

Karakteristik Tes yang Baik dan Proses Penyusunan Instrumen Tes untuk Pembelajaran di Sekolah

Rama Armedi¹

¹ Universitas Islam Negeri Maulana Malik Ibrahim Malang, Indonesia
Email: rarmedi8@gmail.com

Tersedia Online di

<https://jurnal.educ3.org/index.php/pendagogia>

Sejarah Artikel

Diserahkan : 28 Desember 2024
Disetujui : 12 April 2025
Dipublikasikan : 25 April 2025

Kata Kunci:

ciri-ciri, tahap-tahap, instrumen tes

Abstrak: *Evaluation in learning aims to see the abilities and understanding of students in accordance with the objectives that have been set. Tests as a data collection tool must show the tendencies that exist in their preparation. This paper discusses the characteristics of a good test and the stages of preparing test instruments with a library research approach. The results of this paper state that the characteristics of a good test are tests that meet requirements such as validity, reliability, objectivity, practicality, and economy. And through the steps of its development, namely by determining objectives, developing instrument specifications, choosing the type of instrument, preparing the type of instrument, and testing the instrument. A good test will provide an overview of what needs to be improved regarding evaluation. And a good test can be a powerful tool to improve education.*

Keywords: *characteristics, stages, test instruments*

Abstrak: Evaluasi dalam pembelajaran bertujuan untuk melihat kemampuan dan pemahaman peserta didik sesuai dengan tujuan yang telah ditetapkan. Tes menjadi alat pengumpulan data haruslah menunjukkan kecenderungan yang ada dalam penyusunannya. Tulisan ini membahas bagaimana ciri-ciri tes yang baik dan tahap-tahap penyusunan instrumen tes dengan pendekatan *library research*. Hasil tulisan ini mengemukakan bahwa ciri-ciri tes yang baik adalah tes yang memenuhi persyaratan seperti validitas, reliabilitas, objektif, praktis, dan ekonomis. Dan melalui langkah-langkah pengembangannya yaitu dengan menentukan tujuan, mengembangkan spesifikasi instrumen, memilih jenis instrumen, mempersiapkan jenis instrument, uji coba instrument. Dengan tes yang baik akan memberikan gambaran terkait apa yang perlu ditingkatkan terkait evaluasi. Dan tes yang baik dapat menjadi alat yang ampuh untuk meningkatkan pendidikan.

PENDAHULUAN

Dalam dunia pendidikan, tes merupakan salah satu alat evaluasi yang paling umum digunakan untuk mengukur pemahaman dan kemampuan siswa. Evaluasi menjadi penentu seberapa sukses atau tidaknya suatu program dalam mencapai tujuan yang ditetapkan di awal (Adom et al., 2020). Evaluasi juga menjadi salah satu aspek penting dalam pembelajaran. Suatu perencanaan pembelajaran yang telah disusun sebaik mungkin pasti akan menentukan perbedaan dalam realita pelaksanaannya (Rahman & Nasryah, 2019). Perencanaan pembelajaran yang baik membutuhkan instrumen tes yang tepat yang disusun dan dianalisis kualitasnya sebagai bentuk evaluasi dalam proses pembelajaran.

Sejumlah penelitian sebelumnya telah menunjukkan berbagai kecenderungan dalam menyusun tes. Pertama, untuk keperluan evaluasi diperlukan teknik evaluasi

seperti kuesioner, tes, skala, format observasi, dan lainnya (Susanto, 2023). Kedua, kecenderungan mengedepankan validitas karena validitas tes mempengaruhi interpretasi hasil tes dan pengambilan keputusan (Asiva Noor Rachmayani, 2015). Ketiga, reliabilitas juga menjadi fokus karena hasil tes harus konsisten jika diulang dalam kondisi yang sama (Khumaedi, 2012). Kecenderungan-kecenderungan ini menunjukkan bahwa kualitas tes bergantung kepada bagaimana tes tersebut diterapkan dalam konteks pembelajaran maupun pendidikan.

Untuk mencapai pemahaman yang mendalam mengenai penyusunan instrument tes yang baik, diperlukan eksplorasi dari setiap aspek yang mempengaruhi kualitas tes. Ini mencakup analisis mendalam terhadap tujuan tes. Pemahaman mendalam ini penting agar instrument tes yang disusun dapat memenuhi kebutuhan evaluasi dengan tepat dan akurat.

Tujuan dari tulisan ini adalah untuk memberikan panduan yang komprehensif mengenai ciri-ciri tes yang baik serta tahapan-tahapan dalam penyusunan instrument tes. Adapun pembahasan diperoleh berdasarkan eksplorasi dari berbagai *paper*, artikel hasil penelitian, maupun buku referensi yang berkaitan dengan penilaian utamanya penyusunan tes untuk asesmen di sekolah.

METODE

Penelitian ini ditulis menggunakan metode *library reseach* yang dengan kepustakaan sebagai teknik pengumpulan data, seperti buku-buku, artikel-artikel, dan sumber-sumber dari internet. *Library reseach* adalah penelitian yang berkaitan erat dengan pengumpulan data berupa teks, membaca, mencatat, dan mengolah data tersebut menjadi bahan penelitian pengumpulan data berupa teks dan diolah menjadi bahan penelitian (Spiritual, 2024). Pengumpulan data berupa teks dan diolah menjadi bahan penelitian.

HASIL

Tes yang baik memiliki beberapa ciri penting yang harus diperhatikan dalam penyusunannya. Pertama, validitas adalah aspek yang sangat penting, di mana tes harus mengukur apa yang seharusnya diukur. Validitas ini dapat berupa validitas isi, validitas konstruk, dan validitas empiris atau kriteria. Kedua, reliabilitas adalah kemampuan tes untuk memberikan hasil yang konsisten jika dilakukan berulang kali. Reliabilitas ini dapat berupa reliabilitas stabilitas, reliabilitas konsistensi internal, dan reliabilitas ekuivalen. Ketiga, objektivitas adalah kemampuan tes untuk dinilai secara objektif, dengan hasil yang sama meskipun dinilai oleh orang yang berbeda. Keempat, praktis adalah kemudahan dalam pelaksanaan, pemeriksaan, dan dilengkapi dengan petunjuk yang jelas.

Terakhir, ekonomis adalah kemampuan tes untuk dilaksanakan dengan biaya yang terjangkau. Dalam penyusunan instrumen tes, langkah-langkah yang perlu diperhatikan meliputi menentukan tujuan, mengembangkan spesifikasi instrumen, memilih jenis dan tipe instrumen, mempersiapkan kisi-kisi instrumen, dan melakukan uji coba instrumen. Dengan memperhatikan ciri-ciri dan langkah-langkah tersebut, kita dapat menyusun tes yang valid, reliabel, objektif, praktis, dan ekonomis, serta memastikan bahwa hasil tes dapat dipertanggungjawabkan.

PEMBAHASAN

Ciri-Ciri Tes yang Baik

Evaluasi pembelajaran tidak akan pernah lepas dari pelaksanaan proses belajar. Evaluasi pembelajaran dilakukan guna mengetahui *Student Achievement* selama pembelajaran, dan berguna untuk menilai apakah efektif atau tidaknya kegiatan belajar mengajar yang sudah dilakukan, juga sebagai laporan dan kemajuan peserta didik, serta menjadi indikator untuk menentukan kelulusan peserta didik (Maulani et al., 2024).

Hasan mengemukakan bahwa tes adalah “alat pengumpulan data dari konstruksi butir (soal) yang dipergunakan (Magdalena, 2022).” Rumusan Hamid ini memfokuskan bahwa tes adalah alat pengumpulan data. Dalam prosedurnya, pengumpulan data tidak hanya satu-satunya yang ada dalam evaluasi, tetapi ada juga prosedur lainnya. Namun, untuk pengumpulan data tentu saja diperlukan suatu alat, diantaranya adalah tes.

Ciri-ciri atau kualitas tes yang baik sangat dibutuhkan dalam mengevaluasi capaian belajar peserta didik. Sehingga hasil yang diinginkan bisa tercapai, dan juga dapat digunakan sebagai identifikasi terhadap pemahaman peserta didik dalam proses pembelajaran (Zuhriyah, 2024). Tes sebagai alat pengumpulan data yang baik tidak boleh sembarangan, yakni pada instrumen data harus melengkapi sejumlah karakteristik penting agar dapat dikatakan sebagai alat ukur yang baik (Sumaryanta, 2021). Jadi, dapat dipahami bahwa tes yang baik adalah tes yang memperhatikan karakteristik penting dalam syarat pembuatannya. Sejalan dengan penjelasan di atas, Arikunto mengatakan tes dikatakan baik apabila memiliki karakteristik seperti validitas, reliabilitas, objektif, praktis, dan ekonomis (Rudini, 2020).

Pertama, validitas berasal dari kata valid, yang artinya benar, tepat, absah, atau sah. Scarvia B Anderson dkk menyebutkan “sebuah tes dikatakan valid apabila tes tersebut mengukur apa yang hendak diukur.” Artinya tidak ada validitas yang berlaku untuk umum. Namun, berbicara terkait validitas ini tidak menekankan kepada tes itu sendiri, melainkan kepada skor atau hasil pengetesannya (Arikunto, 2018). Validitas dapat dipahami sebagai proses dalam pengumpulan bukti-bukti yang mendukung jenis kesimpulan yang akan diambil dari skor tes (Suharman, 2018). Sederhanya validitas tes akan memberi tahu bagaimana skor-skor tes akan disimpulkan.

Sebelum digunakan sebagai alat untuk mengumpulkan data, instrumen tes terlebih dulu diuji cobakan sampai memenuhi persyaratan yaitu valid. Dinyatakan valid apabila butir-butir tes benar-benar mengukur tujuan tes dalam kemampuan hal tertentu, bukan kemampuan yang lain. Validitas suatu tes menggunakan indeks angka yang akan mengarahkan sampai mana hasil dari pengukuran dapat mencerminkan secara tepat kemampuan yang diperoleh seseorang dalam berbagai proses (Mukhlisa, 2023).

Pada konsepnya ada tiga macam validitas tes, diantaranya validitas isi (*content validity*), validitas konstruk (*construct validity*), dan validitas empiris atau kriteria (Ramadhan et al., 2024):

Pertama, validitas isi adalah validitas yang berfokus kepada elemen-elemen apa yang ada dalam ukur (Ihsan, 2015), sehingga analisis rasional adalah proses utama yang dilakukan dalam analisis validitas isi (Islamiani Safitri, et al., 2024). Validitas isi dari sebuah instrument ditentukan dengan cara mencocokkan apakah butir-butir yang ada diinstrumen sudah mewakili komponen yang akan diukur atau belum. Kedua, validitas konstruksi adalah validitas yang dilihat dari segi kerangka, susunan, atau rekaannya (Zuhriyah, 2024). Proses pengujian validitas konstruk yaitu dengan menghubungkan

alat ukur dengan alat ukur yang lain yang memiliki kesamaan konsep (Ihsan, 2015). Apabila susunan tes sudah memenuhi syarat penyusunan tes maka tes tersebut sudah memenuhi syarat validitas konstruks. Apabila tidak berarti belum memenuhi persyaratan validitas konstruks.

Ketiga, validitas kriteria adalah pengaitan alat ukur dengan alat ukur yang lain sebagai kriteria dengan teori yang ada. Validitas yang dikaitkan dengan kriteria mengarahkan pada sejauh mana hubungan antara skor tes yang dikembangkan dengan kriteria luar yang dipercaya dapat menggambarkan ciri-ciri atau tingkah laku yang diukur. Adapun validitas yang dikaitkan dengan kriteria, yaitu: (a) validitas pengukuran setara (*congruent validity*), jenis validitas ini ditetapkan dengan cara mengkorelasikan antara skor tes yang dikembangkan dengan skor dari tes yang sejenis atau setara; (b) validitas pengukuran serentak (*concurrent validity*), validitas ini ditetapkan dengan cara mengkorelasikan hasil skor tes yang akan disusun dengan skor tes lain (tidak sejenis) yang saat pengetesannya dilakukan bersamaan atau berdekatan waktunya; dan (c) validitas ramalan (*predictive validity*), adalah suatu kondisi yang menunjukkan seberapa jauhkah sebuah tes telah dapat dengan secara tepat menunjukkan kemampuan untuk meramalkan apa yang bakal akan terjadi pada masa yang akan datang.

Kedua, memiliki reliabilitas yang baik. Sebuah tes dikatakan baik apabila mempunyai reliabilitas yang tinggi jika tes tersebut memberikan data hasil yang tetap (Setiyawan, 2014). Reliabilitas adalah suatu perangkat ukur yang dapat dipercaya, jika hasil tes relatif sama atau tidak berubah apabila dilakukan tes secara berulang-ulang. Alat ukur yang demikian dinamakan reliabel (Susetyo, 2015). Suatu perangkat ukur jika digunakan dua kali untuk mengukur gejala yang sama memberikan hasil yang relatif sama, maka alat ukur itu reliabel. Namun, jika pengukuran tidak memberikan hasil yang berbeda atau relatif sama maka hasil tes tersebut stabil meskipun dilakukan tes berulang-ulang.

Budi Susetyo dalam bukunya mengatakan ada tiga jenis reliabilitas yang dapat digunakan untuk menentukan reliabilitas (Susetyo, 2015), yaitu: (1) reliabilitas Stabilitas. Menggunakan satu perangkat alat ukur dan dikenal dengan uji ulang (*test retest*). Pelaksanaannya, peserta tes dikenai ujian dua kali diselingi selang waktu dengan menggunakan perangkat ukur yang sama. Selang waktu tesnya dianjurkan oleh para ahli sekitar tiga minggu antara ujian pertama dengan ujian kedua. Reliabilitas perangkat ukur didasarkan kepada kestabilan jawaban yang diberikan oleh peserta tes. Jika jawaban peserta relatif sama maka perangkat tes reliabel; (2) reliabilitas Konsistensi Internal. Reliabilitas ini berdasarkan pada skor yang diperoleh dari satu perangkat ukur dengan satu kali pengukuran pada peserta tes. Ketetapan butir tes yang dijawab benar oleh peserta tes menggambarkan reliabilitas seluruh tes. Perangkat tes dikatakan reliabel jika skor setiap butir atau pasangan butir tes memiliki kaitan erat satu sama lainnya; dan (3) reliabilitas ekuivalen atau uji paralel, adalah pengujian alat ukur yang dibuat dengan alat ukur yang setara. Pengujian ini digunakan untuk melihat ekuivalensi (kesetaraan) dari kedua perangkat ukur pada suatu pengukuran. Prosedur pelaksanaan yang digunakan dalam pengujian reliabilitas, yaitu responden menempuh dua kali pengukuran dengan menggunakan alat setara dan ada jeda selang waktu atau tidak ada selang waktu. Cara pengujian reliabilitas hampir sama dengan tes ulang hanya yang membedakannya perangkat ukur yang digunakan. Pada pengujian reliabilitas ekuivalen, ada dua perangkat ukur yang setara, sedangkan pada tes ulang perangkat ukurnya hanya ada satu perangkat.

Ketiga, memiliki objektivitas yang baik. Tes objektif adalah tes yang dalam pemeriksaannya dapat dilakukan secara objektif. Hal ini dimaksudkan untuk mengatasi kelemahan-kelemahan dari tes bentuk essay. Tes objektif sering disebut tes dikotomi karena jawabannya antara benar atau salah dan skornya antara 1 atau 0. Disebut tes objektif karena penilaiannya objektif. Siapapun yang mengoreksi jawaban tes objektif hasilnya akan sama karena kunci jawabannya sudah jelas dan pasti (Ismail, 2020). dalam penggunaan tes objektif jumlah soal yang diajukan cenderung lebih banyak dari pada tes essay. Terkadang untuk tes yang berlangsung selama 60 menit dapat diberikan 30-40 soal (Mulyadi, 2010).

Keempat, teruji kepraktisannya. Praktibilitas atau praktis pada tes merujuk kepada kemudahan dan kepraktisan tes dalam proses administrasi. Pada praktibilitas ini tes bersifat simple dan jelas, tes mudah untuk diperiksa, mudah untuk dilaksanakan, dan dilengkapi dengan petunjuk rinci dan jelas. Sebuah tes disebut memiliki praktibilitas yang tinggi apabila tes tersebut bersifat praktis, seperti (Zuhriyah, 2024): (1) mudah dilaksanakan, tidak menuntut peralatan yang banyak dan memberi kebebasan kepada peserta didik mengerjakan terlebih dahulu bagian yang dianggap mudah. Karena bersifat sederhana dalam arti tidak memerlukan peralatan yang sulit pengadaannya; (2) mudah untuk diperiksa, artinya tes itu dilengkapi kunci jawaban maupun pedoman skoringnya. Dilengkapi dengan petunjuk-petunjuk yang jelas sehingga dapat diberikan atau diawali orang lain; dan (3) dilengkapi dengan petunjuk-petunjuk yang jelas sehingga dapat diberikan atau diawasi orang lain.

Kelima, tes bersifat ekonomis. Ciri ekonomis pada tes bermaksud bahwa tes yang akan dilakukan tidak memiliki biaya yang mahal, dan biaya yang dikeluarkan masih bisa dijangkau sehingga tidak memberatkan dalam pelaksanaannya nanti. Kelima ciri-ciri dalam penjelasan di atas merupakan ciri-ciri tes yang baik untuk dilakukan ketika mengevaluasi capaian pembelajaran peserta didik. Namun, dari kelima ciri-ciri tersebut terdapat dua ciri yang sangat perlu diperhatikan dalam pelaksanaan sebuah tes yakni validitas dan reliabilitas. Karena kedua ciri ini sering dijadikan dasar dalam menentukan keakuratan sebuah tes sebagai sebuah alat ukur, baik dalam penggunaan sebagai instrumen keberhasilan belajar maupun dalam instrument. Jadi, sebuah tes harus memiliki keabsahan (*valid*) dan keterpercayaan (*reliable*) sehingga nantinya hasil dari tes tersebut dapat dipertanggungjawabkan.

Tahap Penyusunan Instrumen

Terdapat beberapa langkah yang diperlukan sebelum menyusun sebuah tes, supaya tes yang diberikan tidak berbeda dengan tujuan pelaksanaan tes. Langkah dalam menyusun tes diantaranya sebagai berikut (Ajat Rukajat, 2018):

1. Merumuskan atau menentukan tes.
2. Mengidentifikasi hasil belajar yang hendak diukur dalam tes tersebut.
3. Menandai hasil belajar yang spesifik, yang merupakan tingkah laku atau aktivitas yang bisa diamati dan sesuai dengan Tujuan Pembelajaran Khusus (TPK).
4. Merinci bahan atau mata pelajaran yang akan diukur.
5. Menyiapkan tabel spesifikasi.
6. Menggunakan tabel spesifikasi, sebagai dasar penyusunan tes.

Adapun agar dihasilkan instrumen yang valid dan berkualitas, terdapat langkah-langkah pengembangan sebagai berikut (Zuhriyah, 2024):

a. Menentukan Tujuan

Ada dua tujuan yang ditentukan dalam hal ini, yaitu: (1) Tujuan pembelajaran yang diukur dan dinilai dimensi pertama ini merujuk pada berbagai macam ranah dan sub ranah, yang menjadi tujuan pembelajaran selama jangka waktu tertentu. Jika tujuan-tujuan itu telah dirumuskan sebelumnya, seperti dalam satuan pelajaran, langkah penentuan tujuan ini berarti memilih Tujuan Pembelajaran Khusus yang telah diupayakan untuk dicapai dalam proses pembelajaran, untuk dijadikan objek pengukuran dan kriteria keberhasilan dalam penilaian; dan (2) tujuan dilaksanakannya evaluasi harus mengetahui tujuan evaluasinya secara jelas. Misalnya apakah untuk formatif, diagnostik, atau seleksi. Dengan tujuan evaluasi yang berbeda, meskipun dengan lingkup ranah dan materi yang sama, instrument yang dikembangkan juga berbeda.

b. Mengembangkan Spesifikasi Instrumen

Tujuan utama mengembangkan spesifikasi instrument adalah sebagai pembuatan kisi-kisi instrument/tes. Kisi-kisi ini dibuat untuk merumuskan setepat mungkin ruang lingkup dan tekanan tes dan bagian-bagiannya, sehingga perumusan tersebut dapat menjadi petunjuk yang efektif bagi penyusunan tes.

c. Memilih Jenis dan Tipe Instrumen

Jenis dan tipe instrument yang dapat dipakai dalam pengukuran dan penilaian pendidikan memang sangat beragam. Pemilihan dan jenis instrument harus dilakukan dengan hati-hati sehingga tujuan evaluasi dapat diacapai dengan baik.

d. Mempersiapkan Kisi-Kisi Instrumen

Mempersiapkan dan membuat instrumen pengukuran bukanlah pekerjaan ringan, bahkan membutuhkan berbagai kombinasi kemampuan khusus. Dalam hal ini Sumadi Suryabrata ada lima kemampuan khusus yang harus dimiliki, yaitu: (a) penguasaan materi yang diteskan; (b) kesadaran mengenai tata-tata nilai yang mendasari pendidikan; (c) pemahaman tentang karakteristik peserta didik yang diukur; (d) kemampuan membahasakan gagasan; (e) penguasaan instrument penelitian

e. Uji Coba Instrumen

Instrumen yang sudah disusun perlu diuji cobakan untuk diperbaiki, direvisi agar supaya kualitasnya semakin baik. Semakin banyak frekuensi uji coba dan revisinya, semakin bagus kualitas instrument yang dikembangkan itu.

KESIMPULAN

Evaluasi dalam pembelajaran tidak akan pernah lepas dari proses belajar mengajar. Evaluasi dilakukan untuk melihat kemajuan dan pencapaian peserta didik selama pembelajaran. Tes adalah alat pengumpulan data dalam evaluasi. Pemahaman mendalam sangat diperlukan dalam menyusun tes yang baik. Tes dikatakan baik apabila sudah memenuhi karakteristik seperti validitas, reliabilitas, objektif, praktis, dan ekonomis. Tes yang diberikan atau diujikan haruslah sesuai dengan tujuan pelaksanaan tes yaitu dengan memperhatikan langkah-langkah dalam pembuatan instrument tes seperti menentukan tujuan tes, mengembangkan spesifikasi instrument, memilih jenis dan tipe instrumen, mempersiapkan kisi-kisi instrument, dan melakukan uji coba instrumen.

Dengan membuat tes yang baik serta mengikuti langkah-langkahnya dalam pembuatan tes diharapkan mampu mengukur secara akurat kemampuan dan pemahaman peserta didik, memberikan gambaran yang jelas apa yang harus ditingkatkan, dan membantu dalam pengembangan pengembangan kurikulum yang lebih baik.

SARAN

Secara praktis, pendidik perlu meningkatkan keterampilan dalam menyusun dan mengimplementasikan tes yang valid, reliabel, objektif, praktis, dan ekonomis. Pelatihan dan workshop yang berfokus pada pengembangan instrumen evaluasi dapat membantu pendidik memahami dan menerapkan langkah-langkah yang tepat dalam pembuatan tes. Penggunaan teknologi dalam evaluasi juga dapat meningkatkan efisiensi dan akurasi dalam pengumpulan dan analisis data.

Secara teoritis, penelitian lebih lanjut diperlukan untuk mengembangkan dan memvalidasi teori-teori yang mendasari pembuatan instrumen tes. Penelitian ini dapat mencakup studi tentang validitas dan reliabilitas berbagai jenis tes, serta pengaruh faktor-faktor seperti budaya, bahasa, dan konteks pendidikan terhadap hasil tes. Pengembangan teori yang lebih komprehensif akan membantu dalam menciptakan instrumen evaluasi yang lebih akurat dan relevan.

Untuk penelitian lanjutan, disarankan untuk melakukan studi empiris yang mengevaluasi efektivitas berbagai metode dan teknik dalam pembuatan dan pelaksanaan tes. Penelitian ini dapat mencakup uji coba instrumen pada berbagai populasi peserta didik, analisis data hasil tes, dan pengembangan model evaluasi yang lebih adaptif dan responsif terhadap kebutuhan peserta didik. Penelitian tentang penggunaan teknologi dalam evaluasi, seperti aplikasi berbasis komputer dan analisis data otomatis, juga dapat memberikan wawasan baru tentang cara meningkatkan kualitas dan efisiensi evaluasi pembelajaran.

DAFTAR RUJUKAN

- Adom, D., Mensah, J. A., & Dake, D. A. (2020). Test, measurement, and evaluation: Understanding and use of the concepts in education. *International Journal of Evaluation and Research in Education*, 9(1), 109–119. <https://doi.org/10.11591/ijere.v9i1.20457>
- Ajat Rukajat. (2018). *Teknik Evaluasi Pembelajaran* (1st ed.). Deepublish.
- Arikunto, S. (2018). *Dasar-Dasar Evaluasi Pendidikan*, Edisi 3 (R. Damayanti (ed.)). Bumi Aksara.
- Asiva Noor Rachmayani. (2015). *No Title*. 6.
- Ihsan, H. (2015). Validitas Isi Alat Ukur Penelitian: Konsep Dan Panduan Penilaiannya. *PEDAGOGIA Jurnal Ilmu Pendidikan*, 13(3), 173. <https://doi.org/10.17509/pedagogia.v13i3.6004>
- Islamiani Safitri, Dewi Lestarani, Rahmah Dwi Nor Wita Imtikhanah, Nur Rahmi Akbarini, Meida Wulan Sari, Muh. Fitrah, Ilyas, Taufik Rizki Sista, Ikhsan Dwi Setyono, Rizki Fitria Setyaningtyas, O. N. (2024). *Teori Pengukuran dan Evaluasi* (W. Edi Istiyono (ed.); 1st ed.). CV Ruang Tentor.
- Ismail, M. I. (2020). *Asesmen dan Evaluasi Pembelajaran* (Syarifudin (ed.); 1st ed.). Cendekia Publisher.
- Khumaedi, M. (2012). Reliabilitas Instrumen Penelitian Pendidikan. In *Jurnal Pendidikan Teknik Mesin* (Vol. 12, Issue 1, pp. 25–30). <https://journal.unnes.ac.id/nju/index.php/JPTM/article/view/5273>
- Magdalena, I. (2022). *Dasar-Dasar Evaluasi Pendidikan* (H. Wijayanti (ed.); 1st ed.). CV Jejak.
- Maulani, G., Septiani, S., Rusmayani, N. S. N. G. A. L., Sukamdi, Evenddy, S. S., Lasri, Nababan, H. S., Setiadi, K., Rahayu, I., Simanungkalit, L. N., Edi, S., Sari, N., Atin, S., Thoif, M., Raju, M. J., Rosalinda, Saptadi, N. T. S., Nurlely, L., &

- Syafriansyah, M. (2024). *Evaluasi Pembelajaran* (A. C. Purnomo (ed.)). Sada Kurnia Pustaka.
- Mukhlisa, N. (2023). Validitas Tes Nurul. *JUARA SD: Jurnal Pendidikan Dan Pembelajaran Sekolah Dasar*, 2(1), 142–147.
- Mulyadi. (2010). *Evaluasi Pendidikan Pengembangan Model Evaluasi Pendidikan Agama Islam di Sekolah* (A. N. Kawakip (ed.); 1st ed.). UIN Maliki Press.
- Rahman, A. A., & Nasryah, C. E. (2019). Evaluasi Pembelajaran. In *Uwais Inspirasi Indonesia*.
- Ramadhan, M. F., Siroj, R. A., & Afgani, M. W. (2024). Validitas and Reliabilitas. *Journal on Education*, 6(2), 10967–10975. <https://doi.org/10.31004/joe.v6i2.4885>
- Rudini, M. (2020). Efektivitas Analisis Butir Soal Mata Pelajaran Matematika pada Siswa Kelas IV dalam Meningkatkan Kualitas Guru di SDN Sabang. *Tolis Ilmiah: Jurnal Penelitian*, 2(1), 17–27. https://ojs.umada.ac.id/index.php/Tolis_Ilমiah/article/view/90
- Setiyawan, A. (2014). Faktor- faktor yang Mempengaruhi Reliabilitas Tes. *Jurnal An Nûr*, VI(2), 341–354.
- Spiritual, P. D. N. (2024). 1,2,3 ,. 197–206. <https://doi.org/10.15642/jkpi.2024.14.2.197-206>
- Suharman. (2018). Tes Sebagai Alat Ukur Prestasi Akademik. *Jurnal Ilmiah Pendidikan Agama Islam*, 10(1), 93–115.
- Sumaryanta. (2021). *Model Pengembangan Tes* (R. O. Akbar (ed.); 1st ed.). CV Convident.
- Susanto, S. (2023). Pengembangan Alat Dan Teknik Evaluasi Tes Dalam Pendidikan. *Jurnal Tarbiyah Jamiat Kheir*, 1(1), 51–60.
- Susetyo, B. (2015). *Prosedur Penyusunan dan Analisis Tes untuk Penilaian Hasil Belajar Bidang Kognitif* (Anna (ed.); 1st ed.). Refika Aditama.
- Zuhriyah, I. A. (2024). *Pengembangan Evaluasi Pembelajaran* (E. R. Ananda (ed.); 1st ed.). CV Jejak.