikan Nasional Republik Indonesia (2008) yaitu butir soal yang tidak dapat membedakan peserta didik yang berkemampuan tinggi dan rendah disebabkan oleh beberapa faktor, diantaranya: (1) Kunci jawaban butir soal itu tidak tepat; (2) Butir soal itu memiliki 2 (dua) atau lebih kunci jawaban yang benar; (3) Kompetensi yang diukur tidak jelas; (4) Pengecohnya tidak berfungsi; (5) Materi yang ditanyakan terlalu sulit sehingga banyak peserta didik yang menebak; (6) Sebagian besar siswa yang memahami materi yang ditanyakan berfikir ada yang salah informasi dalam butir soalnya.

Dari penjelasan diatas yang mempengaruhi daya pembeda berkategori jelek disebabkan karena beberapa pengecoh tiap butir soal tidak berfungsi yaitu no 9, 11, 12, 18, 23, 25, dan pengecoh yang tidak berfungsi secara keseluruhan no 26, sehingga perlu dilakukan perbaikan. Oleh karena itu soal yang memiliki daya pembeda yang baik seperti nomor 24 dan 35 sebaiknya soal tersebut disimpan dibank soal karena fungsi pengecoh berfungsi secara keseluruhan. Sedangkan untuk daya pembeda dalam kategori baik sebanyak 16 butir soal (nomor 2, 5, 6, 14, 15, 16, 17, 21, 28, 29, 31, 32, 34, 36, 39, 40) dan daya pembeda berkategori cukup sebanyak 15 butir (nomor 1, 3, 4, 7, 8, 10, 13, 19, 20, 22, 27, 30, 33, 37, 38) sebaiknya soal tersebut diperbaiki karena sebagian fungsi pengecoh tidak berfungsi atau tiap butir soal pengecoh tidak berfungsi secara keseluruhan.

Fungsi Pengecoh

Analisis pengecoh (*distractor*) pada butir soal Ulangan Akhir Semester genap mata pelajaran kearsipan menggunakan program ANATES V.4.0.9. Hasil analisis fungsi pengecoh pada butir soal Ulangan Akhir Semester genap mata pelajaran Kearsipan kemudian dikategorikan seperti pada Tabel 6.

Tabel 6. Hasil Analisis Fungsi Pengecoh Butir Soal UAS Kearsipan

| Fungsi pengecoh | | Nomor soal | Jumlah | (%) |
|-----------------|-----|---------------------------------------------------------|--------|------|
| Jumlah | % | | | |
| 4 | 100 | 24, 35 | 2 | 5 |
| 3 | 75 | 11,12, 27, 37 | 4 | 10 |
| 2 | 50 | 1, 3, 4, 8, 9, 16, 17, 23, 31, 34, 36, 39 | 12 | 30 |
| 1 | 25 | 2, 6, 7, 13, 14, 15, 18, 20, 21, 22, 25, 28, 29, 30, 33 | 15 | 37,5 |
| 0 | 0 | 5, 10, 19, 26, 32, 38, 40 | 7 | 17,5 |
| Jumlah | | | 40 | 100 |

Tabel 6 memiliki fungsi pengecoh sebanyak 5 opsi, dengan 4 fungsi pengecoh dan 1 soal kunci jawaban. Sebanyak 2 (5%) butir soal fungsi pengecoh berfungsi secara menyeluruh di tiap butir soal, sebanyak 4 (10%) butir soal 3 fungsi pengecohnya berfungsi dengan baik di tiap butir soal, sebanyak 12 (30%) butir soal 2 fungsi pengecohnya berfungsi dengan baik tiap butir soal, kemudian sebanyak 15 (37,5%) butir soal 1 fungsi pengecoh yang berfungsi dengan baik tiap butir soal, dan sebanyak 7 (17,5%) butir soal fungsi pengecohnya tidak berfungsi menyeluruh tiap butir soal. Sehingga secara keseluruhan Ulangan Akhir Semester genap mata pelajaran Kearsipan kelas X secara keseluruhan opsi berjumlah 200 (5 opsi pada tiap 40 butir soal) dengan fungsi pengecoh 160 opsi. Pengecoh yang berfungsi dengan baik berjumlah 59 (36,88%) dan pengecoh tidak berfungsi berjumlah 101 (63,13%).

Menurut Arikunto (2013) salah satu penyebab fungsi pengecoh (*distractor*) tidak berfungsi dengan baik yaitu terletak pada rumusan kalimatnya, maka perlu dilakukannya penulisan kembali dengan perubahan yang seperlunya. Berdasarkan hasil telaah yang telah dilakukan faktor yang mempengaruhi tidak terpilih-

nya pengecoh disebabkan karena, opsi jawaban kurang homogen, dan rumusan soal yang belum jelas dengan kunci jawaban. Faktor lain yang tidak berfungsinya suatu pengecoh (*distractor*) berdasarkan penelitian yang dilakukan oleh Ismiyati et al. (2015) penyebab distraktor tidak dipilih oleh peserta tes karena distraktor tersebut sangat terlihat menyesatkan. Sedangkan Rahayu dan Djazari (2016) menyatakan bahwa fungsi pengecoh tidak memiliki daya tarik bagi peserta didik yang kurang menguasai konsep atau materi. Oleh karena itu, pengecoh yang berfungsi dengan baik sebaiknya disimpan dan pengecoh yang tidak berfungsi sebaiknya diperbaiki agar soal dapat berfungsi, sehingga soal tersebut dapat digunakan kembali untuk tes berikutnya.

Validitas, tingkat kesukaran, daya pembeda, dan fungsi pengecoh

Setelah dilakukan analisis validitas, reliabilitas, tingkat kesukaran, daya pembeda dan fungsi pengecoh. Penentuan kualitas soal yang baik dan diterima apabila butir soal valid, tingkat kesukaran berada pada kategori sedang, daya pembeda dikatakan baik, dan fungsi pengecoh berfungsi dengan baik. Seperti penelitian yang dilakukan oleh Rahayu dan Djazari (2016) menyatakan bahwa butir soal dapat disimpan di bank soal apabila memenuhi empat kriteria yaitu butir soal termasuk valid; tingkat kesukaran sedang; daya pembeda dalam kategori sedang, baik dan sangat baik; efektivitas pengecoh termasuk kategori baik dan sangat baik. Sehingga hasil persentase kualitas butir soal Ulangan Akhir Semester genap mata pelajaran Kearsipan secara keseluruhan dapat dilihat pada Tabel 7.

Tabel 7. Distribusi Presentase Kualitas Butir Soal UAS Kearsipan

| Kriteria | Nomor Soal | Jumlah | (%) |
|-------------------------------------|----------------------------------------------------------------------------------------------------------------------------|--------|------|
| Berkualitas (Diterima) | 24, 35 | 2 | 5 |
| Kurang Berkualitas (Diperbaiki) | 1, 2, 3, 4, 5, 6, 7, 8, 10, 11, 12, 13, 14, 15, 16, 17, 19, 20, 21, 23, 27, 28, 29, 30, 31, 32, 33, 34, 36, 37, 38, 39, 40 | 33 | 82,5 |
| Tidak Berkualitas (Dibuang/Diganti) | 9, 18, 22, 25, 26 | 5 | 12,5 |
| Jumlah | | 40 | 100 |

Berdasarkan Tabel 7, distribusi presentase kualitas butir soal UAS Kearsipan sebanyak 2 soal atau 5% yang dapat diterima, sebanyak 33 soal atau 82,5% yang perlu dilakukan perbaikan serta sebanyak 5 soal atau 12,5% yang dibuang dan diganti. Butir soal dapat disimpan di bank soal atau butir soal juga dapat digunakan kembali apabila sudah memenuhi dari empat kriteria yaitu validitas, tingkat kesukaran, daya pembeda dan fungsi pengecoh. Senada dengan penelitian yang dilakukan oleh Astuti (2020) Soal dikatakan berkualitas apabila soal memenuhi karakteristik dalam penilaian butir soal yang meliputi: validitas, reliabilitas, tingkat kesukaran, daya pembeda, dan keefektifan penggunaan distraktor. Dan diperkuat dengan pendapat Supratiknya (2014) bahwa syarat tes yang baik salah satunya menggunakan tahap empiris-statistik, terdapat empat aspek esensial kualitas yang menentukan kinerja tes sebagai alat ukur yaitu validitas, reliabilitas, statistik item tes, dan diskriminasi tes.

Sehingga jika soal Ulangan Akhir Semester genap mata pelajaran kearsipan disesuaikan dengan kriteria validitas butir soal yang valid, dari reliabilitas butir soal dikatakan reliabel, dari tingkat kesukaran butir soal termasuk kriteria sedang atau sesuai dengan proporsinya, kemudian daya pembeda termasuk kriteria baik, dan sangat baik, dan fungsi pengecoh yang berfungsi secara menyeluruh, maka dengan begitu akan menghasilkan kualitas soal yang baik, seperti penelitian Irawati et al. (2020) mengingat pentingnya penilaian akhir semester (pas) maka dalam pelaksanaannya dibutuhkan sebuah instrument soal yang berkualitas sehingga dapat menjamin kualitas butir soal yang diberikan kepada peserta didik.

Berdasarkan hasil penelitian yang telah dilakukan dapat disimpulkan bahwa soal Ulangan Akhir Semester genap mata pelajaran kearsipan tahun ajaran 2019/2020 belum dikatakan baik karena soal yang diterima 2 butir soal atau 5%, dan hampir banyak soal yang harus dilakukan perbaikan serta beberapa soal harus diganti. Soal ulangan akhir semester genap yang dikatakan baik dan diterima sebaiknya soal disimpan di bank soal, sedangkan Soal yang belum dikatakan baik hanya memenuhi dua hingga tiga kriteria dari empat kriteria (validitas, tingkat kesukaran, daya pembeda, dan fungsi pengecoh) sebaiknya diperbaiki dan butir soal yang memenuhi satu kriteria atau tidak memenuhi kriteria secara keseluruhan sebaiknya dibuang kemudian diganti. Butir soal termasuk empat kriteria nomor 24 dan 25. Butir soal termasuk dua dan tiga kriteria yaitu nomor 1, 2, 3, 4, 5, 6, 7, 8, 10, 11, 12, 13, 14, 15, 16, 17, 19, 20, 21, 23, 27, 28, 29, 30, 31, 32,

33, 34, 36, 37, 38, 39, dan 40. Sedangkan butir soal termasuk satu dan tidak memenuhi kriteria secara keseluruhan kriteria yaitu nomor 9, 18, 22, 25, dan 26. Senada dengan penelitian Prastika (2021) bahwa butir soal yang memiliki kualitas kurang baik sebaiknya dilakukan revisi sesuai dengan penyebab soal tersebut menjadi kurang baik dan soal yang tidak baik sebaiknya dibuang atau diganti jika tidak memiliki kemungkinan untuk dirubah.

Penyebab kegagalan dapat dilihat dari validitas, tingkat kesukaran, daya pembeda dan fungsi pengecoh pada soal. Penyebab kegagalan soal yang paling besar dapat dilihat dari fungsi pengecoh dimana dari 40 butir soal, 15 (37,5%) butir soal 1 fungsi pengecoh yang berfungsi dengan baik tiap butir soal, dan sebanyak 7 (17,5%) butir soal fungsi pengecohnya tidak berfungsi menyeluruh tiap butir soal, artinya sebanyak 22 (55%) soal tersebut belum memiliki pengecoh yang berfungsi dengan baik. Penyebab kegagalan yang kedua dapat dilihat dari tingkat kesukaran dimana sebanyak 20 (50%) butir soal termasuk kriteria mudah, dan kriteria sukar sebanyak 1 butir soal (2,5%), artinya sebanyak 21 (52%) butir soal belum sesuai dengan proporsinya sehingga soal tersebut belum mampu mengukur kemampuan siswa. Kegagalan ketiga dapat dilihat dari validitas soal, sebanyak 12 (30%) butir soal tidak valid, artinya soal tersebut belum mengukur apa yang seharusnya diukur. Dari penelitian Muluki (2020) bahwa butir soal yang tidak valid berarti menunjukkan tidak memiliki kesesuaian atau kesejajaran arah dengan skor total. Kegagalan yang terakhir dilihat dari daya pembeda, sebanyak 7 (17,5%) butir soal kriteria tidak baik, artinya soal tersebut belum dapat membedakan siswa yang pandai materi dan yang belum pandai materi.

SIMPULAN

Berdasarkan hasil penelitian dan pembahasan analisis yang telah dilakukan bahwa butir soal pilihan ganda Ulangan Akhir Semester Genap mata pelajaran Kearsipan kelas X OTKP tahun ajaran 2019/2020 di SMK Negeri 1 Salatiga kurang baik, karena pertama, validitas dikatakan valid karena mayoritas soal dalam kategori valid sebanyak 28 soal (70%). Kedua reliabilitas memiliki tingkat reliabilitas yang tinggi karena sebesar $0,87 > 0,70$. Ketiga, tingkat kesukaran belum dikatakan baik karena mayoritas soal termasuk kategori mudah sebanyak 20 soal (50%) dan kategori sedang sebanyak 19 soal (47,5%) artinya belum proporsional antara mudah:sedang:sukar. Keempat, daya pembeda dikatakan baik karena mayoritas dalam kategori baik sebanyak 18 soal (45%) dan kategori cukup 15 soal (37,5%) serta untuk fungsi pengecoh belum dikatakan baik karena butir soal yang berfungsi sangat baik berjumlah 2 soal (5%), baik berjumlah 4 soal (10%), cukup berjumlah 12 soal (30%), tidak baik berjumlah 15 soal (37,5%), dan sangat tidak baik berjumlah 7 soal (17,5%). Kualitas soal Ulangan Akhir Semester ditinjau secara keseluruhan dari validitas, tingkat kesukaran, daya pembeda dan fungsi pengecoh hanya 2 soal (5%) butir soal yang memenuhi empat kriteria artinya soal tersebut diterima, sehingga masih banyak butir soal yang perlu diperbaiki karena butir soal belum memenuhi empat kriteria yaitu validitas, proporsi tingkat kesukaran, daya pembeda dan fungsi pengecoh

Berdasarkan simpulan tersebut, saran yang dapat diberikan yaitu memperbaiki soal yang dianggap tidak baik, kemudian apabila soal sudah diperbaiki kemudian soal diujicobakan sebelum disimpan ke bank soal sekolah, hal ini untuk mengetahui kualitas soal setelah dilakukan perbaikan soal. Untuk soal yang sudah baik dapat disimpan di bank soal untuk diujikan pada tes yang akan datang serta untuk guru sebaiknya melakukan analisis butir soal agar mengetahui kualitas butir soal yang diujikan.

DAFTAR PUSTAKA

- Anderson, S. B., Ball, S., & Murphy, R. T. (1975). *Encyclopedia of educational evaluation: concepts and techniques for evaluating education and training programs*. Jossey-Bass.
- Arifin, Z., & Retnawati, H. (2017). Pengembangan instrumen pengukur higher order thinking skills matematika siswa SMA kelas X. *PYTHAGORAS: Jurnal Pendidikan Matematika*, 12(1), 98. <https://doi.org/10.21831/pg.v12i1.14058>
- Arikunto, S. (2013). *Prosedur penelitian suatu pendekatan praktik*. Rineka Cipta.
- Arikunto, S. (2016). *Dasar-dasar evaluasi pendidikan* (5th ed.). Bumi Aksara.
- Astuti, A. (2020). Analisis soal ujian akhir semester genap mata pelajaran matematika siswa SD Negeri 005 Binuang ajaran 2017/2018. *Jurnal Pendidikan Tambusai*, 4(1), 867-880. <https://doi.org/10.31004/jptam.v4i1.539>

- Bahar, R., Istiyono, E., Widiastuti, W., Munadi, S., Nuryana, Z., & Fajaruddin, S. (2021). Analisis karakteristik soal ujian sekolah hasil musyawarah guru matematika di Tasikmalaya. *AKSIOMA: Jurnal Program Studi Pendidikan Matematika*, 10(4), 2660. <https://doi.org/10.24127/ajpm.v10i4.4359>
- D'Sa, J. L., & Visbal-Dionaldo, M. L. (2017). Analysis of multiple choice questions: item difficulty, discrimination index and distractor efficiency. *Amarjeet Kaur Sandhu*, 9(3), 109.
- Departemen Pendidikan Nasional Republik Indonesia. (2008). *Panduan penulisan butir soal*. Departemen Pendidikan Nasional Republik Indonesia, Direktorat Jenderal Manajemen Pendidikan Dasar dan Menengah.
- Fajaruddin, S., Retnawati, H., Yusron, E., & Sofyaningsih, V. (2021). Exploring the final examination test item characteristics of Pancasila and civic education. *International Journal of Education and Learning*, 3(3), 232–240. <https://doi.org/10.31763/ijele.v3i3.435>
- Halik, A. S., Mania, S., & Nur, F. (2019). Analisis butir soal ujian akhir sekolah (UAS) mata pelajaran matematika pada tahun ajaran 2015/2016 SMP Negeri 36 Makassar. *Al Asma: Journal of Islamic Education*, 1(1), 11. <https://doi.org/10.24252/asma.v1i1.11249>
- Irawati, R., Ekawati, E. Y., & Budiawanti, S. (2020). Analisis butir soal ujian akhir semester gasal menggunakan program Anbuso di SMA Negeri 1 Boyolali tahun ajaran 2019/2020. *Jurnal Materi Dan Pembelajaran Fisika*, 10(1), 11. <https://doi.org/10.20961/jmpf.v10i1.42084>
- Iskandar, A., & Rizal, M. (2017). Analisis kualitas soal di perguruan tinggi berbasis aplikasi tap. *Jurnal Penelitian Dan Evaluasi Pendidikan*, 21(2), 12–23. <https://doi.org/10.21831/pep.v22i1.15609>
- Ismiyati, I., Pramusinto, H., & Korosando, F. (2015). Analisis butir soal ulangan akhir semester mata pelajaran kearsipan. *Majalah Ilmiah Indikator*, 17(2), 64–77.
- Kusaeri, K., & Suprananto, S. (2012). *Pengukuran dan penilaian pendidikan*. Graha Ilmu.
- Matondang, Z. (2009). Validitas dan reliabilitas suatu instrumen penelitian. *Jurnal Tabularasa*, 6(1), 87–97.
- Muluki, A. (2020). Analisis kualitas butir tes semester ganjil mata pelajaran IPA Kelas IV MI Radhiatul Adawiyah. *Jurnal Ilmiah Sekolah Dasar*, 4(1), 86. <https://doi.org/10.23887/jisd.v4i1.23335>
- Negara, H. R. P., & Kurniawati, K. R. A. (2018). Karakteristik butir soal try out ujian nasional SMA mata pelajaran matematika se-Kota Mataram dengan menggunakan program IteMan. *Jurnal VARIAN*, 1(2), 11–21. <https://doi.org/10.30812/varian.v1i2.67>
- Nunnally, J. C. (1981). *Psychometric theory*. Tata Mcgraw Hill Publishing Company Limited.
- Pasi, S. N., & Yusrizal, Y. (2018). Analisis butir soal ujian bahasa Indonesia buatan guru MTsN di Kabupaten Aceh Besar. *Master Bahasa: Jurnal Ilmiah Pendidikan Bahasa Dan Sastra Indonesia*, 6(2), 195–202. <https://doi.org/10.24173/mb.v6i2.11666>
- Prastika, Y. D. (2021). Hubungan minat belajar dan hasil belajar pada mata pelajaran matematika di SMK Yadika Bandar Lampung. *Jurnal Ilmiah Matematika Realistik*, 2(1), 26–32. <https://doi.org/10.33365/ji-mr.v2i1.772>
- Rahayu, R., & Djazari, M. (2016). Analisis kualitas soal pra ujian nasional mata pelajaran ekonomi akuntansi. *Jurnal Pendidikan Akuntansi Indonesia*, 14(1), 85–94. <https://doi.org/10.21831/jpai.v14i1.11370>
- Sari, M. R., & Surtani, S. (2020). Analisis kualitas butir soal ujian akhir semester genap pada mata pelajaran geografi kelas X IIS tahun ajaran 2018/2019 di SMA Pertiwi 1 Padang. *JURNAL BUANA*, 4(2), 273. <https://doi.org/10.24036/student.v4i2.830>
- Selvia, S., Rochmatin, T., & Zanthi, L. S. (2019). Analisis kemampuan pemecahan masalah dan kemampuan penalaran matematik siswa SMP pada materi SPLDV. *JPMI (Jurnal Pembelajaran Matematika Inovatif)*, 2(5), 261. <https://doi.org/10.22460/jpmi.v2i5.p261-270>
- Setiawan, A., Fajaruddin, S., Harun, M. H., Sumiyati, Y., Nartani, C. I., & Fitriana, F. (2020). Character values in the thematic learning of elementary school. *Sekolah Dasar: Kajian Teori Dan Praktik Pendidikan*, 29(2), 158–169. <https://doi.org/10.17977/um009v29i22020p158>
- Siyoto, S., & Sodik, M. A. (2015). *Dasar metodologi penelitian*. Literasi Media Publishing.

- Solichin, M. (2017). Analisis daya beda soal, taraf kesukaran, validitas butir tes, interpretasi hasil tes dan validitas ramalan dalam evaluasi pendidikan. *Dirasat: Jurnal Manajemen Dan Pendidikan Islam*, 2(2), 192–213. <https://doi.org/10.26594/DIRASAT.V2I2.879>
- Sudijono, A. (2005). *Pengantar statistik pendidikan*. PT Raja Grafindo Persada.
- Sudjana, N. (2002). *Penilaian hasil proses belajar mengajar*. PT Remaja Rosdakarya.
- Sudjana, N. (2010). *Penilaian hasil proses belajar mengajar*. PT Remaja Rosdakarya.
- Sugiyono. (2010). *Metode penelitian kuantitatif kualitatif dan R & D*. Alfabeta.
- Sugiyono, S. (2012). *Metode penelitian administrasi dilengkapi dengan metode R dan D*. Alfabeta.
- Sukardi, H. M. (2008). *Evaluasi pendidikan: Prinsip dan operasionalnya*. Bumi Aksara.
- Supandi, S., & Farikhah, L. (2016). Analisis butir soal matematika pada instrumen uji coba materi segitiga. *JIPMat*, 1(1). <https://doi.org/10.26877/jipmat.v1i1.1085>
- Supratiknya, A. (2014). *Pengukuran psikologis*. Universitas Sanata Dharma.
- Warju, W., Ariyanto, S. R., Soeryanto, S., & Trisna, R. A. (2020). Analisis kualitas butir soal tipe HOTS pada kompetensi sistem REM di sekolah menengah kejuruan. *Jurnal Pendidikan Teknologi Dan Kejuruan*, 17(1), 95. <https://doi.org/10.23887/jptk-undiksha.v17i1.22914>
- Yuslita, H., Zulfan, Z., & Arifin, A. (2016). Analisis tingkat kesukaran soal dan daya pembeda soal mata pelajaran sejarah kelas XI semester ganjil di SMA Negeri 5 Banda Aceh tahun pelajaran 2015–2016. *JIM: Jurnal Ilmiah Mahasiswa Pendidikan Sejarah*, 1(1), 131–138. <https://doi.org/10.24815/jimps.v1i1.1784>

Conflict of Interest Statement: The Author(s) declares that the research was conducted in the absence of any commercial or financial relationship that could be construed as a potential conflict of interest.

Copyright: ©Measurement in Educational Research. This is an open access article distributed under the terms of the Creative Commons Attribution 4.0 International Licence (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Measurement in Educational Research is an open access and peer-reviewed journal published by Research and Social Study Institute, Indonesia

Open Access 