

ANALISIS BUTIR SOAL *SPREADSHEET* PADA SISWA SMK SE-JAWA TIMUR KELAS 10 DENGAN PROGRAM ANATES VERSI 4.0 FOR WINDOWS

Aldi Saputro¹, Dimas Virgananda Prakoso², Luqman Hakim³, Vivi Pratiwi⁴
aldi.23070@mhs.unesa.ac.id¹, dimasvirgananda.23074@mhs.unesa.ac.id², luqmanhakim@unesa.ac.id³, vivipratiwi@unesa.ac.id⁴
Universitas Negeri Surabaya

ABSTRAK

Tujuan dari penelitian ini adalah untuk mengevaluasi kualitas soal spreadsheet siswa kelas 10 SMK se-Jawa Timur dengan menggunakan aplikasi ANATES versi 4.0 for Windows. Penelitian ini termasuk dalam kategori deskriptif kuantitatif dan datanya diperoleh dari hasil tes pilihan ganda dan dianalisis secara statistik. Berdasarkan temuan, daya pembeda soal sebagian besar berada pada kategori baik hingga sangat baik, dan tingkat kesukaran berada pada rentang sedang hingga mudah. Reliabilitas pertanyaan berada pada kategori cukup, dan kualitas pengecoh menunjukkan efektivitas yang bervariasi, sebagian besar berada pada kategori baik hingga sangat baik. Analisis ini memberikan wawasan untuk meningkatkan kualitas soal serta meningkatkan validitas dan reliabilitas soal dalam mengukur keterampilan siswa.

Kata Kunci: analisis soal, identifikasi, kesukaran, reliabilitas, ANATES

ANALYSIS OF SPREADSHEET QUESTIONS ON VOCATIONAL HIGH SCHOOL STUDENTS IN EAST JAVA GRADE 10 WITH ANATES PROGRAM VERSION 4.0 FOR WINDOWS

ABSTRACT

The aim of this research is to evaluate the quality of spreadsheet questions for grade 10 vocational school students in East Java using the ANATES version 4.0 for Windows application. This research is included in the quantitative descriptive category and the data was obtained from the results of multiple choicetests and analyzed statistically. Based on the findings, the discriminating power of the questions is mostly in the good to very good category, and the level of difficulty is in the medium to easy range. The reliability of the questions was in the adequate category, and the quality of the distractors showed varying effectiveness, most of which were in the good to very good category. This analysis provides insight into improving the quality of the questions and increasing the validity and reliability of the questions in measuring student skills.

Keywords: question analysis, identification, difficulty, reliability, ANATES

PENDAHULUAN

Menurut Mujiburrahman et al (2023) penilaian merupakan bagian terpenting dari proses pembelajaran, tidak hanya menilai seberapa baik siswa mencapai tujuan pembelajaran, tidak hanya memberikan umpan balik yang

bermanfaat bagi guru dan siswa. Penilaian yang efektif membantu siswa memahami tingkat pencapaian mereka terhadap kompetensi yang diharapkan, serta mengidentifikasi area yang perlu ditingkatkan (IPutu Suardipa et al 2023). Bagi guru, penilaian memberikan data

yang dapat digunakan untuk mengevaluasi keberhasilan metode pengajaran yang diterapkan, mengidentifikasi kekuatan dan kelemahan strategi pembelajaran, serta melakukan perbaikan untuk meningkatkan kualitas pengajaran.

Suatu alat yang paling umum digunakan oleh guru untuk menilai siswa adalah tes Butir et al (2021), yang bisa berupa soal pilihan ganda, esai, atau bentuk evaluasi lainnya. Tujuan dari tes atau ujian adalah untuk mengetahui seberapa memahami siswa materi pelajaran dan kemampuan mereka. Namun, tidak semua soal yang dirancang oleh pendidik memiliki kualitas yang memadai. Terkadang, ada beberapa soal dibuat yang tidak sesuai dengan tujuan pembelajaran yang mana soal itu harus dibuat sesuai dengan tujuan pembelajaran, tidak mampu mengukur kompetensi siswa secara akurat, atau bahkan menimbulkan ambiguitas yang dapat membingungkan siswa. Akibatnya, proses penilaian menjadi kurang efektif dan berpotensi menghasilkan informasi yang kurang sesuai dengan tujuan pembelajaran. Dengan hal ini, diperlukan langkah sistematis untuk memastikan bahwa soal-soal yang digunakan dalam evaluasi benar-benar berkualitas dan sesuai dengan standar yang diharapkan.

Analisis soal adalah proses evaluasi terhadap kualitas butir soal yang bertujuan untuk memastikan bahwa setiap soal dapat mengukur kemampuan siswa secara efektif, adil, dan sesuai dengan apa tujuan pembelajaran (Musyarofah et al 2021; Yusuf et al 2021). Proses ini melibatkan berbagai aspek teknis, seperti daya pembeda, tingkat kesukaran, reliabilitas, dan kualitas pengecoh.

Pengertian daya pembeda terletak pada

kemampuan soal untuk membedakan siswa yang sangat memahami materi dengan siswa yang sangat memahaminya (Ananda Setiyawan et al 2020). Jumlah peserta didik yang mampu menyelesaikan soal dengan benar disebut tingkat kesukaran, yang menunjukkan apakah soal tersebut terlalu mudah, terlalu sulit, atau sesuai dengan tingkatnya. Sementara itu, reliabilitas mengacu pada sejauh mana suatu pertanyaan menghasilkan hasil yang konsisten dalam kondisi serupa. Kualitas pengecoh yang efektif adalah opsi jawaban salah yang dirancang untuk mengarahkan siswa yang kurang memahami materi agar memilih jawaban yang keliru, sehingga membantu memastikan siswa yang menjawab benar benar-benar memahami isinya.

Melalui analisis soal, pendidik dapat mengidentifikasi kelemahan atau kekurangan dalam soal yang telah digunakan (Achmad et al 2022). Sebagai contoh, soal yang terlalu gampang tidak memiliki tantangan yang cukup bagi peserta didik, soal yang terlalu sulit dapat mengurangi dorongan siswa. Soal yang ambigu atau memiliki distraktor yang tidak berfungsi dengan baik dapat menyebabkan siswa menjawab secara tidak konsisten, sehingga hasil evaluasi menjadi kurang valid. Dengan menganalisis dan merevisi soal-soal ini, pendidik dapat meningkatkan kualitas penilaian dan memastikan bahwa setiap soal memenuhi standar kualitas yang diharapkan.

Lebih lanjut, analisis soal tidak hanya bermanfaat untuk meningkatkan kualitas alat evaluasi, tetapi juga memiliki dampak jangka panjang pada proses pembelajaran secara keseluruhan. Soal-soal yang berkualitas memungkinkan pendidik untuk mendapatkan gambaran yang lebih akurat tentang kompetensi

siswa, sehingga dapat merancang strategi pembelajaran yang lebih efektif di masa depan (Mukhlis et al 2024). Bagi siswa, soal yang terstruktur dengan baik akan membantu mereka memahami standar kompetensi yang diharapkan dan mendorong mereka untuk belajar dengan cara yang lebih terarah. Dengan demikian, analisis soal merupakan langkah penting yang tidak boleh diabaikan dalam upaya menciptakan proses pembelajaran yang bermutu tinggi.

Selain itu, dalam konteks pengembangan pendidikan secara umum, analisis soal juga dapat menjadi alat untuk meningkatkan standar mutu pendidikan. Dengan melakukan analisis yang konsisten, institusi pendidikan dapat memastikan bahwa penilaian yang dilakukan tidak hanya mencerminkan pencapaian individu siswa, tetapi juga menjadi indikator keberhasilan kurikulum secara keseluruhan. Proses ini mendukung peningkatan kualitas pendidikan secara berkelanjutan, menciptakan budaya evaluasi yang konstruktif, dan mendorong pengembangan kemampuan berpikir kritis di kalangan siswa. Sebagai hasilnya, analisis soal bukan sekadar langkah teknis, melainkan bagian strategis dalam menciptakan pendidikan yang adil, efektif, dan bermutu tinggi (Al Faraby et al 2024; Moh Rudiniet al 2020; Svensberg et al 2024).

Untuk mencapai tujuan perbaikan proses belajar mengajar dan untuk mengetahui seberapa tinggi kualitas soal, terlebih dahulu harus melakukan analisis terhadap setiap soal. Untuk mengevaluasi kualitas secara kualitatif atau kuantitatif, survei dapat digunakan. Tujuan dari proses menganalisis item kuesioner adalah untuk memperoleh informasi tentang fitur masing-masing item kuesioner. Hal ini

dilakukan untuk menilai kualitas alat dan hasil pembelajaran responden.

Di era digital, kemajuan teknologi telah memperkenalkan berbagai aplikasi yang mempermudah proses analisis soal. Salah satu aplikasi yang paling populer adalah ANATES (Analisis Tes). ANATES adalah perangkat lunak yang dirancang untuk membantu pendidik menganalisis data nilai ujian dengan cepat dan akurat. Aplikasi ini dapat memberikan informasi detail mengenai daya pembeda soal, tingkat kesukaran soal, validitas pada soal, reabilitas soal, dan efektifitas pengecoh. Pendidik dapat membuat

keputusan berdasarkan data saat membuat dan menyempurnakan soal tes. Banyak keuntungan menggunakan ANATES, antara lain kemudahan penggunaan, kecepatan pengolahan data, dan kemudahan memahami hasil analisis. Hal ini menjadikan ANATES alat yang sangat berguna, terutama bagi para pendidik yang ingin meningkatkan kualitas soal penilaian mereka secara sistematis. Namun, untuk mewujudkan potensi penuh dari aplikasi ini, pendidik harus memahami cara kerja ANATES dan langkah-langkah yang diperlukan untuk memanfaatkannya sepenuhnya.

Dalam menyiapkan instrumen pengukuran kemampuan siswa, guru harus mampu menyiapkan instrumen yang sesuai dan baik (Setya Mustafa et al 2022). Penelitian ini menggunakan aplikasi ANATES untuk menguji soal spreadsheet siswa SMK se-Jawa Timur yang disusun oleh guru berdasarkan item spreadsheet siswa SMK se-Jawa Timur untuk memastikan daya pembeda, reliabilitas, tingkat kesukaran, dan kualitas untuk memperoleh informasi tentang kekuatan

dan efektivitas daya pembeda, reliabilitas, tingkat kesukaran dan kualitas pengecoh. Rumusan masalah dalam penelitian ini adalah sebagai berikut. Bagaimana daya pembeda, reliabilitas, tingkat kesukaran, dan kualitas pengecoh soal spreadsheet siswa kelas 10 SMK se-Jawa Timur dengan program aplikasi ANATES versi

4.0 for Windows? Tujuan penelitian ini untuk mengetahui daya pembeda, reliabilitas, tingkat kesukaran, dan kualitas pengecoh soal

spreadsheet pada siswa kelas 10 SMK se-Jawa Timur dengan aplikasi ANATES versi 4.0 for Windows. Penelitian ini tergolong penelitian deskriptif kuantitatif.

METODE PENELITIAN

Penelitian deskriptif kuantitatif bertujuan untuk mendeskripsikan atau mengukur suatu fenomena dengan menggunakan data numerik yang dianalisis secara statistik (Ardiansyah et al 2023). Dengan menggunakan aplikasi ANATES versi 4.0 untuk Windows, metode deskriptif kuantitatif digunakan dalam penelitian ini untuk mengevaluasi kualitas kuesioner spreadsheet yang diberikan kepada siswa di kelas 10 SMK di Jawa Timur.

Fokus penelitian ini adalah soal spreadsheet pada elemen spreadsheet kelas 10 SMK, yang terdiri dari 10 butir soal dengan tes pilihan ganda dengan 5 opsi, yaitu a, b, c, d, dan e. Tujuan penelitian ini adalah untuk menyelidiki, membahas, dan mempelajari seluruh soal berdasarkan hasil analisis dari 10 butir spreadsheet kelas 10 SMK, yang terdiri dari 20 siswa. Penelitian ini juga menggunakan teknik dokumentasi untuk mengumpulkan data. Item dianalisis secara kuantitatif untuk

memberikan informasi guna menentukan daya pembeda, reliabilitas, tingkat kesukaran, dan kualitas pengecoh pada butir soal elemen spreadsheet.

HASIL DAN PEMBAHASAN

Analisis butir soal adalah metode sistematis untuk mengevaluasi kualitas soal untuk mendapatkan informasi yang digunakan untuk mengevaluasi dan menyempurnakan ujian dan soal (Zainal et al 2020). Hasil analisis ini memberikan umpan balik yang bermanfaat bagi pendidik untuk memastikan soal-soal yang dirancang mampu sebagaimana mestinya untuk dapat berfungsi sebagai alat yang akurat dan sah untuk mengukur hasil belajar. Soal yang berfungsi dengan baik harus mampu mengukur pencapaian siswa dengan ketepatan yang tinggi, sesuai dengan tujuan pembelajaran (Khasanah et al 2023; Santosa et al 2022).

Menilai seberapa baik suatu pertanyaan dapat memenuhi fungsinya memerlukan analisis terhadap pertanyaan tersebut. Hal ini dapat dilakukan dengan menggunakan dua pendekatan: analisis kualitatif dan kuantitatif. Dalam penelitian ini, kami mengumpulkan dokumen hasil pengukuran dan melakukan analisis kuantitatif untuk menunjukkan secara empiris kualitas pertanyaan. Sejauh mana setiap item pertanyaan dapat membedakan siswa berkemampuan rendah dan tinggi, tujuannya dari analisis kuantitatif ini adalah untuk mengumpulkan bukti empiris.

ANATES versi 4.0 untuk Windows adalah program analisis soal yang dimaksudkan untuk membantu guru menganalisis soal esai dan reliabilitas, tingkat kesukaran, dan kualitas pengecoh.

Dengan memakai ANATES, pendidik bisa memastikan bahwa soal yang digunakan untuk pengukuran memenuhi standar yang diharapkan. Oleh karena itu, proses analisis ini tidak hanya berkontribusi terhadap peningkatan kualitas soal, namun juga keberhasilan penilaian pembelajaran secara keseluruhan.

Daya Pembeda

Menurut (Ardila Suci Qurrota et al 2022) daya pembeda adalah ukuran seberapa baik suatu pertanyaan agar bisa mengklasifikasikan antara siswa yang memiliki pemahaman rendah dan berkemampuan tinggi. Ini adalah indikator penting dalam evaluasi kualitas soal, karena soal yang baik seharusnya dapat mengidentifikasi perbedaan tingkat pemahaman di antara kelompok siswa. Daya pembeda adalah ukuran seberapa baik suatu pertanyaan agar supaya dapat membedakan antara siswa berintelektual rendah dan berkemampuan tinggi. Hasil yang diperoleh pada tes soal pilihan ganda pada siswa SMK se-Jawa Timur dengan aplikasi ANATES versi 4.0 for Windows sebagai berikut:

Tabel 1. Hasil Daya Pembeda

No Soal	Beda Soal	Indeks DP (%)
1	3	60.00
2	1	20.00
3	0	0.00
4	1	20.00
5	4	80.00
6	0	0.00
7	5	100.00
8	4	80.00
9	1	20.00
10	3	60.00

Hasil ini dari daya pembeda soal ini diperoleh dari jawaban siswa SMK kelas 10 se-Jawa Timur pada soal elemen spreadsheet yang sudah diolah pada Aplikasi ANATES versi 4.0 for Windows. Kriteria indeks daya pembeda dapat dijelaskan pada table berikut:

Tabel 2. Kriteria Indeks Daya Pembeda

No	Indeks Daya Pembeda	Kriteria Daya Pembeda
1	0,00 – 0,20	Jelek
2	0,20 – 0,40	Cukup
3	0,40 – 0,70	Baik
4	0,70 – 1,00	Sangat Baik
5	Negatif	Semuanya Tidak Baik

Tabel di atas menunjukkan dari hasil analisis butir soal spreadsheet kelas 10 SMK dengan ANATES versi 4.0 for Windows. Hasil menunjukkan bahwa daya pembeda dapat dikategorikan dengan baik. Ada tiga soal yang memenuhi kriteria yang sangat baik, yaitu nomor 5, 7, dan 8, dua soal yang memenuhi kriteria yang baik, yaitu nomor 1 dan 10, tiga soal yang memenuhi kriteria yang cukup, yaitu nomor 2, 4, dan 9, dan dua soal yang memenuhi kriteria yang buruk, yaitu nomor 3 dan 6, yang berarti bahwa bagian ini perlu diperbaiki.

Hasil analisis daya pembeda terhadap sepuluh soal menunjukkan bahwa tiga soal (nomor 5, 7, dan 8) memiliki kualitas sangat baik dan dapat mengklasifikasikan terhadap siswa dengan berpengetahuan tinggi dan rendah secara optimal, sehingga dapat digunakan tanpa revisi. Dua soal lainnya (nomor 1 dan

10) termasuk dalam kategori baik, meskipun dapat ditingkatkan sedikit, dan tiga soal (nomor 2, 4, dan 9) tergolong ke dalam kategori cukup, yang memerlukan revisi untuk meningkatkan efektivitasnya, sementara dua soal (nomor 3 dan 6) dengan kualitas jelek harus diperbaiki secara menyeluruh atau diganti karena tidak mampu membedakan siswa secara efektif. Oleh karena itu, diperlukan upaya perbaikan pada soal-soal yang kurang berkualitas, sementara soal yang sudah baik tetap digunakan dengan pemantauan lebih lanjut. Proses analisis ini penting dilakukan secara berkala untuk memastikan kualitas soal yang digunakan dapat mendukung proses penilaian yang valid, adil, dan akurat.

Reliabilitas

Reliabilitas adalah seberapa baik suatu alat pengukuran, seperti tes atau soal pilihan ganda, dapat menghasilkan hasil yang sama setiap kali digunakan dalam kondisi atau kelompok yang sebanding (Wahyu et al 2020). Dalam konteks pengujian, reliabilitas memastikan bahwa soal dapat memberikan hasil yang stabil dan tidak dipengaruhi oleh faktor acak seperti kondisi fisik atau emosional peserta, waktu pelaksanaan, atau lingkungan tes. Tes yang reliabel mencerminkan tingkat kepercayaan terhadap hasil yang diperoleh, sehingga dapat digunakan sebagai dasar pengambilan keputusan yang objektif dan adil. Dengan reliabilitas yang tinggi, hasil evaluasi menjadi lebih konsisten dan dapat diandalkan untuk mengukur kemampuan yang diharapkan (Ndiung et al 2020). Klasifikasi hasil koefisien reliabilitas dapat dilihat pada tabel berikut:

Tabel 3. Klasifikasi Tingkat Reliabilitas

Koefisien Korelasi	Tes Kategori Reliabilitas
0,00 – 0,20	Sangat Rendah
0,20 – 0,40	Rendah
0,40 – 0,70	Cukup
0,70 – 0,90	Tinggi
0,90 – 1,00	Sangat Tinggi

Hasil analisis dari soal spreadsheet pilihan ganda kelas 10 SMK se-Jawa Timur di program aplikasi ANATES versi 4.0 for Windows adalah 0,55, dan berdasarkan tabel klasifikasi tingkat reliabilitas, itu masuk dalam kategori cukup dengan rentang 0,40–0,70. Hasil analisis soal spreadsheet pilihan ganda kelas 10 SMK se-Jawa Timur menunjukkan tingkat reliabilitas yang cukup untuk mengukur kemampuan siswa.

Tingkat Kesukaran

Tingkat kesukaran adalah ukuran yang menunjukkan sejauh mana sebuah soal dapat dijawab dengan benar oleh peserta tes (Wigati et al 2022). Secara umum, tingkat kesukaran menggambarkan kemudahan atau kesulitan suatu soal, yang dinyatakan dalam bentuk indeks atau persentase. Menurut Hendrayadi et al (2024) indeks kesulitan ini umumnya dinyatakan dalam persentase dengan besaran antara 0,00 dan 1,00. Jika indeks kesulitan soal dihitung dari hasil perhitungan yang lebih besar, soal akan lebih mudah. Soal dengan indeks kesulitan yang terlalu rendah atau terlalu tinggi cenderung tidak efektif untuk mengukur

sejauh mana siswa memahami materi yang telah diberikan kepada guru. Oleh karena itu, tingkat kesukaran digunakan untuk mengevaluasi kualitas soal dan memastikan bahwa soal dapat membedakan kemampuan peserta secara tepat (Ramadhan et al 2023).

Tabel berikut menunjukkan tingkat kesukaran soal berdasarkan sedang, sukar, dan mudah. Banyak responden yang mampu menjawab soal pilihan ganda dengan benar.

Tabel 4. Indeks Tingkat Kesukaran Butir Soal

Indeks Tingkat Kesukaran	Kategori
$0,00 \leq TK \leq 0,30$	Sukar
$0,31 \leq TK \leq 0,70$	Sedang
$0,71 \leq TK \leq 1,00$	Mudah

Hasil analisis butir soal pilihan ganda yang dilakukan pada kelas 10 SMK di Jawa Timur menggunakan program aplikasi ANATES versi 4.0 untuk Windows menunjukkan ada empat soal memenuhi kriteria sangat mudah, tiga butir soal memenuhi kriteria mudah, dan tiga butir soal memenuhi kriteria sedang. Hasil analisis tugas standar menengah menunjukkan bahwa tugas tersebut tidak terlalu rumit atau mudah, sehingga siswa dapat menyelesaikannya. Siswa menyerah ketika mereka merasa mereka tidak dapat memecahkan masalah.

Kualitas Pengecoh

Sebagai komponen pendukung utama dalam soal pilihan ganda, pengecoh adalah pilihan jawaban yang salah, tetapi dirancang dengan hati-hati agar masuk akal dan masuk akal bagi siswa yang kurang memahami materi

(Ramadhan et al 2023). Hal ini bertujuan untuk mengurangi kemungkinan jawaban sebenarnya diperoleh melalui keberuntungan atau tebakan semata.

Keberadaan pengecoh yang berkualitas sangat penting karena berkontribusi langsung pada validitas dan reliabilitas soal (Aspriyanti et al 2022). Pengecoh yang efektif tidak hanya dapat mengolompokkan peserta didik yang dapat memahami materi dengan baik dari siswa yang belum memahami materi dengan baik, tetapi juga memberikan gambaran yang lebih mendalam tentang pola kesalahan atau miskonsepsi yang umum terjadi.

Kualitas pengecoh juga memengaruhi daya pembeda soal; dengan kata lain, soal dapat membedakan siswa dari berbagai tingkat pemahaman. (Immanuel et al 2024). Jika pengecoh terlalu mudah diidentifikasi sebagai jawaban yang salah, soal tersebut akan kehilangan kemampuan untuk membedakan siswa yang benar-benar memahami dari siswa yang hanya menebak. Sebaliknya, pengecoh yang dirancang dengan cermat akan membuat peserta didik berpikir lebih mendalam sebelum memilih jawaban.

Dalam praktiknya, kualitas pengecoh juga berperan dalam memastikan keadilan tes (Muhammad Tauhed Supratman et al 2023). Pengecoh yang baik harus relevan, bebas dari bias, dan disesuaikan dengan tingkat kesulitan soal, sehingga tidak memberikan keuntungan atau kerugian kepada kelompok tertentu. Dengan demikian, tes pilihan ganda tidak hanya menjadi alat evaluasi yang obyektif tetapi juga mencerminkan kemampuan peserta didik secara nyata. Hal ini menjadikan perancangan pengecoh yang berkualitas sebagai salah satu elemen penting

dalam pengembangan instrumen tes yang valid, reliabel, dan bermutu tinggi.

Hasil analisis soal spreadsheet kelas 10 SMK se-Jawa Timur yang dilakukan dengan program aplikasi ANATES versi 4.0 for Windows menunjukkan bahwa soal spreadsheet mempunyai kualitas pengecoh dengan kriteria baik. Dengan didukung oleh lima butir soal yang memiliki kualitas pengecoh yang sangat baik terletak pada nomor 1, 7, 8, 9, dan 10.

SIMPULAN

Dalam penelitian ini, program aplikasi ANATES versi 4.0 untuk Windows digunakan untuk mengevaluasi kualitas soal spreadsheet yang diberikan kepada siswa di kelas 10 sekolah menengah kejuruan di provinsi Jawa Timur. Sebagian besar pertanyaan memiliki kualitas yang baik dalam hal kekhasan, kesulitan, dan efek mengganggu, menurut analisis, meskipun beberapa pertanyaan memerlukan penyesuaian untuk meningkatkan kualitasnya. Sangat penting bahwa proses analisis ini dilakukan secara berkala. Ini dilakukan untuk memastikan bahwa pertanyaan yang digunakan memiliki hasil penilaian yang sah dan bahwa mereka dapat membantu meningkatkan kualitas pembelajaran yang berkelanjutan.

UCAPAN TERIMA KASIH

Peneliti mengucapkan terima kasih kepada Universitas Negeri Surabaya karena telah mendukung dan memungkinkan penelitian ini dilakukan. Kami juga mengucapkan terima kasih kepada guru dan siswa profesional dari seluruh Jawa Timur yang berpartisipasi dalam penelitian ini dan

memberikan data yang diperlukan.

DAFTAR PUSTAKA

- Achmad, G. H., Ratnasari, D., Amin, A., Yuliani, E., & Liandara, N. (2022). Penilaian Autentik pada Kurikulum Merdeka Belajar dalam Pembelajaran Pendidikan Agama Islam di Sekolah Dasar. *EDUKATIF : JURNAL ILMU PENDIDIKAN*, 4(4), 5685–5699. <https://doi.org/10.31004/edukatif.v4i4.3280>
- Al Faraby, S., Romadhony, A., & Adiwijaya. (2024). Analysis of LLMs for educational question classification and generation. *Computers and Education: Artificial Intelligence*, 7, 100298. <https://doi.org/10.1016/J.CAEAI.2024.100298>
- Ananda Setiyawan, R., & Sri Wijayanti, P. (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), 2020. <https://doi.org/10.46306/lb.v1i2>
- Ardiansyah, P., Jailani, Ms., Negeri, S., Provinsi Jambi, B., & Sulthan Thaha Saifuddin Jambi, U. (2023). Teknik Pengumpulan Data Dan Instrumen Penelitian Ilmiah. *IHSAN : Jurnal Pendidikan Islam*. <http://ejournal.yayasanpendidikandzurriyatulquran.id/index.php/ihsan>
- Ardila Suci Qurrota, A. A., Styo Siskawati, F., Novita Irawati, T., Studi Pendidikan Matematika, P., Keguruan dan Ilmu Pendidikan, F., Islam Jember, U., Kyai Mojo No, J., Kidul, K., Kaliwates, K., Jember, K., & Timur, J. (2022). Analisis Kelayakan Butir Soal pada Media INTERMATHLY (Interesting Mathematic Monopoly). *Jurnal Cendekia: Jurnal Pendidikan Matematika*.
- Aspriyanti, L., Nopi, R. A., Wangiran, & Naryatmojo, D. L. (2022). Reconstruction and Development of Indonesian Language PAT Question Items Aspects of Class VII Literary Expression MTs N 1 Banjarnegara Academic Year 2021/2022. *Journal of Educational Analytics*, 1(3), 255–266.

- <https://doi.org/10.55927/jeda.v1i3.1716>
- Butir, A., Evaluasi, S., & Akhmadi, M. N. (2021). ANALISIS BUTIR SOAL EVALUASI TEMA 1 KELAS 4 SDN PLUMBUNGAN MENGGUNAKAN PROGRAM ANATES. *Ed-Humanistic*.
- Hendrayadi, Kustati Martin, & Amelia Rezki. (2024). ANALISIS ULANGAN HARIAN MATA PELAJARAN PAI DI SMA NEGERI 10 PADANG TAHUN PELAJARAN 2023/2024 (TELAAH TERHADAP RELIABILITAS, DAYA BEDA DAN TINGKAT KESUKARAN MENGGUNAKAN SOFTWARE ANATES). *Jurnal Review Pendidikan Dan Pengajaran*.
- I Putu Suardipa, K. H. P. (2023). PERAN DESAIN EVALUASI PEMBELAJARAN UNTUK MENINGKATKAN KUALITAS PEMBELAJARAN. *Jurnal Pendidikan, Agama, Dan Budaya*.
- Immanuel, C., Dwi, S., Manik, P., Nababan, A. P., Sianturi, B. Y., & Hasnah, A. (2024). ANALISIS BUTIR SOAL PILIHAN GANDA MATA PELAJARAN MATEMATIKA MATERI BANGUN DATAR BERBASIS BUDAYA LOKAL DI SDN 106163 BANDAR KLIPPA. In *Jurnal Ilmiah Multidisiplin Terpadu* (Vol.8, Issue 6).
- Khasanah, I., Fuady, A., & Islam Malang, U. (2023). Analisis Soal Ulangan Harian Matematika Bentuk Pilihan Ganda. In *MATHEMA JOURNAL E-ISSN* (Vol. 5, Issue 2).
- Moh Rudini. (2020). EFEKTIVITAS ANALISIS BUTIR SOAL MATA PELAJARAN MATEMATIKA PADA SISWA KELAS IV DALAM MENINGKATKAN KUALITAS GURU DI SDN SABANG. *TOLIS ILMIAH: JURNAL PENELITIAN*.
- Muhammad Tauhed Supratman, Harsono Harsono, Ainur Rahman, Achmad Fardza Yunsa N, & Rina Riskiyah. (2023). Peningkatan Kapasitas dan Kompetensi Guru SMP dan MA Syekh Abdurrahman Rabah Pamekasan melalui Pelatihan Asesmen Pembelajaran. *Jurnal Informasi Pengabdian Masyarakat*, 1(1), 01–13.
- <https://doi.org/10.47861/jipm-nalanda.v1i2.179>
- Mujiburrahman, M., Kartiani, B. S., & Parhanuddin, L. (2023). ASESMEN PEMBELAJARAN SEKOLAH DASAR DALAM KURIKULUM MERDEKA. *Pena Anda: Jurnal Pendidikan Sekolah Dasar*, 1(1), 39–48. <https://doi.org/10.33830/penaanda.v1i1.5019>
- Mukhlis, M. (2024). Persepsi Guru terhadap Pemanfaatan ChatGPT dalam Mengembangkan Soal Literasi Membaca: Studi Kasus pada Sekolah Menengah di Provinsi Riau. *Jurnal Pendidikan Dan Kebudayaan*, 9(1), 1–19. <https://doi.org/10.24832/jpnk.v9i1.4873>
- Musyarafah, A., Abdurrahman Siddik Bangka Belitung, S., & Syaikh Abdurrahman Siddik Bangka Belitung, I. (2021). Validitas dan Reliabilitas dalam Analisis Butir Soal. *Al-Mu'arrif: Jurnal Pendidikan Bahasa Arab*, 1(1), 34–44. <https://jurnal.lp2msasbabel.ac.id/index.php/AL-MUARRIB>
- Ndiung, S., & Jediut, M. (2020). Pengembangan instrumen tes hasil belajar matematika peserta didik sekolah dasar berorientasi pada berpikir tingkat tinggi. *Premiere Educandum : Jurnal Pendidikan Dasar Dan Pembelajaran*, 10(1), 94. <https://doi.org/10.25273/pe.v10i1.6274>
- Ramadhan, W., Malahati, F., Romadhon, K., & Ramadhan, S. (2023). Analisis Butir Soal Tipe Multiple Choice Questions pada Penilaian Harian Sekolah Dasar. *Tarbiyah Wa Ta'lim: Jurnal Penelitian Pendidikan Dan Pembelajaran*, 10(2), 93–105. <https://doi.org/10.21093/twt.v10i2.6155>
- Santosa, S., & Badawi, J. A. (2022). Analisis Butir Soal Pilihan Ganda Tema Pertumbuhan dan Perkembangan Makhluk Hidup Kelas III Madrasah Ibtidaiyah. *Jurnal Basicedu*, 6(2), 1678–1686. <https://doi.org/10.31004/basicedu.v6i2.2206>
- Setya Mustafa, P., & Kukuh Masgumelar, N. (2022). Kajian Review: Pengembangan Instrumen Penilaian Sikap, Pengetahuan, dan Keterampilan dalam Pendidikan Jasmani dan Olahraga. *Biormatika: Jurnal Ilmiah Fakultas Keguruan Dan Ilmu*

- Pendidikan*, 8(1), 31–49. <http://ejournal.unsub.ac.id/index.php/FKIP>
- Svensberg, K., Kaae, S., Mottelson, N. B., & Persson, C. L. (2024). Identifying critical elements in using Question Prompt Lists at the pharmacy counter to induce patient activation—using principles of conversation analysis. *Research in Social and Administrative Pharmacy*. <https://doi.org/10.1016/J.SAPHARM.2024.10.008>
- Wahyu, A., & Kemenkes Surakarta Jurusan Fisioterapi, P . (2020). Uji Validitas dan Reliabilitas Alat Ukur SG Posture Evaluation. *Jurnal Keterampilan Fisik*, 2.
- Wigati, I., & Ali, M. (2022). Instrumen Pembelajaran Daring dengan Rasch Model. *Proceedings Raden Patah*.
- Yusuf, R. (2021). ANALISIS KUALITAS BUTIR SOAL UJIAN NASIONAL MATA PELAJARAN MATEMATIKA SEKOLAH MENENGAH PERTAMA. *Journal of Didactic Mathematics*, 1(3). <https://doi.org/10.34007/jdm.v1i3.417>
- Zainal, N. F. (2020). Pengukuran, Assessment dan Evaluasi dalam Pembelajaran Matematika. *Laplace : Jurnal Pendidikan Matematika*, 3(1), 8–26. <https://doi.org/10.31537/laplace.v3i1.310>