



STRATEGI PENENTUAN SOAL PENILAIAN HASIL BELAJAR OLEH GURU SEKOLAH DASAR

Lalu Hamdian Affandi¹, H. Husniati², N. Nurhasanah³, Khairun Nisa⁴

^{1,2,3,4} Pendidikan Guru Sekolah Dasar, Universitas Mataram,

¹⁾hamdian.fkip@unram.ac.id, ²⁾husniati_fkip@unram.ac.id, ³⁾nurhasanah_fkip@unram.ac.id,

⁴⁾khairun_nisa@unram.ac.id

Histori artikel

Received:
22 Juli 2024

Accepted:
8 September 2024

Published:
3 November 2024

Abstrak

Penilaian adalah bagian integral dari sistem pembelajaran dalam Kurikulum Merdeka. Salah satu keterampilan yang dibutuhkan guru adalah keterampilan memilih dan menyusun instrumen penilaian. Penelitian ini bertujuan untuk mengidentifikasi strategi guru dalam menentukan kualitas soal penilaian hasil belajar siswa Sekolah Dasar (SD). Data dalam penelitian ini dikumpulkan melalui angket terbuka dari 93 orang guru SD di Kota Mataram, Nusa Tenggara Barat. Data dianalisis dengan pengkodean, kategorisasi, dan penghitungan prosentase kategori. Penelitian menghasilkan gambaran bahwa 70.18% guru menggunakan pertimbangan tujuan dan materi pelajaran dalam memilih soal, dan 64.52% guru menguji kualitas soal dengan melakukan analisis kesesuaian soal dengan kisi-kisi, kemampuan yang diukur, serta analisis tingkat kesulitan soal. Penelitian ini menemukan bahwa strategi yang digunakan guru dalam menentukan kualitas soal terfokus pada upaya untuk memastikan akurasi soal dalam mengukur ketercapaian tujuan pembelajaran berdasarkan materi yang telah dipelajari oleh siswa. Urgensi penelitian ini ditunjukkan oleh temuan bahwa guru kurang memperhatikan isu reliabilitas dan keadilan (*fairness*) dan hanya mempertimbangkan kriteria validitas ketika menentukan soal. Penelitian ini mengimplikasikan pentingnya melakukan kajian untuk mengeksplorasi faktor penyebab guru mengabaikan kriteria reliabilitas. Selain itu, penelitian ini merekomendasikan pentingnya pengembangan pemberdayaan professional guru yang diarahkan untuk meningkatkan keterampilan guru menyusun atau memilih soal.

Kata-kata Kunci: penilaian hasil belajar siswa, penentuan kualitas soal, validitas dan reliabilitas, penilaian di sekolah dasar

*Coresponding author: Lalu Hamdian Affandii (hamdian.fkip@unram.ac.id)

Abstract: Assessment is an integral part of instructional system in Kurikulum Merdeka. Teacher as instructional designer needs to master those skills related to select and construct assessment instruments. Current research is aimed at identifying the strategies used by elementary school teachers in determining quality of test for assessing students learning. Data in this research was collected through open ended questionnaire from 93 elementary school teachers in Mataram, West Nusa Tenggara. Data then analyzed by coding, categorizing, and finally calculating percentage of each category. This research shows that 70.18% teachers considered learning goals and learning materials in selecting tests, and 64.52% teachers assess test quality by reviewing the compatibility of test with test outline, measured competencies or skills, and level of test difficulties. Present research found that the strategy being used by teachers focused mainly on the effort to assure test accuracy in measuring learning goals based on material being learned by students. Those findings contribute to understanding that teachers tend to overlook reliability and fairness issues and only consider validity issues when selecting the test. Present research implies the necessary for investigating factors affecting teachers to overlook test's reliability. In addition, this research also suggests for development of teacher professional training directed toward improvement of teachers' skills in constructing or selecting classroom tests.

Keywords: assessment of student learning, assessment of test quality, validity and reliability

Latar Belakang

Kurikulum Merdeka (KM) menempatkan penilaian sebagai bagian integral pembelajaran. Integrasi penilaian ke dalam sistem pembelajaran secara utuh membutuhkan kemampuan guru untuk melakukan asesmen. Secara umum, kemampuan tersebut tercakup dalam istilah literasi asesmen, yaitu kemampuan guru untuk memahami dan menerapkan konsep dan prosedur dasar penilaian dalam rangka pengambilan keputusan yang tepat (Popham, 2017; 25). Literasi asesmen tercermin dari kemampuan guru memilih instrumen penilaian, mengumpulkan informasi dengan menggunakan prosedur yang obyektif, reliabel, adil, dan tidak bias, serta memanfaatkan hasil asesmen untuk merencanakan, melakukan perbaikan, dan menilai kualitas pembelajaran.

Dalam faktanya, banyak guru yang terindikasi kesulitan melaksanakan asesmen sebagai proses yang terintegrasi dengan pembelajaran. Penelitian Lam (2019) menunjukkan bahwa pemahaman guru tentang asesmen masih bersifat parsial. Hal ini terlihat dari banyaknya guru yang memosisikan asesmen hanya dalam konteks penarikan kesimpulan tentang penguasaan siswa terhadap tujuan pembelajaran (Suciati & Amirullah, 2017). Hal lain yang digambarkan dari beberapa penelitian menunjukkan kesulitan guru dalam mengembangkan instrumen penilaian, melakukan penskoran, dan melakukan penilaian yang autentik (Sudianto & Kisno, 2021; Suwandani, Karma, & Affandi, 2020; Nabilah, Karma, & Husniati, 2021). Hasil penelitian tersebut menunjukkan bahwa literasi asesmen guru belum memadai (Levy-Vered & Alhija, 2015; Yamtim & Wongwanich, 2014).

Literasi asesmen yang baik merupakan salah satu faktor penentu kualitas pendidikan (Popham, 2009). Melalui asesmen, guru mendapatkan informasi penting yang

dibutuhkan untuk merencanakan pembelajaran yang sesuai dengan kebutuhan siswa (Tomlinson, 2001). Ketika pembelajaran dilaksanakan, asesmen dilakukan guru untuk mengidentifikasi kondisi yang menghambat pembelajaran yang efektif (Popham, 2008). Pada intinya, asesmen dibutuhkan guru untuk menciptakan pembelajaran yang efektif (Hattie, 2012; Popham, 2017). Pembelajaran yang efektif hanya bisa diwujudkan jika guru memiliki informasi tentang apa yang sudah diketahui dan apa yang belum diketahui siswa (Erskine, 2014). Informasi tersebut hanya bisa didapatkan melalui kegiatan asesmen, baik asesmen awal maupun asesmen formatif. Dengan demikian, literasi asesmen yang rendah dapat menjadi penghambat upaya penciptaan pembelajaran yang efektif.

Penelitian yang pernah dilakukan hanya sampai pada gambaran tentang kualitas pemahaman guru terhadap konsep asesmen (Lam, 2019; Suciati & Amirullah, 2017; Levy-Vered & Alhija, 2015; Yamtim & Wongwanich, 2014) dan kesulitan yang dialami guru ketika melakukan asesmen (Sudianto & Kisno, 2021; Suwandani, Karma, & Affandi, 2020; Nabilah, Karma, & Husniati, 2021). Penelitian lainnya hanya memberikan gambaran tentang strategi yang digunakan guru untuk melakukan asesmen (Berry, 2010). Aspek yang belum banyak mendapatkan perhatian guru adalah bagaimana guru menyusun atau memilih instrumen penilaian di kelas.

Kajian empiris tentang strategi penentuan soal dapat memberikan informasi berharga bagi perbaikan kualitas pendidikan di masa yang akan datang. Informasi mengenai strategi pemilihan instrumen tentunya dapat dijadikan pijakan bagi para pengambil kebijakan dalam menentukan dukungan yang dibutuhkan guru. Selain itu, deskripsi detail tentang bagaimana guru memilih asesmen bisa dijadikan pijakan bagi peneliti lain, terutama yang memiliki focus pada pengembangan professional guru.

Penelitian ini adalah penelitian eksploratif yang difokuskan pada strategi yang digunakan guru dalam memilih dan menentukan kualitas soal untuk penilaian hasil belajar siswa. Penelitian ini diarahkan untuk menghasilkan deskripsi tentang pertimbangan yang digunakan guru SD ketika menyusun atau memilih soal serta prosedur yang diterapkan guru SD dalam menentukan kualitas soal.

Metode

Penelitian ini adalah penelitian deskriptif-eksploratif, yaitu penelitian yang menghasilkan gambaran tentang kondisi sebuah variabel melalui penggalian informasi dari subyek penelitian. Penelitian deskriptif mendukung ketercapaian tujuan penelitian yang berkaitan dengan mengeksplorasi kecenderungan subyek dalam menjalankan aktifitas tertentu. Dimensi eksploratif dibutuhkan sebagai cara pengumpulan informasi yang memberi ruang bagi subyek untuk mengemukakan perspektif dan praktik yang telah mereka lakukan

selama menjalani aktifitas tertentu. Karena penelitian ini bertujuan untuk menggambarkan strategi dalam memilih dan menentukan kualitas soal yang digunakan untuk menilai hasil belajar siswa melalui eksplorasi perspektif dan praktik guru, maka pendekatan deskriptif-eksploratif adalah pendekatan yang tepat. Subjek dalam penelitian ini adalah guru SD yang bertugas di Kota Mataram, NTB, dan sekitarnya. Subjek dipilih secara *convenience* (berdasarkan kesediaan), yaitu guru yang dengan sukarela menyediakan waktu untuk mengisi angket. Kesediaan guru mengisi angket ditentukan berdasarkan konfirmasi yang mereka berikan ketika peneliti menginformasikan tujuan penelitian. Data dalam penelitian ini dikumpulkan melalui angket terbuka yang dikembangkan berdasarkan teori literasi asesmen yang dirumuskan oleh Popham (2017). Menurut Popham, salah satu aspek literasi asesmen yang harus dikuasai guru adalah perumusan dan evaluasi kualitas soal. Aspek ini berkenaan dengan kemampuan guru memilih dan mengevaluasi soal berdasarkan kriteria validitas, reliabilitas, dan keadilan (*fairness*). Karena sifatnya yang eksploratif, angket dalam penelitian ini menggunakan pertanyaan terbuka yang memberikan ruang bagi guru untuk mendeskripsikan pengalaman, pemikiran, dan pendapatnya terkait proses penentuan soal yang akan digunakan dalam melakukan asesmen di kelas. Data dalam penelitian ini kemudian dianalisis dengan pendekatan tematik. Langkah analisis data dengan pendekatan tematik adalah pengkodean, kategorisasi, dan kalkulasi frekwensi kategori. Pengkodean merupakan tahapan di mana peneliti menentukan kata kunci yang terkandung di dalam jawaban guru terhadap pertanyaan yang ada di angket. Kategorisasi dilakukan dengan mengelompokkan kata-kata kunci. Setiap kelompok kata kunci kemudian diberikan label yang merupakan kategori jawaban. Setelah kategori terbentuk, peneliti mengkalkulasi jumlah jawaban untuk menemukan frekwensi dari setiap kategori. Pada akhirnya, frekwensi setiap kategori akan menggambarkan strategi guru dalam menentukan soal untuk mengukur hasil belajar siswa.

Hasil Penelitian

Dari 100 guru yang menerima angket, hanya 93 orang yang mengisi dan mengembalikan kepada peneliti. Tabel di bawah ini menggambarkan karakteristik subjek penelitian penelitian.

Tabel 1. Karakteristik Subjek Penelitian

Karakteristik	Jumlah	%
Jenis Kelamin		
Perempuan	75	80.65
Laki-laki	18	19.35
Kelas yang diajar		
Tinggi	47	50.54
Rendah	46	49.49

Status sertifikasi			
Belum	39	41.94	
Sudah	54	58.06	
Total	93		

Pertanyaan penelitian yang pertama berkaitan dengan pertimbangan guru dalam memilih soal. Pertanyaan ini meminta guru untuk mendeskripsikan hal apa saja yang mereka prioritaskan ketika menyusun atau memilih soal yang akan digunakan dalam mengukur hasil belajar siswa. Grafik di bawah ini menggambarkan pertimbangan yang digunakan guru.



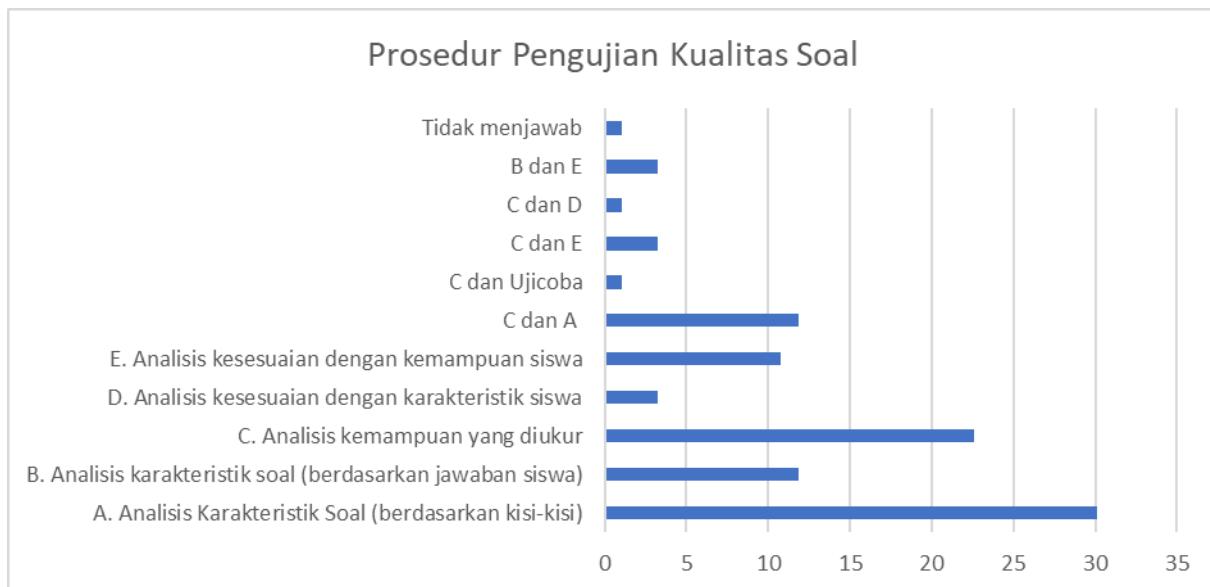
Grafik 1. Pertimbangan Guru Dalam Memilih Soal Hasil Belajar

Pertimbangan guru dalam memilih soal sangat variatif. Secara garis besar, pertimbangan guru dapat dikelompokkan menjadi pertimbangan tujuan, materi, buku teks, karakteristik siswa, dan kemampuan siswa. Pertimbangan tujuan mengindikasikan guru memilih soal karena kesesuaiannya dengan tujuan pembelajaran. Pertimbangan materi menunjukkan pilihan guru pada soal yang sesuai dengan materi yang sudah dipelajari siswa. Pertimbangan buku teks menggambarkan bahwa soal yang digunakan guru mencerminkan isi buku teks. Pertimbangan karakteristik siswa menunjukkan pilihan soal guru ditentukan dari sejauh mana soal memiliki kesesuaian dengan perkembangan kognitif, kematangan emosional, dan kemampuan siswa memahami soal. Sedangkan pertimbangan kemampuan siswa mengindikasikan pilihan soal guru yang berpatokan pada estimasi kemampuan siswa dalam memahami materi yang sudah diajarkan guru.

Data dalam penelitian ini menunjukkan bahwa sebagian guru (30.11%) memilih soal berdasarkan kesesuaian dengan tujuan, diikuti oleh pertimbangan materi (22.58%) dan kesesuaian soal dengan buku teks (10.75%). Selain itu, terdapat guru yang pertimbangannya dalam memilih soal didasarkan pada kombinasi dari berbagai aspek.

Dalam hal ini 17.2% guru menggunakan pertimbangan kesesuaian soal dengan tujuan dan materi pelajaran dan 7.53% menggunakan pertimbangan kesesuaian soal dengan materi dan kemampuan siswa.

Penelitian ini juga menghasilkan data tentang prosedur pengujian kualitas soal yang dilakukan guru. Prosedur pengujian kualitas soal berkaitan dengan tindakan yang dilakukan guru untuk memastikan bahwa soal yang digunakan di dalam asesmen adalah soal yang layak. Grafik 2 di bawah ini menggambarkan sejumlah pertimbangan dan tindakan guru dalam menentukan kualitas soal.



Grafik 2. Prosedur Pengujian Kualitas Soal

Secara garis besar, guru menilai kualitas soal dengan melakukan analisis terhadap soal sebelum dan sesudah digunakan. Pertimbangan yang digunakan guru adalah karakteristik soal, kemampuan yang diukur, kesesuaian soal dengan karakteristik siswa, dan kesesuaian soal dengan kemampuan siswa. Karakteristik soal berkaitan dengan kesesuaian item soal dengan kisi-kisi dan tingkat kesulitan soal. Penentuan kesesuaian soal dengan kisi-kisi dilakukan sebelum soal digunakan, sedangkan penentuan tingkat kesulitan soal dilakukan setelah soal diberikan kepada siswa. Pertimbangan kesesuaian soal dengan kemampuan yang diukur mencerminkan upaya guru untuk memastikan bahwa soal yang nantinya akan digunakan tidak melenceng dari target yang hendak diukur, yakni kemampuan yang tercantum di dalam tujuan pembelajaran. Pertimbangan kesesuaian soal dengan karakteristik siswa mencerminkan asumsi guru bahwa soal yang digunakan adalah soal yang sesuai dengan tahapan kognitif siswa, kemampuan siswa untuk mencerna maksud soal, atau relevansi soal dengan kehidupan sehari-hari siswa. Pertimbangan kesesuaian soal dengan kemampuan siswa berkaitan dengan upaya guru untuk memastikan bahwa soal yang digunakan adalah soal-soal yang potensial bisa dijawab

dengan benar oleh siswa. Pertimbangan-pertimbangan tersebut kemudian dijadikan dasar untuk melakukan penelaahan lebih dalam terhadap soal.

Penelaahan soal sebagai cara penentuan kualitas soal merupakan cara paling umum yang dilakukan oleh guru. 30.11% guru menyandarkan kualitas soal berdasarkan kisi-kisi yang dibuatnya, sedangkan 11.83% guru menentukan kualitas soal berdasarkan jawaban siswa. Selain itu, 22.58% guru menilai kualitas soal berdasarkan kemampuan yang diukur. Pada sisi lain, terdapat guru yang terkesan menilai kualitas soal dengan mempertimbangkan kondisi siswa dengan memastikan kesesuaian soal dengan karakteristik siswa (3.22%) dan kesesuaian soal dengan kemampuan siswa (10.75%). Selain menggunakan pertimbangan tunggal, terdapat guru yang menggunakan pertimbangan ganda dalam menilai kualitas soal. 11.83% guru menilai kualitas soal berdasarkan kesesuaian antara soal dengan kemampuan yang diukur dan kesesuaian soal dengan kisi-kisi soal.

Analisis yang lebih dalam menunjukkan bahwa terdapat indikasi konsistensi pilihan guru dalam menentukan soal yang digunakan dalam asesmen hasil belajar siswa. Hal ini bisa dilihat dari proporsi kesesuaian jawaban pada pertanyaan yang pertama (pertimbangan guru dalam memilih soal) dan pertanyaan kedua (prosedur pengujian kualitas soal). Pertanyaan pertama menghasilkan deskripsi awal yang mencerminkan kecenderungan orientasi penilaian yang disasar oleh soal. Setelah soal didapatkan, guru kemudian menganalisis kelayakan soal berdasarkan kriteria yang mereka anggap penting untuk menghasilkan informasi yang baik dari proses penilaian. Konsistensi pilihan guru bisa dilihat dari kesesuaian antara pilihan soal dengan prosedur pengujinya. Penelitian ini mengungkapkan bahwa 70.18% guru menggunakan pertimbangan tujuan dan materi pelajaran dalam memilih soal. Ketika melakukan pengujian kualitas soal, 64.52% guru melakukan analisis kesesuaian soal dengan kisi-kisi, kemampuan yang diukur oleh soal, dan tingkat kesulitan soal. Biasanya, kisi-kisi soal yang disusun guru didasarkan pada cakupan materi, tujuan pembelajaran, serta tingkat kognitif yang diukur. Dengan demikian, penelitian ini menghasilkan pemahaman bahwa strategi penentuan soal yang digunakan sebagian besar guru terfokus pada pemilihan soal yang mencerminkan keterwakilan tujuan dan materi pembelajaran.

Pembahasan

Secara garis besar, penelitian ini mengungkap fakta bahwa sebagian besar guru (70.18%) menggunakan pertimbangan tujuan dan materi pelajaran dalam memilih soal dalam penilaian hasil belajar siswa. Dalam kaitan dengan pengujian kualitas soal, sebagian besar guru (64.52%) melakukan analisis kesesuaian soal dengan kisi-kisi, kemampuan yang diukur, serta analisis tingkat kesulitan soal. Dari data tersebut, penelitian ini

menghasilkan pemahaman bahwa strategi yang digunakan guru dalam menentukan soal terfokus pada upaya untuk memastikan akurasi soal dalam mengukur ketercapaian tujuan pembelajaran berdasarkan materi yang telah dipelajari oleh siswa.

Hasil penelitian ini mengkonfirmasi hasil penelitian sebelumnya yang menunjukkan bahwa aspek validitas soal menjadi fokus utama guru dalam merencanakan penilaian. Misalnya, penelitian terhadap perencanaan penilaian yang dilakukan guru menunjukkan bahwa soal yang digunakan guru disesuaikan dengan standar capaian pembelajaran yang ditetapkan di dalam kurikulum (Zulminiati & Hartati, 2021). Khusus untuk penilaian terhadap aspek sikap, guru dilaporkan menyusun perencanaan penilaian berdasarkan silabus yang telah disusun sebelumnya. Perencanaan penilaian ini memuat teknik dan instrumen yang akan digunakan. Untuk menentukan aspek yang akan dinilai, guru merujuk kepada indikator pembelajaran (Ramadhani & Ramadan, 2022). Namun demikian, hanya sebagian kecil guru yang memilih butir instrumen berdasarkan data empirik (Setiadi, 2016).

Pada prinsipnya, soal adalah salah satu alat yang bisa digunakan untuk mengumpulkan informasi tentang kemampuan akademik siswa. Proses pengumpulan informasi, di samping interpretasi informasi, dan pemanfaatan informasi untuk mengidentifikasi dan melayani kebutuhan siswa dalam pembelajaran merupakan tahapan penting pelaksanaan asesmen di kelas (Ferrara, Maxey-Moore, & Brookhart, 2020). Informasi yang didapatkan melalui penilaian di kelas adalah pijakan penting guru dalam upayanya menghadirkan pembelajaran yang efektif melalui pemberian umpan balik, pengenalan kekuatan dan kelemahan siswa, serta evaluasi kemampuan yang dimiliki siswa setelah menjalani serangkaian aktifitas belajar (Kane & Wools, 2020).

Penilaian oleh guru di kelas menghendaki terpenuhinya kriteria keterpercayaan (*trustworthiness*), yaitu tingkatan yang menunjukkan sejauh mana informasi yang didapatkan bisa dijadikan acuan dalam pengambilan keputusan untuk menciptakan pembelajaran yang efektif (Alonzo, 2020). Untuk itu, soal yang digunakan haruslah memenuhi setidaknya 2 kriteria, yaitu validitas dan reliabilitas. Pada awalnya, validitas berkaitan dengan akurasi pengukuran (*measurement accuracy*) dan kepantasan pengukuran (*measurement appropriateness*). Akurasi pengukuran merupakan kesesuaian antara alat ukur dengan target yang hendak diukur, sedangkan kepantasan pengukuran berkaitan dengan kesesuaian antara kesimpulan yang diambil berdasarkan hasil pengukuran dengan tujuan dilakukannya pengukuran (Hughes, 2018). Validitas sebagai akurasi berhubungan dengan pengukuran dan validitas sebagai kepantasan berkaitan dengan keberfungsian. Oleh sebab itu, pembuktian validitas bisa dilakukan dengan menggunakan perspektif pengukuran maupun dengan perspektif fungsional (Kane & Wools, 2020). Pembuktian validitas melalui analisis kesesuaian item dengan target ukur biasanya

dilakukan dengan menelaah keterwakilan tujuan pembelajaran dan materi yang diajarkan dalam item-item soal. Penelaahan tersebut bisa dilakukan sendiri oleh guru dengan mengundang guru lain untuk memeriksa secara teliti setiap item soal berdasarkan indikator dan tujuan pembelajaran yang dirumuskan berdasarkan ketentuan kurikulum (Popham, 2017). Secara praktis, pakar lebih merekomendasikan penggunaan perspektif fungsional ketika melakukan pembuktian validitas (Kane & Wools, 2020). Hal ini disebabkan karena pemanfaatan hasil asesmen sebagai dasar perencanaan, perbaikan, dan evaluasi pembelajaran.

Reliabilitas, bersama dengan validitas, adalah kriteria yang harus dipenuhi oleh soal yang baik. Reliabilitas berkaitan dengan isu konsistensi dan presisi. Dalam makna konsistensi, reliabilitas merujuk pada keseragaman informasi yang didapatkan dari soal yang diberikan kepada sekelompok siswa dari satu kesempatan ke kesempatan yang lain. Dengan kata lain, reliabilitas dalam makna konsistensi berarti jawaban siswa terhadap sebuah soal akan relatif sama ketika soal itu diberikan lebih dari satu kali. Sedangkan reliabilitas dalam makna presisi merujuk pada kesesuaian antara informasi yang dihasilkan dari soal dengan kemampuan yang sesungguhnya dimiliki oleh siswa (Popham, 2017). Secara teknis, reliabilitas dalam makna presisi bisa dilihat dari proporsi antara skor yang sesungguhnya (*true score*) dengan *error*. Semakin tinggi proporsi skor yang sesungguh, semakin tinggi presisi pengukuran, semakin reliabel informasi yang didapatkan (Revelle & Condon, 2018). Pembuktian reliabilitas bisa dilakukan dengan 3 cara, yaitu *test-retest*, *alternate forms*, dan internal konsistensi. Ketiga cara tersebut menghendaki agar soal diujicoba dulu sebelum digunakan.

Kedua aspek tersebut (validitas dan reliabilitas) menjadi semakin penting dibuktikan sebelum soal digunakan di era kurikulum merdeka saat ini. Aspek validitas memberikan kepastian tentang kesesuaian soal dengan kemampuan yang hendak dinilai, sedangkan aspek reliabilitas memberikan kepastian tentang kesesuaian informasi yang didapat melalui soal dengan kemampuan siswa yang sebenarnya. Penekanan pada aspek validitas saja berresiko menghasilkan penilaian yang tidak mencerminkan keadaan yang sesungguhnya.

Penggunaan soal yang teruji valid dengan mengesampingkan isu reliabilitas berimplikasi luas pada pengambilan keputusan pendidikan. Kajian terhadap praktik penilaian yang dilakukan guru menunjukkan bahwa penilaian guru terhadap hasil belajar siswa mengandung berbagai pertimbangan yang bercampur dengan kemampuan dan pengetahuan siswa (Brookhart, Guskey, Bowers, McMillan, Smith, Smith, Stevens, & Welsh, 2016). Pertimbangan yang variatif tersebut menghasilkan penilaian yang tidak sepenuhnya akurat sehingga pengambilan keputusan berdasarkan hasil penilaian tersebut berpotensi mengandung bias. Oleh sebab itu, penilaian yang reliabel perlu dire-interpretasi dari presisi

pengukuran menjadi kecukupan informasi yang digunakan guru dalam mengambil keputusan (Popham, 2017).

Kurikulum merdeka yang menempatkan penilaian sebagai bagian integral sistem pembelajaran akan terganggu penerapannya tanpa kehadiran informasi yang akurat. Pemanfaatan hasil penilaian untuk mendiagnosis kemampuan awal siswa tentunya akan berpeluang mendorong guru melakukan kekeliruan ketika merencanakan pembelajaran. Bukan hanya itu, asesmen nonkognitif juga memainkan peran penting dalam menentukan kesiapan belajar siswa, terutama dalam konteks pembelajaran yang disesuaikan dengan kebutuhan siswa (Watu, Lawe, Sayangan, & Laksana, 2024). Dalam konteks asesmen formatif, perbaikan pembelajaran yang dilakukan guru tidak akan tepat sasaran jika hasil penilaian kemampuan siswa tidak mencerminkan keadaan yang sesungguhnya (Brookhart, 2020). Resiko yang lebih besar adalah ketika informasi yang tidak valid dijadikan dasar untuk menyimpulkan hasil akhir pembelajaran.

Dimensi praktis dari penelitian ini mengimplikasikan pentingnya pengambil kebijakan dalam dunia pendidikan untuk bukan hanya mempertimbangkan isu validitas dalam penilaian. Isu reliabilitas juga mesti mendapatkan porsi perhatian yang memadai dalam kaitan dengan pengambilan keputusan berdasarkan hasil penilaian guru. Isu reliabilitas bisa dipertimbangkan dengan memperhatikan kecukupan informasi yang disajikan guru dalam penilaian. Kecukupan informasi tersebut dicerminkan oleh bukan hanya oleh skor kuantitatif, melainkan juga oleh deskripsi yang lebih rinci tentang informasi yang terkandung di dalam skor seperti kemampuan dan pengetahuan yang diukur oleh penilaian, kriteria keberhasilan yang digunakan, dan lain-lain. Kecukupan informasi tersebut menjadi penting agar para pengambil kebijakan memiliki kesadaran pada makna dari skor yang terdapat di dalam penilaian hasil belajar yang dibuat oleh guru. Hal ini menunjukkan pentingnya peningkatan kemampuan guru memilih dan menyusun instrumen yang tidak hanya valid, namun juga reliabel. Area penting yang harus dipahami guru dalam pemilihan dan penentuan soal adalah konsep tentang reliabilitas sebagai kecukupan informasi yang didapatkan melalui kegiatan penilaian. Makna reliabilitas sebagai *information sufficiency* perlu dilatihkan sehingga sistem penilaian yang diterapkan guru bisa memenuhi ekspektasi pengambil kebijakan lainnya.

Kesimpulan

Penelitian ini menyimpulkan bahwa strategi yang digunakan guru dalam menentukan kualitas soal terfokus pada kesesuaian antara soal dengan tujuan pembelajaran dan materi yang diajarkan kepada siswa. Strategi ini mencerminkan pertimbangan guru yang terfokus pada isu validitas soal. Sedangkan isu lain seperti reliabilitas tidak cukup mendapatkan

perhatian yang cukup dari guru. Dalam praktiknya, hal ini berpotensi memunculkan masalah yang berkaitan dengan keputusan yang diambil berdasarkan informasi yang tidak reliabel. Penelitian ini mengindikasikan pentingnya melakukan kajian lanjutan terhadap faktor penyebab terbaikannya aspek reliabilitas dari soal yang digunakan guru. Selain itu, penelitian ini merekomendasikan agar program pemberdayaan guru bisa difokuskan pada upaya untuk meningkatkan kemampuan guru dalam memilih dan menyusun soal yang tidak hanya terbukti valid, namun juga reliabel. Bagi pengambil kebijakan, penelitian ini merekomendasikan agar para pengambil kebijakan tidak hanya mengambil keputusan berdasarkan nilai yang diperoleh dari soal, melainkan melengkapi pertimbangannya dengan memperhatikan aspek kualitatif dari proses pembelajaran yang dilakukan guru.

Daftar Pustaka

- Alonso, A.C. (2020). Defining Trustworthiness for Teachers' Multiple Uses of Classroom Assessment Results. In Susan M. Brookhart & James H. McMillan (eds.). *Classroom Assessment and Educational Measurement*. p.120-145. New York: Routledge.
- Berry, R. (2010). Teachers' Orientations towards Selecting Assessment Strategies. *New Horizons in Education*, 58(1), 96-107.
- Brookhart, S. M., Guskey, T. R., Bowers, A. J., McMillan, J. H., Smith, J. K., Smith, L. F., Stevens, M.T., & Welsh, M. E. (2016). A century of grading research: Meaning and value in the most common educational measure. *Review of educational research*, 86(4), 803-848.
- Brookhart, S.M. (2020). Feedback and Measurement. In Susan M. Brookhart & James H. McMillan (eds.). *Classroom Assessment and Educational Measurement*. p.63-78. New York: Routledge.
- Erskine, J. L. (2014). It changes how teachers teach: How testing is corrupting our classrooms and student learning. *Multicultural Education*, 21(2), 38-40.
- Ferrara, S., Maxey-Moore, K., & Brookhart, S.M. (2020). Guidance in the Standards for Classroom Assessment: Useful or Irrelevant? In Susan M. Brookhart & James H. McMillan (eds.). *Classroom Assessment and Educational Measurement*. p.97-119. New York: Routledge.
- Hartati, S., & Zulminiati, Z. (2020). Fakta-Fakta Penerapan Penilaian Otentik di Taman Kanak-Kanak Negeri 2 Padang. *Jurnal Obsesi: Jurnal Pendidikan Anak Usia Dini*, 5(2), 1035-1044.
- Hattie, J. (2012). *Visible Learning for Teachers: Maximizing Impact on Learning*. London and New York: Routledge.
- Hughes, D.J. (2018). Psychometric Validity: Establishing the Accuracy and Appropriateness of Psychometric Measures. In Paul Irving, Tom Booth, & David J. Hughes (eds.). *The Wiley Handbook of Psychometric Testing: A Multidisciplinary Reference on Survey, Scale, and Test*. p. 751-779. West Sussex, UK: John Wiley & Sons Ltd.
- Kane, M.T., & Wools, S. (2020). Perspectives on the Validity of Classroom Assessment. In Susan M. Brookhart & James H. McMillan (eds.). *Classroom Assessment and Educational Measurement*. p.11-26. New York: Routledge.
- Lam, R. (2019). Teacher assessment literacy: Surveying knowledge, conceptions and practices of classroom-based writing assessment in Hong Kong. *System*, 81, 78-89.

- Levy-Vered, A., & Alhija, F.N.A. (2015). Modelling Beginning Teachers' Assessment Literacy: the Contribution of Training, Self-Efficacy, and Conceptions of Assessment. *Educational Research and Evaluation*, 21(5-6), 378-406, DOI:10.1080/13803611.2015.1117980.
- Nabilah, N., Karma, I. N., & Husniati, H. (2021). Identifikasi Kesulitan Guru dalam Melaksanakan Penilaian Autentik pada Kurikulum 2013 di SDN 50 Cakranegara. *Jurnal Ilmiah Profesi Pendidikan*, 6(4), 617-622.
- Popham, W.J. (2008). *Transformative Assessment*. Alexandria, VA: ASCD.
- Popham, W.J. (2009). Assessment Literacy for Teachers: Faddish or Fundamental? *Theory into Practice*, 48, 4-11.
- Popham, W.J. (2017). *Classroom Assessment: What Teachers Need to Know*. Boston, USA: Pearson Education Inc.
- Ramadhani, R. H. D., & Ramadan, Z. H. (2022). Implementasi Penilaian Ranah Sikap dalam Kurikulum 2013 pada Mata Pelajaran Pendidikan Agama Islam di Sekolah Dasar. *Mimbar PGSD Undiksha*, 10(1), 17-25.
- Revelle, W., & Condon, D.M. (2018). Reliability. In Paul Irwing, Tom Booth, & David J. Hughes (eds.). *The Wiley Handbook of Psychometric Testing: A Multidisciplinary Reference on Survey, Scale, and Test*. p. 709-749. West Sussex, UK: John Wiley & Sons Ltd.
- Setiadi, H. (2016). Pelaksanaan penilaian pada Kurikulum 2013. *Jurnal Penelitian dan Evaluasi Pendidikan*, 20(2), 166-178.
- Suciati, R., & Amirullah, G. (2017). Literasi Asesmen IPA. *JKKP (Jurnal Kesejahteraan Keluarga dan Pendidikan)*, 4(02), 110-118.
- Sudianto, S., & Kisno, K. (2021). Potret Kesiapan Guru Sekolah Dasar dan Manajemen Sekolah dalam Menghadapi Asesmen Nasional. *Jurnal Akuntabilitas Manajemen Pendidikan*, 9(1), 85-97. <https://doi.org/10.21831/jamp.v9i1.39260>.
- Suwandani, R. A., Karma, I. N., & Affandi, L. H. (2020). Analisis Faktor-Faktor Kesulitan Guru Dalam Melaksanakan Penilaian Autentik Di SDN Gugus 1 Kecamatan Janapria. *Jurnal Ilmiah Pendidikan Indonesia*, 2(1), 24-30.
- Tomlinson, C.A. (2001). *How to Differentiate Instruction in Mixed-Ability Classroom. 2nd Edition*. Alexandria, VA: ASCD.
- Watu, M. F., Lawe, Y. U., Sayangan, Y. V., & Laksana, D. N. L. (2024). Penerapan Asesmen Diagnostik Non Kognitif pada Aspek Kesiapan dan gaya belajar siswa SD. *Jurnal Ilmiah Pendidikan Citra Bakti*, 11(2), 615–625. <https://doi.org/10.38048/jipcb.v11i2.3660>
- Yamtim, V., & Wongwanich, S. (2014). A study of classroom assessment literacy of primary school teachers. *Procedia-Social and Behavioral Sciences*, 116, 2998-3004.