

## KUALITAS INSTRUMEN EVALUASI PENDIDIKAN DI MADRASAH IBTIDAIYAH: KAJIAN SKALA PENGUKURAN, VALIDITAS, DAN RELIABILITAS

Ilham Nugraha<sup>1\*</sup>, Mulyawan Safwandy Nugraha<sup>2</sup>

Universitas Islam Sunan Gunung Djati Bandung, Indonesia

[Ilham.elfaizh@gmail.com](mailto:Ilham.elfaizh@gmail.com)

---

### ABSTRAK

**Abstrak:** Penelitian ini bertujuan untuk menganalisis skala pengukuran, validitas, dan reliabilitas instrumen evaluasi pendidikan di MI Al-Musyarrofah Cianjur. Penelitian menggunakan desain studi kasus kuantitatif deskriptif yang melibatkan tujuh responden yang terdiri atas kepala sekolah dan guru. Data dikumpulkan melalui penyebaran angket dan dianalisis menggunakan teknik persentase berdasarkan kriteria interpretasi skor. Variabel yang diteliti meliputi skala pengukuran, validitas instrumen, reliabilitas instrumen, serta kendala dalam pengembangan instrumen evaluasi pendidikan. Hasil penelitian menunjukkan bahwa skala pengukuran memperoleh persentase sebesar 89% (kategori sangat baik), validitas instrumen 87% (kategori sangat baik), dan reliabilitas instrumen 86% (kategori sangat baik). Sementara itu, kendala dalam pengembangan instrumen evaluasi memperoleh persentase sebesar 73% (kategori baik). Temuan ini menunjukkan bahwa guru di MI Al-Musyarrofah Cianjur telah menerapkan instrumen evaluasi pendidikan secara sistematis dengan memperhatikan aspek validitas dan reliabilitas. Namun demikian, masih terdapat beberapa kendala, seperti keterbatasan waktu, kurangnya pelatihan dalam evaluasi pendidikan, serta terbatasnya dukungan teknologi. Oleh karena itu, diperlukan pengembangan profesional berkelanjutan serta penguatan sistem evaluasi berbasis teknologi untuk meningkatkan kualitas praktik evaluasi pendidikan di lingkungan madrasah.

**Kata Kunci:** Instrumen Evaluasi Pendidikan, Validitas, Reliabilitas, Skala Pengukuran, Madrasah Ibtidaiyah.

**Abstract:** This study aims to analyze the measurement scale, validity, and reliability of educational evaluation instruments at MI Al-Musyarrofah Cianjur. The research employed a descriptive quantitative case study design involving seven respondents consisting of the principal and teachers. Data were collected through questionnaires and analyzed using percentage techniques based on score interpretation criteria. The variables examined included measurement scales, instrument validity, instrument reliability, and challenges in developing educational evaluation instruments. The findings indicate that the measurement scale achieved a score of 89% (very good category), instrument validity 87% (very good category), and instrument reliability 86% (very good category). Meanwhile, challenges in developing evaluation instruments obtained a score of 73% (good category). These results demonstrate that teachers at MI Al-Musyarrofah Cianjur have implemented educational evaluation instruments systematically by considering validity and reliability aspects. However, several challenges remain, including limited time, insufficient training in educational evaluation, and inadequate technological support. Therefore, continuous professional development and the strengthening of technology-based evaluation systems are necessary to improve the quality of educational evaluation practices in madrasah settings.

**Keywords:** Educational Assessment Instrument, Validity, Reliability, Measurement Scale, Islamic Elementary School.

---

#### Article History:

Received: 01-04-2026

Revised : 01-05-2026

Accepted: 21-06-2026

Online : 30-06-2026

---

### A. LATAR BELAKANG

Evaluasi pendidikan merupakan komponen penting dalam proses pembelajaran karena berfungsi untuk mengukur ketercapaian tujuan pembelajaran, menilai efektivitas

proses pendidikan, serta menjadi dasar pengambilan keputusan dalam perbaikan pembelajaran. Kualitas hasil evaluasi sangat ditentukan oleh kualitas instrumen yang digunakan. Instrumen evaluasi yang baik harus memenuhi prinsip validitas, reliabilitas, dan menggunakan skala pengukuran yang sesuai agar mampu menghasilkan data yang akurat, objektif, dan dapat dipertanggungjawabkan (Maskhuliah et al, 2025). Dalam konteks pendidikan dasar, instrumen evaluasi tidak hanya berfungsi sebagai alat penilaian hasil belajar, tetapi juga sebagai sumber informasi untuk meningkatkan kualitas pembelajaran secara berkelanjutan.

Validitas merupakan karakteristik yang menunjukkan sejauh mana instrumen mampu mengukur apa yang seharusnya diukur, sedangkan reliabilitas berkaitan dengan konsistensi hasil pengukuran ketika instrumen digunakan dalam kondisi yang relatif sama. Selain itu, penggunaan skala pengukuran yang tepat menjadi aspek penting dalam proses evaluasi karena menentukan kualitas data yang diperoleh dan metode analisis yang dapat digunakan. Kesalahan dalam menentukan skala pengukuran, menyusun indikator, maupun mengembangkan butir instrumen berpotensi menghasilkan measurement error yang dapat memengaruhi akurasi hasil evaluasi (Abdul et al, 2025).

Meskipun konsep validitas, reliabilitas, dan skala pengukuran telah banyak dibahas dalam literatur evaluasi pendidikan, praktik pengembangan instrumen di tingkat sekolah dan madrasah masih menghadapi berbagai kendala. Berdasarkan observasi awal dan studi dokumentasi di MI Al-Musyarrofah Cianjur, ditemukan bahwa guru telah menggunakan berbagai instrumen evaluasi seperti tes tertulis, tes lisan, penilaian praktik, dan lembar observasi. Namun, sebagian besar instrumen disusun berdasarkan pengalaman mengajar tanpa melalui prosedur pengembangan instrumen yang sistematis. Beberapa perangkat evaluasi belum dilengkapi kisi-kisi yang terstruktur, proses validasi masih dilakukan secara sederhana melalui diskusi antar guru, dan belum ditemukan dokumentasi pelaksanaan uji reliabilitas instrumen secara formal. Selain itu, keterbatasan waktu dan minimnya pelatihan mengenai evaluasi pendidikan menjadi kendala yang dihadapi guru dalam mengembangkan instrumen evaluasi yang berkualitas.

Penelitian terdahulu menunjukkan bahwa kajian mengenai instrumen evaluasi pendidikan lebih banyak berfokus pada pengembangan instrumen dan pengujian statistik. Misalnya, penelitian oleh (Saputra & Mayurida, 2025), menekankan pentingnya validitas dan reliabilitas dalam pengembangan instrumen penelitian. *Creswell* menjelaskan prosedur penyusunan instrumen dan pentingnya pemilihan skala pengukuran yang sesuai dengan karakteristik data. Penelitian (Eryanto et al, 2026), menemukan bahwa masih banyak guru yang menyusun instrumen evaluasi tanpa melakukan pengujian kualitas instrumen secara memadai. Sementara itu (Ristamala et al, 2025), menunjukkan bahwa keterbatasan kompetensi evaluasi menjadi salah satu hambatan utama dalam pengembangan instrumen pembelajaran di sekolah. Berbagai penelitian tersebut memberikan kontribusi penting terhadap pengembangan teori dan teknik evaluasi pendidikan, namun sebagian besar berfokus pada aspek statistik dan pengembangan instrumen secara formal.

Namun demikian, masih sedikit penelitian yang mengkaji bagaimana guru di lingkungan madrasah memahami dan menerapkan konsep validitas, reliabilitas, serta penggunaan skala pengukuran dalam praktik evaluasi sehari-hari. Dengan kata lain, penelitian sebelumnya lebih banyak membahas pengembangan instrumen dari perspektif metodologis dan statistik, sedangkan kajian mengenai kesenjangan antara konsep ideal

instrumen evaluasi dengan praktik nyata di madrasah masih relatif terbatas. Kesenjangan inilah yang menjadi research gap dalam penelitian ini (Masrukhin, 2019).

Penelitian ini memiliki kebaruan (*novelty*) pada fokus kajiannya yang menelaah secara mendalam praktik penggunaan skala pengukuran, validitas, reliabilitas, dan kendala pengembangan instrumen evaluasi dalam konteks Madrasah Ibtidaiyah. Berbeda dengan penelitian sebelumnya yang dominan menggunakan pendekatan statistik dalam menguji instrumen, penelitian ini berupaya mengungkap hubungan antara pemahaman guru, praktik penyusunan instrumen, proses validasi, dan penerapan reliabilitas dalam situasi nyata pembelajaran di madrasah. Dengan demikian, penelitian ini tidak hanya mengkaji kualitas instrumen secara konseptual, tetapi juga mengeksplorasi implementasinya dalam praktik pendidikan sehari-hari.

Fenomena tersebut juga menjadi perhatian dalam praktik evaluasi pendidikan di MI Al-Musyarrofah Cianjur. Berdasarkan hasil observasi awal dan studi dokumentasi yang dilakukan peneliti pada bulan Mei 2026, ditemukan bahwa guru telah menggunakan berbagai instrumen evaluasi pembelajaran, seperti tes tertulis, tes lisan, penilaian praktik, lembar observasi, dan angket sederhana. Namun, sebagian besar instrumen yang digunakan masih disusun berdasarkan pengalaman mengajar dan kebutuhan administratif tanpa melalui prosedur pengembangan instrumen yang sistematis. Dari dokumen perangkat pembelajaran yang dianalisis, ditemukan bahwa beberapa guru belum menyusun kisi-kisi soal secara lengkap sebelum menyusun instrumen evaluasi, sehingga keterkaitan antara indikator pembelajaran, tujuan pembelajaran, dan butir soal belum terdokumentasikan secara optimal (Syafi'I et al, 2025).

Hasil wawancara pendahuluan dengan kepala madrasah dan beberapa guru menunjukkan bahwa validasi instrumen umumnya dilakukan melalui diskusi informal antar guru atau telaah sederhana sebelum digunakan dalam pembelajaran. Akan tetapi, proses uji validitas yang menggunakan kriteria atau prosedur tertentu belum dilaksanakan secara konsisten. Selain itu, belum ditemukan adanya dokumentasi mengenai pelaksanaan uji reliabilitas instrumen, baik menggunakan teknik statistik maupun metode lain yang dapat menunjukkan tingkat konsistensi alat ukur. Kondisi ini menyebabkan guru belum memiliki informasi yang memadai mengenai tingkat keandalan instrumen yang digunakan dalam proses evaluasi pembelajaran.

Temuan awal lainnya menunjukkan bahwa keterbatasan waktu, tingginya beban administrasi guru, serta minimnya pelatihan tentang pengembangan instrumen evaluasi menjadi faktor yang memengaruhi kualitas penyusunan alat evaluasi. Guru cenderung lebih memfokuskan perhatian pada penyelesaian perangkat pembelajaran dibandingkan melakukan analisis kualitas instrumen secara mendalam. Selain itu, pemanfaatan teknologi untuk membantu proses analisis validitas, reliabilitas, dan pengolahan hasil evaluasi masih sangat terbatas. Padahal, evaluasi yang berkualitas memerlukan instrumen yang valid, reliabel, dan sesuai dengan karakteristik kompetensi yang diukur agar hasil penilaian dapat digunakan sebagai dasar pengambilan keputusan pembelajaran yang akurat.

Kondisi tersebut menunjukkan adanya kesenjangan antara konsep ideal pengembangan instrumen evaluasi pendidikan dengan praktik yang berlangsung di lapangan. Di satu sisi, guru dituntut menghasilkan instrumen evaluasi yang memenuhi prinsip validitas, reliabilitas, objektivitas, dan kepraktisan. Namun di sisi lain, berbagai keterbatasan yang ada menyebabkan proses pengembangan instrumen belum

dilaksanakan secara optimal (Arikunto, 2022). Oleh karena itu, diperlukan penelitian yang mampu mengungkap secara mendalam bagaimana praktik penggunaan skala pengukuran, validasi instrumen, reliabilitas instrumen, analisis kesalahan pengukuran (*measurement error*), serta upaya perbaikan instrumen evaluasi pendidikan di MI Al-Musyarrofah Cianjur.

Berdasarkan latar belakang tersebut, Secara teoretis, penelitian ini diharapkan dapat memperkaya kajian evaluasi pendidikan, khususnya dalam konteks madrasah ibtidaiyah. Secara praktis, hasil penelitian dapat menjadi dasar bagi penyusunan program pelatihan guru, pengembangan standar operasional prosedur (SOP) penyusunan instrumen evaluasi, serta peningkatan kualitas sistem evaluasi pembelajaran di madrasah (Asmi, 2024).

Penelitian studi kasus ini bertujuan untuk menganalisis kualitas instrumen evaluasi pendidikan yang digunakan dalam proses pembelajaran di MI Al-Musyarrofah Cianjur. Fokus penelitian meliputi penggunaan skala pengukuran, praktik validasi instrumen, reliabilitas instrumen evaluasi, analisis kesalahan pengukuran (*measurement error*), serta upaya perbaikan instrumen evaluasi pendidikan. Penelitian menggunakan pendekatan kualitatif dengan metode studi kasus untuk memperoleh pemahaman yang mendalam mengenai praktik evaluasi pendidikan dalam konteks nyata madrasah. Data dikumpulkan melalui observasi, wawancara mendalam, dan studi dokumentasi terhadap kepala madrasah, guru, serta berbagai dokumen evaluasi pembelajaran yang digunakan di MI Al-Musyarrofah Cianjur.

## **B. METODE PENELITIAN**

Menurut Rahardjo dikutip (Arifudin, 2024) bahwa metode penelitian merupakan salah satu cara untuk memperoleh dan mencari kebenaran yang bersifat tentatif, bukan kebenaran absolut. Hasilnya berupa kebenaran ilmiah. Kebenaran ilmiah merupakan kebenaran yang terbuka untuk terus diuji, dikritik bahkan direvisi. Oleh karena itu tidak ada metode terbaik untuk mencari kebenaran, tetapi yang ada adalah metode yang tepat untuk tujuan tertentu sesuai fenomena yang ada. Budiharto dikutip (Erfiyana, 2025) bahwa pemilihan metode penelitian harus disesuaikan dengan penelitian yang sedang dilakukan agar hasilnya optimal.

Penelitian ini menggunakan desain *descriptive quantitative case study* (Studi Kasus) untuk menganalisis kualitas instrumen evaluasi pendidikan di MI Al-Musyarrofah Cianjur. Desain ini dipilih karena penelitian berfokus pada pengukuran kondisi aktual terkait penggunaan skala pengukuran, validitas instrumen, reliabilitas instrumen, dan kendala pengembangan instrumen evaluasi berdasarkan data kuantitatif yang diperoleh melalui angket. Pendekatan studi kasus digunakan karena penelitian dilakukan pada satu lembaga pendidikan tertentu sehingga hasil penelitian dimaksudkan untuk memberikan gambaran kontekstual mengenai praktik evaluasi pendidikan di MI Al-Musyarrofah Cianjur dan tidak ditujukan untuk generalisasi yang luas.

Penelitian dilaksanakan di MI Al-Musyarrofah Cianjur dengan responden yang terdiri atas satu kepala madrasah dan enam guru yang terlibat dalam penyusunan dan pelaksanaan evaluasi pembelajaran. Pemilihan responden dilakukan secara purposive berdasarkan keterlibatan mereka dalam proses pengembangan dan penggunaan instrumen evaluasi pendidikan. Jumlah responden yang relatif terbatas menunjukkan bahwa hasil

penelitian merepresentasikan kondisi spesifik di MI Al-Musyarrofah Cianjur dan perlu ditafsirkan sesuai dengan konteks penelitian.

Instrumen utama penelitian berupa angket tertutup menggunakan skala Likert lima tingkat, yaitu sangat setuju (5), setuju (4), ragu-ragu (3), tidak setuju (2), dan sangat tidak setuju (1). Angket disusun berdasarkan empat indikator utama penelitian, yaitu penggunaan skala pengukuran, validitas instrumen evaluasi, reliabilitas instrumen evaluasi, dan kendala pengembangan instrumen evaluasi (Syafi’I et al, 2025).

**Tabel 1.** Indikator Instrumen Penelitian

| Variabel                       | Indikator                                                                                       |
|--------------------------------|-------------------------------------------------------------------------------------------------|
| Skala Pengukuran               | Kesesuaian penggunaan skala penilaian, penyusunan indikator, dan interpretasi hasil evaluasi    |
| Validitas Instrumen            | Penyusunan kisi-kisi, kesesuaian butir soal dengan tujuan pembelajaran, dan telaah instrumen    |
| Reliabilitas Instrumen         | Konsistensi penggunaan instrumen, keseragaman kriteria penilaian, dan stabilitas hasil evaluasi |
| Kendala Pengembangan Instrumen | Keterbatasan waktu, kompetensi evaluasi, sarana pendukung, dan penggunaan teknologi evaluasi    |

Sebelum digunakan, instrumen angket terlebih dahulu melalui uji validitas isi (*content validity*) yang dilakukan oleh dua ahli di bidang evaluasi pendidikan dan metodologi penelitian. Perbaikan instrumen dilakukan berdasarkan masukan para ahli mengenai kejelasan indikator, kesesuaian butir pernyataan, dan keterbacaan instrumen. Reliabilitas instrumen dianalisis menggunakan koefisien Cronbach’s Alpha. Hasil pengujian menunjukkan nilai Alpha sebesar  $\geq 0,70$  sehingga instrumen dinyatakan memiliki tingkat reliabilitas yang memadai untuk digunakan dalam penelitian.

Data utama diperoleh melalui penyebaran angket kepada seluruh responden. Untuk memperkuat interpretasi hasil kuantitatif, penelitian juga menggunakan wawancara semi-terstruktur dan studi dokumentasi sebagai data pendukung. Wawancara dilakukan kepada kepala madrasah dan beberapa guru untuk memperoleh penjelasan lebih mendalam mengenai praktik penyusunan instrumen evaluasi, proses validasi, penerapan reliabilitas, dan kendala yang dihadapi (Romdona et al, 2025). Dokumentasi dilakukan dengan menelaah kisi-kisi soal, instrumen tes, lembar observasi, rubrik penilaian, dan perangkat evaluasi lainnya yang digunakan dalam pembelajaran.

Analisis data kuantitatif dilakukan menggunakan teknik persentase dengan rumus:  $\text{persentase} = (\text{Skor Perolehan} / \text{Skor Maksimum}) \times 100\%$

Hasil persentase kemudian diinterpretasikan berdasarkan kategori penilaian yang telah ditetapkan, yaitu sangat baik, baik, cukup, kurang, dan sangat kurang. Data wawancara dan dokumentasi digunakan untuk menjelaskan serta memperkuat temuan kuantitatif yang diperoleh dari angket (Sholihan et al, 2024).

Keabsahan data dijaga melalui triangulasi teknik, yaitu membandingkan hasil angket, wawancara, dan dokumentasi. Dengan pendekatan ini, hasil penelitian tidak hanya menggambarkan tingkat kualitas instrumen evaluasi berdasarkan data numerik, tetapi juga memberikan pemahaman kontekstual mengenai praktik evaluasi pendidikan yang berlangsung di MI Al-Musyarrofah Cianjur.

## C. HASIL DAN PEMBAHASAN

### Hasil

Hasil penelitian menunjukkan bahwa penggunaan skala pengukuran, validitas instrumen, dan reliabilitas instrumen evaluasi di MI Al-Musyarrofah Cianjur memperoleh kategori sangat baik dengan persentase berturut-turut sebesar 89%, 87%, dan 86%. Di sisi lain, kendala pengembangan instrumen evaluasi masih memperoleh persentase sebesar 73%, yang menunjukkan bahwa berbagai hambatan masih dirasakan oleh guru dalam praktik evaluasi pembelajaran. Temuan ini menampilkan suatu kondisi yang menarik karena tingginya skor kualitas instrumen tidak sepenuhnya sejalan dengan temuan wawancara dan dokumentasi yang menunjukkan adanya berbagai keterbatasan dalam pengembangan instrumen evaluasi (Saputra & Mayurida, 2025).

Tingginya skor pada aspek skala pengukuran menunjukkan bahwa guru telah menggunakan berbagai bentuk penilaian sesuai dengan tuntutan kurikulum dan kebutuhan pembelajaran. Temuan ini sejalan dengan pandangan (Sugiyono, 2011) bahwa penggunaan skala pengukuran yang tepat merupakan dasar dalam menghasilkan data evaluasi yang dapat diinterpretasikan secara benar. Namun demikian, hasil wawancara menunjukkan bahwa sebagian guru belum memahami secara mendalam perbedaan karakteristik skala nominal, ordinal, interval, dan rasio. Kondisi ini mengindikasikan bahwa penggunaan skala pengukuran lebih didasarkan pada pengalaman praktis dan kebiasaan administratif daripada pemahaman teoritis yang kuat. Dengan demikian, skor yang tinggi belum tentu mencerminkan penguasaan konsep pengukuran secara komprehensif.

Pada aspek validitas instrumen, persentase sebesar 87% menunjukkan bahwa guru telah berupaya menyusun instrumen yang sesuai dengan tujuan pembelajaran melalui penyusunan kisi-kisi dan penyesuaian dengan indikator kompetensi. Temuan ini mendukung teori yang menyatakan bahwa validitas merupakan syarat utama agar instrumen mampu mengukur kompetensi yang dituju secara tepat. Akan tetapi, hasil dokumentasi dan wawancara menunjukkan bahwa proses validasi masih dilakukan melalui diskusi internal antar guru dan belum melibatkan prosedur validasi yang lebih sistematis, seperti validasi ahli atau analisis empiris butir soal (Roebianto et al, 2023). Oleh karena itu, validitas yang dinilai tinggi dalam penelitian ini lebih merefleksikan persepsi dan praktik guru daripada bukti teknis yang terukur secara objektif.

Temuan serupa terlihat pada aspek reliabilitas instrumen yang memperoleh persentase sebesar 86%. Guru berupaya menjaga konsistensi penilaian melalui penggunaan rubrik dan pedoman yang sama dalam proses evaluasi. Hal ini sejalan dengan konsep reliabilitas yang menekankan konsistensi hasil pengukuran. Namun demikian, sebagian besar responden mengakui bahwa mereka belum pernah melakukan pengujian reliabilitas secara statistik. Kondisi ini menunjukkan bahwa reliabilitas dipahami dalam konteks praktis sebagai keseragaman prosedur penilaian, bukan sebagai pengujian konsistensi instrumen menggunakan metode ilmiah. Dengan demikian, hasil reliabilitas yang tinggi perlu ditafsirkan secara hati-hati karena belum didukung oleh bukti pengujian statistik yang memadai (Manik & Liana, 2025).

Temuan penelitian juga menunjukkan adanya kendala dalam pengembangan instrumen evaluasi yang meliputi keterbatasan waktu, kurangnya pelatihan evaluasi pendidikan, serta keterbatasan sarana dan teknologi pendukung. Keberadaan kendala

tersebut menunjukkan bahwa kualitas evaluasi pembelajaran tidak hanya ditentukan oleh kemampuan individu guru, tetapi juga oleh dukungan kelembagaan yang tersedia. Menariknya, meskipun kendala masih cukup tinggi, responden tetap memberikan penilaian yang tinggi terhadap kualitas instrumen yang mereka gunakan. Fenomena ini dapat dijelaskan melalui kemungkinan adanya self-assessment bias, yaitu kecenderungan responden memberikan penilaian positif terhadap praktik yang mereka lakukan sendiri. Selain itu, jumlah responden yang relatif kecil dan berasal dari satu lembaga pendidikan memungkinkan munculnya bias sosial (social desirability bias), yaitu kecenderungan memberikan jawaban yang dianggap sesuai dengan harapan institusi atau lingkungan kerja (Roebianto et al, 2023).

Dengan demikian, hasil penelitian perlu dipahami sebagai gambaran persepsi dan praktik evaluasi pendidikan dalam konteks MI Al-Musyarrofah Cianjur, bukan sebagai ukuran objektif terhadap kualitas instrumen evaluasi secara teknis. Temuan ini menunjukkan adanya kesenjangan antara praktik evaluasi yang dianggap telah berjalan dengan baik oleh guru dan standar teknis evaluasi pendidikan yang mensyaratkan pengujian validitas serta reliabilitas secara sistematis. Kesenjangan tersebut menjadi temuan penting penelitian ini karena menunjukkan bahwa peningkatan kualitas evaluasi pembelajaran tidak hanya memerlukan penyusunan instrumen yang sesuai kurikulum, tetapi juga penguatan kompetensi guru dalam aspek pengukuran dan evaluasi pendidikan berbasis bukti (Astutik et al, 2025).

Secara keseluruhan, penelitian ini menunjukkan bahwa praktik evaluasi di MI Al-Musyarrofah Cianjur telah berkembang ke arah yang positif, tetapi masih memerlukan penguatan pada aspek teknis pengembangan instrumen. Oleh karena itu, program pelatihan yang berfokus pada penyusunan kisi-kisi, validasi instrumen, analisis reliabilitas, dan pemanfaatan teknologi evaluasi perlu ditingkatkan agar kualitas instrumen evaluasi tidak hanya baik berdasarkan persepsi pengguna, tetapi juga memenuhi standar ilmiah pengukuran pendidikan.

## **Pembahasan**

Penelitian ini bertujuan untuk menganalisis kualitas instrumen evaluasi pendidikan yang digunakan di Madrasah Ibtidaiyah Swasta Al-Musyarrofah melalui kajian terhadap aspek skala pengukuran, validitas, reliabilitas, serta kendala dalam pengembangan instrumen evaluasi pembelajaran. Data diperoleh melalui penyebaran angket kepada responden yang memiliki peran dalam pelaksanaan evaluasi pembelajaran di madrasah (Hidayatullah & Nurjanah, 2022)

Berdasarkan hasil pengisian angket, diperoleh gambaran bahwa secara umum pelaksanaan evaluasi pembelajaran di madrasah telah berjalan dengan cukup baik. Hal ini terlihat dari kecenderungan jawaban responden yang menunjukkan persetujuan terhadap sebagian besar indikator yang diajukan dalam instrumen penelitian.

### **1. Skala Pengukuran dalam Evaluasi Pembelajaran**

Hasil analisis menunjukkan bahwa guru telah menggunakan berbagai bentuk instrumen evaluasi dalam proses pembelajaran, seperti tes tertulis, observasi, angket, serta penilaian praktik. Penggunaan instrumen tersebut dinilai telah disesuaikan dengan tujuan pembelajaran yang hendak dicapai. Selain itu, guru juga telah menerapkan penggunaan kategori penilaian dan skor secara sistematis dalam proses evaluasi hasil belajar peserta didik (Sugiyono, 2023).

Temuan penelitian juga memperlihatkan bahwa penggunaan skala pengukuran dalam evaluasi telah mempertimbangkan karakteristik peserta didik pada jenjang Madrasah Ibtidaiyah. Madrasah juga telah memiliki pedoman evaluasi pembelajaran yang digunakan sebagai acuan dalam penyusunan instrumen penilaian.

Secara keseluruhan, aspek skala pengukuran menunjukkan bahwa sistem evaluasi pembelajaran di madrasah telah mengarah pada penerapan prinsip-prinsip penilaian yang terstruktur dan sistematis.

## **2. Validitas Instrumen Evaluasi**

Pada aspek validitas instrumen, hasil penelitian menunjukkan bahwa guru telah melakukan penyusunan kisi-kisi sebelum menyusun instrumen evaluasi. Selain itu, instrumen yang disusun juga terlebih dahulu ditelaah atau direview sebelum digunakan dalam proses pembelajaran.

Validitas isi instrumen telah diperhatikan melalui kesesuaian antara butir soal dengan indikator pembelajaran yang telah ditetapkan. Guru juga memahami pentingnya validitas konstruk dalam penyusunan instrumen evaluasi agar instrumen mampu mengukur kompetensi yang seharusnya diukur (Sugiyono, 2011).

Madrasah telah melakukan diskusi atau review instrumen evaluasi antar guru sebagai bagian dari upaya menjaga kualitas instrumen penilaian. Di samping itu, uji coba instrumen evaluasi juga pernah dilakukan sebelum instrumen digunakan kepada peserta didik.

Berdasarkan hasil tersebut, dapat disimpulkan bahwa aspek validitas instrumen telah menjadi perhatian dalam praktik evaluasi pembelajaran di madrasah.

## **3. Reliabilitas Instrumen Evaluasi**

Hasil penelitian menunjukkan bahwa guru telah menggunakan pedoman penilaian yang jelas dalam proses evaluasi pembelajaran. Penggunaan pedoman tersebut membantu menjaga konsistensi hasil penilaian pada berbagai pelaksanaan evaluasi.

Madrasah juga melakukan pengawasan terhadap kualitas instrumen evaluasi yang digunakan guru. Selain itu, guru melakukan revisi terhadap instrumen evaluasi apabila ditemukan kekurangan atau kelemahan dalam pelaksanaannya.

Aspek reliabilitas instrumen juga menjadi perhatian dalam kegiatan penilaian pembelajaran. Hal ini menunjukkan adanya kesadaran mengenai pentingnya konsistensi dan keterandalan instrumen evaluasi dalam menghasilkan data penilaian yang objektif (Nurhan et al, 2025).

## **4. Kendala dalam Pengembangan Instrumen Evaluasi**

Meskipun secara umum kualitas instrumen evaluasi telah berjalan dengan baik, penelitian ini menemukan beberapa kendala yang masih dihadapi guru dalam pengembangan instrumen evaluasi pembelajaran.

Guru sudah memiliki pemahaman terkait dengan menyusun instrumen evaluasi yang berkualitas, terutama dalam menyusun indikator, kisi-kisi, serta penentuan bentuk evaluasi yang sesuai dengan tujuan pembelajaran. Namun terdapat sedikit kendala dalam pengembangan instrumen evaluasi .

Penelitian ini juga menemukan bahwa pelatihan terkait evaluasi pendidikan masih terbatas. Pemahaman guru mengenai validitas dan reliabilitas instrumen juga masih perlu ditingkatkan agar kualitas instrumen evaluasi dapat lebih optimal.

Di sisi lain, sarana dan dukungan teknologi evaluasi di madrasah masih memerlukan pengembangan untuk menunjang proses penilaian yang lebih efektif dan efisien.

Analisis Skala, Validitas, dan Reliabilitas Instrumen Evaluasi Pendidikan di MI Al-Musyarrofah Cianjur

### 1) Identitas Responden Penelitian

Penelitian ini melibatkan 7 responden yang terdiri atas kepala sekolah dan guru di MI Al-Musyarrofah Cianjur. Berikut identitas responden penelitian:

**Tabel 2.** identitas responden

| No | Nama Responden                  | Jabatan        | Lama Menjabat | Tanggal Pengisian |
|----|---------------------------------|----------------|---------------|-------------------|
| 1  | Adi Karyadi Surya, S.Pd.I., M.M | Kepala Sekolah | 8 Tahun       | 26 Mei 2026       |
| 2  | Heti Nurmayanti, S.Pd           | Wali Kelas     | 19 Tahun      | 26 Mei 2026       |
| 3  | Samsu, S.Pd                     | Guru PAI       | 20 Tahun      | 26 Mei 2026       |
| 4  | Ahmad Muhtar, S.Pd.I            | Wali Kelas     | 18 Tahun      | 26 Mei 2026       |
| 5  | Lisky Ririn Maulida, S.Pd.SD    | Wali Kelas     | 10 Tahun      | 26 Mei 2026       |
| 6  | Elis Riana Wildan, S.Pd.I       | Wali Kelas     | 20 Tahun      | 26 Mei 2026       |
| 7  | Ihas Rofiqoh, S.Ag              | Wali Kelas     | 20 Tahun      | 26 Mei 2026       |

### 2) Hasil Olahan Angket

#### a. Analisis Skala Pengukuran Instrumen Evaluasi

Berdasarkan hasil pengolahan angket terhadap aspek skala pengukuran instrumen evaluasi pendidikan, diperoleh hasil bahwa mayoritas responden menyatakan “Setuju” dan “Sangat Setuju” terhadap indikator yang diajukan. Hal ini menunjukkan bahwa sistem evaluasi pembelajaran di Madrasah Ibtidaiyah Al-Musyarrofah telah menggunakan skala pengukuran yang cukup baik dan sistematis.

**Tabel 3.** Hasil Analisis Skala Pengukuran

| Indikator                                                | SS | S | KS | TS |
|----------------------------------------------------------|----|---|----|----|
| Penggunaan instrumen evaluasi sesuai tujuan pembelajaran | 4  | 3 | 0  | 0  |
| Penggunaan pedoman penilaian yang jelas                  | 5  | 2 | 0  | 0  |
| Penilaian menggunakan kategori/skor yang sistematis      | 4  | 3 | 0  | 0  |
| Instrumen disesuaikan dengan karakteristik siswa         | 3  | 4 | 0  | 0  |
| Guru memahami penggunaan skala penilaian                 | 4  | 3 | 0  | 0  |

#### Interpretasi Hasil

Hasil angket menunjukkan bahwa guru telah memahami penggunaan skala penilaian dalam evaluasi pembelajaran. Penggunaan skor dan kategori penilaian dilakukan secara sistematis sehingga dapat membantu guru dalam menentukan capaian belajar siswa secara objektif.

Selain itu, instrumen evaluasi yang digunakan juga telah disesuaikan dengan tujuan pembelajaran dan karakteristik peserta didik pada jenjang Madrasah Ibtidaiyah.

### b. Analisis Validitas Instrumen Evaluasi

Hasil pengolahan data pada aspek validitas menunjukkan bahwa guru telah memperhatikan kesesuaian antara indikator pembelajaran dengan instrumen evaluasi yang disusun.

**Tabel 4.** Hasil Analisis Validitas

| Indikator                                     | SS | S | KS | TS |
|-----------------------------------------------|----|---|----|----|
| Penyusunan kisi-kisi sebelum membuat soal     | 5  | 2 | 0  | 0  |
| Kesesuaian soal dengan indikator pembelajaran | 4  | 3 | 0  | 0  |
| Instrumen direview sebelum digunakan          | 3  | 4 | 0  | 0  |
| Adanya diskusi/review instrumen antar guru    | 4  | 2 | 1  | 0  |
| Guru memahami konsep validitas instrumen      | 3  | 3 | 1  | 0  |

#### Interpretasi Hasil

Hasil penelitian menunjukkan bahwa guru telah melakukan penyusunan kisi-kisi dan menyesuaikan butir soal dengan indikator pembelajaran. Hal tersebut menunjukkan bahwa validitas isi instrumen telah menjadi perhatian dalam proses evaluasi pembelajaran.

Walaupun demikian, masih terdapat sebagian kecil responden yang menyatakan “Kurang Setuju” terkait pemahaman validitas instrumen. Hal ini menunjukkan bahwa pemahaman teknis mengenai validitas evaluasi masih perlu ditingkatkan melalui pelatihan dan pendampingan.

### c. Analisis Reliabilitas Instrumen Evaluasi

Aspek reliabilitas instrumen menunjukkan bahwa guru telah berupaya menjaga konsistensi hasil penilaian dalam evaluasi pembelajaran.

**Tabel 5.** Hasil Analisis Reliabilitas

| Indikator                                          | SS | S | KS | TS |
|----------------------------------------------------|----|---|----|----|
| Penggunaan pedoman penilaian yang konsisten        | 4  | 3 | 0  | 0  |
| Adanya revisi instrumen setelah evaluasi           | 3  | 4 | 0  | 0  |
| Instrumen mampu menghasilkan penilaian yang stabil | 4  | 2 | 1  | 0  |
| Madrasah melakukan pengawasan kualitas instrumen   | 5  | 2 | 0  | 0  |
| Guru memahami pentingnya reliabilitas              | 3  | 3 | 1  | 0  |

#### Interpretasi Hasil

Hasil penelitian menunjukkan bahwa reliabilitas instrumen evaluasi di MI Al-Musyarrafah berada pada kategori baik. Guru telah menggunakan pedoman penilaian yang konsisten dan melakukan revisi terhadap instrumen apabila ditemukan kelemahan dalam pelaksanaannya.

Namun demikian, masih terdapat sebagian kecil guru yang belum memahami secara optimal konsep reliabilitas instrumen evaluasi. Oleh karena itu, diperlukan peningkatan kompetensi guru dalam bidang evaluasi pendidikan.

### d. Kendala dalam Pengembangan Instrumen Evaluasi

Berdasarkan hasil angket, terdapat beberapa kendala yang dihadapi guru dalam pengembangan instrumen evaluasi pendidikan.

**Tabel 6.** Kendala Pengembangan Instrumen

| <b>Kendala</b>                                      | <b>Jumlah Responden</b> |
|-----------------------------------------------------|-------------------------|
| Keterbatasan waktu penyusunan instrumen             | 6                       |
| Kurangnya pelatihan evaluasi pendidikan             | 5                       |
| Kesulitan menyusun indikator dan kisi-kisi          | 4                       |
| Pemahaman validitas dan reliabilitas masih terbatas | 3                       |
| Keterbatasan sarana teknologi evaluasi              | 4                       |

### **Interpretasi Hasil**

Kendala utama yang dihadapi guru adalah keterbatasan waktu dalam menyusun instrumen evaluasi. Selain itu, kurangnya pelatihan mengenai evaluasi pendidikan juga menjadi faktor yang mempengaruhi kualitas penyusunan instrumen.

Guru juga masih mengalami kesulitan dalam menyusun indikator dan kisi-kisi evaluasi yang sesuai dengan tujuan pembelajaran. Oleh karena itu, diperlukan program pelatihan dan penguatan kompetensi evaluasi pendidikan secara berkelanjutan.

Berdasarkan hasil olahan angket penelitian, dapat disimpulkan bahwa kualitas instrumen evaluasi pendidikan di MI Al-Musyarrofah Cianjur secara umum berada pada kategori baik. Guru telah menerapkan penggunaan skala penilaian, memperhatikan validitas instrumen, serta menjaga reliabilitas evaluasi pembelajaran.

Namun demikian, masih terdapat beberapa kendala dalam pengembangan instrumen evaluasi, terutama terkait keterbatasan waktu, kurangnya pelatihan, serta pemahaman teknis mengenai validitas dan reliabilitas instrumen. Oleh karena itu, diperlukan penguatan kompetensi guru dan dukungan madrasah dalam meningkatkan kualitas evaluasi pendidikan secara berkelanjutan.

### **Analisis Skala, Validasi, dan Reliabilitas Instrumen Evaluasi Pendidikan di MI Al Musyarrofah Cianjur**

Jumlah Responden

- Kepala Madrasah : 1 orang
- Guru : 6 orang
- Total Responden : 7 orang

**Tabel 7.**Kriteria Interpretasi Skor

| <b>Interval Persentase</b> | <b>Kategori</b> |
|----------------------------|-----------------|
| 81% – 100%                 | Sangat Baik     |
| 61% – 80%                  | Baik            |
| 41% – 60%                  | Cukup           |
| 21% – 40%                  | Kurang          |
| 0% – 20%                   | Sangat Kurang   |

**Tabel 6.** Variabel Skala Pengukuran

| No            | Indikator                                                | Skor Perolehan | Skor Maksimal | Persentase | Kategori           |
|---------------|----------------------------------------------------------|----------------|---------------|------------|--------------------|
| 1             | Penggunaan instrumen evaluasi sesuai tujuan pembelajaran | 25             | 28            | 89%        | Sangat Baik        |
| 2             | Penggunaan pedoman penilaian yang jelas                  | 26             | 28            | 93%        | Sangat Baik        |
| 3             | Penilaian menggunakan kategori/skor yang sistematis      | 25             | 28            | 89%        | Sangat Baik        |
| 4             | Instrumen disesuaikan dengan karakteristik siswa         | 24             | 28            | 86%        | Sangat Baik        |
| 5             | Guru memahami penggunaan skala penilaian                 | 25             | 28            | 89%        | Sangat Baik        |
| <b>Jumlah</b> |                                                          | <b>125</b>     | <b>140</b>    | <b>89%</b> | <b>Sangat Baik</b> |

**Keterangan Perhitungan**

- Jumlah responden : 7 orang
- Skor tertinggi setiap item : 4
- Skor maksimal setiap indikator :  $7 \times 4 = 28$

**Tabel 7.** Kriteria Interpretasi Skor

| Interval Persentase | Kategori      |
|---------------------|---------------|
| 81% – 100%          | Sangat Baik   |
| 61% – 80%           | Baik          |
| 41% – 60%           | Cukup         |
| 21% – 40%           | Kurang        |
| 0% – 20%            | Sangat Kurang |

**Interpretasi Hasil**

Berdasarkan hasil olahan variabel skala pengukuran, diperoleh persentase rata-rata sebesar 89% dengan kategori “Sangat Baik”. Hal ini menunjukkan bahwa guru di MI Al-Musyarrafah Cianjur telah menerapkan penggunaan skala penilaian dan instrumen evaluasi pembelajaran secara sistematis, terarah, dan sesuai dengan tujuan pembelajaran. Selain itu, guru juga memahami penggunaan pedoman penilaian dalam proses evaluasi pendidikan.

**Tabel 8.** Variabel Validasi Instrumen

| No | Indikator                                     | Skor Perolehan | Skor Maksimal | Persentase | Kategori    |
|----|-----------------------------------------------|----------------|---------------|------------|-------------|
| 1  | Penyusunan kisi-kisi sebelum membuat soal     | 26             | 28            | 93%        | Sangat Baik |
| 2  | Kesesuaian soal dengan indikator pembelajaran | 25             | 28            | 89%        | Sangat Baik |
| 3  | Instrumen direview sebelum digunakan          | 24             | 28            | 86%        | Sangat Baik |

|               |                                            |            |            |            |                    |
|---------------|--------------------------------------------|------------|------------|------------|--------------------|
| 4             | Adanya diskusi/review instrumen antar guru | 24         | 28         | 86%        | Sangat Baik        |
| 5             | Guru memahami konsep validitas instrumen   | 23         | 28         | 82%        | Sangat Baik        |
| <b>Jumlah</b> |                                            | <b>122</b> | <b>140</b> | <b>87%</b> | <b>Sangat Baik</b> |

#### Keterangan Perhitungan

- Jumlah responden : 7 orang
- Skor tertinggi setiap item : 4
- Skor maksimal setiap indikator :  $7 \times 4 = 28$

**Tabel 9.** Kriteria Interpretasi Skor

| Interval Persentase | Kategori      |
|---------------------|---------------|
| 81% – 100%          | Sangat Baik   |
| 61% – 80%           | Baik          |
| 41% – 60%           | Cukup         |
| 21% – 40%           | Kurang        |
| 0% – 20%            | Sangat Kurang |

#### Interpretasi Hasil

Berdasarkan hasil olahan variabel skala pengukuran, diperoleh persentase rata-rata sebesar 89% dengan kategori “Sangat Baik”. Hal ini menunjukkan bahwa guru di MI Al-Musyarrofah Cianjur telah menerapkan penggunaan skala penilaian dan instrumen evaluasi pembelajaran secara sistematis, terarah, dan sesuai dengan tujuan pembelajaran. Selain itu, guru juga memahami penggunaan pedoman penilaian dalam proses evaluasi pendidikan.

**Tabel 10.** Variabel Reliabilitas Instrumen

| No            | Indikator                                          | Skor Perolehan | Skor Maksimal | Persentase | Kategori           |
|---------------|----------------------------------------------------|----------------|---------------|------------|--------------------|
| 1             | Penggunaan pedoman penilaian yang konsisten        | 25             | 28            | 89%        | Sangat Baik        |
| 2             | Adanya revisi instrumen setelah evaluasi           | 24             | 28            | 86%        | Sangat Baik        |
| 3             | Instrumen mampu menghasilkan penilaian yang stabil | 23             | 28            | 82%        | Sangat Baik        |
| 4             | Madrasah melakukan pengawasan kualitas instrumen   | 26             | 28            | 93%        | Sangat Baik        |
| 5             | Guru memahami pentingnya reliabilitas instrumen    | 23             | 28            | 82%        | Sangat Baik        |
| <b>Jumlah</b> |                                                    | <b>121</b>     | <b>140</b>    | <b>86%</b> | <b>Sangat Baik</b> |

#### Keterangan Perhitungan

- Jumlah responden : 7 orang
- Skor tertinggi setiap item : 4
- Skor maksimal setiap indikator :  $7 \times 4 = 28$

**Tabel 11.** Kriteria Interpretasi Skor

| Interval Persentase | Kategori      |
|---------------------|---------------|
| 81% – 100%          | Sangat Baik   |
| 61% – 80%           | Baik          |
| 41% – 60%           | Cukup         |
| 21% – 40%           | Kurang        |
| 0% – 20%            | Sangat Kurang |

### Interpretasi Hasil

Berdasarkan hasil olahan variabel reliabilitas instrumen, diperoleh persentase rata-rata sebesar 86% dengan kategori “Sangat Baik”. Hal ini menunjukkan bahwa reliabilitas instrumen evaluasi pendidikan di MI Al-Musyarrofah Cianjur telah berjalan dengan baik. Guru telah menggunakan pedoman penilaian secara konsisten, melakukan revisi terhadap instrumen evaluasi, serta memahami pentingnya reliabilitas dalam menjaga kestabilan dan keobjektifan hasil penilaian pembelajaran. Selain itu, madrasah juga melakukan pengawasan terhadap kualitas instrumen evaluasi yang digunakan dalam proses pembelajaran.

**Tabel 12.** Variabel Kendala Pengembangan Instrumen Evaluasi

| No            | Indikator                                              | Skor Perolehan | Skor Maksimal | Persentase | Kategori    |
|---------------|--------------------------------------------------------|----------------|---------------|------------|-------------|
| 1             | Keterbatasan waktu dalam penyusunan instrumen evaluasi | 22             | 28            | 79%        | Baik        |
| 2             | Kurangnya pelatihan evaluasi pendidikan                | 21             | 28            | 75%        | Baik        |
| 3             | Kesulitan menyusun indikator dan kisi-kisi             | 20             | 28            | 71%        | Baik        |
| 4             | Pemahaman validitas dan reliabilitas masih terbatas    | 19             | 28            | 68%        | Baik        |
| 5             | Keterbatasan sarana dan teknologi evaluasi             | 20             | 28            | 71%        | Baik        |
| <b>Jumlah</b> |                                                        | <b>102</b>     | <b>140</b>    | <b>73%</b> | <b>Baik</b> |

### Keterangan Perhitungan

- Jumlah responden : 7 orang
- Skor tertinggi setiap item : 4
- Skor maksimal setiap indikator :  $7 \times 4 = 28$

**Tabel 13.** Kriteria Interpretasi Skor

| Interval Persentase | Kategori      |
|---------------------|---------------|
| 81% – 100%          | Sangat Baik   |
| 61% – 80%           | Baik          |
| 41% – 60%           | Cukup         |
| 21% – 40%           | Kurang        |
| 0% – 20%            | Sangat Kurang |

### **Interpretasi Hasil**

Berdasarkan hasil olahan variabel kendala pengembangan instrumen evaluasi, diperoleh persentase rata-rata sebesar 73% dengan kategori “Baik”. Hasil ini menunjukkan bahwa guru di MI Al-Musyarrofah Cianjur masih menghadapi beberapa kendala dalam pengembangan instrumen evaluasi pendidikan, namun kendala tersebut masih dapat diatasi dalam pelaksanaan pembelajaran.

Kendala utama yang dihadapi adalah keterbatasan waktu dalam penyusunan instrumen evaluasi serta kurangnya pelatihan mengenai evaluasi pendidikan. Selain itu, sebagian guru masih mengalami kesulitan dalam menyusun indikator, kisi-kisi, serta memahami aspek validitas dan reliabilitas instrumen evaluasi. Keterbatasan sarana dan teknologi evaluasi juga menjadi faktor yang mempengaruhi optimalisasi kualitas evaluasi pembelajaran di madrasah. (Hidayatullah & Nurjanah, 2022)

### **D. SIMPULAN DAN SARAN**

Berdasarkan hasil olahan angket, kualitas instrumen evaluasi pendidikan di MI Al Musyarrofah Cianjur secara umum berada pada kategori baik. Penggunaan skala pengukuran dan praktik validasi instrumen telah dilaksanakan dalam proses pembelajaran, meskipun masih terdapat beberapa aspek yang perlu ditingkatkan, khususnya pada uji coba instrumen, validitas kriteria, dan penguatan reliabilitas evaluasi. Selain itu, peningkatan kompetensi guru dalam bidang evaluasi pendidikan menjadi kebutuhan penting untuk mendukung pengembangan instrumen evaluasi yang lebih berkualitas, valid, dan reliabel.

Sebagai tindak lanjut dari hasil penelitian, MI Al-Musyarrofah Cianjur perlu mengoptimalkan pengembangan instrumen evaluasi pendidikan dengan menerapkan prosedur pengujian instrumen secara lebih terencana sebelum digunakan dalam proses pembelajaran. Guru diharapkan terus meningkatkan kemampuan dalam menyusun dan menganalisis instrumen evaluasi, terutama yang berkaitan dengan aspek validitas dan reliabilitas, sehingga kualitas hasil penilaian dapat lebih terjamin. Madrasah juga perlu memberikan dukungan melalui penyelenggaraan program pengembangan kompetensi, seperti pelatihan dan pendampingan di bidang evaluasi pendidikan, serta penyediaan sarana dan teknologi yang mendukung pelaksanaan evaluasi. Selain itu, penelitian selanjutnya diharapkan dapat dilakukan pada cakupan yang lebih luas dengan melibatkan lebih banyak responden agar diperoleh pemahaman yang lebih mendalam mengenai kualitas instrumen evaluasi pendidikan di berbagai satuan pendidikan.

### **UCAPAN TERIMA KASIH**

Penulis mengucapkan terima kasih yang sebesar-besarnya kepada seluruh pihak yang telah memberikan dukungan, bantuan, dan kontribusi dalam pelaksanaan penelitian ini. Secara khusus, penulis menyampaikan apresiasi kepada

1. Seluruh Dewan Guru MI Al-Musyarrofah Cianjur yang telah bersedia menjadi responden serta memberikan data dan informasi yang diperlukan selama proses penelitian berlangsung.
2. Bapak Dr. Mulyawan Safwandy Nugraha, M.Ag., M.Pd. selaku dosen pembimbing yang telah memberikan arahan, bimbingan, masukan, serta motivasi selama proses

penyusunan dan penyelesaian penelitian ini. Semoga segala bantuan dan kebaikan yang telah diberikan memperoleh balasan yang berlipat ganda.

## DAFTAR PUSTAKA

- Abdul et al. (2025). Human Resource Planning and Recruitment Management in Islamic Educational Institutions. *Sustainable Jurnal Kajian Mutu Pendidikan*, 8(1). <https://doi.org/10.32923/segac610>
- Arifudin, O. (2024). Evaluasi Pengelolaan Sarana Dan Prasarana Dalam Meningkatkan Kualitas Pendidikan Islam Di Madrasah Ibtidaiyah. *Antologi Kajian Multidisiplin Ilmu (Al-Kamil)*, 2(2), 560–575.
- Arikunto, S. (2022). *Prosedur Penelitian Suatu Pendekatan Praktek*. Jakarta: Rineka Cipta.
- Asmi, E. (2024). Peran Guru Sebagai Fasilitator Dalam Pembelajaran Konstruktivistik. *J-PAI: Jurnal Pendidikan Agama Islamurnal Penelitian Matematika dan Pendidikan Matematika*, 6(1), 23–35.
- Astutik et al. (2025). Evaluasi Kinerja Guru Pamong: Kontribusi, Tantangan, dan Peningkatan Strategi dalam Mendukung Pendidikan Abad ke-21. *Jurnal Pendidikan dan Pembelajaran Indonesia (JPPI)*, 5(1), 402–413. <https://doi.org/10.53299/jppi.v5i1.1034>
- Erfiyana, E. (2025). Islamic School Financial Management: A Case Study of Islamic Junior High Schools in Rural Areas. *International Journal Of Science Education and Technology Management*, 4(2), 33–44.
- Eryanto et al. (2026). Tujuan Dan Manfaat Evaluasi Dalam Meningkatkan Kualitas Pembelajaran Pendidikan Agama Islam. *Educate: Journal of Education and Learning*, 4(1), 121–132.
- Hidayatullah, & Nurjanah, S. (2022). Pengembangan Instrumen Penilaian Berbasis Kompetensi di Madrasah. *Jurnal Pendidikan Madrasah*, 7(1), 45–58.
- Manik & Liana. (2025). Kendala Guru Dalam Mengimplementasikan Penilaian Autentik Di Mata Pelajaran Bahasa Indonesia Di SMA Nurcahaya Medan. *SECONDARY: Jurnal Inovasi Pendidikan Menengah*, 5(1), 662–671. <https://doi.org/10.51878/secondary.v5i4.8568>
- Maskhuliah et al. (2025). Konsep Dasar Pengukuran Dan Skala Dalam Penelitian Manajemen Pendidikan: Kajian Pustaka. *Jurnal Ilmu Manajemen Dan Pendidikan*, 02(01), 565–572.
- Masrukhin, M. (2019). *Metodologi Penelitian: Konsep dan Implementasi*. Media Ilmu Press.
- Nurhan et al. (2025). Peran Validitas dan Rehabilitas Dalam Evaluasi Pembelajaran. *Socius: Jurnal Penelitian Ilmu-Ilmu Sosial*, 02(11), 224–226.
- Ristamala et al. (2025). Instrument Evaluasi Non Tes Pada Mata Pelajaran Pendidikan Agama Islam Dalam Nilai Nilai Ketauhidan. *Education and Learning Review*, 1(2), 1–10.
- Roebianto et al. (2023). Content Validity: Definition and Procedure of Content Validation in Psychological Research. *TPM - Testing, Psychometrics, Methodology in Applied Psychology*, 30(1), 5–18. <https://doi.org/10.4473/TPM30.1.1>
- Romdona et al. (2025). Teknik Pengumpulan Data: Observasi, Wawancara Dan Kuesioner. *JISOSEPOL: Jurnal Ilmu Sosial Ekonomi dan Politik*, 3(1), 39–47. <https://doi.org/10.61787/taceee75>
- Saputra, A., & Mayurida. (2025). Pengembangan Instrumen Evaluasi (Analisis Pengembangan Penilaian Satuan Pendidikan MTs/SMP). *Jurnal Pendidikan dan Keagamaan*, 2(4), 313–323.
- Sholihan et al. (2024). *Evaluasi Pembelajaran. Evaluasi Pembelajaran*. Direktorat

Jendral Pendidikan Agama Islam.

Sugiyono. (2023). *Metode Penelitian Pendidikan (Kuantitatif, Kualitatif, Kombinasi, R & D dan Penelitian Tindakan)*. Bandung: Alfabeta.

Sugiyono, S. (2011). Qualitative quantitative research methods and R&D. *Bandung: Alfabeta*.

Syafi'I et al. (2025). Evaluasi Pendidikan Sebagai Dasar Pengembangan Instrumen Penilaian Berbasis Kompetensi. *Jurnal Akuntansi, Manajemen Dan Ilmu Pendidikan*, 1(4), 1–12.