

DAYA PEMBEDA SOAL, INDEKS KESUKARAN SOAL DAN EFEKTIFITAS OPSI

Murni Ati¹, Devina Santri², Mariam Nasution³

^{1, 2, 3} Universitas Islam Syekh Ali Hasan Ahmad Addary

¹murniati0225@gmail.com, ²devinasantri@gmail.com, ³mariam@uinsyahada.ac.id

ABSTRACT

This study aims to examine the concepts and roles of item discrimination index, item difficulty index, and distractor effectiveness as indicators of item quality in educational assessment through a literature review approach. The research was conducted by collecting and analyzing various relevant sources, including books, scientific journals, research articles, and academic documents discussing item analysis. The obtained data were analyzed descriptively to provide a comprehensive understanding of the functions and characteristics of these three aspects. The findings indicate that the discrimination index reflects an item's ability to distinguish between high- and low-achieving students, the difficulty index measures the level of item difficulty, and distractor effectiveness evaluates the functionality of incorrect answer choices in multiple-choice tests. These three components complement one another in determining the quality of assessment instruments, enabling more valid, reliable, and objective measurement of student learning outcomes. Therefore, understanding and applying the analysis of item discrimination, item difficulty, and distractor effectiveness are essential in developing high-quality assessment instruments that support the improvement of teaching and learning processes as well as student achievement.

Keywords: *Item discrimination index, item difficulty index, distractor effectiveness, item analysis, literature review.*

ABSTRAK

Penelitian ini bertujuan untuk menelaah konsep dan peranan daya pembeda soal, indeks kesukaran soal, serta efektivitas opsi sebagai indikator kualitas butir soal dalam kegiatan evaluasi pembelajaran melalui metode studi literatur. Penelitian dilakukan dengan mengumpulkan dan mengkaji berbagai sumber pustaka yang relevan, seperti buku, jurnal ilmiah, artikel penelitian, dan dokumen akademik yang membahas analisis butir soal. Data yang diperoleh dianalisis secara deskriptif untuk memberikan gambaran yang mendalam mengenai fungsi dan karakteristik ketiga aspek tersebut. Hasil kajian menunjukkan bahwa daya pembeda merupakan kemampuan suatu soal untuk membedakan peserta didik berdasarkan tingkat kemampuannya, indeks kesukaran berfungsi untuk mengukur tingkat kesulitan suatu butir soal, sedangkan efektivitas opsi digunakan untuk menilai keberfungsian pengecoh dalam soal pilihan ganda. Ketiga komponen tersebut memiliki peran yang saling melengkapi dalam menentukan kualitas instrumen penilaian sehingga dapat

menghasilkan pengukuran yang lebih valid, reliabel, dan objektif. Dengan demikian, pemahaman dan penerapan analisis daya pembeda, indeks kesukaran, dan efektivitas opsi menjadi hal yang penting dalam pengembangan instrumen evaluasi yang berkualitas guna mendukung peningkatan mutu pembelajaran dan hasil belajar peserta didik.

Kata Kunci: Daya pembeda soal, indeks kesukaran soal, efektivitas opsi, analisis butir soal, studi literatur.

A. Pendahuluan

Evaluasi pembelajaran merupakan bagian esensial dalam proses pendidikan yang berperan untuk menilai tingkat pencapaian hasil belajar peserta didik. Melalui kegiatan evaluasi, pendidik dapat memperoleh data yang akurat terkait sejauh mana tujuan pembelajaran telah tercapai, sekaligus menjadikannya sebagai landasan dalam menentukan langkah perbaikan proses pembelajaran (Mubarok et al., 2024). Salah satu alat evaluasi yang sering digunakan adalah tes, terutama dalam bentuk soal pilihan ganda.

Meskipun demikian, mutu suatu tes tidak hanya ditentukan oleh jumlah soal yang disajikan, tetapi juga oleh kualitas masing-masing butir soal. Butir soal yang baik harus mampu mengukur kemampuan peserta didik secara tepat, membedakan antara peserta didik dengan tingkat kemampuan tinggi

dan rendah, serta memiliki tingkat kesukaran yang seimbang. Hal ini sejalan dengan pendapat Qadir dkk (2024) yang menyatakan bahwa soal yang baik harus memiliki tingkat kesukaran yang ideal yaitu tidak terlalu mudah atau tidak terlalu sulit dan mampu membedakan peserta didik yang mampu dan tidak mampu. Selain itu, dalam soal pilihan ganda, alternatif jawaban atau pengecoh juga harus dirancang secara efektif agar dapat meminimalkan kemungkinan peserta didik menjawab secara acak.

Dalam kenyataannya, masih banyak soal yang belum memenuhi standar kualitas yang baik. Terdapat soal yang terlalu mudah atau terlalu sulit sehingga kurang mampu memberikan gambaran yang akurat tentang kemampuan peserta didik. Selain itu, ada pula soal yang memiliki daya pembeda rendah sehingga tidak dapat membedakan tingkat penguasaan materi antar

peserta didik. Di samping itu, efektivitas pengecoh sering kali kurang diperhatikan, padahal hal tersebut berpengaruh terhadap keabsahan hasil tes.

Oleh sebab itu, diperlukan analisis butir soal yang meliputi daya pembeda, indeks kesukaran, dan efektivitas opsi. Menurut Dianti et al., (2025) analisis ini bertujuan untuk menilai kualitas setiap butir soal sehingga dapat ditentukan apakah soal tersebut layak digunakan, perlu diperbaiki, atau sebaiknya tidak digunakan. Dengan demikian, hasil evaluasi yang diperoleh menjadi lebih akurat, objektif, dan dapat dipertanggungjawabkan. Berdasarkan urgensi tersebut, makalah ini disusun untuk membahas mengenai daya pembeda soal, indeks kesukaran, dan efektivitas opsi agar mampu menyusun serta menggunakan instrumen evaluasi yang berkualitas dalam proses pembelajaran.

B. Metode Penelitian

Penelitian ini menerapkan metode studi literatur dengan pendekatan deskriptif kualitatif untuk mengkaji berbagai konsep dan teori yang berkaitan dengan daya pembeda

soal, indeks kesukaran soal, dan efektivitas opsi dalam analisis butir soal. Data yang digunakan merupakan data sekunder yang diperoleh dari berbagai sumber ilmiah, seperti buku, jurnal nasional maupun internasional, artikel penelitian, skripsi, tesis, disertasi, serta dokumen akademik lainnya yang relevan dengan topik penelitian. Pengumpulan data dilakukan melalui teknik dokumentasi dengan menelusuri, memilih, dan mengumpulkan berbagai referensi yang membahas kualitas instrumen evaluasi pembelajaran berdasarkan ketiga indikator tersebut. Pemilihan sumber dilakukan dengan mempertimbangkan relevansi, kredibilitas, dan keterkinian informasi agar data yang diperoleh dapat mendukung tujuan penelitian secara optimal.

Data yang telah terkumpul kemudian dianalisis menggunakan teknik analisis isi (content analysis) untuk mengidentifikasi, menafsirkan, dan menyintesis berbagai informasi yang terdapat dalam sumber-sumber literatur. Proses penelitian diawali dengan penentuan fokus kajian, dilanjutkan dengan pencarian dan seleksi literatur yang sesuai,

pengelompokan data berdasarkan tema pembahasan, serta analisis dan perbandingan hasil penelitian terdahulu. Selanjutnya, temuan-temuan yang diperoleh disusun secara sistematis untuk memberikan gambaran yang komprehensif mengenai fungsi dan peranan daya pembeda soal, indeks kesukaran soal, dan efektivitas opsi dalam menentukan kualitas butir soal. Melalui pendekatan ini, penelitian diharapkan mampu menghasilkan pemahaman yang lebih mendalam mengenai pentingnya analisis butir soal sebagai upaya meningkatkan validitas, reliabilitas, dan objektivitas instrumen penilaian dalam proses pembelajaran

C. Hasil Penelitian dan Pembahasan

1. Pengertian daya pembeda soal, indeks kesukaran soal, dan efektivitas opsi

Dalam analisis butir soal, terdapat sejumlah komponen utama yang digunakan untuk menilai kualitas suatu soal, yaitu daya pembeda, indeks kesukaran, dan efektivitas opsi (pengecoh). Ketiga komponen tersebut merupakan indikator penting dalam menentukan kelayakan suatu

butir soal untuk digunakan dalam evaluasi pembelajaran.

a. Daya pembeda soal

Daya pembeda soal adalah kemampuan suatu butir soal dalam membedakan peserta didik yang memiliki tingkat kemampuan tinggi dengan peserta didik yang berkemampuan rendah. Soal dengan daya pembeda yang baik akan mampu memperlihatkan perbedaan penguasaan materi antara kelompok peserta didik yang unggul dan yang kurang unggul. Hal ini sejalan dengan penelitian Mubarak et al., (Mubarak et al., 2024) yang menyatakan bahwa daya pembeda merupakan kemampuan soal dalam mengklasifikasikan peserta didik berdasarkan tingkat kemampuannya melalui indeks diskriminasi.

Daya pembeda menjadi indikator penting dalam menentukan kualitas soal karena berhubungan langsung dengan ketepatan pengukuran hasil belajar. Hal ini sejalan dengan penelitian Mustaqim (2024) yang menyatakan bahwa daya pembeda digunakan untuk membedakan peserta didik dengan kemampuan tinggi dan rendah melalui indeks diskriminasi. Soal yang memiliki daya pembeda yang baik akan

menunjukkan bahwa peserta didik dengan kemampuan tinggi cenderung menjawab dengan benar, sedangkan peserta didik dengan kemampuan rendah cenderung menjawab salah. Dengan demikian, semakin tinggi nilai daya pembeda, maka semakin baik pula kualitas soal dalam mengidentifikasi kemampuan peserta didik.

b. Indeks kesukaran

Indeks kesukaran merupakan ukuran yang digunakan untuk menunjukkan tingkat kesulitan suatu butir soal, apakah termasuk kategori mudah, sedang, atau sulit. Indeks ini umumnya dinyatakan dalam bentuk proporsi dengan rentang nilai antara 0 sampai 1. Indeks kesukaran berfungsi untuk menentukan taraf kesulitan soal, di mana nilai yang mendekati 0 menunjukkan soal yang sangat sulit, sedangkan nilai yang mendekati 1 menunjukkan soal yang sangat mudah (Qadir et al., 2024).

Dalam penyusunan instrumen evaluasi, soal yang baik umumnya berada pada tingkat kesukaran sedang, karena mampu memberikan gambaran yang lebih akurat mengenai kemampuan peserta didik dibandingkan dengan soal yang terlalu mudah atau terlalu sulit. Oleh

karena itu, indeks kesukaran menjadi salah satu aspek penting dalam analisis butir soal guna menjamin kualitas tes yang digunakan.

c. Efektivitas opsi (pengecoh)

Efektivitas opsi atau pengecoh adalah tingkat keberfungsian alternatif jawaban yang salah dalam soal pilihan ganda. Menurut Qadir dkk (Qadir et al., 2024) pengecoh dikatakan efektif apabila mampu menarik perhatian peserta didik yang belum memahami materi, sehingga tidak semua peserta langsung memilih jawaban yang benar.

Pengecoh yang baik adalah yang dipilih oleh sebagian peserta didik, minimal sekitar 5%, sehingga dapat berfungsi sebagai pengalih yang efektif dalam mengukur kemampuan siswa. Apabila suatu opsi tidak pernah dipilih oleh peserta didik, maka opsi tersebut dianggap tidak berfungsi dan perlu diperbaiki. Dengan demikian, efektivitas pengecoh menjadi aspek penting dalam meningkatkan kualitas soal, terutama dalam memastikan bahwa soal mampu mengukur kemampuan peserta didik secara objektif.

2. Cara menghitung daya pembeda soal

Daya pembeda soal digunakan untuk mengetahui sejauh mana suatu butir soal mampu membedakan peserta didik yang memiliki kemampuan tinggi dan yang berkemampuan rendah. Daya pembeda atau disingkat dengan DB berhubungan dengan derajat kemampuan butir membedakan dengan baik perilaku pengambil tes dalam tes yang di kembangkan (Qadir et al., 2024). Tujuan mencari DB adalah untuk menentukan apakah butir soal tersebut memiliki kemampuan membedakan kelompok dalam aspek yang diukur, sesuai dengan perbedaan yang ada pada kelompok tersebut.

Pada umumnya, proses perhitungannya dilakukan dengan membagi peserta tes menjadi dua kelompok, yaitu kelompok atas yang mewakili peserta yang tergolong pandai atau mencapai skor total hasil belajar yang tinggi dan kelompok bawah yang mewakili peserta yang bodoh atau memperoleh skor total hasil belajar yang rendah. Setelah itu, dihitung banyaknya peserta dari masing-masing kelompok yang menjawab soal dengan benar. Perhitungan daya pembeda dapat

dilakukan dengan menggunakan rumus:

$$D = (BA - BB) / N$$

Keterangan:

D = daya pembeda

BA = jumlah peserta dari kelompok atas yang menjawab benar

BB = jumlah peserta dari kelompok bawah yang menjawab benar

N = jumlah peserta dalam masing-masing kelompok

Sumber Nilai daya pembeda berada pada rentang -1,00 hingga +1,00. Nilai yang mendekati 1 menunjukkan bahwa soal memiliki daya pembeda yang sangat baik, sedangkan nilai yang mendekati 0 menandakan bahwa daya pembedanya rendah. Adapun nilai negatif menunjukkan bahwa soal tersebut kurang baik karena lebih banyak dijawab benar oleh kelompok bawah dibandingkan kelompok atas.

Menurut Qadir et al. (2024), kriteria besarnya koefisien daya beda diklasifikasikan menjadi empat kategori. Secara lebih rinci dijelaskan dalam tabel di bawah ini:

Tabel 1. Kategori Daya Pembeda

Kategori daya beda	Koefisien korelasi
Baik	0,40–1,00
Sedang (tidak perlu revisi)	0,30–0,39
Perlu direvisi	0,20–0,29
Tidak baik	-1,00–0,19

Adapun tahapan dalam menghitung daya pembeda meliputi beberapa langkah, yaitu: mengurutkan hasil tes dari nilai tertinggi hingga terendah, menetapkan kelompok atas dan kelompok bawah, menghitung jumlah peserta yang menjawab benar pada masing-masing kelompok, kemudian memasukkan data tersebut ke dalam rumus daya pembeda. Hal ini sejalan dengan pendapat Mustaqim (Mustaqim & Sulisti, 2024) yang menyatakan bahwa daya pembeda diperoleh dengan membandingkan proporsi jawaban benar antara kelompok atas dan kelompok bawah, sehingga dapat diketahui kemampuan soal dalam membedakan tingkat penguasaan peserta didik.

Berdasarkan uraian diatas, dapat disimpulkan bahwa perhitungan daya pembeda dilakukan dengan cara membandingkan jumlah atau proporsi peserta didik yang menjawab benar pada kelompok berkemampuan tinggi dengan kelompok berkemampuan rendah. Tahapannya dimulai dari mengurutkan hasil tes, menetapkan kelompok atas dan kelompok bawah,

lalu menghitung selisih jumlah jawaban benar dari kedua kelompok tersebut menggunakan rumus daya pembeda. Nilai yang diperoleh mencerminkan kemampuan suatu butir soal dalam membedakan tingkat kemampuan peserta didik, sehingga semakin tinggi nilainya, semakin baik pula kualitas soal tersebut.

3. Cara menentukan indeks kesukaran soal

Tingkat kesukaran soal merupakan salah satu aspek penting dalam menilai kualitas alat evaluasi pembelajaran. Menurut Lestari dan Yudhanegara dalam Saputri et al. (2023), konsep ini menggambarkan tingkat mudah atau sulitnya suatu butir soal yang dinyatakan dalam bentuk angka suatu bilangan atau angka.

Soal dapat dikatakan baik apabila tidak tergolong terlalu mudah maupun terlalu sulit. Soal yang terlalu sulit cenderung membuat siswa merasa putus asa karena berada di luar kemampuannya, sedangkan soal yang terlalu mudah kurang mampu mendorong siswa untuk berpikir lebih dalam dan ia berusaha menyelesaikannya (Arikunto, 2020). Tingkat kesukaran

menunjukkan sejauh mana suatu soal dianggap mudah atau sulit oleh siswa. Jika semakin banyak siswa yang mampu menjawab dengan benar, maka soal tersebut tergolong mudah. Sebaliknya, jika hanya sedikit siswa yang dapat menjawab dengan benar, maka soal tersebut termasuk dalam kategori sulit (Setiyawan & Wijayanti, 2020).

Soal yang berkualitas adalah soal yang memiliki tingkat kesukaran yang seimbang, tidak terlalu mudah dan tidak terlalu sulit. Soal yang terlalu mudah cenderung tidak mendorong siswa untuk berpikir lebih jauh, sedangkan soal yang terlalu sulit dapat menurunkan semangat belajar karena berada di luar kemampuan mereka. Untuk mengetahui tingkat mudah atau sulitnya suatu soal, digunakan suatu ukuran yang disebut indeks kesukaran (*difficulty index*). Oleh karena itu, dalam penyusunan soal perlu diperhatikan adanya variasi tingkat kesukaran, mulai dari mudah, sedang, hingga sukar secara proporsional. Penentuan tingkat kesukaran ini didasarkan pada kemampuan siswa dalam menjawab soal, bukan dari sudut pandang guru sebagai penyusun soal.

Bermutu atau tidaknya butir-butir item soal dapat diketahui dari derajat kesukaran yang dimiliki oleh masing-masing butir soal tersebut (Sundayana, 2016). Rumus untuk menentukan indeks kesukaran soal yaitu :

$$TK = \frac{JB_A + JB_B}{2 \cdot JS_A}$$

Keterangan :

TK : Tingkat Kesukaran

JB_A : Jumlah peserta didik kelompok atas yang menjawab benar

JB_B : Jumlah peserta didik kelompok bawah yang menjawab benar

JS_A : Jumlah peserta didik kelompok atas

Adapun kriteria dari indeks kesukaran soal yaitu:

Tabel 2. Kriteria Indeks Kesukaran

Indeks kesukaran soal	Kriteria
TK = 0,00	Terlalu sukar
0,00 < TK ≤ 0,30	Sukar
0,30 < TK ≤ 0,70	Sedang
0,70 < TK < 1,00	Mudah
TK = 1,00	Terlalu mudah

Sumber: (Sundayana, 2016)

4. Cara menganalisis efektivitas opsi (pengecoh)

Efektivitas opsi (pengecoh) dianalisis untuk mengetahui sejauh mana alternatif jawaban yang salah dalam soal pilihan ganda mampu berfungsi dengan baik dalam mengelabui peserta didik yang belum memahami materi. Analisis ini dilakukan berdasarkan data hasil jawaban peserta didik dengan melihat sebaran pilihan pada setiap opsi. Agar dapat melakukan fungsinya untuk mengecoh maka pengecoh harus dibuat semirip mungkin dengan kunci jawaban, baik dari segi bentuk, bahasa, maupun keterkaitan dengan materi yang diujikan.

Berdasarkan hal tersebut, efektivitas pengecoh dapat dilihat dari banyaknya peserta didik yang memilih setiap alternatif jawaban. Pengecoh dikatakan berfungsi dengan baik apabila mampu menarik perhatian sebagian peserta tes, khususnya yang belum memahami materi, yang ditunjukkan dengan adanya peserta yang memilih opsi tersebut. Secara umum, pengecoh dianggap efektif jika dipilih oleh minimal sekitar 5% dari jumlah peserta (Fatimah & Alfath, 2019).

Sebaliknya, jika suatu pengecoh tidak dipilih atau hanya dipilih oleh sangat sedikit peserta, maka pengecoh tersebut belum berfungsi dengan baik dan perlu diperbaiki atau diganti.

Dengan demikian, distraktor dapat diperlakukan dengan tiga cara, yaitu dipertahankan jika sudah berfungsi dengan baik, diganti jika tidak berfungsi, atau diperbaiki jika masih kurang efektif. Oleh karena itu, analisis efektivitas pengecoh sangat penting dalam menilai kualitas soal pilihan ganda, karena dapat membantu penyusun soal memperbaiki dan meningkatkan kualitas instrumen agar lebih tepat dalam mengukur serta membedakan kemampuan siswa.

5. Kriteria soal yang baik ditinjau dari daya pembeda, tingkat kesukaran, dan efektivitas opsi

Kualitas suatu butir soal dapat dilihat dari tiga indikator utama, yaitu daya pembeda, tingkat kesukaran, dan efektivitas opsi (pengecoh). Ketiga aspek ini saling berkaitan dan menjadi dasar dalam menentukan apakah suatu soal layak digunakan atau perlu diperbaiki.

Dilihat dari daya pembeda, soal yang baik adalah soal yang mampu

membedakan dengan jelas antara siswa yang memiliki kemampuan tinggi dan siswa yang berkemampuan rendah. Soal dengan daya pembeda yang tinggi menunjukkan bahwa instrumen tersebut efektif dalam mengukur perbedaan kemampuan peserta didik. Sebaliknya, soal yang memiliki daya pembeda rendah atau bahkan bernilai negatif sebaiknya tidak digunakan atau perlu direvisi, karena tidak dapat menunjukkan perbedaan kemampuan secara tepat.

Selanjutnya, dari segi tingkat kesukaran, soal yang baik adalah soal yang memiliki tingkat kesulitan yang seimbang. Artinya, soal tidak terlalu mudah dan tidak terlalu sulit, sehingga tetap dapat menantang siswa tanpa membuat mereka merasa kesulitan berlebihan. Tingkat kesukaran ini ditentukan berdasarkan kemampuan siswa dalam menjawab soal, bukan dari sudut pandang penyusun soal.

Selain itu, efektivitas opsi pengecoh juga menjadi bagian penting dalam penilaian kualitas soal pilihan ganda. Setiap pilihan jawaban yang salah harus mampu berfungsi sebagai pengecoh yang baik, yaitu menarik perhatian siswa

yang belum memahami materi. Pengecoh yang efektif biasanya dipilih oleh sebagian peserta didik, sedangkan pengecoh yang tidak pernah dipilih menunjukkan bahwa opsi tersebut kurang berfungsi dan perlu diperbaiki.

Dengan memperhatikan ketiga indikator tersebut secara bersama-sama, soal yang disusun akan memiliki kualitas yang lebih baik. Soal tidak hanya mampu mengukur kemampuan siswa secara akurat, tetapi juga dapat membedakan tingkat pemahaman siswa serta mendukung proses evaluasi pembelajaran yang lebih efektif.

D. Kesimpulan

Analisis butir soal yang mencakup daya pembeda, indeks kesukaran, dan efektivitas opsi merupakan tahapan penting dalam menilai mutu instrumen evaluasi pembelajaran. Daya pembeda menunjukkan sejauh mana suatu soal mampu membedakan peserta didik berdasarkan tingkat kemampuannya, indeks kesukaran menggambarkan kategori mudah atau sulitnya soal, sedangkan efektivitas opsi mencerminkan berfungsi atau tidaknya pengecoh dalam soal

pilihan ganda. Proses perhitungan dan analisis ketiga aspek tersebut didasarkan pada data hasil tes peserta didik sehingga menghasilkan informasi yang objektif. Oleh karena itu, soal yang berkualitas adalah soal yang memiliki daya pembeda tinggi, tingkat kesukaran yang sedang, serta pengecoh yang berfungsi dengan baik, sehingga dapat mengukur kemampuan peserta didik secara tepat dan akurat.

DAFTAR PUSTAKA

- Arikunto, S. (2020). Prosedur penelitian suatu pendekatan praktik. In *Jakarta: Rineka Cipta* (p. 172).
- Dianti, K., Ulfah, M., Salam, A., Gunawan, G., & Luthfiyah, L. (2025). Analisis Asesmen Diagnostik, Formatif dan Sumatif Serta Implikasinya terhadap Efektivitas Sistem Evaluasi Pendidikan. 5(2), 555–565.
- Fatimah, L. U., & Alfath, K. (2019). Analisis Kesukaran Soal, Daya Pembeda, dan Fungsi Distraktor. *Jurnal Komunikasi Dan Pendidikan Islam*, 8(2), 37–64.
- Mubarok, A., Khoerotunnisa, N., & Sopyan, A. (2024). Peran Evaluasi Pembelajaran dalam Meningkatkan Hasil Belajar. 8(April 2015), 28286–28290.
- Mustaqim, & Sulisti, H. (2024). Analisis butir soal PAS matematika peminatan: daya pembeda, tingkat kesukaran, dan kualitas pengecoh. *Jurnal Tadris Matematika Fakultas*, 3(2018), 44–56.
- Qadir, A., Huda, N., & Hermina, D. (2024). Analisis Butir Tes: Tingkat Kesukaran, Daya Pembeda dan Efektifitas Pengecoh. *Jurnal Agama, Sosial, Dan Budaya*, 3(3), 1450–1467.
- Saputri, H. A., Zulhijrah, Larasati, N. J., & Shaleh. (2023). Analisis Instrumen Asesmen: Validitas, Reliabilitas, Tingkat Kesukaran, dan Daya Beda Butir Soal. *Didaktik : Jurnal Ilmiah PGSD FKIP Universitas Mandiri*, 9(5), 2986–2995. <https://journal.stkipsubang.ac.id/index.php/didaktik/article/view/2268>
- Setiyawan, R. A., & Wijayanti, P. S. (2020). Analisis kualitas instrumen untuk mengukur kemampuan pemecahan masalah siswa selama pembelajaran daring di masa pandemi. *Lebesgue : Jurnal Ilmiah Pendidikan Matematika, Matematika Dan Statistika*, 1(2), 130–139. <https://l1nq.com/h76j66g>
- Sundayana, R. (2016). *Statistika Penelitian Pendidikan* (ketiga). Alfabeta