

Kajian Literatur Mengenai Reliabilitas Soal Sebagai Instrumen Evaluasi

Siti Alya Lubis¹, Zulpan², Nur Azizah³, Hilya Sahri Ramadhani⁴, Fadia Hannum⁵, Naila Azmi Rangkuti⁶

Sekolah Tinggi Agama Islam Negeri Mandailing Natal, Indonesia¹⁻⁶

Email Korespondensi: sitialyalubis440@gmail.com , zulpan30122@gmail.com , nurazizahnst0512@gmail.com , hilyasahri@gmail.com , fadiahannum@gmail.com , nailahrangkuti84@gmail.com

*Article received: 22 Januari 2026, Review process: 11 Februari
Article Accepted: 25 April 2026, Article published: 07 Mei 2026*

ABSTRACT

The purpose of this study to give comprehensive, clearly explanation and constructive a concept of reliability. The method used qualitative approach especially literature review. It did do from identifying, comparing, analyzing, and creating constructs and concepts of reliability. Data sources took from scientific articles and books relevant for topic. Finding of the study are a structured concept related to the understanding of reliability, factors of impact of reliability, relationship reliability and validity, and reliability testing techniques. Hopefully this article can be contributed to a more comprehensive knowledge about the concept reliability and can be supporting to increase quality of learning evaluation.

Key Words: *reliability test, reliability testing techniques*

ABSTRAK

Tujuan penelitian ini untuk memberikan penjelasan lebih komprehensif, jelas, lebih konstruktif tentang pemahaman reliabilitas secara utuh. Penelitian ini menggunakan pendekatan kualitatif dengan metode kajian literatur, yaitu mengidentifikasi, membandingkan, serta menganalisis dan menciptakan konstruksi dan konsep tentang pemahaman reliabilitas soal. Sumber data diambil dari artikel ilmiah dan buku relevan dengan topik pembahasan. Hasil penelitian telah terbentuk konsep terstruktur terkait pemahaman reliabilitas, faktor-faktor mempengaruhi reliabilitas, hubungan reliabilitas dan validitas dan teknik pengujian reliabilitas. Kajian ini diharapkan dapat memberikan sumbangan pemahaman lebih komprehensif mengenai pentingnya reliabilitas soal untuk mendukung kualitas evaluasi pembelajaran.

Kata Kunci: *pemahaman reliabilitas, teknik pengujian reliabilitas.*

PENDAHULUAN

Elemen terpenting evaluasi pembelajaran yakni untuk memperoleh informasi mengenai kemampuan, perkembangan, serta tingkat pemahaman murid. Hasil tersebut dijadikan dasar menentukan tindak lanjut pembelajaran, baik dalam bentuk perbaikan maupun penguatan kepada murid. Evaluasi pembelajaran tidak hanya dipahami sebagai proses simbolis pemberian nilai melainkan sebagai alat bantu untuk meningkatkan kualitas pembelajaran dan perkembangan murid (Phafiandita & Permadani, 2022). Keputusan hasil belajar peserta dianggap penting untuk melakukan refleksi maupun perbaikan terhadap proses pembelajaran serta pencapaian hasil belajar.

Aktivitas evaluasi tentu menggunakan alat tes sesuai dengan kondisinya untuk mengetahui hasilnya, dipertegas oleh Husein et al. (2025) bahwa asesmen merupakan hasil dari pengukuran menggunakan alat tes untuk menghasilkan sebuah nilai. Bahasa alat tes pada pembelajaran secara umum yang mudah diketahui oleh guru-guru adalah "Soal". Pembuatan soal tentunya sudah melakukan tahapan-tahapan; 1)kajian teori dan analisis kurikulum,2)memperhatikan indikator materi serta validasi ahli,3)revisi berdasarkan masukan validator dan menyusun draf soal,4)pengujian, seterusnya menganalisis hasil uji validitas, reliabilitas, daya beda dan tingkat kesukaran soal dan pengecoh soal (Nubatonis & Jupri, 2024).

Hal yang jarang dibahas serta disajikan datanya untuk menentukan kualitas soal adalah hasil uji validitas dan reliabilitas soal, sehingga minim pengetahuan guru-guru terhadap hal tersebut serta tidak ada tindak lanjut untuk perbaikan. Soal yang baik harus memenuhi beberapa kriteria tertentu agar hasil pengukuran yang diperoleh dapat dipercaya (Prasetyo, 2025). Beberapa kriteria tersebut meliputi validitas, reliabilitas, objektivitas, dan kepraktisan. Di antara berbagai kriteria tersebut, reliabilitas menjadi salah satu aspek penting karena berkaitan dengan konsistensi hasil pengukuran yang dilakukan, jika tes tidak reliabel, maka murid dengan kemampuan sama akan memperoleh nilai sangat berbeda, pada akhirnya dapat menimbulkan ketidakadilan pada proses penilaian.

Pada praktiknya masih ditemukan berbagai permasalahan terkait penyusunan soal lebih berfokus pada isi materi yang diujikan tanpa memperhatikan kualitas soal, termasuk aspek reliabilitasnya. Akibatnya, hasil evaluasi yang diperoleh belum tentu mampu menggambarkan kemampuan murid secara objektif dan konsisten. Kondisi ini menunjukkan bahwa pemahaman mengenai reliabilitas soal masih perlu mendapatkan perhatian yang lebih mendalam dalam proses evaluasi pembelajaran (Saputra, 2025). Ditambahkan lagi guru masih kurang memperhatikan aspek reliabilitas sejak tahap perencanaan tes, bukan hanya saat pelaksanaan (Muharromah et al., 2025).

Perkembangan pendidikan pada era modern menuntut adanya sistem evaluasi yang semakin berkualitas, objektif, dan akurat. Pendidik tidak hanya dituntut untuk mampu menyusun soal, tetapi juga memahami bagaimana kualitas instrumen tersebut dapat diukur dan dianalisis (Boroallo & Purnamasari, 2025). Dengan demikian, pemahaman mengenai reliabilitas soal menjadi sangat penting sebagai dasar dalam menghasilkan instrumen evaluasi yang mampu memberikan

hasil pengukuran yang dapat dipercaya. Reliabilitas soal mengacu pada kemampuan tes menghasilkan hasil pengukuran yang tetap atau konsisten ketika digunakan dalam kondisi yang serupa (Zayrin et al., 2025). Tes yang memiliki reliabilitas tinggi akan menghasilkan data yang stabil dan dapat dipercaya, sehingga mampu memberikan gambaran lebih akurat mengenai kemampuan murid. Sebaliknya, Tes memiliki reliabilitas rendah menyebabkan hasil evaluasi menjadi tidak konsisten berpotensi menimbulkan kesalahan pengambilan keputusan (Saputri et al., 2023).

Pengetahuan tentang reliabilitas soal dikalangan pendidik masih kurang khususnya konstruksi secara mendasar sehingga membuat kebingungan, terutama konstruksi pengertian, pengaplikasian dan penganalisisan. Maka hal tersebut mendorong penulis untuk memberikan konstruksi yang jelas terkait reliabilitas soal. Kajian penelitian ini diharapkan dapat memberikan pemahaman lebih komprehensif mengenai pentingnya reliabilitas soal untuk mendukung kualitas evaluasi pembelajaran.

METODE

Penelitian ini menggunakan pendekatan kualitatif dengan metode kajian literatur (*literature review*), yakni mengidentifikasi, membandingkan, serta menganalisis dari berbagai teori dan hasil penelitian yang berkaitan dengan konsep reliabilitas soal. Pengumpulan data diambil dari berbagai literatur ilmiah, buku, dan serta sumber akademik lain membahas tentang evaluasi pembelajaran dan reliabilitas tes. Teknik pengumpulan data dilakukan melalui studi dokumentasi yakni mengidentifikasi, membaca, mencatat, dan mengelompokkan informasi penting dari berbagai sumber yang berkaitan dengan reliabilitas soal. Data yang telah diperoleh kemudian dianalisis secara deskriptif-kualitatif melalui beberapa tahapan, yaitu reduksi data, penyajian data, dan penarikan kesimpulan. Reduksi data dilakukan dengan memilih informasi sesuai dengan fokus penelitian, kemudian data disajikan secara sistematis untuk mempermudah proses analisis. Tahap akhir dilakukan dengan menyusun interpretasi dan konsep pemikiran berdasarkan hasil kajian berbagai teori yang telah dianalisis. Melalui metode ini, diharapkan dapat memberikan pemahaman yang lebih komprehensif mengenai konsep reliabilitas soal, faktor-faktor yang memengaruhi, serta teknik pengukurannya.

PEMBAHASAN DAN HASIL

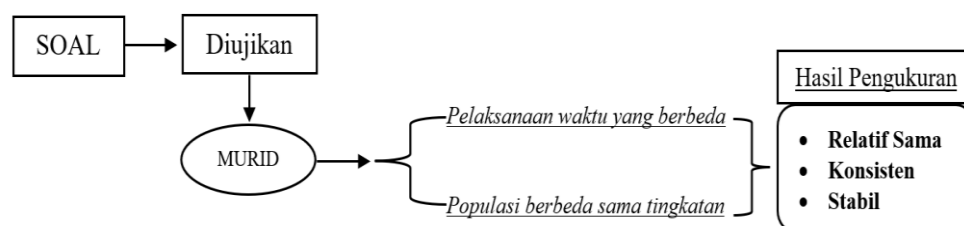
Pemahaman Reliabilitas

Pada evaluasi pembelajaran reliabilitas memiliki peran yang sangat penting karena hasil evaluasi pendidikan sering dijadikan pengambilan keputusan. Tes yang memiliki reliabilitas tinggi akan menghasilkan data yang objektif dan konsisten, sedangkan reliabilitas rendah lebih menghasilkan data tidak tetap atau mengalami perubahan di setiap pengukurannya. Kondisi tersebut dapat menyebabkan ketidakakuratan saat penggambaran kemampuan murid yang sebenarnya (Kaliwanto, 2023). Konsep reliabilitas mengacu pada konsistensi hasil skor pada item yang terdapat pada tes, sehingga uji reliabilitas sesungguhnya

menguji ketepatan hasil skor pengukuran terhadap tes (Ridha & Al Ghifari, 2025). Reliabilitas memastikan hasilnya konsisten dari waktu ke waktu, faktor penting ketika instrumen digunakan berulang kali atau di populasi yang berbeda (Siregar, 2025). Reliabilitas soal mengacu pada kemampuan suatu instrumen menghasilkan pengukuran yang relatif konsisten apabila digunakan berulang kali pada kondisi yang sama (Fatayah & Yuliana, 2022).

Reliabilitas mempunyai beberapa ciri penting penentu kualitas suatu alat ukur, yakni konsistensi hasil pengukuran adalah salah satu ciri utamanya. Tes yang reliabel dapat memberikan hasil hampir sama pada saat dilakukan pada kondisi yang hampir sama (Armedi et al. 2025). Hal ini membuktikan bahwa hasil evaluasi tidak mudah berubah hanya karena faktor lain seperti kebetulan. Selain konsistensi, tes yang reliabel juga memiliki hasil pengukuran yang stabil, berarti tes bisa menghasilkan hasil atau nilai yang tetap walaupun digunakan di waktu berbeda. Stabilitas hasil pengukuran dapat dijadikan dasar penilaian yang lebih terpercaya karena tes mempertahankan kualitas pengukurannya dari waktu ke waktu (Muharromah 2025).

Karakteristik selanjutnya, tingkat kesalahan pengukuran sangat rendah, kesalahan pengukuran muncul dikarenakan keambiguan soal, petunjuk pengerjaan yang kurang tepat, ataupun suasana pelaksanaan tes yang kurang mendukung. Tes yang reliabel mampu meminimalkan kesalahan tersebut sehingga data yang diperoleh menjadi lebih akurat dan objektif serta kemampuan sebenarnya dari murid. Keberadaan berbagai ciri atau karakteristik tersebut menandakan bahwa reliabilitas tidak hanya berhubungan dengan hasil akhir tes, melainkan juga berhubungan erat terhadap kualitas instrumen secara menyeluruh (Kunanti, 2025). Instrumen yang reliabel mampu membantu guru atau peneliti dalam memperoleh hasil evaluasi yang lebih adil dan objektif serta dapat dipertanggung jawabkan. Terciptanya pemahaman utuh tentang pengertian reliabilitas penulis memberikan gambaran dasar berupa peta konsep sebagai berikut:



Gambar 1. Konsep Pemahaman Reliabilitas

Dari gambar tersebut kita dapat memahami bahwa reliabilitas adalah hasil pengukuran relatif sama, konsisten, stabil walaupun dilaksanakan pada waktu yang berbeda dan populasi berbeda sama tingkatan. Reliabilitas dapat dijadikan indikator penting untuk mengetahui apakah suatu soal dapat digunakan sebagai alat ukur yang mampu memberikan hasil yang akurat. Terdapat faktor yang mempengaruhi reliabilitas tes seperti; kejelasan soal kurang, keambiguan soal, petunjuk pengerjaan kurang tepat, dan suasana pelaksanaan tes yang kurang mendukung. Selain itu, reliabilitas berkaitan erat terhadap besar kecilnya kesalahan pengukuran, semakin tinggi tingkat reliabilitas maka semakin kecil kemungkinan

terjadinya kesalahan pengukuran. Oleh karena itu, reliabilitas menjadi salah satu syarat utama yang perlu diperhatikan pada penyusunan tes pada evaluasi pembelajaran.

Faktor Mempengaruhi Reliabilitas

Melengkapi paparan dari penulis mengenai faktor- faktor yang dapat dipengaruhi reliabilitas soal antara lain; faktor pertama bisa seperti kualitas butir soal. Soal yang disusun dengan jelas, terstruktur dan tidak ambigu dan jelas maksud soalnya, serta sesuai dengan tujuan pembelajaran akan meningkatkan konsistensi hasil pengukuran. Sebaliknya soal yang tidak jelas atau sulit dipahami dapat menyebabkan perbedaan pemahaman setiap individu sehingga hasil yang dihasilkan tidak konsisten (Solichin, 2021).

Selain kualitas soal, panjang dari tes juga memengaruhi reliabilitas. Tes yang menggunakan soal dengan jumlah yang lebih banyak akan lebih menghasilkan reliabilitas yang tinggi karena soal yang banyak bisa mencakup banyak materi pembelajaran yang lebih luas. Tetapi soal yang terlalu banyak juga dapat memengaruhi kualitas jawaban karena beban murid terlalu berat.

Faktor lain yang memengaruhi reliabilitas soal adalah tingkat kesulitan, soal yang terlalu sulit dan terlalu mudah juga dapat memengaruhi hasil reliabilitas hasil yang didapatkan nantinya bisa tidak akurat, maka soal yang digunakan seharusnya adalah soal yang seimbang antara kesulitan dan daya pembedanya karena akan menghasilkan hasil yang lebih akurat. Selain hal-hal teknis, kondisi pelaksanaan tes juga dapat memengaruhi hasil reliabilitas. Lingkungan yang kurang nyaman bagi siswa, adanya keterbatasan waktu, keadaan mental murid. Dapat menyebabkan hasil evaluasi menjadi kurang konsisten (Alwi, 2015). Penulis memberi konsep rangkuman untuk memudahkan pembaca apa saja faktor-faktor yang memengaruhi reliabilitas soal dapat di gambarkan sebagai berikut pada tabel 1:

Tabel 1. Faktor Mempengaruhi Reliabilitas

FAKTOR	DESKRIPSI
Kejelasan Soal	Soal disusun harus jelas, terstruktur, tidak ambigu dengan tujuan pembelajaran.
Kuantitas Soal	Jumlah soal harus dapat menggambarkan cakupan materi tidak terlalu sedikit dan tidak terlalu banyak fraksional.
Tingkatan Kesukaran Soal	Soal tidak terlalu susah dan tidak terlalu mudah, adanya soal harus seimbang.
Daya Pembeda Soal	Daya pembeda soal dapat membedakan kemampuan peserta.
Psikologis Murid	Keadaan peserta pada saat menjawab soal akan mempengaruhi hasil.
Lingkungan	Lingkungan pada saat mengerjakan soal akan mempengaruhi hasil.
Waktu	Waktu Pengerjaan memiliki promosi yang terlalu cepat dan terlalu lama.

Paparan tersebut kita memahami bahwa terdapat beberapa faktor yang mempengaruhi tingkat reliabilitas soal. Hal tersebut menjadi pertimbangan kita sebelum melakukan pengujian soal untuk mendapatkan hasil yang hampir sama, tidak jauh berbeda dan stabil walaupun di ujikan dengan kondisi sama dan pada waktu yang berbeda.

Hubungan Reliabilitas dan Validitas

Reliabilitas dan validitas merupakan dua komponen penting saling berhubungan pada evaluasi pembelajaran. Validitas mengarah pada ketepatan alat ukur, sedangkan reliabilitas berkaitan hasil pengukuran memiliki skor yang hampir sama, kestabilan dan konsisten. Apabila kedua aspek ini terpenuhi maka tes hasil evaluasi akan mampu menggambarkan kemampuan murid secara objektif. Instrumen evaluasi yang reliabel akan menghasilkan skor konsisten pada beberapa pelaksanaan tes yang memiliki kondisi yang sama (Arbeni et al., 2025). Sedangkan validitas soal memastikan setiap soal benar sesuai, materi, konstruksi dan tujuan pembelajaran terhadap kompetensi murid (Zulpan et al., 2024). Oleh karena itu, reliabilitas serta validitas saling melengkapi guna menciptakan soal yang memiliki kualitas yang baik.

Suatu instrumen mungkin saja mempunyai reliabilitas tinggi, tetapi validitas rendah, kondisi ini muncul saat hasil tes selalu konsisten, namun isi soal tidak sesuai dengan tujuan pengukuran. Di samping itu ada juga instrumen yang valid tetapi reliabilitasnya rendah karena menghasilkan skor tidak konsisten sehingga tingkat kepercayaan hasil tes ikut menurun. Hal ini menunjukkan hubungan reliabilitas dan validitas tidak hanya bertumpu kepada satu aspek saja (Nurhan et al., 2025). Guru juga perlu memperhatikan tingkat kesukaran, pemilihan indikator, penyusunan soal serta kejelasan bahasa agar mampu menghasilkan nilai yang konsisten. Jadi, dapat dipahami reliabilitas dan validitas mempunyai hubungan yang erat karena keduanya sama menentukan kualitas instrumen evaluasi. Jika reliabilitas dan validitas terpenuhi secara seimbang maka hasil evaluasi akan lebih objektif untuk penentuan hasil belajar murid.

Teknik Pengukuran Reliabilitas

Teknik menentukan reliabilitas dapat dilakukan dengan berbagai cara antara lain; 1) metode *test-retest* yakni uji kestabilan dilakukan dengan perbedaan kondisi atau waktu antara tes pertama dan tes kedua (Suyiono, 2010). Sementara Pallent (2010) menyatakan pemberian tes yang sama sebanyak 2 kali kepada responden dalam kurun waktu yang berbeda. Hal ini diperkuat oleh Susanto et al. (2015) untuk menentukan reliabilitas tes hasil belajar bentuk objektif dengan cara membela tes menjadi 2 bagian dengan menggunakan formula *Spearman-Brown* yakni Model Gasal-Genap dan Model Kiri-Kanan. Sedangkan formula Flanagan & Rulon juga menggunakan teknik belah dua namun sasaran menekankan pada variasi dan perbedaan hasil skor untuk menentukan reliabilitas. 2) metode akuivalen Saidah et al. (2026) menjelaskan tes diberikan 2 versi yang setara kepada responden pada

waktu bersamaan, jadi hasil kedua tes tersebut dinilai korelasinya apakah memperoleh koefisien yang ekuivalen.

Selanjutnya, metode paling sering digunakan 3) metode konsistensi internal yakni menekankan hanya satu kali memberikan tes kepada satu kelompok. Teknik ini sering digunakan karena hasil pengukurannya dinilai cukup akurat dan mudah digunakan untuk menghitung reliabilitas. Teknik yang digunakan adalah *Cronbach Alpha*, *Kuder-Ricardson* dan *Spilt-Half*. *Cronbach Alpha* digunakan pada tes penilaian berbentuk uraian maupun angket dan kuesioner (Asmita & Fitriani, 2022). *Kuder-Ricardson* sering dipakai dalam pengukuran tes objektif karena teknik ini lebih cocok digunakan pada instrumen objektif dengan jawabannya hanya benar dan salah, seperti soal pilihan ganda. Terakhir teknik *spilt-half* ini dibagi menjadi dua bagian hasil skor biasa dengan menggunakan ganjil-genap atau sebaliknya, cara ini tergolong lebih sederhana dibandingkan teknik-teknik lainnya maka dijadikan alternatif untuk pengukuran reliabilitas. Peneliti mencoba memberikan konsep untuk lebih memudahkan sebagai berikut:

	<i>Test-Retest</i>	<i>Akuivalen</i>	<i>Konsistensi Internu</i>
Tujuan	Stabilitas hasil tes waktu ke waktu	Melihat kesetaraan kedua soal	Konsistensi hasil antara soal
Soal	Satu soal yang sama	Dua soal yang set	Satu soal yang sama
Pengujian	Dua kali pengujian	Satu kali pengujian	Satu kali pengujian
Waktu	Ada jeda waktu	Dilakukan pada waktu yang sama	Dilakukan pada waktu yang sama
Teknik Pengujian	<i>Spearman-Brown Flanagan & Rulonnt</i>	<i>Pearson Product</i>	<i>Cronbach Alpha</i> <i>Kuder-Ricardson</i> <i>Spilt-Half</i>

Metode dan teknik pengukuran reliabilitas dilakukan untuk melihat apakah suatu instrumen tes evaluasi memberikan hasil yang konsisten. Jika instrumen tes digunakan memberikan hasil tidak jauh berbeda maka soal dapat dikatakan reliabel. Maka dengan itu reliabilitas menjadi salah satu hal penting untuk menentukan kualitas tes pembelajaran.

SIMPULAN

Reliabilitas soal merupakan unsur penting dalam menentukan kualitas instrumen evaluasi pembelajaran karena menghasilkan pengukuran yang relatif sama, stabil, konsisten. Tes yang reliabel mampu menghasilkan data lebih stabil dan objektif sehingga dapat dilakukan untuk dasar pengambilan keputusan pada pembelajaran. Beberapa faktor yang mempengaruhi tingkat reliabilitas soal antara lain; kejelasan soal, kuantitas, soal, tingkat kesukaran, daya pembeda, psikologis murid, lingkungan dan waktu. Reliabilitas juga mempunyai hubungan yang erat dengan validitas karena keduanya sama menentukan kualitas tes evaluasi pembelajaran. Jika reliabilitas dan validitas terpenuhi secara seimbang maka hasil evaluasi akan lebih objektif untuk penentuan hasil belajar murid. Pada tataran pembuktian reliabilitas dapat diujikan dengan menggunakan tiga metode antara

lain; 1) *Test-Retest* bertujuan untuk stabilitas tes dari waktu ke waktu dengan teknik pengujian menggunakan *Spearman-Brown* dan *Flanagan & Rulon*, 2) *Akuivalen* bertujuan untuk melihat hasil kesetaraan dari dua tes atau soal dengan teknik pengujian menggunakan *Pearson Product Moment*, 3) *Konsistensi Internal* bertujuan untuk melihat konsistensi hasil antar-butir soal dengan teknik pengujian menggunakan *Cronbach Alpha*, *Kuder-Ricardson* dan *Spilt-Half*. Berdasarkan hasil pemikiran penulis, reliabilitas soal tidak hanya bergantung pada teknik analisis, tetapi juga pada kemampuan guru dalam menyusun instrumen evaluasi yang jelas, terukur serta sesuai dengan tujuan pembelajaran. Oleh karena itu, peningkatan kualitas evaluasi pembelajaran perlu didukung oleh kompetensi guru dan penggunaan instrumen yang lebih terencana agar hasil penilaian menjadi lebih akurat dan dapat dipercaya.

DAFTAR PUSTAKA

- Alwi. (2015). Pengaruh jumlah alternatif jawaban tes obyektif bentuk pilihan ganda terhadap reliabilitas. *Nama Jurnal Belum Diketahui*, 3(2).
- Arbeni, W., Windiani, A., Sariyani, D., & Sihotang, B. (2025). Test reliability analysis in educational evaluation: A quantitative approach to consistency and validity. *Nama Jurnal Belum Diketahui*, 5(1), 59–64.
- Armedi, R. (2025). Karakteristik tes yang baik dan proses penyusunan instrumen tes untuk pembelajaran di sekolah. *Nama Jurnal Belum Diketahui*, 5, 10–17.
- Boroallo, R. P., & Purnamasari, D. I. (2025). Pentingnya evaluasi pembelajaran dalam meningkatkan kualitas pengajaran di era modern. *Nama Jurnal Belum Diketahui*, 3(4), 2632–238.
- Daulay, A., Ramadhani, S., Rafli, M., Zulpan, Z., Husein, A., & Wulandari, M. N. (2026). Pemahaman konsep validitas serta pengujiannya dalam penelitian. *Teaching and Learning Research Journal*, 2(1), 173-178.
- Fatayah, F., & Yuliana, I. F. (2022). Analisis validitas dan reliabilitas dalam mendukung ketuntasan belajar model STEM. *Nama Jurnal Belum Diketahui*, 18(1), 49–60.
- Hariyati, T., & Zulpan, Z. (2018). Validity and reliability of performance assessment of acetic acid visual inspection. *Jurnal Literasiologi*, 1(2), 8-8.
- Husein, A., Sahrul, S., & Nugraheni, F. W. (2025). Penilaian Proyek Penguatan Profil Pelajar Pancasila Pada Kurikulum Merdeka. *Teaching and Learning Journal of Mandalika (Teacher)* e-ISSN, 2721-9666.
- Kaliwanto, B. (2023). Pelatihan pengukuran teknik bagi siswa SMA Islam Assa'adah Puri Serpong Tangerang Selatan. *Cakrawala: Jurnal Pengabdian Masyarakat Global*, 2(4), 429–439.
- Kunanti, O. P. (2025). Penerapan *Partial Credit Model (PCM)* dalam meningkatkan akurasi pengukuran keterampilan berpikir kritis matematis siswa (Skripsi/Tesis/Disertasi tidak diterbitkan). Nama Perguruan Tinggi.
- Muharromah, F. (2025). Reliabilitas alat ukur: Jenis-jenis, cara pengukuran, dan faktor-faktor yang mempengaruhinya dalam pembelajaran. *Nama Jurnal Belum Diketahui*, 4(1), 63–73.
- Nurhan, M. A. J., Kholis, N., & Rasyid, N. A. (2025). Peran validitas dan rehabilitas

- dalam evaluasi pembelajaran. *Nama Jurnal Belum Diketahui*, 02, 224–227.
- Pallant, J. (2010). *SPSS Survival Manual: A step by step guide to data analysis using the SPSS program* (4th Editio). Allen & Unwin
- Phafiandita, A. N., & Permadani, A. (2022). Urgensi evaluasi pembelajaran di kelas. *Jurnal Inovasi dan Riset Akademik*, 3(2), 111–121.
- Prasetyo, Edi. (2025). Evaluasi perencanaan desain pembelajaran, pelaksanaan proses. *Jurnal Permapendis Sumut*, 5(2), 156–65.
- Ridha, A. R., Sudarto, & Ghifari, F. A. (2025). Reliabilitas alat ukur, jenis-jenis, cara pengukuran dan faktor-faktor. *Nama Jurnal Belum Diketahui*, 4(1), 121–131.
- Saidah, N., Mawaddah, N., Hidayah, N., Zulpan, Z., Husein, A., & Arianto, A. (2026). Konsep reliabilitas dalam penelitian dan teknik pengujianya. *Educational Journal*, 1(3), 775-784.
- Saputra, A. (2025). Pengembangan instrumen evaluasi (Analisis pengembangan penilaian satuan pendidikan MTs / SMP). *Nama Jurnal Belum Diketahui*, 2(4), 1–14.
- Saputri, H. A., Zuhijrah, Larasati, N. J., & Shaleh. (2023). Analisis instrumen assesmen: Validitas, reliabilitas, tingkat kesukaran dan daya beda butir soal. *Nama Jurnal Belum Diketahui*, 09, 2986–2995.
- Sari, G. R. M., Pariha, L., Nugraha, P. L., & Iskandar, S. (2025). Analisis pentingnya evaluasi pembelajaran dalam meningkatkan pemahaman siswa sekolah dasar. *Nama Jurnal Belum Diketahui*, 10, 448–459.
- Solichin, M. (2021). Analisis daya beda soal, taraf kesukaran, validitas butir tes, interpretasi hasil tes dan validitas ramalan dalam evaluasi pendidikan. *Nama Jurnal Belum Diketahui*, 2, 192–213.
- Sugiyono. (2010). *Statistika untuk Penelitian*. Bandung: Alfabeta. Bandung: Alfabeta.
- Waruwu, M. (2024). Pendekatan penelitian kualitatif: Konsep, prosedur, kelebihan dan peran di bidang pendidikan. *Nama Jurnal Belum Diketahui*, 5, 198–211.
- Zayrin, A. A., Nupus, H., Maizia, K. K., & Marsela, S. (2025). Analisis instrumen penelitian pendidikan (Uji validitas dan reliabilitas instrumen penelitian). *Nama Jurnal Belum Diketahui*, 780–789.
- Zulpan, Z., Alfansuri, D. U., Salsabila, S., Hasibuan, S. W., Azizah, N., & Mustikawati, D. L. (2025). Kualitas soal: Tingkat kesukaran daya pembeda pengecoh soal. *Teaching and Learning Research Journal*, 1(4), 93-101.
- Zulpan, Z., & Rusli, A. (2020). Validitas dan reliabilitas instrumen penilaian membaca *short functional text* pada siswa SMP kelas VIII. *Jurnal Pendidikan Guru*, 1(1).