



REKONSEPTULISASI VALIDITAS PENILAIAN PAI MELALUI PERSPEKTIF *CONSTRUCTIVE ALIGNMENT* PADA KURIKULUM MERDEKA

Akhsani Ulvatun Ni'mah¹, Moh. Sahlan²

¹ Magister Pendidikan Agama Islam, Universitas Al-Falah As-Sunniah, Indonesia

² Magister Pendidikan Agama Islam, Universitas Al-Falah As-Sunniah, Indonesia

e-mail: ¹akhsaniulva@uas.ac.id , ²mohsahlan1963@gmail.com

ABSTRACT

In the field of Islamic Religious Education (PAI), discussions of assessment validity have often focused narrowly on the technical quality of assessment instruments, with limited attention paid to how these tools integrate with teaching practices. This study aims to reframe the understanding of assessment validity by applying the concept of Constructive Alignment within the context of the Merdeka Curriculum. Employing a qualitative conceptual approach, the research engaged in a careful analysis of theoretical frameworks, policy documents, and current academic literature, using document analysis and thematic synthesis as primary methods. The findings highlight four key points. First, the prevailing view of assessment validity in PAI remains centered on the attributes of assessment instruments, emphasizing their psychometric properties. Second, this approach falls short in addressing the multifaceted nature of learning outcomes in PAI, particularly those involving cognitive, affective, and psychomotor domains. Third, Constructive Alignment provides a more comprehensive perspective by conceptualizing validity as the alignment among intended learning outcomes, instructional activities, and assessment processes. Finally, this research calls for a shift from focusing solely on the technical validity of assessment instruments to prioritizing alignment validity, where the quality of assessment is judged not only by the accuracy of the tools but also by the coherence of the entire educational experience. This conceptual framework offers a new theoretical foundation for developing more integrated and holistic assessment practices in Islamic Religious Education.

Keywords: *Assessment Validity; Constructive Alignment; Islamic Religious Education; Merdeka Curriculum; Alignment Validity.*

ABSTRAK

Dalam ranah Pendidikan Agama Islam (PAI), pembahasan mengenai validitas asesmen kerap kali terfokus secara sempit pada kualitas teknis instrumen asesmen, dengan perhatian yang terbatas terhadap integrasi instrumen tersebut dalam praktik pembelajaran. Studi ini bertujuan untuk merekonstruksi pemahaman tentang validitas asesmen dengan menerapkan konsep Constructive Alignment dalam konteks Kurikulum Merdeka. Dengan pendekatan konseptual kualitatif, penelitian ini melakukan analisis mendalam terhadap model konseptual, dokumen kebijakan, serta literatur akademik terkini melalui metode analisis dokumen dan sintesis tematik. Temuan penelitian ini menyoroti empat poin utama. Pertama, dominasi pemahaman validitas asesmen dalam PAI masih berpusat pada karakteristik instrumen asesmen, khususnya aspek psikometrik. Kedua, paradigma ini dinilai belum mampu menjawab kompleksitas capaian pembelajaran PAI, terutama pada ranah kognitif, afektif, dan psikomotorik. Ketiga, Constructive Alignment menawarkan perspektif yang lebih komprehensif dengan memaknai validitas sebagai keselarasan antara tujuan pembelajaran, aktivitas pembelajaran, dan proses asesmen. Terakhir, penelitian ini mengusulkan pergeseran fokus dari validitas teknis instrumen asesmen menuju validitas keselarasan, di mana kualitas asesmen dinilai tidak hanya dari ketepatan instrumen, melainkan juga dari koherensi keseluruhan pengalaman pembelajaran. Kerangka konseptual ini memberikan landasan baru bagi pengembangan praktik asesmen yang lebih terintegrasi dan holistik dalam Pendidikan Agama Islam.

Kata Kunci: *Validitas Asesmen; Konstruktif Alignment; Pendidikan Agama Islam; Kurikulum Merdeka; Capaian Pembelajaran*

PENDAHULUAN

Salah satu persoalan mendasar dalam implementasi Kurikulum Merdeka pada mata pelajaran Pendidikan Agama Islam (PAI) adalah belum optimalnya validitas penilaian yang digunakan untuk mengukur ketercapaian kompetensi peserta didik secara autentik dan bermakna (Nanda Hanafi & Bashith, 2025; Hafiz et al., 2024). Asesmen tidak lagi dipahami hanya sebagai mekanisme untuk menentukan capaian akhir peserta didik (Sulastri, 2025), tetapi juga berfungsi sebagai sarana untuk memberikan umpan balik (*feedback*) (Lestari et al., 2025), memantau perkembangan belajar (Havista et al., 2025), dan mendukung pengambilan keputusan pedagogis yang berkelanjutan (Hanafi et al., 2025). Dalam konteks Indonesia, Kurikulum Merdeka semakin memperkuat fungsi asesmen sebagai instrumen yang mendukung pembelajaran berdiferensiasi, refleksi belajar, dan pencapaian kompetensi secara holistik (Sangadah et al., 2025). Penilaian PAI tidak hanya berorientasi pada penguasaan aspek kognitif berupa pengetahuan keagamaan, tetapi juga mencakup dimensi afektif dan psikomotorik yang berkaitan dengan sikap, akhlak, penghayatan nilai-nilai Islam, serta praktik ibadah peserta didik. Karakteristik tersebut sejalan dengan tujuan pendidikan Islam yang tidak hanya menekankan aspek *knowing*, tetapi juga *being* dan *doing* (Black & Wiliam, 1998).

Meskipun demikian, berbagai penilaian menunjukkan bahwa diskursus validitas penilaian PAI masih didominasi oleh pendekatan psikometrik yang berfokus pada kualitas instrumen. Penelitian yang dilakukan oleh (Mulyani et al., 2024). Menempatkan validitas isi, validitas konstruk, reliabilitas, tingkat kesukaran, dan daya pembeda sebagai indikator utama kualitas instrumen penilaian PAI. Temuan serupa juga ditunjukkan oleh (Alfiyah et al., 2025) yang menegaskan bahwa akurasi dan konsistensi instrumen merupakan dasar utama dalam menentukan kualitas evaluasi pembelajaran PAI. Selain itu, penelitian (Boini et al., 2024) Mengenai analisis instrumen tes PAI, masih menggunakan kerangka teori tes klasik yang berorientasi pada pengujian karakteristik butir soal. Berbagai temuan tersebut dipahami sebagai karakteristik instrumen sehingga perhatian terhadap aspek lain yang memengaruhi kualitas penilaian relatif terbatas.

Dominasi paradigma validitas instrumen telah memberikan kontribusi penting dalam menjamin kualitas teknis pengukuran. Namun, instrumen yang dinyatakan valid dan reliabel secara statistik belum tentu mampu merepresentasikan kompetensi yang benar-benar menjadi tujuan pembelajaran. Ketika kemampuan praktik salat dinilai hanya melalui tes pilihan ganda, asesmen yang dihasilkan tidak sepenuhnya mencerminkan kemampuan peserta didik dalam melaksanakan salat secara benar. Hal tersebut berpotensi menyebabkan ketidaksesuaian antara tujuan pembelajaran, aktivitas pembelajaran, dan bentuk asesmen yang digunakan. John Biggs, dengan Teori *Constructive Alignment*, menekankan bahwa keselarasan antara tujuan, proses, dan asesmen menentukan hasil akhir pembelajaran (J. Biggs & Tang, 2011). Suatu asesmen dianggap valid apabila secara konsisten mengukur capaian yang ditargetkan dan selaras dengan pengalaman belajar yang dialami peserta didik.

Argumen tersebut memperoleh dukungan dari perkembangan literatur asesmen modern yang mulai menempatkan alignment sebagai salah satu sumber bukti validitas. Cizek, Kosh, dan Toutkooushian menegaskan bahwa keselarasan antara tujuan pembelajaran dan asesmen merupakan komponen penting dalam evaluasi validitas suatu sistem penilaian pendidikan (Cizek et al., 2018).

Meskipun *Constructive Alignment* telah banyak digunakan dalam kajian desain kurikulum dan pembelajaran, penggunaannya sebagai perspektif untuk mengkaji validitas penilaian dalam Pendidikan Agama Islam masih relatif terbatas. Sebagian besar penelitian PAI lebih berfokus pada pengembangan instrumen, pengukuran reliabilitas, validitas isi, serta analisis kualitas butir soal. Sementara itu, kajian yang menempatkan validitas sebagai persoalan keselarasan antara tujuan pembelajaran, aktivitas pembelajaran, dan asesmen dalam konteks Kurikulum Merdeka masih jarang ditemukan.

Berdasarkan kondisi tersebut, penelitian ini menawarkan kontribusi konseptual berupa rekonseptualisasi validitas penilaian PAI melalui perspektif *Constructive Alignment*. Penelitian ini tidak bertujuan mengembangkan instrumen penilaian baru maupun menguji efektivitas model pembelajaran tertentu, melainkan memperluas cara pandang terhadap validitas penilaian yang selama ini lebih banyak dipahami sebagai kualitas instrumen. Melalui perspektif *Constructive Alignment*, validitas diposisikan

sebagai keselarasan antara tujuan pembelajaran, aktivitas pembelajaran, dan asesmen sehingga penilaian dapat lebih merepresentasikan tujuan pendidikan Islam secara utuh. Oleh karena itu, penelitian ini bertujuan untuk: (1) menganalisis keterbatasan paradigma validitas instrumen dalam penilaian PAI; (2) mengkaji *Constructive Alignment* sebagai perspektif evaluatif dalam memahami validitas penilaian; dan (3) merumuskan rekonseptualisasi validitas penilaian PAI pada implementasi Kurikulum Merdeka.

METODE

Penelitian ini menggunakan pendekatan kualitatif dengan desain *conceptual paper* berbasis studi kepustakaan (*library research*). Desain ini dipilih untuk mengkaji secara kritis pengembangan konsep validitas penilaian dalam Pendidikan Agama Islam (PAI) serta menyusun rekonseptualisasi validitas melalui perspektif *Constructive Alignment*. Pembahasan difokuskan pada pengembangan argumentasi konseptual berdasarkan sintesis berbagai literatur ilmiah yang relevan (Jaakkola, 2020).

Data penelitian berupa artikel jurnal, buku ilmiah, dan dokumentasi kebijakan pendidikan yang membahas validitas penilaian, *Constructive Alignment*, asesmen PAI, dan Kurikulum Merdeka. Literatur diperoleh melalui penelusuran basis data akademis seperti Scopus, ERIC, Google Scholar, dan Crossref, kemudian dipilih secara purposif berdasarkan relevansi topik, kredibilitas sumber, serta kontribusinya terhadap fokus penelitian.

Analisis data dilakukan melalui *content analysis* untuk mengidentifikasi konsep, pola pemikiran dan kecenderungan paradigma validitas penilaian dalam berbagai literatur (Krippendorff, 2018). Selanjutnya, hasil analisis disintesis menggunakan *conceptual synthesis* dengan mengintegrasikan teori validitas penilaian, *Constructive Alignment*, dan karakteristik asesmen PAI. Proses ini bertujuan untuk menemukan keterbatasan paradigma validitas instrumen sekaligus merumuskan hubungan konseptual antara capaian pembelajaran, pengalaman belajar, dan asesmen (J. Biggs & Tang, 2011).

Keabsahan argumentasi konseptual dijaga melalui proses triangulasi sumber, yaitu dengan membandingkan teori, hasil penelitian terdahulu, dan dokumen kebijakan yang relevan. Sintesis dilakukan secara terukur untuk memastikan bahwa kerangka konseptual yang dihasilkan dibangun atas keterkaitan antarkonsep yang konsisten dan dapat dipertanggungjawabkan.

HASIL DAN PEMBAHASAN

Hasil

1. Dominasi Paradigma Validitas Instrumen dalam Penilaian PAI

Hasil sintesis literatur menunjukkan bahwa pemahaman mengenai validitas penilaian dalam PAI masih didominasi oleh paradigma yang berorientasi pada instrumen. Kualitas penilaian umumnya ditentukan melalui indikator psikometrik, seperti validitas isi, validitas konstruk, reliabilitas, tingkat kesukaran, dan daya pembeda. Fokus tersebut menempatkan instrumen sebagai komponen utama dalam menentukan mutu asesmen.

Analisis juga menunjukkan bahwa pengembangan penilaian PAI mayoritas diarahkan pada penyempurnaan karakteristik teknis instrumen agar menghasilkan pengukuran (*measurement*), sehingga keberhasilan asesmen lebih banyak diukur berdasarkan kualitas instrumen yang digunakan.

Secara keseluruhan, hasil penelitian menunjukkan bahwa validitas penilaian PAI masih diposisikan sebagai karakteristik teknis instrumen. Temuan ini menjadi dasar untuk menelaah lebih lanjut sejauh mana paradigma tersebut mampu merepresentasikan capaian pembelajaran PAI secara utuh.

2. Keterbatasan Paradigma Validitas Instrumen dalam Mengukur Capaian PAI

Hasil analisis menunjukkan bahwa paradigma validitas instrumen belum sepenuhnya mampu merepresentasikan karakteristik capaian pembelajaran PAI yang bersifat multidimensional. Kompetensi PAI tidak hanya mencakup aspek kognitif, tetapi juga meliputi dimensi afektif dan psikomotorik yang diwujudkan melalui sikap, nilai, serta praktik keagamaan.

Sintesis literatur menunjukkan adanya ketidaksesuaian antara kompetensi yang menjadi tujuan pembelajaran dan bentuk asesmen yang digunakan. Dalam beberapa konteks, kompetensi praktik dan pembentukan sikap masih dinilai menggunakan instrumen yang berorientasi pada penguasaan pengetahuan. Kondisi tersebut menunjukkan adanya fenomena *misalignment* antara tujuan pembelajaran dan asesmen.

Temuan penelitian juga menunjukkan bahwa instrumen yang memenuhi persyaratan psikometrik belum tentu mampu menggambarkan ketercapaian seluruh kompetensi PAI. Oleh karena itu, kualitas teknis instrumen belum sepenuhnya mencerminkan kualitas representasi terhadap pembelajaran yang ingin dicapai.

3. *Constructive Alignment* sebagai Perspektif Evaluatif dalam Memahami Validitas Penilaian

Hasil sintesis menunjukkan bahwa validitas penilaian lebih menyeluruh apabila dipahami melalui hubungan antara capaian pembelajaran, aktivitas pembelajaran, dan asesmen sebagai satu kesatuan yang saling berkaitan. Ketiga komponen tersebut membentuk sistem yang menentukan kualitas informasi yang dihasilkan melalui proses penilaian.

Analisis juga menunjukkan bahwa keterkaitan antara tujuan pembelajaran, pengalaman belajar, dan asesmen memberikan dasar yang lebih utuh dalam menjabarkan ketercapaian kompetensi peserta didik. Dalam kerangka tersebut, asesmen tidak berdiri sebagai aktivitas yang terpisah, melainkan menjadi bagian dari keseluruhan proses pembelajaran.

Berdasarkan temuan yang dihasilkan, penelitian ini menempatkan *Constructive Alignment* sebagai perspektif evaluatif yang memperluas cara pandang terhadap validitas penilaian. Perspektif ini menjadi landasan untuk mengembangkan pemahaman validitas yang tidak hanya berorientasi pada instrumen, tetapi juga pada keterpaduan seluruh komponen pembelajaran.

4. Rekonseptualisasi Validitas Penilaian PAI Berbasis *Constructive Alignment*

Temuan menunjukkan bahwa konstruk validitas mengenai penilaian Pendidikan Agama Islam ditetapkan melalui interaksi antara prestasi belajar, pembelajaran pengalaman, dan ukuran evaluatif. Ketiga aspek ini bersama-sama menciptakan kerangka kerja sinergis yang berhasil merangkum realisasi kompetensi pelajar.

Sintesis konseptual memuncak dalam identifikasi tiga dimensi utama yang berkaitan dengan validitas penilaian berbasis penyelarasan konstruktif, khususnya penyelarasan hasil, penyelarasan pengalaman belajar, dan penyelarasan penilaian. Masing-masing dimensi ini menjelaskan koherensi antara kompetensi yang dimaksud, proses pembelajaran yang dilaksanakan, dan penilaian yang diberikan.

Berdasarkan ketiga dimensi ini, penyelidikan menghasilkan rekonseptualisasi validitas penilaian PAI, menyatakan bahwa validitas harus dianggap tidak hanya sebagai atribut instrumen, tetapi juga sebagai ukuran kualitas penyelarasan dalam sistem pendidikan. Kerangka konseptual ini muncul sebagai hasil penting dari penelitian dan berfungsi sebagai dasar untuk wacana mengenai peningkatan validitas penilaian dalam konteks implementasi Kurikulum Independen.

Pembahasan

1. Memahami Dominasi Validitas dalam Tradisi Penilaian PAI

Validitas penilaian dalam Pendidikan Agama Islam (PAI) masih didominasi oleh paradigma yang menitikberatkan pada aspek teknis instrumen. Dalam kerangka ini, kualitas penilaian sering diukur berdasarkan kemampuan instrumen menghasilkan data yang akurat, konsisten, dan kredibel (NurAzizi et al., 2025). Pemahaman validitas cenderung dipersempit sebagai karakteristik alat ukur, bukan sebagai bagian integral dari sistem pembelajaran yang lebih luas (Ulfa & Pashayev, 2025). Akibatnya, penilaian dianggap berkualitas apabila instrumen yang digunakan memenuhi standar pengukuran tertentu, meskipun keterkaitan instrumen dengan proses pembelajaran sering kali belum menjadi perhatian utama.

Kondisi ini berkaitan erat dengan sejarah perkembangan penilaian modern yang berakar pada tradisi psikometrik. Sejak awal abad ke-20, penilaian berkembang dalam kerangka ilmu pengukuran pendidikan yang berorientasi pada penyediaan data objektif mengenai kemampuan peserta didik (Brennan, 2006). Dalam tradisi ini, keberhasilan penilaian diukur dari kemampuan instrumen menghasilkan skor yang stabil dan dapat dipertanggungjawabkan secara ilmiah (Schellekens et al., 2023). Implikasi pendekatan ini adalah penilaian lebih sering diposisikan sebagai aktivitas pengukuran semata, bukan sebagai bagian integral dari proses pembelajaran yang bersifat manusiawi. Tradisi ini membentuk persepsi bahwa validitas identik dengan kualitas teknis instrumen yang digunakan dalam evaluasi.

Pandangan ini semakin diperkuat oleh perkembangan teori validitas modern, yang menempatkan validitas sebagai landasan utama dalam menafsirkan hasil asesmen. Messick (Messick, 1995), misalnya, memandang validitas sebagai konsep holistik yang berkaitan dengan ketepatan interpretasi dan penggunaan hasil tes. Kane menekankan pentingnya argumen yang mendasari inferensi dari skor yang diperoleh. Meskipun pemikiran kedua tokoh tersebut memperluas pemahaman tentang validitas, penerapannya di lapangan masih sering terfokus pada pembuktian kualitas instrumen.

Akibatnya, indikator seperti validitas isi, konstruk, kriteria, dan reliabilitas tetap menjadi acuan utama dalam penelitian dan praktik penilaian pendidikan.

Dominasi paradigma ini juga tercermin dalam berbagai kajian mengenai penilaian PAI. Sebagian besar penelitian menilai kualitas asesmen berdasarkan validitas isi, konstruk, reliabilitas, tingkat kesukaran, dan daya pembeda (Subhaktiyasa, 2024; Nursichati et al., 2025). Fokus pengembangan umumnya diarahkan pada penciptaan instrumen yang memenuhi standar psikometrik agar data yang dihasilkan akurat dan dapat dipertanggungjawabkan (Sunandar et al., 2024). Bahkan dalam pengembangan model atau bentuk asesmen baru, kualitas instrumen tetap menjadi tolok ukur utama keberhasilan. Temuan-temuan ini mengindikasikan bahwa validitas masih dipandang sebagai atribut utama instrumen penilaian, sehingga dimensi lain yang lebih humanis sering terabaikan.

Dalam konteks PAI, dominasi paradigma instrumen juga dipengaruhi oleh karakteristik sistem evaluasi pendidikan yang berlaku di sekolah (Ridho et al., 2025). Praktik penilaian diarahkan untuk menghasilkan data capaian belajar yang dapat dipertanggungjawabkan secara objektif kepada berbagai pihak (Park, 2025). Tuntutan akuntabilitas ini mendorong guru memilih instrumen yang mudah diukur, dibandingkan, dan dianalisis (Tóth & Csapó, 2022; Hoeve, 2022). Konsekuensinya, pendekatan yang berfokus pada kualitas instrumen menjadi pilihan yang dianggap paling rasional dan efisien. Pandangan ini membentuk persepsi bahwa validitas identik dengan ketepatan alat ukur, tanpa selalu mempertimbangkan dimensi proses dan pengalaman belajar siswa secara menyeluruh (Atchison et al., 2022). Menegaskan bahwa dominasi paradigma validitas instrumen membawa sejumlah keterbatasan. Penekanan berlebihan pada aspek teknis alat ukur berisiko mempersempit makna validitas hanya sebatas akurasi pengukuran (Meijer et al., 2022). Padahal, instrumen yang memenuhi kriteria psikometrik belum tentu mampu merefleksikan seluruh kompetensi yang menjadi tujuan pembelajaran. Ketika perhatian lebih difokuskan pada instrumen, keterkaitan antara tujuan, aktivitas pembelajaran, dan asesmen bisa terabaikan. Akhirnya, validitas lebih dipandang sebagai masalah teknis semata, bukan sebagai bagian dari kualitas pembelajaran yang menyeluruh dan bermakna bagi perkembangan peserta didik.

Dengan demikian, dominasi paradigma validitas instrumen menunjukkan bahwa kualitas penilaian PAI masih banyak dilihat dari sisi teknis alat ukur (Nasrullah, 2025). Meskipun paradigma ini penting untuk menjamin akurasi dan konsistensi hasil pengukuran, orientasi yang terlalu sempit pada instrumen berisiko membatasi makna validitas hanya pada aspek pengukuran. Kondisi ini menegaskan perlunya sudut pandang yang lebih menyeluruh dan humanis dalam memahami validitas penilaian, khususnya dalam konteks PAI yang tidak hanya menargetkan penguasaan pengetahuan, tetapi juga pembentukan sikap dan praktik keagamaan siswa. Permasalahan ini menjadi alasan utama untuk menelaah secara kritis keterbatasan paradigma validitas instrumen dalam menilai capaian pembelajaran PAI secara holistik.

2. Ketika Validitas Instrumen Tidak Lagi Memadai untuk Menilai Kompetensi PAI

Temuan penelitian menunjukkan bahwa paradigma validitas instrumen hanya memberikan penjelasan parsial terhadap pencapaian tujuan pembelajaran multidimensional dalam Pendidikan Agama Islam (PAI) (Rahma, 2025). Keterbatasan ini sangat penting mengingat PAI tidak hanya bertujuan untuk menguasai pengetahuan agama, tetapi juga untuk menumbuhkan sikap religius dan menerapkan nilai-nilai Islam dalam kehidupan sehari-hari (Nurmasihin et al., 2025; Qolbi, 2025). Oleh karena itu, keberhasilan belajar seharusnya tidak diukur hanya dari kemampuan siswa dalam menjawab soal dengan benar, melainkan juga mencerminkan perubahan holistik dalam pengetahuan, sikap, dan perilaku siswa. Dengan demikian, kualitas asesmen perlu dinilai tidak hanya dari ketepatan instrumennya, tetapi juga dari kemampuannya merepresentasikan seluruh kompetensi yang dirumuskan dalam tujuan pembelajaran.

Keterbatasan paradigma validitas instrumen semakin nampak ketika asesmen diarahkan untuk mengukur kompetensi afektif dan psikomotorik (Dewi et al., 2025). Keterampilan kognitif memang sering kali dapat diukur melalui tes tertulis, namun kompetensi afektif—seperti internalisasi nilai, sikap, dan karakter—berkembang secara bertahap dan sulit diukur dengan instrumen semacam itu (Fikri et al., 2025; Rohimah et al., 2025). Demikian pula, kompetensi psikomotorik memerlukan observasi langsung ketika siswa melakukan keterampilan tertentu atau terlibat dalam praktik keagamaan (Sunandar et al., 2024). Dalam konteks ini, asesmen yang secara statistik valid dan reliabel tetap dapat gagal menangkap kompleksitas dan luasnya perkembangan siswa. Akibatnya, hasil asesmen hanya dapat mencerminkan sebagian kecil capaian belajar yang sesungguhnya.

Analisis lebih lanjut menunjukkan adanya ketidakselarasan yang berkelanjutan antara tujuan pembelajaran yang telah ditetapkan dan metode asesmen yang diterapkan dalam praktik (Utari et al., 2024). Sering kali, tujuan yang diharapkan adalah menumbuhkan keterampilan praktis atau membina karakter religius siswa, namun praktik asesmen masih banyak berfokus pada evaluasi pengetahuan melalui ujian tertulis (Irmayanti et al., 2024; Riaz & Ghuffar, 2025). Ketidaksesuaian ini menyebabkan kesenjangan antara kompetensi yang ingin dikembangkan oleh pendidik dan data yang diperoleh dari asesmen. Ketika asesmen hanya mengukur pengetahuan siswa, sementara tujuan menekankan kemampuan dan perilaku yang dapat ditunjukkan secara nyata, validitas asesmen dalam merepresentasikan hasil belajar autentik menjadi berkurang.

Fenomena ini terlihat jelas dalam pembelajaran Pendidikan Agama Islam (PAI). Kompetensi seperti melaksanakan salat, membaca Al-Qur'an, atau menunjukkan perilaku jujur menuntut siswa untuk menunjukkan kinerja nyata, bukan sekadar pemahaman teoretis. Jika penguasaan kompetensi tersebut hanya diukur melalui tes pilihan ganda atau esai, maka hasilnya cenderung hanya merefleksikan pengetahuan siswa, bukan kemampuan mereka dalam menerapkan praktik tersebut di kehidupan nyata. Dalam evaluasi pendidikan, hal ini menggarisbawahi perlunya memperluas pemahaman tentang validitas asesmen. Validitas harus dipandang sebagai

hubungan dinamis antara tujuan pembelajaran, pengalaman belajar, dan praktik asesmen. Jika ketiga elemen ini tidak saling mendukung, asesmen berisiko menghasilkan gambaran yang tidak akurat tentang capaian belajar siswa. Oleh karena itu, kritik terhadap paradigma validitas instrumen bukan untuk mengurangi pentingnya kualitas instrumen, melainkan untuk menunjukkan bahwa kualitas instrumen saja tidak cukup untuk menjelaskan keberhasilan pembelajaran PAI yang holistik.

Sebagai kesimpulan, temuan ini menunjukkan bahwa keterbatasan paradigma validitas instrumen bukan berasal dari konsep validitas itu sendiri, melainkan dari kecenderungan untuk mereduksinya hanya pada aspek teknis alat ukur. Dalam konteks PAI dan Kurikulum Merdeka, validitas perlu dipahami secara lebih luas, yaitu dengan memperhatikan keterkaitan antara kompetensi yang dituju, pengalaman belajar, dan praktik asesmen. Pengakuan atas keterbatasan ini menegaskan perlunya perspektif yang lebih terintegrasi, di mana validitas menjadi bagian yang tak terpisahkan dari desain pembelajaran. Pada bagian berikutnya akan diperkenalkan pendekatan *Constructive Alignment* sebagai kerangka evaluasi yang menawarkan pemahaman validitas asesmen yang lebih holistik dan bernuansa.

3. Memperluas Mekan Validitas Penilaian melalui Perspektif *Constructive Alignment*

Pendidikan Agama Islam (PAI) tidak hanya terletak pada kualitas instrumen, tetapi juga pada pemaknaan validitas itu sendiri. Selama ini, validitas sering dipahami secara sempit sebagai aspek pengukuran, sehingga fokus utama tertuju pada ketepatan instrumen dalam menghasilkan skor objektif (Rachmaningtyas et al., 2025). Namun, penelitian sebelumnya membuktikan bahwa instrumen yang valid secara psikometrik belum tentu merepresentasikan ketercapaian tujuan pembelajaran secara menyeluruh (Schellekens et al., 2023; Hoeve, 2022; Atchison et al., 2022). Temuan ini menegaskan perlunya pemahaman validitas yang lebih komprehensif sebagai bagian integral dari sistem pembelajaran. Dalam hal ini, pendekatan *Constructive Alignment* menawarkan perspektif baru dengan memandang validitas sebagai kualitas hubungan yang harmonis antara tujuan pembelajaran, aktivitas pembelajaran, dan asesmen. Konsep yang dikembangkan oleh John Biggs dan Catherine Tang ini menekankan bahwa pembelajaran bermakna hanya tercapai jika seluruh komponen pembelajaran bergerak selaras menuju tujuan yang sama.

Dalam *Constructive Alignment*, penilaian dimulai dari capaian pembelajaran (*intended learning outcomes*), bukan dari instrumen (J. B. Biggs & Tang, 2015). Artinya, sebelum merancang asesmen, pendidik harus memastikan kompetensi yang ingin dicapai peserta didik. Pendekatan ini berbeda dengan paradigma tradisional yang memulai dari pembuatan instrumen (Ridho et al., 2025). Jika capaian pembelajaran menjadi dasar perencanaan, seluruh aktivitas dan asesmen dapat dirancang secara terarah untuk mendukung pencapaian kompetensi. Dengan demikian, validitas tidak hanya diukur dari kemampuan instrumen mengukur konstruk, tetapi juga dari sejauh mana asesmen menilai kompetensi yang telah dirumuskan sebagai tujuan pembelajaran.

Perubahan paradigma ini sangat relevan untuk PAI, di mana pembelajaran tidak hanya menargetkan pemahaman konsep keislaman, tetapi juga pengalaman menghayati dan mengamalkan nilai-nilai Islam dalam kehidupan sehari-hari. Jika tujuan pembelajaran menekankan pembentukan akhlak, keterampilan ibadah, atau kemampuan membaca Al-Qur'an, maka aktivitas pembelajaran harus memberi ruang bagi peserta didik untuk mengalami, mempraktikkan, dan merefleksikan kompetensi tersebut secara nyata (Sulastri, 2025). Asesmen juga perlu dirancang untuk mengukur kemampuan yang benar-benar diaktualisasikan dalam proses pembelajaran (Ulfa & Pashayev, 2025). Dengan demikian, validitas asesmen sangat bergantung pada kualitas dan relevansi pengalaman belajar yang diberikan.

Konsep *alignment* dapat mengatasi fenomena *misalignment* yang telah diidentifikasi sebelumnya. Jika tujuan pembelajaran menekankan keterampilan praktik, tetapi asesmen hanya mengukur pengetahuan deklaratif, maka terjadi ketidaksesuaian yang memengaruhi validitas (C. R. Brown & Moore, 1994). Sebaliknya, jika tujuan, aktivitas pembelajaran, dan asesmen dirancang secara selaras, data asesmen akan lebih representatif terhadap kompetensi yang diharapkan. Dalam kerangka ini, validitas tidak hanya soal ketepatan pengukuran, tetapi juga sejauh mana asesmen mewakili dan merefleksikan tujuan pembelajaran secara autentik.

Dengan demikian, *Constructive Alignment* memberikan landasan konseptual yang lebih berkelanjutan dalam memahami validitas penilaian PAI. Jika paradigma validitas instrumen berfokus pada optimalisasi alat ukur, *Constructive Alignment* memastikan seluruh sistem pembelajaran berjalan harmonis dan terintegrasi. Temuan ini menjadi tonggak penting, karena menegaskan bahwa perluasan makna validitas memerlukan perubahan paradigma tentang hubungan antara tujuan, aktivitas pembelajaran, dan asesmen, bukan hanya perbaikan instrumen. Oleh karena itu, perlu dirumuskan kerangka baru validitas penilaian PAI yang berorientasi pada keselarasan pembelajaran secara menyeluruh. Kerangka konseptual ini akan dibahas lebih lanjut pada subbab berikutnya tentang rekonseptualisasi validitas penilaian PAI berbasis *Constructive Alignment*.

4. Menuju Kerangka Baru Validitas Penilaian PAI pada Kurikulum Merdeka

Mutu penilaian sangat bergantung pada kemampuan instrumen dalam menghasilkan data yang akurat, konsisten, dan dapat diandalkan (Syamsi et al., 2025; NurAzizi et al., 2025). Temuan ini menunjukkan bahwa validitas lebih dipahami sebagai sifat yang melekat pada alat ukur, bukan bagian integral dari keseluruhan sistem pembelajaran. Dengan kata lain, suatu penilaian dianggap berkualitas apabila instrumen yang digunakan sudah memenuhi standar pengukuran yang telah ditetapkan, meskipun keterkaitannya dengan proses pembelajaran sering kali kurang menjadi perhatian utama.

Situasi ini tidak dapat dipisahkan dari perkembangan historis penilaian modern yang berakar pada tradisi psikometrik. Sejak awal abad ke-20, penilaian berkembang sebagai bagian dari ilmu pengukuran pendidikan, dengan tujuan menghasilkan data yang objektif tentang kemampuan peserta didik (G. T. L. Brown, 2022). Dalam tradisi ini, keberhasilan penilaian ditentukan oleh kemampuan instrumen untuk menghasilkan

skor yang stabil dan dapat diinterpretasikan secara ilmiah (Newton, 2023). Akibatnya, penilaian lebih dipandang sebagai aktivitas pengukuran daripada sebagai bagian integral dari proses pembelajaran. Tradisi ini kemudian membentuk perspektif bahwa validitas terutama merupakan fungsi dari kualitas teknis instrumen yang digunakan dalam proses evaluasi.

Pandangan ini semakin diperkuat oleh perkembangan teori validitas modern yang memposisikan validitas sebagai dasar utama dalam menganalisis hasil penilaian. Messick menjelaskan bahwa validitas merupakan konsep terpadu yang berkaitan dengan ketepatan interpretasi dan penggunaan hasil tes, sedangkan Kane menekankan pentingnya penalaran yang mendukung inferensi yang dihasilkan dari skor yang diperoleh (Messick, 1995). Meskipun kedua tokoh tersebut telah memperluas pemahaman kita tentang validitas, implementasinya dalam praktik pendidikan sering kali masih berfokus pada pembuktian kualitas instrumen.

Beberapa penelitian menunjukkan bahwa kualitas penilaian umumnya diukur menggunakan validitas isi, validitas konstruk, reliabilitas, tingkat kesukaran, dan daya pembeda (Ridho et al., 2025; Chong et al., 2021; Rust, 2023). Penelitian lebih banyak berfokus pada pengembangan instrumen yang memenuhi standar psikometrik untuk menghasilkan data yang akurat. Bahkan ketika penelitian mengembangkan model atau format penilaian baru, pengujian kualitas instrumen tetap menjadi indikator utama keberhasilan penelitian tersebut. Data ini menunjukkan bahwa validitas masih diposisikan sebagai atribut inheren dari instrumen penilaian.

Dominasi paradigma validitas instrumen juga menghadirkan beberapa keterbatasan. Fokus yang berlebihan pada kualitas teknis alat ukur berpotensi mempersempit makna validitas hanya pada aspek ketepatan pengukuran semata. Dalam praktiknya, instrumen yang memenuhi berbagai kriteria psikometrik belum tentu merepresentasikan seluruh kompetensi yang menjadi tujuan pembelajaran. Ketika perhatian lebih banyak diarahkan pada instrumen, hubungan antara tujuan pembelajaran, kegiatan pembelajaran, dan penilaian berisiko terabaikan. Akibatnya, validitas cenderung dipahami sebagai persoalan teknis daripada sebagai bagian dari kualitas pembelajaran secara menyeluruh.

Dengan demikian, dominasi paradigma validitas instrumen menunjukkan bahwa kualitas penilaian PAI masih banyak dipahami dalam kerangka karakteristik teknis alat ukur. Paradigma ini berperan penting dalam memastikan ketepatan dan konsistensi hasil pengukuran sehingga tidak dapat diabaikan dalam praktik evaluasi pembelajaran. Namun, fokus yang terlalu kuat pada instrumen berisiko membatasi pemahaman validitas hanya pada aspek pengukuran semata. Situasi ini menegaskan perlunya perspektif yang lebih komprehensif mengenai validitas penilaian, khususnya dalam konteks Pendidikan Agama Islam (PAI) yang tidak hanya bertujuan pada penguasaan pengetahuan, tetapi juga pembentukan sikap dan perilaku keagamaan. Persoalan inilah yang menjadi dasar untuk mengkaji lebih lanjut keterbatasan paradigma validitas instrumen dalam mengukur hasil belajar PAI secara holistik.

KESIMPULAN

Penelitian ini menyoroti bahwa permasalahan utama validitas penilaian dalam Pendidikan Agama Islam (PAI) tidak hanya bergantung pada kualitas instrumen semata, melainkan juga pada pemahaman komprehensif mengenai validitas dalam keseluruhan sistem pembelajaran. Temuan penelitian mengungkapkan bahwa meskipun sebuah instrumen telah memenuhi standar psikometrik, instrumen tersebut belum tentu mampu mencerminkan capaian pembelajaran secara menyeluruh jika tidak dirancang secara selaras dengan tujuan pembelajaran dan pengalaman belajar peserta didik. Berdasarkan hasil tersebut, penelitian ini mengajukan pemikiran ulang tentang validitas penilaian PAI dengan mengadopsi perspektif *Constructive Alignment*, yang memandang validitas sebagai tingkat keselarasan antara capaian pembelajaran, pengalaman belajar, dan asesmen (*alignment validity*). Dengan demikian, makna validitas diperluas dari sekadar karakteristik instrumen menjadi kualitas sistem pembelajaran secara keseluruhan. Kontribusi utama penelitian ini adalah menghadirkan kerangka konseptual yang dapat digunakan sebagai dasar dalam mengembangkan asesmen PAI yang lebih holistik, autentik, dan terintegrasi dengan implementasi Kurikulum Merdeka. Namun, perlu dicatat bahwa penelitian ini bersifat konseptual sehingga model yang diajukan belum diuji melalui implementasi empiris di kelas. Oleh karena itu, penelitian lanjutan diperlukan untuk mengembangkan instrumen operasional ataupun menguji model *alignment validity* pada berbagai jenjang dan konteks pembelajaran PAI, agar dapat memperoleh bukti empiris terkait efektivitas dan penerapannya dalam praktik pendidikan.

DAFTAR PUSTAKA

- Alfiah, N., Zuhriyah, I. A., & Chalishah, N. D. (2025). Validity and Reliability Analysis: Measuring the Accuracy and Consistency of Instruments for the Evaluation of Islamic Religious Education Tests. *JiIP - Jurnal Ilmiah Ilmu Pendidikan*, 8(3), 2672–2679. <https://doi.org/10.54371/jiip.v8i3.7230>
- Atchison, D., Garet, M. S., Smith, T. M., & Song, M. (2022). The Validity of Measures of Instructional Alignment With State Standards Based on Surveys of Enacted Curriculum. *AERA Open*, 8(1), 1–19. <https://doi.org/10.1177/23328584221098761>
- Black, P., & Wiliam, D. (1998). Assessment and Classroom Learning. *Assessment in Education*, 5(1), 7–74.
- Boini, B., Sahrul, S., Asmanah, A., & Sumin, S. (2024). Analisis Instrumen Tes Penilaian Pendidikan Agama Islam Menggunakan Tes Teori Klasik. *At Turots: Jurnal Pendidikan Islam*, 5(1), 391–396. <https://doi.org/10.51468/jpi.v5i001.445>
- Brown, G. T. L. (2022). The Past, Present and Future of Educational Assessment: A Transdisciplinary Perspective. *Frontiers in Education*, 7. <https://doi.org/10.3389/feduc.2022.1060633>
- Chong, J., Mokshein, S. E., & Mustapha, R. (2021). A Content Validity Study for Vocational Teachers' Assessment Literacy Instrument (VoTAL). *International Journal of Academic Research in Business and Social Sciences*, 11(4), 868–883. <https://doi.org/10.6007/IJARBS/V11-I4/9547>
- Cizek, G. J., Kosh, A. E., & Toutkoushian, E. K. (2018). Gathering and Evaluating Validity Evidence: The Generalized Assessment Alignment Tool. *Journal of Educational Measurement*, 55(4), 477–512. <https://doi.org/10.1111/jedm.12189>
- Dewi, M. N., Masritawati, M., Fatimah, N., & Asma, S. (2025). Integrating Affective–Spiritual Assessment in Education: Instrument Development and Implementation. *El-Rusyd*, 10(2). <https://doi.org/10.58485/elrusyd.v10i2.447>
- Fikri, M., Firdayeni, & Jamilus. (2025). Transformasi Evaluasi Pendidikan Islam dari

- Penilaian Kognitif ke Pembentukan Karakter Peserta Didik. *YASIN: Jurnal Pendidikan dan Sosial Budaya*, 5(6), 7045–7057. <https://doi.org/10.58578/yasin.v5i6.8504>
- Hafiz, S. R., Nasution, A. F., Rahayu, W., Muliadi, & Sitompul, C. M. (2024). Teacher Problems in Implementing the Independent Curriculum in Islamic Religious Education Subjects. *Competitive: Journal of Education*, 3(1). <https://doi.org/10.58355/competitive.v3i1.38>
- Hanafi, S. S. A. N., Fadhila, A. S., & Bashith, A. (2025). Pendekatan Penilaian Autentik dalam Evaluasi Pembelajaran Pendidikan Agama Islam (PAI): The Authentic Assessment Approach in the Evaluation of Islamic Religious Education (PAI) Learning. *EDUCAN: Jurnal Pendidikan Islam*, 9(2), 186–202. <https://doi.org/10.21111/educan.v9i2.14679>
- Havista, N., Zahara, & Rahman, T. (2025). Strategi Guru PAI dalam Mengatasi Kesulitan Belajar dalam Pembelajaran Pendidikan Agama Islam. *Jurnal Manajemen dan Pendidikan Agama Islam*, 3(2), 352–366. <https://doi.org/10.61132/jmpai.v3i2.1009>
- Hoeve, K. B. (2022). A Validity Framework for Accountability: Educational Measurement and Language Testing. *Language Testing in Asia*, 12(1), 1–21. <https://doi.org/10.1186/s40468-021-00153-2>
- Irmayanti, A. P., Remiswal, Khadijah, & Syamsi. (2024). Revealing the Practice of Norm-referenced Assessment and Criterion-referenced Assessment in Islamic Education Learning: A Case Study at SMPN 24 Padang. *At-Ta'lim: Media Informasi Pendidikan Islam*, 23(1), 87–100. <https://doi.org/10.29300/attalim.v23i1.4327>
- Jaakkola, E. (2020). Designing Conceptual Articles: Four Approaches. *AMS Review*, 10(1–2), 18–26. <https://doi.org/10.1007/s13162-020-00161-0>
- Lestari, A. D., Alghifari, R. I., & Zuhriyah, I. A. (2025). Pendekatan Komprehensif Terhadap Umpan Balik Dalam Evaluasi PAI: Antara Capaian, Refleksi, dan Intervensi. *Muróbbi: Jurnal Ilmu Pendidikan*, 9(2), 331–347. <https://doi.org/10.52431/murobbi.v9i2.3909>
- Meijer, H., Brouwer, J., Hoekstra, R., & Strijbos, J.-W. (2022). Exploring Construct and Consequential Validity of Collaborative Learning Assessment in Higher Education. *Small Group Research*, 53(6), 891–925. <https://doi.org/10.1177/10464964221095545>
- Mulyani, R., Karimullah, Wahab, & Sobirin, S. (2024). Characteristics of Islamic Religious Education Assessment Instruments. *At Tuots: Jurnal Pendidikan Islam*, 6(1), 250–258. <https://doi.org/10.51468/jpi.v6i1.439>
- Nanda Hanafi, S. S. A., & Bashith, A. (2025). Pendekatan Penilaian Autentik dalam Evaluasi Pembelajaran Pendidikan Agama Islam (PAI). *Educan*, 9(2), 51–67. <https://doi.org/10.21111/educan.v9i2.14679>
- Nasrullah. (2025). Menggugat Model Evaluasi Kognitif dalam PAI sebagai Penghambat Pendewasaan Iman. *AN-NUR: Jurnal Pendidikan dan Pembelajaran*, 1(2), 443–459. <https://doi.org/10.65802/an-nur.v1i2.67>
- Newton, P. (2023). *Book Announcement: Validity in Educational and Psychological Assessment*. Apollo – University of Cambridge Repository. <https://doi.org/10.17863/CAM.100296>
- NurAzizi, M. H., Alfauzi, A. R., & Bashith, A. (2025). Assessment Strategy Based on Validity and Reliability in Determining the Minimum Completeness Criteria for Islamic Religious Education. *Inteligensi: Jurnal Ilmu Pendidikan*, 8(1), 186–197. <https://doi.org/10.33366/ilg.v8i1.7248>