

Rekonstruksi Uji Validitas Instrumen dalam Penelitian Pendidikan Agama Islam di Era Transformasi Digital

¹Thaufan Aziz

¹Sekolah Tinggi Ilmu Tarbiyah Madani Yogyakarta, Indonesia

Email: 1ofanaziz@gmail.com

Abstrak

Transformasi digital telah membawa perubahan signifikan terhadap metodologi penelitian pendidikan, termasuk dalam pengembangan dan pengujian validitas instrumen pada penelitian Pendidikan Agama Islam (PAI). Namun, praktik validasi instrumen pada penelitian PAI masih didominasi oleh pendekatan konvensional yang berorientasi pada pemenuhan persyaratan statistik sehingga belum sepenuhnya mampu merepresentasikan konstruk keagamaan yang kompleks dan multidimensional. Penelitian ini bertujuan merekonstruksi paradigma uji validitas instrumen yang lebih adaptif terhadap perkembangan psikometrika modern dan teknologi digital. Penelitian menggunakan pendekatan kualitatif dengan jenis studi kepustakaan (library research) melalui analisis konseptual terhadap berbagai literatur ilmiah yang relevan. Data dianalisis menggunakan teknik content analysis dan thematic analysis untuk mengidentifikasi kesenjangan konseptual serta merumuskan model rekonstruksi validitas instrumen. Hasil penelitian menunjukkan bahwa validitas instrumen perlu dipahami sebagai proses ilmiah yang bersifat berkelanjutan melalui integrasi validitas isi, validitas konstruk, validitas empiris, serta pemanfaatan Artificial Intelligence, Rasch Measurement Theory, Confirmatory Factor Analysis, dan learning analytics. Berdasarkan sintesis tersebut, penelitian ini menawarkan **Digital Integrated Validation Framework (DIVF)** sebagai kerangka konseptual baru yang mampu meningkatkan kualitas pengembangan instrumen penelitian PAI. Kerangka ini diharapkan menjadi acuan metodologis bagi peneliti dalam menghasilkan instrumen yang lebih valid, reliabel, adaptif, dan sesuai dengan tuntutan penelitian di era transformasi digital.

Kata Kunci: Validitas Instrumen; Pendidikan Agama Islam; Transformasi Digital; Psikometrika; Artificial Intelligence.

Abstract

Digital transformation has significantly reshaped educational research methodologies, including the development and validation of research instruments in Islamic Religious Education (IRE). Nevertheless, instrument validation practices in IRE research remain largely dominated by conventional statistical approaches that primarily focus on meeting methodological requirements rather than accurately representing complex and multidimensional religious constructs. This study aims to reconstruct the paradigm of instrument validity by integrating contemporary psychometric approaches with digital technologies. The study employed a qualitative approach using library research through conceptual analysis of relevant scholarly literature. Data were analyzed using content analysis and thematic analysis to identify conceptual gaps and formulate a reconstructed framework for instrument validation. The findings indicate that instrument validity should be viewed as a continuous scientific process that integrates content validity, construct validity, empirical validity, Artificial Intelligence, Rasch Measurement Theory, Confirmatory Factor Analysis, and learning analytics. Based on this synthesis, the study proposes the **Digital Integrated Validation Framework (DIVF)** as a novel conceptual framework for validating research instruments in Islamic Religious Education. This framework is expected to strengthen methodological practices by providing a more valid, reliable, adaptive, and evidence-based approach to instrument development in the era of digital transformation.

Keywords: Instrument Validity; Islamic Religious Education; Digital Transformation; Psychometrics; Artificial Intelligence

PENDAHULUAN

Pendidikan Agama Islam (PAI) pada era transformasi digital menghadapi perubahan paradigma yang tidak hanya menyentuh aspek pembelajaran, tetapi juga merambah pada aspek penelitian dan pengembangan ilmu pengetahuan. Perkembangan teknologi digital, kecerdasan buatan (Artificial Intelligence), learning analytics, big data, serta berbagai platform penelitian berbasis daring telah mengubah cara peneliti mengumpulkan, mengolah, dan menafsirkan data. Dalam konteks tersebut, kualitas instrumen penelitian menjadi faktor yang sangat menentukan karena validitas instrumen berfungsi sebagai fondasi utama dalam menghasilkan temuan yang akurat, objektif, dan dapat dipertanggungjawabkan secara ilmiah. Instrumen yang tidak valid berpotensi menghasilkan kesimpulan yang bias sehingga melemahkan kontribusi penelitian terhadap pengembangan keilmuan Pendidikan Agama Islam (Asdar et al., 2022).

Dalam tradisi penelitian Pendidikan Agama Islam, instrumen sering digunakan untuk mengukur berbagai konstruk abstrak dan multidimensional, seperti religiusitas, karakter Islami, akhlak, spiritualitas, moderasi beragama, kompetensi guru PAI, budaya religius sekolah, hingga internalisasi nilai-nilai Islam. Karakteristik konstruk tersebut berbeda dengan pengukuran aspek kognitif yang bersifat lebih konkret karena melibatkan dimensi afektif, keyakinan, sikap, dan perilaku yang kompleks. Oleh karena itu, proses pengembangan dan pengujian validitas instrumen memerlukan pendekatan yang lebih komprehensif agar mampu merepresentasikan konstruk yang diukur secara tepat (Hair et al., 2022).

Namun demikian, berbagai penelitian PAI masih menunjukkan kecenderungan penggunaan pendekatan validitas yang bersifat konvensional. Praktik yang umum dilakukan adalah menguji validitas instrumen menggunakan korelasi Product Moment, meminta penilaian ahli (expert judgment) secara terbatas, kemudian menyimpulkan bahwa instrumen telah valid ketika memenuhi nilai koefisien tertentu. Pendekatan tersebut memang memiliki landasan metodologis yang kuat dan telah lama digunakan dalam

penelitian pendidikan. Akan tetapi, kompleksitas fenomena pendidikan kontemporer dan perkembangan teknologi digital menuntut adanya pengembangan paradigma validitas yang lebih adaptif dan multidimensional.

Permasalahan mendasar yang muncul adalah bahwa sebagian besar penelitian Pendidikan Agama Islam masih memandang validitas sebagai prosedur statistik yang bersifat administratif, bukan sebagai proses ilmiah yang berkelanjutan untuk memastikan ketepatan interpretasi data. Akibatnya, banyak instrumen yang secara statistik dinyatakan valid, tetapi belum tentu mampu menggambarkan realitas konstruk keagamaan yang sebenarnya. Sebagai contoh, instrumen religiusitas yang hanya diuji menggunakan korelasi item-total belum tentu mampu menjelaskan hubungan antar dimensi keyakinan, praktik ibadah, pengalaman keagamaan, dan pengamalan nilai Islam dalam kehidupan sehari-hari. Kondisi ini menunjukkan adanya kesenjangan antara perkembangan teori validitas modern dengan praktik penelitian yang masih dominan menggunakan pendekatan klasik (Kline, 2023).

Pada sisi lain, transformasi digital telah melahirkan berbagai pendekatan baru dalam pengembangan dan validasi instrumen penelitian. Pemanfaatan Artificial Intelligence memungkinkan identifikasi indikator secara lebih sistematis melalui analisis literatur yang luas. Learning analytics mampu menyediakan data perilaku peserta secara real-time. Model Rasch memungkinkan analisis karakteristik butir yang lebih akurat dibandingkan pendekatan korelasional tradisional. Confirmatory Factor Analysis (CFA) dapat digunakan untuk menguji kesesuaian konstruk teoritis dengan data empiris. Bahkan, perkembangan digital assessment memungkinkan validitas instrumen dievaluasi secara berkelanjutan berdasarkan interaksi pengguna dalam lingkungan pembelajaran digital. Berbagai perkembangan tersebut menunjukkan bahwa validitas instrumen telah bergeser dari konsep yang bersifat statis menuju konsep yang lebih dinamis, integratif, dan berbasis data (Chang, 2024).

Meskipun demikian, kajian mengenai rekonstruksi validitas instrumen dalam penelitian Pendidikan Agama Islam masih relatif terbatas. Sebagian besar penelitian terdahulu berfokus pada teknik pengujian validitas tertentu, seperti validitas isi, validitas konstruk, atau validitas empiris, tanpa mengkaji bagaimana ketiga pendekatan tersebut dapat diintegrasikan dengan teknologi digital untuk menjawab tantangan penelitian masa kini. Akibatnya, belum tersedia kerangka konseptual yang secara khusus dirancang untuk mengakomodasi kebutuhan validasi instrumen penelitian Pendidikan Agama Islam pada era transformasi digital.

Berdasarkan kondisi tersebut, artikel ini berangkat dari satu permasalahan utama, yaitu: mengapa pendekatan validitas instrumen yang selama ini digunakan dalam penelitian Pendidikan Agama Islam belum sepenuhnya mampu menjamin ketepatan pengukuran konstruk keagamaan yang kompleks di era transformasi digital, dan bagaimana model rekonstruksi validitas yang lebih komprehensif dapat dikembangkan untuk menjawab tantangan tersebut? Permasalahan tersebut dipilih karena menyentuh akar persoalan metodologis yang secara langsung memengaruhi kualitas penelitian PAI. Ketika instrumen yang digunakan tidak mampu merepresentasikan konstruk secara akurat, maka kualitas data, ketepatan analisis, dan validitas kesimpulan penelitian turut dipertanyakan. Oleh sebab itu, diperlukan upaya rekonstruksi paradigma validitas instrumen yang tidak hanya mengandalkan pendekatan statistik konvensional, tetapi juga mengintegrasikan perkembangan psikometrika modern dan teknologi digital.

Artikel ini bertujuan untuk merekonstruksi konsep uji validitas instrumen dalam penelitian Pendidikan Agama Islam melalui pengembangan kerangka validasi yang mengintegrasikan validitas isi berbasis pakar, analisis konstruk modern, teknologi digital, serta pendekatan psikometrik kontemporer. Rekonstruksi ini diharapkan dapat menjadi landasan metodologis baru bagi peneliti Pendidikan Agama Islam dalam menghasilkan instrumen yang lebih valid, reliabel, adaptif, dan relevan dengan tuntutan era

transformasi digital. Selain memberikan kontribusi teoretis terhadap pengembangan metodologi penelitian PAI, kajian ini juga diharapkan mampu menjadi rujukan praktis bagi akademisi, mahasiswa pascasarjana, dan peneliti dalam meningkatkan kualitas penelitian keagamaan di masa mendatang.

METODE PENELITIAN

Penelitian ini menggunakan pendekatan kualitatif dengan jenis studi kepustakaan (*library research*) yang dikembangkan melalui analisis konseptual (*conceptual paper*). Pendekatan ini dipilih karena tujuan penelitian bukan menguji hipotesis empiris, melainkan merekonstruksi paradigma uji validitas instrumen dalam penelitian Pendidikan Agama Islam di era transformasi digital berdasarkan sintesis berbagai teori, konsep, dan temuan penelitian mutakhir. Sumber data berupa literatur primer yang berasal dari artikel ilmiah bereputasi internasional dan nasional, buku psikometrika, metodologi penelitian pendidikan, serta dokumen ilmiah yang membahas validitas instrumen, pengukuran pendidikan, transformasi digital, Artificial Intelligence (AI), *learning analytics*, *Rasch Measurement Theory*, dan *Confirmatory Factor Analysis* (CFA). Literatur dipilih berdasarkan kriteria relevansi, kemutakhiran, dan kredibilitas dengan mengutamakan publikasi sepuluh tahun terakhir, tanpa mengesampingkan karya-karya klasik yang menjadi landasan teoritis dalam kajian validitas instrumen (Creswell & Creswell, 2023).

Analisis data dilakukan menggunakan analisis isi (*content analysis*) yang dipadukan dengan analisis tematik (*thematic analysis*) melalui tiga tahapan, yaitu reduksi data, kategorisasi konsep, dan sintesis kritis. Pada tahap pertama, berbagai konsep mengenai validitas instrumen diklasifikasikan berdasarkan paradigma validitas konvensional dan paradigma validitas modern berbasis transformasi digital. Tahap kedua dilakukan identifikasi kesenjangan konseptual (*research gap*) antara praktik validasi instrumen dalam penelitian Pendidikan Agama Islam dengan perkembangan psikometrika kontemporer. Selanjutnya, tahap sintesis digunakan untuk merekonstruksi kerangka konseptual baru yang

mengintegrasikan validitas isi, validitas konstruk, validitas empiris, serta pemanfaatan teknologi digital sebagai model validasi instrumen yang lebih komprehensif dan adaptif terhadap kebutuhan penelitian Pendidikan Agama Islam pada era transformasi digital. Untuk menjamin ketelitian ilmiah, proses analisis dilakukan secara sistematis melalui triangulasi sumber pustaka, komparasi antarhasil penelitian, serta interpretasi kritis terhadap berbagai perspektif yang berkembang dalam literatur (DeVellis & Thorpe, 2021).

HASIL DAN PEMBAHASAN

Keterbatasan Paradigma Uji Validitas Instrumen dalam Penelitian Pendidikan Agama Islam

Transformasi digital telah mengubah hampir seluruh aspek penelitian pendidikan, mulai dari proses pengumpulan data, pengembangan instrumen, analisis statistik, hingga interpretasi hasil penelitian. Namun, perubahan tersebut belum sepenuhnya diikuti oleh perkembangan metodologi penelitian Pendidikan Agama Islam (PAI), khususnya dalam aspek pengujian validitas instrumen. Meskipun berbagai perangkat digital, perangkat lunak statistik, serta teknologi berbasis Artificial Intelligence (AI) telah tersedia dan mudah diakses, praktik validasi instrumen dalam penelitian PAI masih didominasi oleh paradigma konvensional yang berorientasi pada pemenuhan persyaratan statistik semata. Kondisi ini menunjukkan adanya kesenjangan antara perkembangan teknologi dengan praktik metodologis yang diterapkan oleh para peneliti.

Dalam berbagai penelitian PAI, validitas instrumen umumnya dipahami sebagai proses menentukan apakah setiap butir pernyataan memiliki koefisien korelasi yang memenuhi nilai tertentu. Akibatnya, keberhasilan validasi sering kali hanya diukur berdasarkan terpenuhinya nilai r hitung yang lebih besar daripada r tabel atau nilai signifikansi tertentu. Pendekatan tersebut memang memiliki dasar ilmiah yang kuat dan masih relevan dalam konteks tertentu. Akan tetapi, validitas pada hakikatnya bukan sekadar persoalan hubungan statistik antarvariabel, melainkan berkaitan dengan

sejauh mana instrumen benar-benar mampu merepresentasikan konstruk teoretis yang hendak diukur. Dengan kata lain, validitas tidak hanya berbicara mengenai angka, tetapi juga mengenai ketepatan makna (*validity of interpretation*) atas data yang dihasilkan (Setyawarno et al., 2025).

Permasalahan tersebut menjadi semakin kompleks ketika penelitian PAI berupaya mengukur konstruk yang bersifat abstrak, multidimensional, dan sarat dengan nilai-nilai keagamaan. Variabel seperti religiusitas, akhlak, spiritualitas, karakter Islami, moderasi beragama, budaya religius sekolah, maupun kompetensi kepribadian guru tidak dapat direduksi menjadi indikator-indikator yang bersifat linier. Konstruk-konstruk tersebut tersusun atas dimensi kognitif, afektif, konatif, dan perilaku yang saling berinteraksi secara dinamis sesuai dengan konteks sosial, budaya, dan perkembangan individu. Oleh sebab itu, penggunaan uji validitas yang hanya berorientasi pada korelasi antarbutir sering kali belum mampu menggambarkan kompleksitas konstruk tersebut secara utuh (Terry et al., 2025).

Paradigma validitas konvensional juga cenderung memosisikan proses validasi sebagai tahapan akhir setelah instrumen selesai disusun. Peneliti terlebih dahulu merancang indikator berdasarkan kajian teori, kemudian melakukan uji coba kepada sejumlah responden untuk memperoleh nilai validitas statistik. Model seperti ini mengandung kelemahan mendasar karena kualitas indikator sejak tahap perancangan belum tentu telah mencerminkan keseluruhan dimensi konstruk yang akan diukur. Dalam konteks Pendidikan Agama Islam, ketidaktepatan dalam merumuskan indikator dapat menyebabkan instrumen kehilangan makna substantif meskipun secara statistik memenuhi kriteria validitas.

Selain itu, praktik *expert judgment* yang selama ini digunakan dalam penelitian PAI juga masih bersifat subjektif dan belum mengikuti prosedur psikometrik modern. Tidak sedikit penelitian yang hanya melibatkan satu atau dua validator tanpa menggunakan teknik kuantifikasi tingkat kesepakatan antar ahli, seperti *Content Validity Ratio* (CVR), *Content Validity Index* (CVI),

maupun koefisien Aiken's V. Akibatnya, keputusan mengenai valid atau tidaknya suatu indikator lebih banyak didasarkan pada pertimbangan individual daripada kesepakatan ilmiah yang terukur. Kondisi ini berpotensi menimbulkan bias konseptual yang pada akhirnya memengaruhi kualitas instrumen penelitian (Terry et al., 2025).

Kelemahan lain terlihat pada dominannya penggunaan pendekatan *Classical Test Theory* (CTT) dalam penelitian Pendidikan Agama Islam. Pendekatan ini menempatkan skor responden sebagai dasar utama dalam menentukan kualitas instrumen sehingga karakteristik butir sangat bergantung pada karakteristik sampel penelitian. Konsekuensinya, suatu butir dapat dinyatakan valid pada satu kelompok responden, tetapi belum tentu menunjukkan karakteristik yang sama ketika digunakan pada kelompok yang berbeda. Hal tersebut menjadi tantangan serius bagi penelitian PAI yang sering kali melibatkan latar belakang sosial, budaya, dan tingkat religiusitas responden yang beragam (Speer et al., 2025).

Perkembangan psikometrika modern sebenarnya telah menawarkan berbagai alternatif yang lebih komprehensif. *Rasch Measurement Theory*, misalnya, memungkinkan peneliti menganalisis tingkat kesulitan butir, kemampuan responden, kesesuaian model (*fit statistics*), serta mendeteksi adanya butir yang bias (*Differential Item Functioning*). Demikian pula *Confirmatory Factor Analysis* (CFA) mampu menguji apakah struktur konstruk yang dirancang benar-benar sesuai dengan data empiris. Sayangnya, kedua pendekatan tersebut masih relatif jarang digunakan dalam penelitian Pendidikan Agama Islam, sehingga kualitas instrumen sering kali hanya dinilai berdasarkan indikator-indikator statistik yang bersifat parsial.

Era transformasi digital sesungguhnya telah memperluas makna validitas instrumen. Validitas tidak lagi dipahami sebagai proses yang dilakukan satu kali sebelum penelitian berlangsung, tetapi sebagai proses yang bersifat dinamis, berkelanjutan, dan berbasis data (*continuous validation*). Pemanfaatan Artificial Intelligence (AI), *learning analytics*, *natural*

language processing, hingga sistem asesmen digital memungkinkan peneliti memperoleh informasi yang jauh lebih kaya mengenai kualitas setiap indikator. Data interaksi pengguna, pola respons, waktu pengerjaan, hingga konsistensi jawaban dapat dianalisis secara otomatis untuk memperkuat interpretasi terhadap validitas instrumen. Dengan demikian, proses validasi tidak hanya menghasilkan keputusan "valid" atau "tidak valid", tetapi juga memberikan informasi mengenai kelemahan setiap butir sehingga dapat dilakukan penyempurnaan secara berkelanjutan (Speer et al., 2025).

Berdasarkan analisis tersebut, dapat dipahami bahwa akar persoalan penelitian Pendidikan Agama Islam bukan terletak pada penggunaan teknik validitas konvensional itu sendiri, melainkan pada paradigma yang masih memandang validitas sebagai prosedur administratif untuk memenuhi persyaratan metodologis. Paradigma tersebut belum mampu mengakomodasi kompleksitas konstruk keagamaan maupun perkembangan teknologi digital yang telah mengubah cara pengembangan instrumen penelitian. Oleh karena itu, diperlukan rekonstruksi paradigma validitas yang lebih integratif, yaitu paradigma yang menggabungkan kekuatan validitas isi, validitas konstruk, validitas empiris, serta teknologi digital dalam satu kerangka metodologis yang komprehensif.

Tabel 1. Perbandingan Paradigma Uji Validitas Instrumen Konvensional dan Paradigma Transformasi Digital

Aspek	Paradigma Konvensional	Paradigma Transformasi Digital
Orientasi validitas	Memenuhi persyaratan statistik	Menjamin ketepatan interpretasi konstruk
Fokus analisis	Korelasi antarbutir	Representasi konstruk secara multidimensional
Peran validator	Penilaian subjektif pakar	Kolaborasi pakar yang terukur melalui CVR, CVI, dan Aiken's V
Pendekatan psikometrik	Classical Test Theory (CTT)	Rasch Measurement Theory, CFA, dan SEM
Pemanfaatan teknologi	Terbatas pada perangkat statistik	Artificial Intelligence, Digital Assessment, Learning Analytics, dan Big Data
Proses validasi	Dilakukan satu kali sebelum penelitian	Berkelanjutan (<i>continuous validation</i>) sepanjang siklus penelitian

Hasil validasi	Status valid atau tidak valid	Informasi komprehensif mengenai kualitas indikator dan rekomendasi perbaikannya
----------------	-------------------------------	---

Temuan konseptual pada subbab ini menunjukkan bahwa permasalahan utama validitas instrumen dalam penelitian Pendidikan Agama Islam bukan semata-mata terletak pada teknik analisis yang digunakan, tetapi pada paradigma validitas yang belum beradaptasi dengan perkembangan psikometrika modern dan transformasi digital. Oleh karena itu, diperlukan suatu model rekonstruksi yang mampu mengintegrasikan berbagai pendekatan validitas ke dalam satu kerangka yang lebih komprehensif. Model tersebut akan dibahas pada subbab berikutnya sebagai solusi konseptual terhadap keterbatasan paradigma validitas konvensional dalam penelitian Pendidikan Agama Islam.

Rekonstruksi Uji Validitas Instrumen dalam Penelitian Pendidikan Agama Islam Berbasis Transformasi Digital

Analisis pada subbab sebelumnya menunjukkan bahwa permasalahan utama dalam pengujian validitas instrumen penelitian Pendidikan Agama Islam (PAI) bukan terletak pada kelemahan teknik statistik yang digunakan, melainkan pada paradigma validitas yang masih bersifat parsial dan linier. Validitas cenderung dipahami sebagai tahapan administratif yang dilakukan setelah instrumen selesai disusun, sehingga proses pengembangan instrumen belum sepenuhnya mempertimbangkan keterkaitan antara landasan teoretis, kualitas indikator, karakteristik responden, serta perkembangan teknologi digital. Paradigma tersebut menyebabkan proses validasi lebih berorientasi pada pemenuhan persyaratan metodologis dibandingkan memastikan bahwa instrumen benar-benar mampu merepresentasikan konstruk keagamaan yang kompleks (Sari et al., 2025).

Dalam perspektif psikometrika modern, validitas merupakan proses pengumpulan berbagai bukti (*validity evidence*) untuk mendukung interpretasi terhadap skor yang dihasilkan oleh instrumen. Dengan demikian, validitas tidak lagi dipandang sebagai sifat yang melekat pada instrumen,

tetapi sebagai hasil dari akumulasi bukti ilmiah yang diperoleh melalui berbagai pendekatan, mulai dari validitas isi, validitas konstruk, validitas berbasis proses respons, validitas hubungan dengan variabel lain, hingga validitas konsekuensial. Paradigma ini menjadi semakin relevan dalam penelitian Pendidikan Agama Islam karena sebagian besar variabel yang dikaji merupakan konstruk laten (*latent constructs*) yang tidak dapat diamati secara langsung.

Transformasi digital membuka peluang baru dalam membangun sistem validasi instrumen yang lebih komprehensif. Kehadiran Artificial Intelligence (AI), *learning analytics*, *Natural Language Processing* (NLP), *Rasch Measurement Theory*, *Confirmatory Factor Analysis* (CFA), dan sistem asesmen digital memungkinkan peneliti melakukan validasi secara lebih objektif, adaptif, dan berkelanjutan. Teknologi tersebut tidak dimaksudkan untuk menggantikan peran peneliti maupun pakar, melainkan memperkuat proses pengambilan keputusan melalui analisis data yang lebih akurat dan sistematis. Oleh karena itu, rekonstruksi validitas instrumen perlu diarahkan pada integrasi antara kekuatan metodologi klasik dengan perkembangan psikometrika modern dan teknologi digital (Hilalludin et al., 2026).

Berdasarkan sintesis berbagai konsep tersebut, penelitian ini menawarkan sebuah kerangka konseptual yang dinamakan Digital Integrated Validation Framework (DIVF). Framework ini merupakan model rekonstruksi uji validitas instrumen yang mengintegrasikan tahapan konseptual, psikometrik, dan digital ke dalam satu siklus validasi yang berkesinambungan. Berbeda dengan pendekatan konvensional yang hanya menempatkan uji validitas sebagai tahap akhir, DIVF memandang validitas sebagai proses yang berlangsung sejak identifikasi konstruk hingga evaluasi berkelanjutan setelah instrumen digunakan (Maska et al., 2026).

Model DIVF terdiri atas tujuh tahapan yang saling berkaitan. Tahap pertama adalah rekonstruksi konstruk, yaitu proses mengidentifikasi dimensi dan indikator berdasarkan teori Pendidikan Agama Islam, hasil penelitian

mutakhir, serta kebutuhan empiris di lapangan. Pada tahap ini, Artificial Intelligence dapat dimanfaatkan untuk membantu peneliti melakukan pemetaan literatur (*literature mapping*), mengidentifikasi kata kunci dominan, dan menemukan hubungan antarkonsep yang mungkin terlewatkan dalam telaah manual. Walaupun demikian, keputusan akhir mengenai penyusunan konstruk tetap berada pada peneliti dan pakar bidang Pendidikan Agama Islam sehingga substansi keilmuan tetap terjaga (Setyawarno et al., 2025).

Tahap kedua adalah validitas isi berbasis kolaborasi pakar. Berbeda dengan praktik yang hanya melibatkan satu atau dua validator, framework ini menekankan keterlibatan pakar dari berbagai disiplin, seperti Pendidikan Agama Islam, evaluasi pendidikan, psikometri, dan teknologi pembelajaran. Penilaian para pakar selanjutnya dianalisis menggunakan indeks kuantitatif, seperti *Content Validity Ratio* (CVR), *Content Validity Index* (CVI), atau koefisien Aiken's V sehingga keputusan mengenai kelayakan setiap indikator memiliki dasar empiris yang lebih kuat dan mengurangi subjektivitas (Maska et al., 2026).

Tahap ketiga adalah uji coba digital (*digital pilot testing*). Pemanfaatan platform survei daring memungkinkan peneliti memperoleh data yang lebih beragam dalam waktu yang relatif singkat. Selain menghasilkan jawaban responden, platform digital juga menyediakan informasi tambahan, seperti durasi pengisian, pola respons, tingkat kelengkapan jawaban, dan konsistensi pengisian. Informasi tersebut menjadi sumber bukti tambahan yang dapat digunakan untuk mengevaluasi kualitas instrumen sebelum diterapkan pada penelitian utama (Maska et al., 2026).

Tahap keempat adalah analisis psikometrik modern. Pada tahap ini, kualitas setiap butir dianalisis menggunakan *Rasch Measurement Theory* untuk mengidentifikasi tingkat kesulitan butir, kemampuan responden, kesesuaian model (*fit statistics*), serta kemungkinan adanya bias butir (*Differential Item Functioning*). Selanjutnya, struktur konstruk diuji menggunakan *Confirmatory Factor Analysis* (CFA) sehingga hubungan antara

indikator dan variabel laten dapat diverifikasi secara empiris. Integrasi kedua pendekatan tersebut menghasilkan informasi yang jauh lebih komprehensif dibandingkan analisis korelasi item-total yang umum digunakan dalam pendekatan konvensional.

Tahap kelima adalah validasi berbasis *learning analytics*. Pada fase ini, data interaksi pengguna terhadap instrumen dianalisis untuk memperoleh gambaran mengenai perilaku responden ketika mengerjakan instrumen. Informasi seperti pola pengisian, tingkat konsistensi jawaban, waktu penyelesaian, maupun perubahan respons dapat menjadi indikator tambahan dalam mengevaluasi kualitas instrumen. Pendekatan ini memperluas konsep validitas dari sekadar hubungan statistik menjadi analisis terhadap proses respons (*response process validity*) (Li et al., 2026).

Tahap keenam adalah evaluasi berkelanjutan (*continuous validation*). Validitas instrumen dipandang sebagai proses yang selalu berkembang sesuai dengan perubahan karakteristik responden, konteks penelitian, maupun perkembangan teknologi. Oleh karena itu, setiap penggunaan instrumen menghasilkan data baru yang dapat dimanfaatkan untuk memperbaiki indikator sehingga kualitas instrumen meningkat secara berkesinambungan.

Tabel 2. Rekonstruksi Uji Validitas Instrumen Berbasis Digital Integrated Validation Framework (DIVF)

Tahapan	Pendekatan Konvensional	Rekonstruksi melalui DIVF	Kontribusi terhadap Penelitian PAI
1. Identifikasi konstruk	Kajian teori terbatas	Sintesis teori, hasil riset, dan pemetaan literatur berbantuan AI	Konstruk lebih komprehensif dan kontekstual
2. Validitas isi	Expert judgment sederhana	Kolaborasi multidisiplin dengan CVR, CVI, dan Aiken's V	Mengurangi subjektivitas penilaian
3. Uji coba instrumen	Uji coba manual	Digital pilot testing berbasis platform daring	Data lebih luas dan efisien
4. Analisis empiris	Korelasi Product Moment	Rasch Measurement Theory dan CFA	Menghasilkan bukti validitas yang lebih kuat

5. Analisis proses respons	Tidak dilakukan	Learning analytics dan analisis perilaku responden	Memahami kualitas instrumen secara lebih mendalam
6. Evaluasi	Sekali pada awal penelitian	Continuous validation berbasis data	Instrumen selalu diperbarui sesuai perkembangan konteks
7. Luaran	Instrumen dinyatakan valid	Instrumen adaptif, akurat, reliabel, dan berkelanjutan	Mendukung peningkatan mutu penelitian PAI

Rekonstruksi melalui DIVF menunjukkan bahwa validitas instrumen bukan lagi dipahami sebagai kegiatan verifikasi statistik semata, tetapi sebagai proses ilmiah yang mengintegrasikan bukti konseptual, psikometrik, dan digital secara berkelanjutan. Dalam konteks penelitian Pendidikan Agama Islam, model ini memberikan peluang untuk menghasilkan instrumen yang lebih sensitif dalam mengukur konstruk-konstruk kompleks seperti religiusitas, akhlak, spiritualitas, karakter Islami, moderasi beragama, maupun budaya religius. Dengan demikian, hasil penelitian tidak hanya memiliki kekuatan metodologis yang lebih tinggi, tetapi juga memberikan dasar yang lebih kuat bagi pengembangan teori dan praktik Pendidikan Agama Islam di era transformasi digital.

Implikasi Rekonstruksi Uji Validitas Instrumen terhadap Pengembangan Penelitian Pendidikan Agama Islam

Rekonstruksi uji validitas instrumen melalui Digital Integrated Validation Framework (DIVF) tidak hanya menawarkan perubahan pada aspek teknis pengujian instrumen, tetapi juga menghadirkan perubahan paradigma dalam penelitian Pendidikan Agama Islam (PAI). Selama ini, validitas instrumen lebih banyak dipahami sebagai tahap verifikasi statistik sebelum penelitian dilaksanakan. Paradigma tersebut mengakibatkan pengembangan instrumen cenderung berorientasi pada pemenuhan persyaratan metodologis tanpa memberikan perhatian yang memadai terhadap kesesuaian antara konstruk teoretis, karakteristik responden, serta dinamika perkembangan teknologi digital. Melalui DIVF, validitas diposisikan sebagai proses ilmiah yang berlangsung secara berkelanjutan mulai dari

perumusan konstruk, penyusunan indikator, pengujian psikometrik, hingga evaluasi setelah instrumen diterapkan. Pergeseran paradigma ini menjadikan validitas bukan sekadar ukuran kualitas instrumen, melainkan mekanisme untuk menjamin ketepatan interpretasi hasil penelitian (Kraus et al., 2024).

Dalam konteks Pendidikan Agama Islam, perubahan paradigma tersebut memiliki implikasi yang sangat signifikan karena sebagian besar objek penelitian berupa konstruk laten yang tidak dapat diamati secara langsung. Variabel seperti religiusitas, akhlak, spiritualitas, karakter Islami, moderasi beragama, integritas, maupun budaya religius sekolah merupakan konsep multidimensional yang dibentuk oleh interaksi antara aspek kognitif, afektif, dan perilaku. Oleh karena itu, kualitas pengukuran sangat bergantung pada kemampuan instrumen dalam merepresentasikan setiap dimensi konstruk secara utuh. Rekonstruksi validitas melalui DIVF memungkinkan peneliti memperoleh bukti validitas yang lebih komprehensif sehingga interpretasi terhadap skor responden menjadi lebih akurat dan memiliki dasar ilmiah yang lebih kuat (Krägeloh et al., 2025).

Implikasi berikutnya terlihat pada proses pengembangan instrumen penelitian. Dalam paradigma konvensional, penyusunan indikator umumnya dilakukan melalui kajian teori dan pengalaman peneliti, kemudian divalidasi oleh sejumlah pakar sebelum diuji secara statistik. Pendekatan tersebut sering kali menghasilkan indikator yang belum sepenuhnya mencerminkan perkembangan fenomena keagamaan di masyarakat digital (Hilalludin et al., 2026). Melalui DIVF, penyusunan indikator dilakukan dengan mengintegrasikan kajian teori, hasil penelitian mutakhir, analisis kebutuhan lapangan, serta pemanfaatan Artificial Intelligence (AI) untuk memetakan konsep-konsep yang berkembang dalam literatur ilmiah. Dengan demikian, indikator yang dihasilkan tidak hanya memiliki dasar teoritis yang kuat, tetapi juga lebih adaptif terhadap perubahan konteks sosial dan perkembangan teknologi (Bond & Fox, 2015).

Pada tahap pengumpulan data, implementasi DIVF mendorong pemanfaatan platform asesmen digital yang mampu menghasilkan data lebih kaya dibandingkan instrumen konvensional. Selain memperoleh jawaban responden, peneliti dapat menganalisis durasi pengisian, pola respons, tingkat konsistensi jawaban, serta kemungkinan terjadinya *response bias*. Informasi tersebut menjadi bukti tambahan yang memperkuat interpretasi terhadap kualitas instrumen. Pendekatan ini menunjukkan bahwa validitas tidak lagi bergantung pada satu jenis analisis statistik, melainkan dibangun melalui integrasi berbagai sumber bukti (*multiple sources of validity evidence*).

Dari sisi analisis data, penggunaan *Rasch Measurement Theory* dan *Confirmatory Factor Analysis* (CFA) memberikan kontribusi penting dalam meningkatkan kualitas penelitian Pendidikan Agama Islam. Rasch memungkinkan peneliti mengidentifikasi karakteristik setiap butir secara lebih rinci, termasuk tingkat kesulitan, tingkat kecocokan model, serta kemungkinan adanya bias berdasarkan karakteristik responden. Sementara itu, CFA memberikan bukti empiris mengenai kesesuaian struktur konstruk dengan model teoretis yang telah dirumuskan sebelumnya. Kombinasi kedua pendekatan tersebut menghasilkan proses validasi yang lebih komprehensif dibandingkan penggunaan korelasi item-total semata, sehingga kualitas instrumen menjadi lebih dapat dipertanggungjawabkan secara ilmiah.

Implementasi DIVF juga berimplikasi terhadap peningkatan mutu publikasi ilmiah. Dalam beberapa tahun terakhir, jurnal nasional terakreditasi maupun jurnal internasional bereputasi semakin menekankan pentingnya kualitas instrumen penelitian sebagai salah satu indikator kredibilitas hasil penelitian. Artikel yang hanya melaporkan nilai validitas berdasarkan korelasi sederhana sering kali dinilai belum memberikan bukti validitas yang memadai, terutama ketika mengukur konstruk yang kompleks. Sebaliknya, penggunaan pendekatan validitas yang mengintegrasikan bukti konseptual, psikometrik, dan digital menunjukkan bahwa proses pengembangan instrumen dilakukan secara sistematis dan mengikuti perkembangan

metodologi penelitian kontemporer. Dengan demikian, penerapan DIVF berpotensi meningkatkan kualitas metodologis artikel sehingga lebih kompetitif untuk dipublikasikan pada jurnal bereputasi (Hilalludin & Rouzi, 2026).

Selain berdampak pada kualitas penelitian, DIVF juga memiliki implikasi praktis bagi pengembangan kebijakan pendidikan. Instrumen yang memiliki validitas tinggi akan menghasilkan data yang lebih akurat sebagai dasar pengambilan keputusan. Dalam konteks Pendidikan Agama Islam, data yang diperoleh melalui instrumen yang valid dapat dimanfaatkan untuk mengevaluasi efektivitas pembelajaran, mengukur keberhasilan program penguatan karakter, merancang kebijakan moderasi beragama, hingga menyusun strategi peningkatan kompetensi guru PAI. Dengan demikian, rekonstruksi validitas instrumen tidak hanya memberikan manfaat akademik, tetapi juga berkontribusi terhadap peningkatan mutu pengelolaan pendidikan berbasis bukti (*evidence-based education*).

Tabel 3. Implikasi Implementasi Digital Integrated Validation Framework (DIVF) terhadap Penelitian Pendidikan Agama Islam

Aspek Penelitian	Praktik Konvensional	Implementasi DIVF	Dampak yang Diharapkan
Pengembangan konstruk	Berdasarkan kajian teori terbatas	Integrasi teori, hasil riset, kebutuhan lapangan, dan AI	Konstruk lebih komprehensif dan relevan
Penyusunan indikator	Berdasarkan interpretasi peneliti	Berbasis pemetaan konsep dan validasi multidisiplin	Indikator lebih representatif
Validasi isi	Expert judgment sederhana	CVR, CVI, Aiken's V, dan kolaborasi pakar	Validitas isi lebih objektif
Pengumpulan data	Angket konvensional	Platform asesmen digital	Data lebih luas, cepat, dan terdokumentasi
Analisis instrumen	Korelasi Product Moment	Rasch Measurement Theory dan CFA	Bukti validitas lebih komprehensif
Evaluasi instrumen	Dilakukan satu kali	Continuous validation berbasis data	Instrumen selalu diperbarui sesuai konteks
Luaran penelitian	Validitas administratif	Validitas konseptual, empiris, dan digital	Kualitas penelitian meningkat

Berdasarkan keseluruhan pembahasan, rekonstruksi uji validitas instrumen melalui Digital Integrated Validation Framework (DIVF) menjadi jawaban atas permasalahan yang dikemukakan pada bagian pendahuluan, yaitu masih dominannya paradigma validitas konvensional dalam penelitian Pendidikan Agama Islam. Framework ini memperluas makna validitas dari sekadar pengujian statistik menuju proses ilmiah yang mengintegrasikan bukti konseptual, psikometrik, dan teknologi digital secara berkelanjutan. Dengan pendekatan tersebut, instrumen penelitian tidak hanya memenuhi persyaratan metodologis, tetapi juga memiliki kemampuan yang lebih baik dalam merepresentasikan konstruk keagamaan yang kompleks, sehingga mampu menghasilkan data yang lebih akurat, interpretasi yang lebih valid, serta rekomendasi yang lebih kredibel bagi pengembangan ilmu dan praktik Pendidikan Agama Islam pada era transformasi digital.

KESIMPULAN

Artikel ini menegaskan bahwa tantangan utama uji validitas instrumen dalam penelitian Pendidikan Agama Islam bukan semata-mata terletak pada pemilihan teknik analisis statistik, melainkan pada perlunya perubahan paradigma dalam memaknai validitas sebagai proses ilmiah yang komprehensif. Kompleksitas konstruk keagamaan yang bersifat multidimensional menuntut pendekatan validasi yang tidak hanya mampu menguji kualitas butir instrumen, tetapi juga menjamin ketepatan representasi konsep, akurasi interpretasi data, dan relevansi instrumen terhadap dinamika transformasi digital. Oleh karena itu, pengembangan metodologi penelitian Pendidikan Agama Islam perlu bergeser dari pendekatan yang bersifat prosedural menuju pendekatan yang integratif, adaptif, dan berbasis bukti (evidence-based).

Kontribusi utama artikel ini terletak pada pengajuan Digital Integrated Validation Framework (DIVF) sebagai kerangka konseptual untuk merekonstruksi uji validitas instrumen dalam penelitian Pendidikan Agama Islam. Kerangka ini menawarkan integrasi antara validitas isi, validitas

konstruk, validitas empiris, dan pemanfaatan teknologi digital sebagai satu kesatuan proses yang berkelanjutan. Ke depan, model ini perlu diuji secara empiris pada berbagai konteks penelitian Pendidikan Agama Islam guna memperoleh bukti implementatif mengenai efektivitasnya dalam meningkatkan kualitas instrumen dan kredibilitas hasil penelitian. Dengan demikian, rekonstruksi validitas instrumen tidak hanya memperkaya pengembangan metodologi penelitian, tetapi juga memperkuat fondasi keilmuan Pendidikan Agama Islam yang responsif terhadap perkembangan teknologi dan kebutuhan masyarakat di era transformasi digital.

DAFTAR PUSTAKA

- Asdar, A. K., Afriadi, B., & Muawanah. (2022). Standardisasi instrumen integritas akademik melalui pendekatan *Confirmatory Factor Analysis* dan Rasch Model. *Jurnal Manajemen Pendidikan*, 13(2). <https://doi.org/10.21009/jmp.v13i2.31446>
- Bond, T. G., & Fox, C. M. (2015). *Applying the Rasch model: Fundamental measurement in the human sciences* (3rd ed.). Routledge.
- Chang, H.-H. (2024). Harnessing AI for educational measurement: Standards and emerging frontiers. *Educational Measurement: Issues and Practice*. <https://doi.org/10.3102/10769986241264033>
- Creswell, J. W., & Creswell, J. D. (2023). *Research design: Qualitative, quantitative, and mixed methods approaches* (6th ed.). Sage Publications.
- DeVellis, R. F., & Thorpe, C. T. (2021). *Scale development: Theory and applications* (5th ed.). Sage Publications.
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2022). *Multivariate data analysis* (9th ed.). Cengage Learning.
- Hilalludin, H. (2026). Internalisasi Nilai Akhlak Dalam Pendidikan Agama Islam Berwawasan Multikultural Di SDN Butuh 02 Semarang. *Ar-Ruhul Ilmi: Jurnal Pendidikan dan Pemikiran Islam*, 2(02), 154-167.
- Hilalludin, H. (2026). Strategi Sekolah Dasar Dalam Menerapkan Pendidikan Multikultural Melalui Kegiatan Pembelajaran Dan Ekstrakurikuler. *SciNusa: Jurnal Ilmu Sosial dan Humaniora*, 2(03), 245-259.
- Hilalludin, H. (2026). Transformasi Digital Dalam Perbankan Syariah: Analisis Implementasi Fintech Berbasis Prinsip Syariah Di Era Society 5.0. *Al-Hilali: Jurnal Perbankan dan Ekonomi Islam*, 2(03), 242-256.
- Hilalludin, H., & Rouzi, K. S. (2026). Membangun Jiwa Tangguh Dan Hati Sabar: Integrasi Konsep Hardiness Dan Sabar Dalam Pembentukan Karakter Islami. *Imanu: Jurnal Hukum dan Peradaban Islam*, 2(01), 104-117.

- Hilalludin, H., Al Nomani, S., & Sulhi, M. (2026). DEVELOPING ISLAMIC RELIGIOUS EDUCATION STRATEGIES TO ENHANCE STUDENTS'LEARNING MOTIVATION IN THE DIGITAL ERA. *INJIES: Journal of Islamic Education Studies*, 3(1), 107-122.
- Kline, R. B. (2023). *Principles and practice of structural equation modeling* (5th ed.). Guilford Press.
- Krägeloh, C. U., Melekhov, V., Alyami, M. M., & Medvedev, O. N. (2025). Artificial Intelligence Attitudes Inventory (AIAI): Development and validation using Rasch methodology. *Current Psychology*, 44. <https://doi.org/10.1007/s12144-025-08009-1>
- Kraus, E. B., Wild, J., & Hilbert, S. (2024). Using interpretable machine learning for differential item functioning detection in psychometric tests. *Applied Psychological Measurement*, 48. <https://doi.org/10.1177/01466216241238744>
- Li, P., Chong, Y.-Y., Man, W., & Li, H. (2026). Measuring attitudes toward artificial intelligence: Validation of the Chinese ATTARI-12 and invariance across genders and academic disciplines. *Frontiers in Psychology*. <https://doi.org/10.3389/fpsyg.2026.1836060>
- Maska, L., Vlachopanou, P., Kalamaras, D., & Tsameti, A. (2026). Validating the AIM-N: An AI-motivation and needs scale with multi-group invariance and MIMIC-DIF evidence in higher education. *PLOS ONE*, 21(3). <https://doi.org/10.1371/journal.pone.0341134>
- Sari, H. P., Arisanti, D., Ilhami, A., Salman, N. F., & Faizah, U. (2025). Validity and reliability of the Artificial Intelligence literacy instrument for prospective Islamic education teachers: A Rasch analysis. *POTENSIA: Jurnal Kependidikan Islam*.
- Setyawarno, D., Maryati, & Natadiwijaya, I. F. (2025). Promoting a valid question model for measuring computational thinking skills based on *Confirmatory Factor Analysis* and Rasch model. *Cogent Education*. <https://doi.org/10.1080/2331186X.2025.2505339>
- Speer, A. M., Oswald, F. L., & Putka, D. J. (2025). Reliability evidence for AI-based scores in organizational contexts: Applying lessons learned from psychometrics. *Organizational Research Methods*. <https://doi.org/10.1177/10944281251346404>
- Terry, J., Strait, G., Alsarraf, S., et al. (2025). Artificial intelligence in scale development: Evaluating AI-generated survey items against gold standard measures. *Current Psychology*. <https://doi.org/10.1007/s12144-025-08240-w>