

VALIDITAS ALAT UKUR DALAM EVALUASI PEMBELAJARAN: STUDI FAKTOR YANG MEMPENGARUHI VALIDITAS ALAT UKUR

Ega Najwa Ismail¹, Denissa Aulia Ramadhani², Ellsya Endany Anie Ramadhan³, Akil⁴,
Abdul Azis⁵

2310631110209@student.unsika.ac.id¹, 2310631110208@student.unsika.ac.id²,
2310631110210@student.unsika.ac.id³, akil@fai.unsika.ac.id⁴, abdulaziz@fai.unsika.ac.id⁵

Universitas Singaperbangsa Karawang

ABSTRAK

Validitas alat ukur adalah, aspek krusial dalam evaluasi pembelajaran untuk menentukan sejauh mana instrumen mampu mengukur secara akurat. Pada penerapannya, masih sangat banyak instrumen evaluasi masih menunjukkan hasil yang tidak sesuai dengan kompetensi peserta didik yang sebenarnya. Ini menunjukkan bahwa adanya permasalahan pada instrumen validitas yang digunakan. Penelitian ini bertujuan untuk menganalisis faktor-faktor yang mempengaruhi validitas alat ukur pada evaluasi pembelajaran, yang mencakup teknis dan non-teknis. Dan menggunakan metode studi literatur review dengan pendekatan kualitatif deskriptif, data yang dikumpulkan dari berbagai jurnal ilmiah, buku akademik, dan sumber terpercaya lainnya. Pada penelitian ini menunjukkan bahwa validitas alat ukur dipengaruhi oleh empat faktor utama: isi tes, konstruksi instrumen, kondisi peserta didik, dan aspek administrasi tes. Ketidaksesuaian pada penyusunan soal, kondisi, psikologis peserta didik saat melakukan tes, dan pelaksanaan tes yang tidak sesuai standar, berkontribusi besar terhadap rendahnya validitas hasil penilaian. Agar dapat meningkatkan validitas alat ukur, sangat disarankan untuk menyusun soal sesuai dengan tujuan pembelajaran, menggunakan bahasa yang mudah dimengerti, validasi oleh ahli, serta pelaksanaan tes yang konsisten. Penelitian ini memiliki andil yang sangat penting dalam pengembangan evaluasi pembelajaran yang lebih objektif, akurat, dan adil dengan menyoroti pendekatan holistik dalam menjaga kualitas alat ukur. Dengan ini, hasil evaluasi bisa dijadikan dasar untuk pengambilan keputusan pendidikan yang tepat dan bertanggung jawab.

Kata Kunci: Validitas Alat Ukur, Evaluasi Pembelajaran, Akurasi Hasil Penilaian, Faktor Pengaruh.

ABSTRACT

The validity of measuring instruments is a crucial aspect in learning evaluation to determine the extent to which the instrument is able to measure accurately. In its application, many evaluation instruments still show results that are not in accordance with the actual competence of students. This indicates a problem with the validity of the instrument used. This study aims to analyze the factors that affect the validity of measuring instruments in learning evaluation, which include technical and non-technical. And using the literature study method with a qualitative approach, data collected from various scientific journals, academic books, and other reliable sources. The results of this study indicate that the validity of measuring instruments is influenced by four main factors, namely test content, instrument construction, learner conditions, and aspects of test administration. Discrepancies in the preparation of questions, conditions, psychological students at the time of the test, and the implementation of tests that are not in accordance with standards, contribute greatly to the low validity of the assessment results. In order to increase the validity of measuring instruments, it is strongly recommended to develop questions that are in accordance with learning objectives, use language that is easy to understand, validation by experts, and consistent test implementation. This research has a very important role in developing learning evaluations that are more objective, accurate and fair by prioritizing a holistic approach in maintaining the quality of measuring instruments. Thus, evaluation results can be used as a basis for making appropriate and responsible educational decisions.

Keywords: Instrument Validity, Learning Evaluation, Assessment Accuracy, Influencing Factors.

PENDAHULUAN

Di dunia pendidikan, evaluasi pembelajaran merupakan komponen penting yang digunakan untuk mengukur pencapaian kompetensi peserta didik serta efektivitas proses pembelajaran. Evaluasi tidak hanya berfungsi sebagai alat ukur keberhasilan siswa, tetapi juga menjadi dasar dalam perbaikan strategi pengajaran, pengambilan keputusan kurikulum, dan peningkatan kualitas pendidikan secara menyeluruh. Agar hasil evaluasi dapat dijadikan dasar pengambilan keputusan yang tepat, diperlukan alat ukur yang valid, yaitu alat yang benar-benar mampu mengukur apa yang seharusnya diukur (Lia, 2021). Validitas alat ukur menjadi kunci utama untuk memastikan bahwa data yang dihasilkan dari proses penilaian benar-benar mencerminkan kemampuan atau pemahaman peserta didik secara objektif dan dapat dipertanggungjawabkan secara ilmiah. Tanpa validitas yang memadai, hasil evaluasi bisa menyesatkan dan menimbulkan bias dalam penilaian, yang pada akhirnya berdampak pada kualitas pendidikan secara keseluruhan, mulai dari kesalahan dalam mengidentifikasi kebutuhan belajar siswa hingga ketidaktepatan dalam menetapkan kebijakan pendidikan. Oleh karena itu, pemahaman terhadap konsep validitas, jenis-jenisnya, serta berbagai faktor yang dapat memengaruhinya menjadi penting bagi setiap pihak yang terlibat dalam perancangan dan pelaksanaan evaluasi pembelajaran, baik itu guru, dosen, peneliti pendidikan, maupun pengembang kurikulum. Validitas bukan sekadar persoalan teknis dalam penyusunan instrumen, melainkan berkaitan erat dengan integritas sistem pendidikan itu sendiri. (Ida & Musyarofah, 2021)

Meskipun validitas alat ukur telah lama diakui sebagai aspek krusial dalam evaluasi pembelajaran, pada praktiknya masih banyak ditemukan permasalahan yang menghambat tercapainya validitas yang optimal. Salah satu masalah utama adalah kurangnya perhatian terhadap proses konstruksi instrumen, di mana banyak alat ukur disusun tanpa melalui tahapan analisis indikator, uji coba, atau validasi empiris yang memadai. Selain itu, ketidaksesuaian antara indikator yang diukur dengan tujuan pembelajaran sering kali menyebabkan hasil evaluasi tidak mencerminkan kompetensi yang sebenarnya. Faktor lain yang juga menjadi kendala adalah minimnya pemahaman pendidik mengenai prinsip-prinsip dasar validitas dan bagaimana menerapkannya dalam konteks kelas. Dalam beberapa kasus, keterbatasan waktu, sumber daya, dan tekanan administratif menyebabkan penyusunan evaluasi dilakukan secara terburu-buru, sehingga aspek validitas menjadi terabaikan. Permasalahan-permasalahan tersebut menunjukkan bahwa isu validitas bukan hanya persoalan teknis, tetapi juga berkaitan dengan kesadaran, kompetensi, dan budaya evaluasi dalam lingkungan Pendidikan (Farida & Musyarofah, 2021).

Penelitian mengenai pengembangan dan validitas alat ukur dalam evaluasi pembelajaran telah banyak dilakukan oleh para akademisi, salah satunya oleh (Zahroh et al., 2025) yang berfokus pada proses pembuatan instrumen evaluasi yang sah dan andal melalui analisis kebutuhan, penyusunan butir soal, serta uji coba instrumen di kelas. Hasil penelitian tersebut menunjukkan bahwa dengan proses pengembangan yang sistematis, alat evaluasi dapat mencerminkan kemampuan peserta didik secara objektif dan memberikan umpan balik yang berguna bagi pengajar. Penelitian ini menekankan pentingnya keterkaitan antara instrumen evaluasi dengan tujuan pembelajaran, serta perlunya pendekatan yang terstruktur dalam menyusun instrumen agar hasil penilaian akurat dan bermanfaat. Namun, penelitian tersebut lebih menitikberatkan pada proses teknis pengembangan instrumen, sementara penelitian ini mengambil sudut pandang yang lebih luas dengan menganalisis faktor-faktor yang memengaruhi validitas alat ukur, baik dari aspek instrumen, responden, maupun faktor eksternal yang mungkin muncul selama proses evaluasi berlangsung.

Meskipun berbagai penelitian sebelumnya telah membahas pentingnya pengembangan instrumen evaluasi pembelajaran yang valid dan andal, sebagian besar masih

berfokus pada aspek teknis penyusunan alat ukur, seperti analisis kebutuhan, pembuatan butir soal, dan uji coba instrumen. Penelitian-penelitian tersebut belum banyak menyoroti secara mendalam faktor-faktor yang secara simultan memengaruhi validitas alat ukur dalam konteks yang lebih luas (Phafiandita et al., 2022). Di sinilah letak kebaruan (novelty) dari penelitian ini, yaitu dengan mengkaji validitas alat ukur dari berbagai faktor yang mempengaruhi secara menyeluruh, meliputi isi tes, konstruksi instrumen, jawaban peserta didik, serta aspek administrasi tes. Pendekatan ini memberikan gambaran yang lebih komprehensif dan aplikatif mengenai bagaimana setiap komponen dapat berkontribusi terhadap tinggi rendahnya validitas dalam evaluasi pembelajaran. Dengan demikian, penelitian ini diharapkan dapat mengisi celah (gap) dalam literatur yang selama ini cenderung memisahkan antara aspek teknis dan non-teknis dalam pembahasan mengenai validitas alat ukur (Nurhalimah et al., 2022).

Urgensi dari penelitian ini terletak pada perlunya pemahaman yang lebih menyeluruh mengenai validitas alat evaluasi pendidikan, yaitu dengan mengkaji berbagai faktor yang mempengaruhinya. Mayoritas penelitian masih berfokus pada aspek teknis, seperti validasi statistik dan analisis butir soal. (Haq, 2022). Penelitian ini bertujuan untuk menunjukkan validitas dari perspektif pengukuran yang lebih luas, termasuk faktor tes, konstruksi instrumen, respon peserta didik, dan aspek administratif. Dengan pendekatan ini, diharapkan penelitian ini dapat memberikan kontribusi yang lebih bermanfaat bagi pengembangan alat penilaian yang tidak hanya akurat dan kuat secara teoritis, tetapi juga relevan dengan konteks pengajar.

METODE PENELITIAN

Pada penelitian ini menggunakan pendekatan kualitatif deskriptif dengan jenis studi kepustakaan (literature review). Peneliti memilih pendekatan ini untuk meninjau secara dalam berbagai perspektif dan teori yang relevan yang berkenaan dengan validitas alat ukur dalam evaluasi pembelajaran. Sumber data pada penelitian ini berasal dari data sekunder, berupa artikel ilmiah, buku akademik, jurnal nasional, dan dokumen-dokumen terpercaya yang membahas konsep validitas, evaluasi pembelajaran, dan faktor yang dapat mempengaruhi akurasi penilain. Karena penelitian ini bersifat studi literatur, maka objek pada penelitian ini adalah konsep-konsep dan temuan ilmiah yang relevan dengan topik pembahasan. Dan subjek penelitian berupa sumber-sumber rujukan tertulis yang dianalisis.

Teknik pengumpulan data dilakukan dengan penelusuran sistematis pada pustaka-pustaka ilmiah yang didapatkan pada jurnal, perpustakaan digital, dan referensi akademik lainnya yang telah terverifikasi. Teknik analisis data yang dipakai adalah dengan analisis isi (content analysis), yaitu dengan cara mengidentifikasi, dan mengelompokkan informasi dari berbagai sumber untuk bisa menemukan pola, tema utama, dan hubungan antara faktor-faktor yang dapat mempengaruhi validitas alat ukur. Kemudian hasil analisis disusun secara deskriptif untuk menjelaskan bagaimana faktor tersebut dapat berperan untuk menentukan akurasi hasil evaluasi pembelajaran.

HASIL DAN PEMBAHASAN

A. Pengertian validitas alat ukur.

Dalam Kamus Besar Bahasa Indonesia (KBBI), validitas adalah sifat benar menurut bahan bukti yang ada, logika berpikir, atau kekuatan hukum; sifat valid; kesahihan. Validitas mengukur sejauh mana hasil bisa mengukur dengan benar apa yang seharusnya diukur. (KBBI, 2025)

Validitas berasal dari istilah validity yang mempunyai arti sejauh mana ketepatan dan kecermatan suatu alat ukur (tes) dalam melakukan fungsi ukurnya. Apabila suatu tes

dinyatakan mempunyai validitas yang tinggi, berarti alat tersebut telah melakukan pengukuran dengan benar atau memberikan hasil yang sesuai dengan tujuan pengukuran. Dengan kata lain, hasil yang diperoleh dari pengukuran tersebut mencerminkan dengan tepat kenyataan dan kondisi sebenarnya dari apa yang diukur. (Ramadhan, 2024)

Validitas suatu tes sebagian besar mencerminkan tingkat efektivitas pengukuran tes, atau tingkat keakuratan ukuran tes tertentu. Validitas tes berfokus pada sejauh mana sebuah tes benar-benar mengukur apa yang seharusnya diukur. Maksudnya seberapa jauh suatu tes mampu mengungkapkan dengan benar ciri atau keadaan yang sesungguhnya dari obyek yang diukur, akan tergantung dari tingkat validitas tes yang berkaitan. (Matondang, 2009)

Validitas berkaitan dengan sejauh mana alat penilaian mampu mengukur dengan tepat konsep yang akan dinilai, sehingga benar-benar mencerminkan apa yang seharusnya diukur. Sebuah tes yang valid untuk satu tujuan pengambilan keputusan tertentu belum tentu valid jika digunakan untuk tujuan atau keputusan yang berbeda. Oleh karena itu, validitas suatu tes harus selalu dikaitkan dengan konteks dan tujuan evaluasi yang spesifik. Sebagai contoh, tes seleksi masuk sekolah menengah atas harus secara akurat mencerminkan prestasi dan hasil belajar calon siswa. Validitas adalah sejauh mana tes itu dapat mengukur apa yang akan diukur. Validitas alat ukur tidak terganggu (Ihsan H, 2025)

Pentingnya validitas dalam evaluasi pendidikan terletak pada kemampuannya untuk memastikan bahwa hasil tes dapat digunakan sebagai dasar untuk membuat keputusan yang tepat dan akurat. Misalnya, dalam menentukan penempatan siswa, menilai keberhasilan program pendidikan, atau mengevaluasi kinerja guru. Tes yang valid juga memainkan peran penting dalam menilai sejauh mana kurikulum dan metode pengajaran efektif. Melalui tes yang valid, kita dapat mengidentifikasi area di mana siswa mengalami kesulitan, sehingga kurikulum dapat disesuaikan untuk memenuhi kebutuhan belajar mereka. Selain itu, validitas tes sangat penting dalam membangun kepercayaan antara siswa, orang tua, dan pendidik terhadap sistem evaluasi yang digunakan. Hasil evaluasi yang valid memberikan dasar yang kuat untuk menegakkan prinsip pemerataan dalam pendidikan. Selain itu, validitas tes mendorong perbaikan dalam pengembangan instrumen evaluasi dan pendekatan penilaian yang lebih tepat untuk meningkatkan kualitas pendidikan secara keseluruhan. (Mustopa & Khofifah Aulia, 2025)

B. Jenis-Jenis Validitas Alat Ukur

Konsep validitas tes bisa dibedakan menjadi tiga macam yaitu validitas isi (content validity), validitas konstruk (construct validity), dan validitas empiris atau validitas kriteria.

1. Validitas isi (Content validity)

Validitas isi (content validity) Validitas isi, yang juga dikenal sebagai validitas kurikulum, mengacu pada tingkat kesesuaian instrumen penilaian dengan materi atau kompetensi yang tercantum dalam kurikulum. Validitas isi merujuk pada kedalaman tes, yang berisi seperangkat soal-soal, dan dilihat dari isinya memang mengukur sesuatu yang akan diukur. Derajat representatif isi untuk mengukur kedalaman ukuran tes itu. Salah satu cara yang biasa digunakan untuk menentukan validitas bisa dengan cara mengkaji isi tes itu. Misalnya untuk mengukur kemampuan matematika bisa menggunakan tes yang terdiri atas 25 soal penjumlahan dan pengurangan, dibandingkan dengan tes yang terdiri atas 10 soal tentang olah raga tapi tidak ada hal yang bersangkutan dengan penjumlahan dan pengurangan. Jika soal-soal yang digunakan sangat menunjukkan sampel atribut yang diukur, maka itu sudah memenuhi ketentuan validitas isi. Oleh karena itu, validitas isi sangat ditentukan oleh dua faktor utama, yaitu instrumen tes itu sendiri serta proses yang mempengaruhi individu ketika merespon tes tersebut. Sebagai contoh, sebuah tes tertulis dalam perekrutan karyawan mungkin tidak dapat sepenuhnya mengukur kemampuan karyawan yang sebenarnya dalam melakukan suatu tugas, meskipun tes tersebut valid dalam

menilai sejauh mana pengetahuan mereka terkait dengan tugas yang akan dilakukan. Validitas isi dapat diperoleh dengan meninjau setiap butir soal dalam tes. Jika semua item benar-benar mengukur aspek yang dituju, maka validitas isi dapat dianggap telah terpenuhi. (Eliyah, 2019)

2. Validitas konstruk (Construct validity)

Konstruk (construct) mengacu pada konsep atau objek yang abstrak, tetapi gejalanya masih dapat diamati dan diukur. Contohnya adalah gravitasi, massa, kemampuan matematika, kemampuan bahasa Inggris, perasaan senang atau sedih. Untuk memahami konsep konstruk, gravitasi dapat digunakan sebagai ilustrasi. Ketika sebuah apel jatuh ke tanah, kita dapat menggunakan konsep gravitasi untuk menjelaskan dan memprediksi perilaku yang terlihat (seperti jatuhnya apel). Artinya, kita tidak dapat secara langsung melihat “gravitasi” itu sendiri, tetapi kita dapat mengenalinya melalui peristiwa seperti jatuhnya apel. Jadi, yang kita bisa lakukan hanya mengukur gravitasi dan mengembangkan teori tentang gravitasi.

Validitas konstruk mengacu pada sejauh mana suatu alat ukur sesuai dengan kerangka teori yang mendasari pembuatannya. Sebuah tes dianggap memiliki validitas konstruk jika butir-butir soal di dalamnya mampu mengukur semua aspek berpikir yang tercantum dalam standar kompetensi dasar dan indikator yang ada di dalam kurikulum.

Konstruksi yang dimaksud pada validitas berupa rekaan psikologis yang bersinggungan dengan aspek-aspek ingatan, pemahaman, aplikasi, analisis, sintesis, dan evaluasi. (Syamsurizal, 2020)

3. Validitas kriteria (Criterion validity)

Dalam validitas berbasis kriteria, validitas suatu alat ukur dipertimbangkan dari seberapa besar hasil yang diperoleh dengan menggunakan alat yang sedang dibahas serupa dengan hasil alat ukur lain yang digunakan sebagai acuan. Kriteria umumnya mengacu pada hasil pengukuran atribut yang serupa dengan alat lain yang diakui serupa (memenuhi standar akademis dan profesional tertentu). Sebagai contoh, seseorang yang merancang tes kecerdasan menggunakan hasil dari tes Wechsler atau tes Stanford-Binet sebagai referensi.

Validitas kriteria dibagi menjadi dua jenis, tergantung pada kapan alat tersebut dapat digunakan. Jenis pertama disebut validitas konkuren, di mana kriteria dapat digunakan dalam waktu dekat. Jenis kedua disebut validitas prediktif jika kriteria dapat digunakan setelah beberapa waktu berlalu. (Muhammad, 2021)

C. Faktor-Faktor yang Bisa Mempengaruhi Akurasi Hasil Penilaian

Akurasi hasil penilaian bisa dipengaruhi oleh berbagai faktor yang saling bersinggungan. Penilaian yang akurat tidak hanya bergantung pada satu aspek saja, melainkan melibatkan kualitas instrumen, kondisi peserta didik, lingkungan sekitar, hingga peran penguji. Berikut ini adalah beberapa faktor-faktor yang mempengaruhi akurasi hasil penilaian:

1. Faktor Isi Tes

Isi tes merupakan salah satu faktor utama yang mempengaruhi validitas alat ukur. Isi tes sangat berpengaruh pada sejauh mana butir-butir soal yang disusun dapat merepresentasikan semua kompetensi atau materi yang akan diukur. Jika isi tes tidak sesuai dengan tujuan, maka hasil tes tersebut tidak dapat dikatakan valid. Validitas isi bisa dikatakan valid apabila seluruh aspek penting dari materi yang diujikan tercermin dalam soal-soal tes. Sebaliknya, jika hanya sebagian kecil dari materi yang diujikan, atau terdapat soal-soal yang keluar dari konteks materi yang telah diajarkan, maka validitas alat ukur tersebut akan menurun. Oleh karena itu, penting bagi penyusun tes untuk memperhatikan kelengkapan dan relevansi materi dalam soal. (Ndiung, 2020)

Dalam proses penyusunan tes, pembuatan kisi-kisi menjadi langkah penting untuk menjamin validitas isi. Kisi-kisi berfungsi sebagai pedoman dalam menyusun soal agar mencakup seluruh aspek materi yang harus diuji, serta untuk menjaga proporsi antar topik sesuai dengan bobot kepentingannya. Selain itu, bahasa dalam soal juga harus jelas dan tidak menimbulkan makna ganda, agar peserta didik dapat memahami soal sesuai dengan maksud yang diinginkan.

Sebagai contoh, dalam pembuatan tes matematika untuk topik "Aljabar" di tingkat SMP, soal-soal yang disusun harus mencakup semua subtopik yang relevan, seperti penjumlahan, pengurangan, perkalian, dan pembagian aljabar. Apabila soal hanya menguji satu subtopik saja, sementara aspek lain yang sama pentingnya diabaikan, maka tes tersebut tidak dapat dikatakan valid secara isi. (Riinawati, 2021)

2. Faktor Konstruksi Instrumen

Konstruksi instrumen berkaitan dengan bagaimana tes tersebut dirancang dan disusun, baik dari segi format, struktur, maupun penyajian soal. Validitas suatu alat ukur akan menurun apabila instrumen disusun dengan format yang tidak sistematis, tata bahasa yang membingungkan, atau petunjuk pengerjaan yang tidak jelas. Dengan instrumen yang tidak terstruktur dengan baik bisa menyebabkan peserta didik salah memahami maksud soal, sehingga jawaban yang diberikan peserta didik tidak merefleksikan kemampuan atau pengetahuan mereka yang asli.

Agar validitas tetap terjaga, konstruksi instrumen harus memperhatikan beberapa hal penting, seperti kejelasan instruksi, keterbacaan soal, pemilihan bentuk soal yang sesuai dengan tujuan pengukuran, dan urutan soal yang logis. Instrumen dilarang memakai kata-kata yang ambigu atau istilah-istilah teknis yang asing bagi peserta didik. (Widiana, 2023)

3. Faktor Jawaban Peserta Didik

Validitas alat ukur akan terganggu bila jawaban yang diberikan oleh peserta didik tidak sepenuhnya mencerminkan kemampuan atau kompetensi yang sebenarnya. Faktor seperti, kondisi fisik dan mental peserta didik saat tes berlangsung, motivasi dalam mengerjakan tes, atau bahkan karena salah memahami soal, bisa menjadi penyebabnya.

Misalnya, peserta didik yang mengalami kecemasan tinggi saat ujian dapat mengalami penurunan performa, sehingga jawabannya tidak menunjukkan kemampuan yang sesungguhnya. Begitu pula peserta didik yang tidak serius atau kurang termotivasi dalam mengerjakan tes, mungkin memberikan jawaban secara asal-asalan, sehingga hasilnya tidak valid. (Magdalena, 2021)

Selain itu, kesalahan peserta didik dalam memahami petunjuk pengerjaan atau interpretasi soal juga dapat menyebabkan jawaban yang salah, walaupun sebenarnya mereka menguasai materi tersebut. Oleh karena itu, kondisi peserta didik pada saat tes perlu diperhatikan agar hasil yang diperoleh benar-benar valid. Dalam rangka meningkatkan validitas dari faktor ini, guru atau penyusun tes perlu memastikan bahwa suasana ujian kondusif, memberikan motivasi kepada peserta didik, serta memastikan bahwa soal-soal disusun dengan bahasa yang sederhana dan mudah dipahami. (Prijuwanto, 2020)

4. Faktor Administrasi Tes

Faktor administrasi ini berkaitan dengan bagaimana tes dilaksanakan, termasuk tata cara pemberian instruksi, pengaturan waktu, dan prosedur pelaksanaan secara keseluruhan.

Jika proses administrasi tes tidak terstandarisasi atau tidak konsisten, maka hasil yang diperoleh bisa tidak valid. Misalnya, perbedaan dalam cara penyampaian instruksi antar kelompok peserta didik dapat menyebabkan interpretasi soal yang berbeda-beda. Atau, apabila ada ketidaktepatan dalam pengaturan waktu misalnya satu kelompok diberi waktu lebih lama daripada yang lain hasilnya tentu akan mempengaruhi tingkat kesulitan yang dirasakan peserta didik. Untuk menjaga validitas, administrasi tes harus dilaksanakan secara

seragam, adil, dan sesuai dengan prosedur yang telah dirancang. Instruksi harus disampaikan dengan jelas, waktu pengerjaan harus diatur dengan ketat, dan pengawasan harus dilakukan dengan baik. Dengan administrasi tes yang rapi dan konsisten, kemungkinan gangguan terhadap validitas alat ukur dapat diminimalisasi, sehingga hasil tes benar-benar merefleksikan kemampuan peserta didik secara objektif. (Fathorrasik, 2016)

D. Upaya Meningkatkan Validitas Alat Ukur

Validitas alat ukur merupakan aspek penting dalam proses evaluasi pembelajaran. Dengan dilakukan berbagai upaya yang sistematis dan terencana dapat memastikan kalau alat ukur bisa digunakan dan mampu untuk mengukur kompetensi yang diinginkan. Langkah yang dapat dilakukan untuk meningkatkan validitas alat ukur dalam penilaian pendidikan.

1. Menyusun Soal Sesuai Tujuan Pembelajaran

Tujuan pembelajaran adalah arah yang ingin dicapai setelah proses belajar mengajar, misalnya siswa mampu menjelaskan proses fotosintesis atau mampu menerapkan rumus tertentu pada kehidupan sehari-harinya. Oleh karena itu, soal yang disusun harus selaras dengan apa yang ingin dicapai tersebut. Jika soal tidak berhubungan dengan tujuan, maka alat ukur tidak akan benar-benar mengukur kemampuan siswa yang sebenarnya, sehingga validitasnya rendah. Penyusunan soal harus mempertimbangkan apa yang ingin diukur dari siswa. Jika guru ingin mengukur pemahaman konsep, maka soal tidak boleh sekadar menuntut hafalan definisi. Sebaliknya, soal harus mendorong siswa untuk menjelaskan, menganalisis, atau memberikan contoh. Misalnya, alih-alih menanyakan "Apa definisi fotosintesis?", guru bisa memberikan pertanyaan seperti "Jelaskan tahapan-tahapan fotosintesis pada tumbuhan hijau dan berikan contohnya dalam kehidupan sehari-hari." Dengan begitu, soal akan lebih sesuai dengan kompetensi yang diharapkan.

Agar penyusunan soal lebih terarah, penting bagi guru untuk membuat kisi-kisi atau blueprint sebelum mulai menulis soal. Blueprint ini berisi daftar kompetensi yang harus diuji, materi yang menjadi fokus, jenis soal yang digunakan, serta tingkat berpikir yang ingin dicapai, apakah hanya mengingat, memahami, atau menganalisis. Dengan adanya blueprint, guru dapat memastikan setiap soal yang dibuat memiliki alasan yang kuat untuk dimasukkan dalam tes dan tidak keluar dari konteks tujuan pembelajaran. (Dakran, 2024)

2. Melakukan Review dan Validasi oleh Ahli

Keterlibatan ahli dalam proses ini menjadi sangat krusial karena ahli memiliki pengalaman, pengetahuan, serta sudut pandang yang luas dalam bidang yang sedang diuji. Review dan validasi oleh ahli dilakukan dengan meminta seorang atau beberapa orang yang kompeten di bidangnya untuk menilai soal-soal atau instrumen yang sudah disusun. Para ahli ini akan menilai dari berbagai aspek, seperti kesesuaian soal dengan tujuan pembelajaran, tingkat kesulitan soal, kejelasan bahasa, dan relevansi materi. Mereka akan memberikan masukan apakah soal itu sudah mengukur apa yang semestinya diukur, atau justru perlu diperbaiki agar lebih tepat sasaran. Proses ini penting untuk mengurangi bias dari pembuat soal, karena sering kali guru atau penyusun instrumen tidak sadar ada kekeliruan dalam soal yang dibuatnya. Melibatkan ahli juga dapat membantu mendeteksi adanya kesalahan-kesalahan kecil namun berpengaruh besar, seperti soal yang ambigu, terlalu sulit, terlalu mudah, atau soal yang jawabannya bisa multitafsir. Dengan adanya masukan dari ahli, soal-soal tersebut dapat direvisi dan disempurnakan sebelum digunakan kepada siswa. Ini tentu akan meningkatkan validitas dan reliabilitas alat ukur, karena soal yang ambigu atau tidak relevan akan menurunkan kualitas hasil pengukuran. (Handako, 2024)

Proses review ini biasanya dilakukan dengan memberikan instrumen kepada ahli untuk dianalisis, lalu ahli memberikan catatan-catatan atau checklist yang berisi evaluasi terhadap soal tersebut. Kadang, dilakukan juga uji validitas isi dengan menggunakan indeks

validitas, di mana para ahli memberi skor terhadap relevansi masing-masing soal terhadap tujuan pembelajaran, lalu hasilnya dihitung untuk menentukan apakah soal itu layak digunakan.

Dengan melibatkan ahli dalam proses review dan validasi, penyusun alat ukur bisa lebih yakin bahwa instrumen yang digunakan benar-benar berkualitas tinggi. Hal ini akan menghasilkan data penilaian yang lebih akurat dan adil terhadap kemampuan siswa. Oleh karena itu, review dan validasi oleh ahli menjadi bagian penting yang tidak boleh dilewatkan dalam rangka meningkatkan kualitas alat ukur dalam pembelajaran. (TAO Testing, n.d.)

3. Menggunakan Bahasa yang Jelas dan Tepat

Dalam meningkatkan kualitas alat ukur, penggunaan bahasa yang jelas dan tepat menjadi salah satu faktor yang sangat penting. Bahasa yang digunakan dalam soal harus dipastikan dapat dipahami dengan mudah oleh seluruh peserta tes, tanpa menimbulkan kebingungan atau multitafsir. Jika soal ditulis dengan bahasa yang membingungkan, ambigu, atau terlalu rumit, maka kemungkinan besar siswa akan salah paham terhadap maksud soal, bukan karena mereka tidak menguasai materi, melainkan karena tidak memahami pertanyaannya. Hal ini tentu akan menurunkan validitas alat ukur.

Bahasa yang jelas artinya kalimat dalam soal disusun secara sederhana, lugas, dan langsung ke inti persoalan. Hindari penggunaan kalimat yang panjang, berbelit-belit, atau terlalu banyak menggunakan istilah teknis yang tidak perlu. Bahasa yang tepat berarti kata-kata yang dipilih benar-benar sesuai dengan konsep yang ingin diuji, sehingga tidak menimbulkan interpretasi yang salah dari siswa. (Kejarcita, n.d.)

Misalnya, pada saat penyusunan soal Pendidikan Agama Islam, soal harus menunjuk pada pemahaman konsep ajaran Islam atau penerapan nilai-nilai seagamaan dalam kehidupan sehari-hari. Dan soal tidak boleh menggunakan bahasa yang sulit untuk dipahami dan terlalu banyak narasi panjang yang tidak relevan, ini mengakibatkan peserta didik sulit memahami dan mengaburkan fokus pertanyaannya.

4. Pelatihan dan Pengembangan untuk Pembuat Soal

Banyak masalah dalam kualitas soal sebenarnya berasal dari kurangnya pemahaman dan keterampilan pembuat soal itu sendiri. Oleh karena itu, membekali para guru atau penyusun instrumen dengan pelatihan yang baik menjadi langkah strategis untuk meningkatkan validitas dan reliabilitas alat ukur.

Pelatihan ini bisa mencakup banyak aspek, mulai dari bagaimana cara menyusun soal yang baik, prinsip-prinsip dalam membuat soal objektif maupun subjektif, teknik menghindari bias dalam soal, hingga cara melakukan analisis butir soal setelah ujian. Dengan pelatihan, pembuat soal akan lebih memahami bagaimana menyusun soal yang sesuai dengan tujuan pembelajaran, menggunakan bahasa yang jelas, menentukan tingkat kesulitan yang tepat, dan memastikan soal benar-benar mengukur kompetensi yang diinginkan.

Pelatihan bisa diwujudkan dengan kegiatan konkret seperti pelatihan komputer, seminar Bahasa Indonesia, serta seminar meningkatkan profesionalisme guru. Kegiatan-kegiatan ini untuk meningkatkan keterampilan teknis, juga untuk memperkuat kompetensi pedagogis dan profesionalisme guru secara keseluruhan. Ini sejalan dengan penelitian oleh Kartika (2014) yang menunjukkan berbagai program pengembangan pada guru, baik dengan melalui pelatihan teknis dan seminar, sangat berperan penting untuk meningkatkan kualitas pendidikan dan profesionalisme guru.

Pelatihan dan pengembangan juga berfungsi untuk membangun kesadaran bahwa pembuatan soal bukan hanya soal membuat pertanyaan, tetapi merupakan bagian penting dari keseluruhan proses pembelajaran yang harus dirancang secara ilmiah. Melalui pelatihan, para guru juga diajarkan tentang pentingnya validitas dan reliabilitas instrumen,

serta bagaimana cara menguji dan mengevaluasi kualitas soal yang sudah dibuat.

Tanpa adanya pelatihan dan pengembangan, maka risiko terjadinya soal-soal yang tidak valid, bias, atau yang tidak sesuai dengan tujuan pembelajaran akan semakin besar. Oleh sebab itu, investasi dalam pelatihan ini bukan hanya meningkatkan kualitas soal, tetapi juga meningkatkan kualitas pembelajaran secara keseluruhan. Dengan adanya pelatihan dan pengembangan yang terus-menerus, pembuat soal akan semakin terampil, kritis, dan profesional dalam menyusun instrumen penilaian, sehingga pada akhirnya kualitas alat ukur yang digunakan di dalam kelas pun menjadi jauh lebih baik. (LinkkedIn, n.d.)

KESIMPULAN

Validitas alat ukur memegang peranan sentral dalam evaluasi pembelajaran, karena berfungsi memastikan instrumen yang dipakai dapat mengukur kompetensi yang dituju secara tepat. Validitas terbagi menjadi validitas isi, validitas konstruk, dan validitas kriteria, yang masing-masing memberikan kontribusi terhadap ketepatan instrumen dalam mengungkap capaian belajar peserta didik.

Faktor-faktor yang mempengaruhi validitas meliputi kualitas isi tes, konstruksi instrumen, kondisi peserta didik, serta aspek administrasi dalam pelaksanaan tes. Ketidak sempurnaan pada salah satu faktor tersebut dapat berimplikasi pada berkurangnya validitas instrumen.

Upaya sistematis untuk menjaga dan meningkatkan validitas perlu dilakukan, antara lain melalui penyusunan soal yang selaras dengan tujuan pembelajaran, penggunaan bahasa yang tepat dan jelas, validasi instrumen oleh ahli, serta penyelenggaraan tes dengan prosedur yang konsisten. Pelatihan berkelanjutan bagi penyusun instrumen juga menjadi langkah penting untuk memastikan kualitas evaluasi. Dengan demikian, evaluasi pembelajaran dapat menghasilkan data yang akurat, objektif, dan dapat dipertanggungjawabkan dalam mendukung peningkatan mutu pendidikan.

DAFTAR PUSTAKA

- Farida, A. M., & Musyarofah, A. (2021). Validitas dan Reliabilitas dalam Analisis Butir Soal. *Al-Mu'arrib: Jurnal Pendidikan Bahasa Arab*, 1.
- Fathorrasik. (2016). Validitas dan Reabilitas Tes. *Kompasiana*. <https://www.kompasiana.com/fathorrasik1/57a0a5e2ae7e611b19e1a4a2/validitas-dan-reliabilitas-tes>.
- Haq, V. A. (2022). Menguji Validitas Dan Reliabilitas Pada Mata Pelajaran Al Qur'an Hadits Menggunakan Korelasi Produk Momen Spearman Brown. *An-Nawa: Jurnal Studi Islam*, 4(1), 11–24.
- Ida F.F, & Musyarofah. (2021). Validitas dan Reabilitas dalam Analisis Butir Soal. *Al-MU'Arrib: Journal of Arabic Education*, 1(1), 34–44. <https://doi.org/10.32923/al-muarrib.v1i1.2100>
- Ihsan H. (2025). Validitas Alat Ukur Penelitian Konsep dan Panduan Penilaiannya. *PEDAGOGIA: Jurnal Ilmu Pendidikan*, 12(1).
- Kejarcita. (n.d.). Tips Meningkatkan Validitas Tugas/Ujian yang Di Kerjakan Siswa Selama Online Learning. *Kejarcita Blog*. <https://blog.kejarcita.id/tips-meningkatkan-validitas-ujian-yang-dikerjakan-siswa-selama-online-learning/>.
- Lia, M. K. (2021). Validitas dan Reliabilitas Instrumen CTTMC pada Pembelajaran IPA Terpadu SMP Marantika Lia Kristyasari. *Pedagogika: Jurnal Ilmu-Ilmu Kependidikan*, 1(1), 76–85.
- LinkkedIn. (n.d.). How Do You Improve Your Skills and Knowledge of Validity an Reability? *Linkedin.Com*. <https://www.linkedin.com/advice/1/how-do-you-improve-your-skills-knowledge-validity-reliability>.
- Mustopa, & Khofifah Aulia. (2025). Pentingnya Validitas dalam Evaluasi Pembelajaran. *Kompasiana*. <https://www.kompasiana.com/khofifahauliamustopa3437/664604f7de948f6a68713202/penti>

- ngnya-validitas-dalam-evaluasi-pembelajaran).
- Nurhalimah, S., Hidayati, Y., Rosidi, I., & Hadi, W. P. (2022). Hubungan antara validitas item dengan daya pembeda dan tingkat kesukaran soal pilihan ganda pas. *Natural Science Education Research (NSER)*, 4(3), 249–257.
- Phafiandita, A. N., Permadani, A., Pradani, A. S., & Wahyudi, M. I. (2022). Urgensi evaluasi pembelajaran di kelas. *JIRA: Jurnal Inovasi Dan Riset Akademik*, 3(2), 111–121.
- Riinawati. (2021). *Pengantar Evaluasi Pendidikan*. Yogyakarta: Thema Publishing.
- TAO Testing. (n.d.). 4 Cara Meningkatkan Validitas Isi Ujian. TAO Testing Blog. <https://www.taotesting.com/blog/4-ways-to-improve-exam-content-validity/>.
- Zahroh, F. L., Muflih, M., & Hilmiyati, F. (2025). PEMBUATAN INSTRUMEN EVALUASI PEMBELAJARAN ; Analisis Literatur Review. 5, 319–328.
- Dakran, D. Z. (2024). Evaluasi Dalam Pross Pembelajaran di SMA Negeri 1 Panyambung Selatan. *Congnoscere: Jurnal Komunikasi dan Media Pendidikan*, 2(2).
- Eliyah. (2019). Validitas Tes Sebagai Media Ukur Ketepatan Evaluasi Mengukur Pencapaian Tujuan Pembelajaran Di SD/MI. *Jurnal Edukatif*, V(1) 45.
- Handako, Y. W. (2024). *Metode Penelitian Kualitatif Panduan Praktis Untuk Penelitian Administrasi Pendidikan*. PT. Sonpedia Publishing Indonesia .
- Kartika, J. (2014). Implementasi Program Pengembangan Guru dalam Meningkatkan Kopetensi di SMP Negeri 131 Jakarta Selatan. Jakarta: Universitas Islam Negeri Syarif Hidayatullah Jakarta.
- KBBI. (2025, Mei 21). Validitas. Retrieved from Kamus Besar Bahasa Indonesia: <https://kbbi.web.id/validitas>
- Magdalena, I. (2021). *Desain Evaluasi Pembelajaran SD*. CV Jejak (Jejak Publisher).
- Matondang, Z. (2009). Validitas dan Reabilitas Suatu Instrumen Penelitian. *Jurnal Tabularasa*, 89.
- Muhammad, A. N. (2021). Validitas Instrumen Assesmen (Analisis Validitas Isi dan Kontruks Instrumen Aseesmen Buku Pelajaran Al-Quran Hadis Kelas 6 Madrasah Ibtidaiyah Materi Surat Ad Dhuha BAB VI). *Jurnal CONTEMPLATE: Jumal Pendidikan Bahasa Arab*, 28.
- Ndiung, S. J. (2020). Pengembangan instrumen hasil belajar matematika peserta didik sekolah dasar berorientasi pada berpikir tingkat tinggi. *Premiere Educandum: Jurnal Pendidikan Dasar dan Pembelajaran*.
- Prijuwanto, S. W. (2020). *Evaluasi Pembelajaran*. Sanata Dharma University Press.
- Ramadhan, M. S. (2024). Validitas and Reablitas. *Journal on Education*.
- Syamsurizal, S. (2020). *Validitas dan Reabilitas ALat Ukur*. 4.
- Widiana, I. W. (2023). *Validasi Penyusunan Instrumen Penelitian Pendidikan*. PT. RajaGrafindo Persada-Rajawali Pers.