

Analisis Butir Soal Tema 3 “Peduli Terhadap Makhluk Hidup” Kelas IV SDN 2 Kertaraharja Tahun Pelajaran 2021/2022

Refa Tiwi Alia Fatin

PGSD, Universitas Pendidikan Indonesia Kampus Serang, Indonesia
refafatin24@upi.edu

Sekarsari Nurdini

PGSD, Universitas Pendidikan Indonesia Kampus Serang, Indonesia
sekarsarinurdini@upi.edu

Siti Farkhanah

PGSD, Universitas Pendidikan Indonesia Kampus Serang, Indonesia
farkhanahsiti02@upi.edu

Bella Belinda

PGSD, Universitas Pendidikan Indonesia Kampus Serang, Indonesia
belindabella@upi.edu

Abstract

This study aims to determine the feasibility of the question items compiled by researchers for multiple choice tests on theme 3 "Caring for Living Things" grade IV elementary school. The feasibility of the question item in question is in terms of the validity of the question, the reliability of the question, the difficulty index of the question, the differentiability, and distractor effectiveness. This research is a descriptive research that is analyzed by quantitative methods. The subjects are grade IV students of SDN 2 Kertaraharja for the 2021/2022 academic year, totaling 40 students. Data collection technique with multiple choice test amounting to 20 items. The data obtained are analyzed with the help of the SPSS 25.0 program. The results showed that (1) The quality of the material aspects and construction aspects in the grain showed that the overall question item was very good. However, the language aspect still does not meet the criteria for writing indicators (2) The quality of the uji validity shows 90% of the items of the question are valid. (3) The reliability tested has an value r_{11} more than 0.70 i.e. 0.763 meaning the question of having high reliability; (4) The difficulty rate of 95% of the questions includes medium question criteria, 5% of difficult questions; (5) The differentiator of the questions shows 95% of the questions are good and 5% are not good; (6) The fecundity of the extractor indicates that 100% of the option is declared to have functioned effectively; and (7) Overall the quality of the

question items is good but the question items number 16 and 19 are not good items.

Keywords: analysis grain of questions, quality, quantity

Pendahuluan

Perkembangan ilmu pengetahuan dan teknologi memberikan dampak yang positif dan negatif dalam berbagai aspek kehidupan. Perkembangan ilmu pengetahuan dan teknologi membawa manusia kedalam persaingan global, sehingga diperlukan sumber daya manusia yang berkualitas agar dapat bersaing dengan negara luar. Kualitas sumber daya manusia dapat ditingkatkan melalui pendidikan.

Menurut Undang-undang No. 20 tahun 2003, pendidikan merupakan usaha sadar dan terencana untuk mengembangkan segala potensi yang dimiliki peserta didik melalui proses pembelajaran. Pendidikan bertujuan untuk mengembangkan potensi anak agar memiliki spiritual keagamaan, pengendalian diri, berkepribadian, memiliki kecerdasan, berakhlak mulia, serta memiliki keterampilan yang diperlukan sebagai anggota masyarakat dan warga negara.

Kualitas pendidikan dipengaruhi oleh peran pendidik dalam proses pembelajaran di sekolah maupun di tingkat universitas (Munthe, 2015). Persepsi yang mengatakan bahwa pendidik hanya sebatas mentransfer ilmu pengetahuan ke peserta didik sudah lama ditinggalkan. Sejalan dengan perkembangan kurikulum, peran pendidik bukan hanya sebatas mengajar (dalam konteks mentrasfer ilmu pengetahuan), tetapi juga membimbing, mengembangkan potensi peserta didik, dan mengevaluasi segala proses dan output pembelajaran. urgensi evaluasi dalam proses dan output pembelajaran ialah untuk mengetahui keberhasilan proses pembelajaran yang telah dilakukan (Arifin, 2016).

Arikunto (2012) mengartikan evaluasi sebagai sebuah proses pengumpulan data untuk menentukan sejauh mana, dalam hal apa, dan menunjukkan bagaimana tujuan pendidikan itu telah tercapai. Dalam evaluasi hasil proses pembelajaran, terdapat dua teknik yang dapat digunakan yaitu tes dan non-tes. Tes merupakan sejumlah pertanyaan yang harus diberikan jawaban dan tanggapan dengan tujuan untuk mengukur tingkat kemampuan seseorang (Widoyoko, 2015). Teknik tes merupakan teknik yang sering digunakan seorang pendidik untuk mengetahui sejauh mana kemampuan pesertadidik atau mahasiswanya dalam menguasai materi. Tes yang dilakukan pendidik dapat berupa ulangan harian, ulangan tengah semester dan ulangan akhir semester.

Ulangan merupakan suatu proses untuk melihat ketercapaian kompetensi yang telah dikuasai oleh peserta didik pada aspek hasil belajar tertentu. Ulangan ini juga dapat dikatakan sebagai evaluasi formatif, hal ini dikarenakan guru ingin mengetahui setiap penguasaan materi setiap bab yang telah diajarkan oleh guru (Fahmi & Hidayat, 2014). Dalam memberikan ulangan kepada peserta didik salah satu instrumen yang diberikan kepada peserta didik yaitu soal pilihan ganda (objektif). Penggunaan instrumen pilihan ganda yang akan diberikan kepada peserta didik harus memiliki kualitas yang baik.

Seorang pendidik harus memiliki kemampuan dalam menyusun instrumen tes, sehingga tes tersebut dapat mengukur kemampuan peserta didik. Tes yang baik akan memberikan gambaran yang valid mengenai hasil belajar peserta didik. Maka dari itu, suatu tes haruslah berkualitas. Suatu tes dikatakan berkualitas apabila tes tersebut telah valid dan reliabel (Jurnal, 2018).

Suatu tes dikatakan valid jika tes yang digunakan dapat dengan tepat mengukur sesuatu yang ingin diukur. Tingkat validitas diperoleh dengan membandingkan probabilitas nilai r_{hitung} dengan r_{tabel} dan Bila tingkat signifikansi atau kesalahan \leq

0,05 maka alat ukur tersebut dikatakan valid (Qomusuddin, 2019). Selain valid, tes juga harus reliabel. Tes yang dikatakan reliabel, ketika tes tersebut diujikan berulang kali dan menunjukkan hasil yang relatif tetap. Suatu instrumen dikatakan reliabel saat nilai koefisien reliabilitas Spearman-Brown lebih dari 0,70, $r_{11} > 0,70$ (Fraenkel, Wallen, & Hyun, 2012). Fatimah & Alfath (2019) mengungkapkan pendidik perlu melakukan analisis butir soal guna mengetahui kualitas soal, layak atau tidak untuk digunakan yaitu dengan analisis butir soal ditinjau dari segi indeks kesukaran soal, daya beda soal, efektivitas pengecoh pada soal pilihan berganda, validitas, dan reliabilitas. Soal yang dianalisis adalah pilihan ganda sebanyak 20 butir soal pada Tema 3 “Peduli Terhadap Makhluk Hidup”.

Metode

Penelitian ini termasuk dalam jenis penelitian deskriptif melalui pendekatan kuantitatif. Penelitian deskriptif kuantitatif merupakan penelitian yang menekankan pada analisis data dengan menggunakan perhitungan angka atau statistik (Daryanto, 2012). Tujuan dari penelitian deskriptif kuantitatif adalah untuk mendeskripsikan suatu objek/kegiatan yang menjadi perhatian peneliti. Penelitian ini dilaksanakan pada bulan April 2022. Prosedur atau tahapan penelitian pada penelitian kali ini sebagai berikut: (1) menyusun instrumen atau hal-hal yang berkaitan dengan pengukuran, (2) mengumpulkan data, (3) mendeskripsikan data yang telah dikumpulkan, (4) menganalisis data, (5) merumuskan simpulan dari hasil analisis. Data yang diperoleh dalam bentuk angka akan diidentifikasi untuk ditarik kesimpulan dengan bantuan SPSS 25.0 untuk memudahkan proses kuantifikasi, Peneliti menggunakan SPSS 25.0

Partisipan

Subjek pada penelitian ini adalah peserta didik kelas IV SDN 2 Kertaraharja tahun pelajaran 2021/2022, yang berjumlah 40 peserta didik.

Instrumen

Instrumen yang digunakan pada penelitian ini adalah tes pilihan ganda sejumlah 20 butir pada tema 3 “Peduli Terhadap Makhluk Hidup” kelas IV sekolah dasar.

Teknik Pengumpulan Data

Teknik pengumpulan data pada penelitian ini berupa tes dalam bentuk 20 butir pilihan ganda yang diujikan ke peserta didik kelas IV SDN 2 Kertaraharja tahun ajaran 2021/2022.

Teknik Analisis Data

Teknik analisis data secara kualitatif dilihat dari segi materi, konstruksi, bahasa. adalah cara peneliti mengolah data sehingga data tersebut menjadi lebih mudah untuk dideskripsikan, dimaknai, serta ditafsirkan.

Analisis Kualitatif

Analisis Soal dari aspek Materi, Konstruksi dan Bahasa

Untuk membantu analisis soal dapat dibuat kartu telaah soal atau tabel rangkuman telaah soal seperti yang terlihat pada Tabel 1.

Tabel 1. Analisis Soal dari Aspek Materi, Konstruksi dan Bahasa

Aspek yang di telaah	Kriteria
Materi	1. Kesesuaian soal dengan indikator
	2. Hanya ada satu kunci jawaban
	3. Pilihan jawaban homogen dan logis
	4. Hanya ada satu kunci jawaban tepat
Konstruksi	1. Pokok soal dirumuskan dengan singkat, jelas dan tegas
	2. Rumusan pokok soal dan pilihan jawaban merupakan pernyataan yang diperlukan saja
	3. Pokok soal tidak memberi petunjuk kunci jawaban
	4. Pokok soal besar dari pernyataan yang bersifat negatif ganda
	5. Gambar, grafik, tabel, diagram atau sejenisnya berfungsi
	6. Panjang pilihan jawaban relatif sama
	7. Pilihan jawaban tidak menggunakan “semua jawaban diatas salah/benar”
	8. Pilihan jawaban yang berbentuk angka/waktu disusun berdasarkan urutan besar kecilnya angka atau kronologisnya
	9. Butir jawaban soal tidak bergantung pada jawaban soal sebelumnya
Bahasa	1. Menggunakan bahasa yang sesuai dengan kaidah bahasa Indonesia
	2. Menggunakan bahasa yang komunikatif
	3. Tidak menggunakan bahasa yang berlaku setempat/tabu
	4. Pilihan jawaban tidak mengulang kata/kelompok kata yang sama, kecuali merupakan satu kesatuan pengertian

Analisis soal dilakukan dengan mencocokkan rumusan soal dengan kriteria yang ada. Apabila rumusan soal sesuai dengan kriteria yang ada maka diberi tanda (√) pada kolom soal. Jika rumusan soal tidak sesuai dengan kriteria yang ada maka diberi tanda silang (X) pada kolom soal dan jika kriteria yang ada tidak terdapat pada soal maka diberi tanda (-).

Analisis Kuantitatif

Validitas

Validitas item instrumen digunakan untuk mengetahui dukungan suatu item terhadap skor total. Untuk menguji validitas setiap butir instrumen, skor-skor yang ada pada butir instrumen yang dimaksud dikorelasikan dengan skor total. Menguji validitas menggunakan bantuan aplikasi IBM SPSS 25.0. Pengolahan data dengan SPSS uji validitas (1) Buka file Uji Coba (2) Dari menu utama SPSS, pilih menu Analyze, kemudian pilih submenu Correlate. Lalu pilih Bivariate. Nampak di layar kotak dialog, dengan pengisian Bivariate Correlations. Masukkan semua item ke kotak variables. Pada bagian Correlation Coefficients beri tanda pada Pearson. Selanjutnya pada Test of Significance pilih two-tailed. Beri tanda pada Flag significant correlations, kemudian klik OK untuk proses data. Output SPSS disajikan pada tabel.

Reliabilitas

Reliabilitas dihitung dengan menggunakan pendekatan konsistensi internal. Formula yang digunakan adalah formula Spearman-Brown yang dikenal dengan istilah teknik belah dua. Model pembelahan yang digunakan adalah model pembelahan awal-akhir. Sebelum menghitung reliabilitas tes, terlebih dahulu dilakukan korelasi hasil pengukuran kedua kelompok soal menggunakan rumus korelasi product momen Pearson, yaitu:

$$r_{xy} = \frac{n \sum xy - (\sum x)(\sum y)}{\sqrt{(n \sum x^2 - (\sum x)^2)(n \sum y^2 - (\sum y)^2)}}$$

Untuk mengetahui tingkat reliabilitas keseluruhan soal, dicari koefisien korelasi dengan menggunakan rumus Spearman brown dengan pembelahan awal-akhir yaitu sebagai berikut:

$$r_{11} = \frac{2r_{xy}}{(1+r_{xy})}$$

r_{xy} = koefisien antara skor-skor setiap belahan tes

r_{11} = koefisien reliabilitas yang sudah disesuaikan

Tingkat Kesukaran Soal

Menghitung indeks kesukaran soal dapat menggunakan rumus sebagai berikut :

$$P_i = \frac{n}{N}$$

P_i = Tingkat kesukaran butir soal ke i

i = nomor butir soal

n = banyaknya peserta didik yang menjawab butir soal dengan benar

N = banyaknya peserta didik yang menjawab butir soal

Kategori: Sukar (indeks < 0,3); Sedang (indeks 0,3 – 0,7); Mudah (indeks > 0,7)

Daya Pembeda Soal

Daya pembeda butir soal dihitung dengan menggunakan rumus:

$$d = \frac{nT}{NT} - \frac{nR}{NR}$$

d = daya beda butir atau indeks diskriminasi butir

nT = jumlah penjawab benar pada kelompok tinggi

NT = jumlah peserta tes dari kelompok berkemampuan tinggi

nR = jumlah penjawab benar pada kelompok kemampuan rendah

NR = jumlah peserta tes dari kelompok berkemampuan rendah

Kategori: Tidak baik (indeks < 0,3); Baik (indeks > 0,3)

Efektivitas Distraktor

Persentase fungsi butir pengecoh (efektivitas distraktor) dihitung dengan menggunakan rumus:

$$\text{Keberfungsian Pengecoh} = \frac{\text{jumlah peserta yang memilih pengecoh}}{\text{jumlah peserta tes keseluruhan}} \times 100\%$$

Hasil

Analisis Kualitatif

Analisis Soal dari aspek Materi, Konstruksi dan Bahasa

Tabel 2. Hasil Analisis Soal dari Aspek Materi, Konstruksi dan Bahasa

No	Aspek yang Ditelaah	Jumlah Soal yang Tidak Sesuai Kriteria	Jumlah Soal yang Dianalisis	Persentase
1.	Materi	-	20	0%
2.	Konstruksi	-	20	0%
3.	Bahasa	3	20	15%

Analisis Kuantitatif

Validitas

Uji validitas butir soal pilihan ganda ini dihitung dengan menggunakan rumus korelasi point biserial dan IBM SPSS 25.0 menghasilkan nilai yang sama .

Tabel 3. Hasil Analisis Soal berdasarkan Validitas

No	Nilai	Butir Soal Nomor	Persentase
1.	> 0,312	1,2,3,4,5,6,7,8,,9,10,11,12,13,14,15,17,18,20	90%
2.	< 0,312	16,19	10%

Reliabilitas

Hasil analisis butir soal tes pilihan ganda sejumlah 20 butir pada tema 3 “Peduli Terhadap Makhluk Hidup” peserta didik kelas IV SDN 2 Kertaraharja tahun pelajaran 2021/2022 yang diujikan memiliki nilai r_{11} lebih besar dari 0,70 yaitu sebesar 0,763 sehingga dapat diartikan bahwa soal tersebut memiliki tingkat reliabilitas yang tinggi atau reliabel.

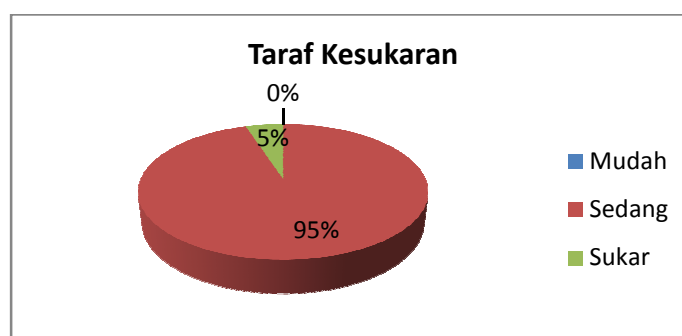
Taraf Kesukaran

Uji taraf kesukaran tes digunakan untuk menjangar banyaknya subjek peserta tes yang dapat mengerjakan tes dengan benar. Taraf kesukaran soal dari 40 butir soal, diperoleh data hasil perhitungan 19 butir soal dalam kategori sedang, 1 butir soal dalam kategori taraf sukar.

Tabel 4. Hasil Analisis Soal berdasarkan Taraf Kesukaran

No	Indeks	Butir Soal Nomor	Persentase
1.	Mudah	-	0%
2.	Sedang	1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,20	95%
3.	Sukar	19	5%

Diagram 1. Persentase Taraf Kesukaran Butir Soal



Daya Pembeda Soal

Hasil daya pembeda soal dari diperoleh data hasil perhitungan 20 butir soal dalam kategori baik, 1 butir soal dalam kategori tidak baik.

Tabel 5. Hasil Analisis Soal berdasarkan Daya Pembeda Soal

No	Daya Pembeda	Butir Soal Nomor	Persentase
1.	Baik	1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,17,18,19,20	95%
2.	Tidak Baik	16	5%

Efektivitas Distraktor

Hasil analisis efektifitas distraktor butir soal tes pilihan ganda sejumlah 20 butir menunjukkan pengecoh pada setiap option butir soal berfungsi semua tanpa terkecuali.

Pembahasan

Analisis Kualitatif

Hasil analisis yang diperoleh menunjukkan bahwa soal telah memenuhi kriteria aspek materi dan konstruksi, namun masih ada soal yang belum memenuhi kriteria aspek bahasa. Butir soal yang tidak menggunakan bahasa yang sesuai dengan kaidah bahasa Indonesia akan membuat peserta didik menjadi bingung maksud dari soal tersebut. Secara keseluruhan, terdapat 17 atau 75% soal yang memenuhi

semua kriteria dari aspek materi, konstruksi dan bahasa. Serta 3 soal lainnya atau 15% soal harus diperbaiki dari segi kualitatifnya.

Analisis Kualitatif

Hasil perhitungan validitas dikonsultasikan ke dalam r_{tabel} pada taraf signifikansi 0,05 atau 5%. Nilai untuk r_{tabel} yakni 0,312. Uji validitas dikatakan baik apabila $r_{hasil} > r_{tabel}$. Dari 20 butir soal yang diuji terdapat 18 butir soal valid karena nilai r_{hasil} lebih besar daripada r_{tabel} . Lalu, 2 butir soal tidak valid yaitu butir soal nomor 16 yang memiliki r_{hasil} 0,254 dan butir soal nomor 19 r_{hasil} 0,115 maka dari itu dikatakan tidak valid karena lebih kecil dari pada r_{tabel} . Hasil analisis validitas butir soal tersebut menunjukkan bahwa 90% butir soal valid dan 10% tidak valid. Artinya butir-butir soal tersebut tidak dapat mengukur apa yang seharusnya diukur oleh soal. Butir soal yang tidak valid harus diperbaiki agar kesalahan pengukuran pada tes selanjutnya tidak terjadi lagi.

Tingkat reliabilitas keseluruhan soal dihitung dengan menggunakan rumus Sperman-Brown, diperoleh hasil yang menunjukkan bahwa soal tes pilihan ganda sejumlah yang diujikan memiliki nilai r_{11} lebih besar dari 0,70 yaitu sebesar 0,763 sehingga dapat diartikan bahwa soal tersebut memiliki tingkat reliabilitas yang tinggi atau reliabel dan dapat dipercaya. Artinya hasil yang diperoleh dari ujian semester genap tersebut dapat menunjukkan kemampuan siswa “yang sebenarnya”.

Hasil analisis yang diperoleh menunjukkan sebanyak 19 soal termasuk kriteria soal sedang, 1 soal yaitu butir soal nomor 19 termasuk kriteria soal sukar. Tingkat kesukaran yang terukur secara kuantitatif merupakan tingkat kesukaran soal berdasarkan tingkat kemampuan siswa. Artinya 19 soal atau 95% soal mampu diselesaikan dengan oleh siswa. Hanya terdapat 1 soal atau 5% soal yang dapat diselesaikan oleh 11 peserta didik dari jumlah peserta didik keseluruhan. Analisis tingkat kesukaran yang diperoleh menunjukkan bahwa sebagian besar soal berada pada kriteria soal sedang. Menurut beberapa ahli sebaiknya sebuah paket soal yang diberikan kepada siswa memiliki keseimbangan antara soal mudah, sedang, dan sukar, sehingga soal pada tema 3 “Peduli Terhadap Makhluk Hidup” peserta didik kelas IV tersebut dapat diperbaiki.

Hasil analisis kuantitatif daya pembeda soal menunjukkan bahwa terdapat 19 soal atau 95% soal dengan daya pembeda yang baik. Serta terdapat 1 soal atau 5% soal dengan daya pembeda yang tidak baik sekali, yaitu soal nomor 16.

Hasil analisis efektivitas distraktor yang telah dilakukan menunjukkan bahwa seluruh option atau 100% option dinyatakan telah berfungsi efektif. Distraktor berfungsi disebabkan karena distraktor yang dibuat dengan baik, sehingga siswa terkecoh dengan option distraktor.

Kesimpulan

Guru yang profesional adalah guru yang menguasai kompetensi pedagogik. Kompetensi pedagogik adalah kemampuan pemahaman terhadap peserta didik, perancangan dan pelaksanaan pembelajaran, evaluasi hasil belajar, dan pengembangan peserta didik untuk mengaktualisasikan berbagai potensi yang dimilikinya. Oleh karena itu, kompetensi seorang guru tidak hanya menyusun alat evaluasi guna mengetahui pencapaian hasil belajar peserta didik akan tetapi juga dapat mengevaluasi apakah evaluasi yang telah disusun sudah dapat menjalankan fungsinya sebagai alat pengukur hasil belajar yang memiliki kualitas yang tinggi sehingga dapat meningkatkan kualitas butir soal, kompetensi guru, dan hasil belajar peserta didik.

Berdasarkan hasil penelitian dan pembahasan, dapat diketahui bahwa: (1) Kualitas aspek materi dan aspek konstruksi pada butir soal pada tema 3 “Peduli Terhadap

Makhluk Hidup”peserta didik kelas IV SDN 2 Kertaraharja tahun pelajaran 2021/2022 menunjukkan bahwa secara keseluruhan butir soal sangat baik. Terdapat 3 butir soal yang masih belum memenuhi kriteria indikator aspek bahasayakni kesalahan kaidah penulisan. Hal ini dianalisis berdasarkan Pedoman Umum Ejaan Bahasa Indonesia (PUEBI) oleh Badan Pengembangan dan Pembinaan Bahasa (BPPB) Kemdikbud;(2) Kualitas uji validitas yang menggunakan rumus korelasi point biserial dan IBM SPSS 25.0 menghasilkan nilai yang sama yaitu menunjukan 18 butir soal valid . Lalu, 2 butir soal tidak valid yaitu butir soal nomor 16 dan 19; (3) Reliabilitas yang diujikan memiliki nilai r_{11} lebih besar dari 0,70 yaitu sebesar 0,763 sehingga dapat diartikan bahwa soal tersebut memiliki tingkat reliabilitas yang tinggi atau reliabel dan dapat dipercaya; (4) Tingkat kesukaran yang terukur secara kuantitatif, hasil analisis yang diperoleh menunjukkan sebanyak 19 soal termasuk kriteria soal sedang, 1 soal termasuk kriteria soal sukar yaitu butir soal nomor 19; (5) Hasil analisis kuantitatif daya pembeda soal menunjukkan bahwa terdapat 19 atau 95% soal dengan daya pembeda yang baik dan 5% tidak baik yakni butir soal nomor 16; (6) Hasil analisis efektivitas distraktor yang telah dilakukan menunjukkan bahwa seluruh option atau 100% option dinyatakan telah berfungsi efektif; dan (7) Secara keseluruhan butir soal yang telah diujikan butir soal 16 dan 19 seharusnya tidak digunakan karena dari hasil analisis dua butir tersebut bisa dikatakan bukan butir-butir soal yang baikharus diganti dengan soal yang baru, sedangkan soal dengan kriteria baik dapat digunakan atau disimpan dalam bank soal dan dapat diperbaharui sesuai dengan kriteria-kriteria berlaku.

Saran

Adapun saran dari tulisan ini adalah (1) Soal tes yang digunakan sebaiknya disesuaikan dengan kriteria aspek materi, konstruksi dan bahasa agar soal tes tersebut memiliki kualitas yang baik dan mampu mengukur kemampuan siswa dengan tepat; (2) Soal tes yang digunakan hendaknya memiliki distribusi jenjang ranah kognitif yang tersebar merata, sehingga soal tersebut dapat mengukur kemampuan siswa dari tingkat yang rendah sampai tingkat yang tinggi; (3) Soal yang digunakan hendaknya disesuaikan dengan tingkat berfikir dan kemampuan siswa, agar soal tersebut dapat berfungsi sebagai alat ukur yang valid dan reliabel, dalam membuktikan validitas lebih tepatnya menggunakan aspek validitas isi, validitas konstruk, dan validitas kriteria; dan (4) Dalam penyusunan soal, sebaiknya dimantapkan dengan melakukan uji coba soal tes yang akan diujikan, melakukan analisis butir soal dan melakukan pengembangan terhadap tes yang disusun berdasarkan analisis butir soal yang telah dilakukan.

Referensi

- Arifin, Z., 2016, *Evaluasi Pembelajaran (Prinsip, Teknik, dan Prosedur)*, Bandung: PT. Remaja Rosdakarya.
- Arikunto, Suharsimi., 2012, *Dasar-Dasar Evaluasi Pendidikan*, Jakarta: Bumi Aksara.
- Daryanto., 2012, *Evaluasi Pendidikan*, Jakarta :Rineka Cipta
- Fahmi, D. A., & Hidayat, M. A., 2014. Korelasi Nilai Ulangan Harian, Nilai Ulangan Tengah Semester, Dan Nilai Ulangan Akhir Semester, Dengan Nilai Raport Semester Gasal Matematika Siswa Kelas Iv Sd Negeri 1 Bae Kudus. *Malih Peddas (Majalah Ilmiah Pendidikan Dasar)*, 4(1), 20-28
<https://doi.org/10.26877/malihpeddas.v4i1.525>

- Fatimah, L. U., & Alfath, K., 2019, Analisis Kesukaran Soal, Daya Pembeda Dan Fungsi Distraktor. AL-MANAR: Jurnal Komunikasi dan Pendidikan Islam, 8(2), 37-64.
- Fraenkel, J. R., Wallen, N. E., & Hyun, H. H., 2012, *How To Design And Evaluate Research In Education*.
- Jurnal, R., 2018, Metode Kuantitatif Dengan Pendekatan Klasik Pada Aplikasi Analisis Butir Soal Sebagai Media Evaluasi Penentuan Soal Yang Berkualitas. *Kilat*, 7(1), 15-23. <https://doi.org/10.33322/kilat.v7i1.104>
- Munthe, A., 2015, Pentingnya Evaluasi Program di Institusi Pendidikan: Sebuah Pengantar, Pengertian, Tujuan dan Manfaat. *Scholaria: Jurnal Pendidikan Dan Kebudayaan*, 5(2), 1-14.
<https://doi.org/10.24246/j.scholaria.2015.v5.i2.p1-14>
- Qomusuddin, I. F., 2019, *Statistik pendidikan (Lengkap dengan aplikasi IMB SPSS Statistic 20.0)*. Deepublish
- Widoyoko, Eko Putro., 2015, *Evaluasi Program Pembelajaran*, Yogyakarta: Pustaka Pelajar.