

Optimalisasi Evaluasi Ranah Kognitif dalam Pembelajaran PAI melalui Analisis Kualitas Instrumen Penilaian

Indina Ulil Fasya¹, Syaifuddin²

¹Universitas Islam Negeri Sunan Ampel Surabaya, Indonesia
indinafasya123@gmail.com, syaifuddin.tarbiyah@uinsa.ac.id

Informasi Artikel	Abstrak
Riwayat artikel: Diterima: 2026-06-02 Direvisi: 2026-06-08 Diterima: 2026-06-27	Abstrak Evaluasi pembelajaran Pendidikan Agama Islam (PAI) memiliki kompleksitas yang lebih tinggi dibandingkan mata pelajaran lain karena tidak hanya mengukur aspek pengetahuan, tetapi juga kemampuan berpikir kritis, reflektif, dan pemahaman kontekstual peserta didik terhadap nilai-nilai keislaman. Permasalahan yang masih ditemukan dalam praktik pembelajaran menunjukkan bahwa instrumen evaluasi PAI cenderung didominasi oleh soal berbasis Lower Order Thinking Skills (LOTS) yang lebih menekankan kemampuan mengingat dan memahami dibanding kemampuan menganalisis, mengevaluasi, dan mencipta. Kondisi tersebut menyebabkan evaluasi pembelajaran belum mampu mengoptimalkan pengembangan kemampuan berpikir tingkat tinggi peserta didik. Penelitian ini bertujuan menganalisis kualitas instrumen penilaian ranah kognitif dalam pembelajaran PAI serta mengidentifikasi kontribusinya terhadap optimalisasi evaluasi pembelajaran. Penelitian menggunakan pendekatan kualitatif dengan jenis library research melalui teknik dokumentasi dan studi literatur. Data dianalisis menggunakan content analysis terhadap berbagai artikel ilmiah, buku, dan penelitian relevan mengenai evaluasi pembelajaran PAI dan kualitas instrumen penilaian. Hasil penelitian menunjukkan bahwa kualitas instrumen evaluasi dipengaruhi oleh validitas, reliabilitas, tingkat kesukaran soal, daya pembeda, dan integrasi Higher Order Thinking Skills (HOTS) dalam penyusunan soal. Temuan penelitian juga memperlihatkan bahwa sebagian besar instrumen evaluasi PAI masih berorientasi pada kemampuan kognitif tingkat rendah sehingga belum mampu mendukung tuntutan pembelajaran abad ke-21 secara optimal. Penelitian ini menegaskan bahwa optimalisasi evaluasi ranah kognitif memerlukan pengembangan instrumen penilaian yang sistematis, autentik, dan berbasis HOTS agar evaluasi pembelajaran PAI menjadi lebih relevan, terukur, dan kontekstual.
Kata kunci: Evaluasi kognitif Pembelajaran PAI Kualitas instrumen HOTS Evaluasi pembelajaran Analisis instrumen	
Jenis artikel: Artikel penelitian	

Karya ini dilisensikan di bawah Lisensi Creative Commons Attribution 4.0 International (CC BY 4.0).



1. Pendahuluan

1.1. Informasi latar belakang

Pendidikan abad ke-21 menempatkan evaluasi pembelajaran sebagai komponen strategis dalam meningkatkan mutu pendidikan global, termasuk dalam pembelajaran Pendidikan Agama Islam (PAI). Transformasi sistem pendidikan internasional menuntut proses evaluasi yang tidak hanya mengukur kemampuan mengingat, tetapi juga kemampuan berpikir kritis, analitis, dan reflektif peserta didik. Organisasi pendidikan dunia menekankan pentingnya asesmen berbasis kompetensi untuk mendukung pengembangan Higher Order Thinking Skills (HOTS) sebagai bagian dari kesiapan menghadapi tantangan global. Kondisi tersebut mendorong perubahan paradigma evaluasi pembelajaran dari sekadar pengukuran hasil menuju pengembangan kemampuan kognitif secara komprehensif. Menurut (Zahra & At-Taqiyyah, 2024), instrumen penilaian yang baik harus

mampu mengukur kemampuan berpikir tingkat tinggi secara objektif dan sistematis. Realitas di lapangan menunjukkan bahwa sebagian besar instrumen evaluasi PAI masih berorientasi pada kemampuan menghafal sehingga belum optimal dalam mengukur kemampuan analisis dan evaluasi peserta didik. Perkembangan evaluasi pembelajaran secara global juga dipengaruhi oleh kemajuan teknologi pendidikan dan perubahan kebijakan kurikulum berbasis kompetensi. Berbagai negara mulai menerapkan sistem penilaian autentik dan asesmen berbasis literasi untuk meningkatkan kualitas pembelajaran. Dalam konteks Indonesia, implementasi Kurikulum Merdeka menempatkan evaluasi kognitif sebagai instrumen penting dalam mengukur capaian pembelajaran peserta didik secara menyeluruh. Menurut (Nuryani et al., 2024), evaluasi hasil belajar dalam pembelajaran PAI harus mampu memberikan informasi akurat mengenai tingkat pemahaman peserta didik terhadap materi keagamaan. Kondisi tersebut diperkuat oleh temuan (Yanti et al., 2022) yang menunjukkan bahwa pelaksanaan penilaian ranah kognitif pada pembelajaran PAI masih menghadapi kendala dalam penyusunan instrumen yang valid dan reliabel. Situasi ini menunjukkan bahwa optimalisasi evaluasi kognitif menjadi kebutuhan mendesak dalam peningkatan mutu pembelajaran PAI. Kajian mengenai evaluasi ranah kognitif dalam pendidikan pada dasarnya berakar pada teori Taksonomi Bloom yang kemudian direvisi oleh Anderson dan Krathwohl menjadi dimensi kognitif yang lebih kompleks. Teori tersebut menjelaskan bahwa kemampuan kognitif peserta didik berkembang mulai dari mengingat, memahami, menerapkan, menganalisis, mengevaluasi, hingga mencipta. Menurut (Astuti, 2021) revisi Taksonomi Bloom memberikan kerangka konseptual yang lebih sistematis dalam menyusun instrumen evaluasi berbasis kemampuan berpikir tingkat tinggi. Perspektif tersebut diperkuat oleh (Rahmawati, 2021) yang menyatakan bahwa instrumen evaluasi PAI perlu dirancang berdasarkan indikator kognitif yang proporsional agar mampu mengukur kompetensi peserta didik secara utuh. Evolusi teori evaluasi pendidikan juga menunjukkan pergeseran dari pendekatan tradisional menuju asesmen autentik yang menekankan validitas proses dan hasil pembelajaran. Penilaian ranah kognitif dalam pembelajaran PAI memiliki karakteristik khusus karena tidak hanya mengukur aspek intelektual, tetapi juga berkaitan dengan pemahaman nilai, moral, dan implementasi ajaran Islam dalam kehidupan sehari-hari. Menurut (Rosyidi, 2020), instrumen asesmen kognitif dalam PAI harus memperhatikan keseimbangan antara aspek konseptual dan kontekstual agar hasil evaluasi benar-benar mencerminkan kemampuan peserta didik. Penelitian (Suliswiyadi, 2020) menunjukkan bahwa pembelajaran PAI memiliki dimensi pembentukan karakter yang menuntut evaluasi lebih komprehensif dibandingkan mata pelajaran umum lainnya. Karakteristik tersebut menyebabkan proses penyusunan instrumen evaluasi PAI memerlukan analisis kualitas yang lebih mendalam. Kondisi ini semakin penting ketika pembelajaran PAI dituntut untuk menghasilkan peserta didik yang mampu berpikir kritis sekaligus memiliki pemahaman keagamaan yang moderat dan aplikatif. Permasalahan utama dalam evaluasi pembelajaran PAI saat ini terletak pada kualitas instrumen penilaian yang belum memenuhi standar validitas, reliabilitas, tingkat kesukaran, dan daya pembeda yang baik. Menurut Zakkiyah et al. (Zakkiyah et al., 2024), banyak instrumen tes kognitif PAI di sekolah dasar masih didominasi soal dengan kategori berpikir tingkat rendah sehingga belum mampu mengukur kemampuan analitis peserta didik. Temuan serupa disampaikan oleh Amalia (Amalia, 2023) yang menyatakan bahwa kualitas instrumen asesmen ranah kognitif PAI di perguruan tinggi masih menunjukkan ketidakseimbangan distribusi level kognitif berdasarkan Taksonomi Bloom. Kondisi tersebut berdampak pada rendahnya akurasi hasil evaluasi pembelajaran dan kurang optimalnya pemetaan kemampuan peserta didik. Menurut Diki et al. (Diki et al., 2022), pengembangan instrumen tes berbasis kemampuan berpikir analitis menjadi solusi penting dalam meningkatkan kualitas evaluasi pembelajaran PAI. Permasalahan tersebut menunjukkan urgensi analisis kualitas instrumen penilaian sebagai langkah strategis dalam optimalisasi evaluasi ranah kognitif.

1.2. Kesenjangan penelitian

Penelitian terdahulu menunjukkan adanya variasi metodologi dalam menganalisis kualitas instrumen penilaian pembelajaran. Sebagian besar penelitian menggunakan pendekatan deskriptif kuantitatif melalui analisis validitas, reliabilitas, tingkat kesukaran, dan daya pembeda soal. Penelitian (Muti'ah & Maemonah, 2022) menggunakan analisis Taksonomi Bloom untuk mengidentifikasi level kognitif soal uraian PAI di sekolah dasar. Penelitian (Rizqillah et al., 2023) menerapkan model evaluasi CIPP dalam menganalisis kualitas butir soal, sedangkan (Pudjiati & Madani, n.d.) mengembangkan pendekatan Rasch Model sebagai metode analisis instrumen autentik yang lebih objektif. Menurut (Erviana et al., 2025), analisis butir soal secara empiris mampu memberikan gambaran kualitas instrumen secara lebih akurat dibandingkan penilaian subjektif guru. Variasi metodologi tersebut menunjukkan bahwa analisis kualitas instrumen evaluasi masih terus berkembang dan memerlukan penguatan konseptual serta implementatif. Perdebatan akademik mengenai evaluasi pembelajaran PAI muncul pada penentuan bentuk instrumen yang paling efektif dalam mengukur kemampuan kognitif peserta didik. Sebagian peneliti menekankan penggunaan soal objektif karena dianggap lebih praktis dan memiliki reliabilitas tinggi. Menurut (Zahra & At-Taqiyyah, 2024), pertanyaan pilihan ganda dapat digunakan untuk mengukur kemampuan berpikir tingkat tinggi apabila dirancang dengan indikator HOTS yang tepat. Perspektif berbeda dikemukakan oleh (Meiliani et al., 2023) yang menilai bahwa instrumen autentik lebih efektif dalam menggambarkan kemampuan nyata peserta didik dibandingkan tes objektif konvensional. Penelitian (Supriadi et al., 2025) juga menegaskan bahwa pengembangan instrumen berbasis HOTS mampu meningkatkan kualitas pembelajaran PAI secara signifikan. Perbedaan pandangan tersebut menunjukkan bahwa analisis kualitas instrumen penilaian masih menjadi ruang diskusi ilmiah yang relevan untuk dikaji lebih lanjut. Research gap dalam penelitian ini terletak pada masih terbatasnya kajian yang secara spesifik mengintegrasikan analisis kualitas instrumen penilaian dengan optimalisasi evaluasi ranah kognitif dalam pembelajaran PAI. Sebagian besar penelitian sebelumnya hanya berfokus pada analisis butir soal atau implementasi penilaian tanpa menghubungkannya secara komprehensif dengan peningkatan kualitas evaluasi pembelajaran. Penelitian (Ramadhan & Subando, 2025) lebih menitikberatkan pada kualitas butir soal fiqih, sedangkan (Fitriana & Mustofa, 2023) hanya mengkaji teknik evaluasi kognitif dalam praktik pembelajaran. Penelitian (Adzillah & Hamami, 2025) membahas pengembangan asesmen kognitif dan psikomotorik, namun belum secara mendalam menguraikan analisis kualitas instrumen sebagai dasar optimalisasi evaluasi. Kesenjangan tersebut menunjukkan perlunya penelitian yang mengintegrasikan analisis kualitas instrumen penilaian dengan efektivitas evaluasi ranah kognitif dalam pembelajaran PAI secara sistematis dan empiris.

1.3. Tujuan penelitian

Penelitian ini bertujuan untuk menganalisis kualitas instrumen penilaian ranah kognitif dalam pembelajaran Pendidikan Agama Islam (PAI) serta mengidentifikasi kontribusinya terhadap optimalisasi evaluasi pembelajaran. Analisis dilakukan terhadap berbagai aspek kualitas instrumen, meliputi validitas, reliabilitas, tingkat kesukaran, daya pembeda, dan integrasi Higher Order Thinking Skills (HOTS) dalam penyusunan instrumen evaluasi. Penelitian ini juga bertujuan mengkaji bagaimana kualitas instrumen penilaian memengaruhi efektivitas evaluasi pembelajaran PAI dalam mengukur kemampuan berpikir kritis, analitis, dan reflektif peserta didik. Selain itu, penelitian ini diarahkan untuk memperkuat konsep evaluasi kognitif berbasis analisis kualitas instrumen dalam pembelajaran PAI. Secara praktis, penelitian ini diharapkan mampu menjadi acuan bagi guru, pengembang kurikulum, dan lembaga pendidikan dalam menyusun instrumen evaluasi yang valid, reliabel, objektif, dan sesuai dengan tuntutan pembelajaran abad ke-21. Penelitian ini juga menawarkan kebaruan melalui integrasi antara analisis kualitas instrumen penilaian dan pendekatan optimalisasi evaluasi pembelajaran secara komprehensif.

1.4. Pertanyaan penelitian

Berdasarkan latar belakang dan kesenjangan penelitian yang telah diuraikan, penelitian ini difokuskan pada beberapa pertanyaan penelitian. Pertama, bagaimana kualitas instrumen penilaian ranah kognitif yang digunakan dalam pembelajaran Pendidikan Agama Islam (PAI) ditinjau dari aspek validitas, reliabilitas, tingkat kesukaran, dan daya pembeda soal. Kedua, bagaimana implementasi Higher Order Thinking Skills (HOTS) diterapkan dalam instrumen evaluasi pembelajaran PAI. Ketiga, bagaimana kontribusi kualitas instrumen penilaian terhadap optimalisasi evaluasi pembelajaran PAI dalam mengukur kemampuan berpikir kritis dan analitis peserta didik. Keempat, mengapa sebagian besar instrumen evaluasi PAI masih didominasi oleh kemampuan berpikir tingkat rendah atau Lower Order Thinking Skills (LOTS). Kelima, bagaimana pengembangan instrumen penilaian yang sistematis, autentik, dan berbasis HOTS dapat mendukung peningkatan kualitas evaluasi pembelajaran PAI secara lebih objektif dan kontekstual..

1.5. Signifikansi penelitian

Penelitian ini memiliki signifikansi teoretis dan praktis dalam pengembangan evaluasi pembelajaran Pendidikan Agama Islam (PAI), khususnya pada ranah kognitif. Secara teoretis, penelitian ini berkontribusi dalam memperkuat kajian mengenai evaluasi kognitif berbasis analisis kualitas instrumen penilaian melalui integrasi antara konsep validitas, reliabilitas, tingkat kesukaran, daya pembeda, dan penerapan Higher Order Thinking Skills (HOTS) dalam pembelajaran PAI. Kajian ini juga memperluas pemahaman bahwa kualitas evaluasi pembelajaran tidak hanya ditentukan oleh hasil belajar peserta didik, tetapi juga oleh kualitas instrumen yang digunakan dalam proses pengukuran. Temuan penelitian ini memberikan penguatan konseptual terhadap teori evaluasi pendidikan yang menempatkan instrumen penilaian sebagai komponen strategis dalam membangun evaluasi yang objektif, autentik, dan kontekstual. Penelitian ini sekaligus mempertegas pentingnya transformasi evaluasi pembelajaran PAI dari pendekatan berbasis hafalan menuju evaluasi yang mampu mengukur kemampuan berpikir kritis, analitis, dan reflektif peserta didik sesuai tuntutan pendidikan abad ke-21.

Secara praktis, penelitian ini diharapkan dapat menjadi acuan bagi guru, pengembang kurikulum, dan lembaga pendidikan dalam menyusun instrumen evaluasi PAI yang lebih valid, reliabel, dan berbasis HOTS. Temuan penelitian menunjukkan bahwa rendahnya kualitas instrumen penilaian berpotensi menghasilkan bias pengukuran dan ketidaktepatan dalam pengambilan keputusan pembelajaran. Oleh karena itu, penelitian ini memberikan implikasi nyata terhadap pentingnya pengembangan instrumen evaluasi yang sistematis dan sesuai dengan karakteristik pembelajaran PAI. Penelitian ini juga relevan dalam mendukung implementasi Kurikulum Merdeka yang menekankan asesmen berbasis kompetensi dan pengembangan kemampuan berpikir tingkat tinggi peserta didik. Selain itu, hasil penelitian diharapkan mampu menjadi dasar pengembangan kebijakan evaluasi pembelajaran yang lebih komprehensif, sehingga evaluasi PAI tidak lagi berorientasi pada kemampuan menghafal semata, tetapi mampu mendorong pembelajaran yang lebih reflektif, kontekstual, dan bermakna.

2. Metodologi

2.1. Desain Penelitian

Penelitian ini menggunakan jenis penelitian kualitatif dengan pendekatan library research yang berorientasi pada pengkajian mendalam terhadap berbagai sumber literatur ilmiah terkait optimalisasi evaluasi ranah kognitif dalam pembelajaran Pendidikan Agama Islam (PAI) melalui analisis kualitas instrumen penilaian. Pendekatan kepustakaan dipilih karena penelitian berfokus pada eksplorasi konseptual, sintesis teoritis, serta interpretasi kritis terhadap temuan penelitian terdahulu yang relevan dengan evaluasi pembelajaran PAI. Sumber data penelitian terdiri atas buku akademik, artikel jurnal nasional terakreditasi, serta dokumen ilmiah yang diterbitkan pada rentang tahun 2020–2025 dan memiliki keterkaitan langsung dengan konsep evaluasi ranah kognitif, kualitas instrumen penilaian, dan asesmen berbasis Higher Order Thinking Skills (HOTS). Teknik

pengumpulan data dilakukan melalui dokumentasi dan studi literatur dengan cara mengidentifikasi, mengklasifikasi, serta menelaah sumber-sumber ilmiah yang memiliki validitas akademik tinggi. Menurut (Sugiyono, 2022) penelitian kepustakaan memungkinkan peneliti memperoleh pemahaman teoritis secara mendalam melalui pengkajian berbagai sumber ilmiah yang relevan, sedangkan (Yusuf, 2020) menjelaskan bahwa analisis isi digunakan untuk mengungkap makna, pola, dan kecenderungan data tekstual secara objektif dan sistematis. Pendekatan tersebut dipandang relevan untuk menghasilkan interpretasi ilmiah yang mendalam mengenai urgensi optimalisasi evaluasi kognitif melalui analisis kualitas instrumen penilaian dalam konteks pendidikan Islam kontemporer.

2.2. Sampel dan Teknik Pengambilan Sampel

Sampel dalam penelitian ini berupa sumber-sumber literatur ilmiah yang relevan dengan kajian optimalisasi evaluasi ranah kognitif dalam pembelajaran Pendidikan Agama Islam (PAI). Sumber tersebut mencakup buku akademik, artikel jurnal nasional terakreditasi, serta dokumen ilmiah lainnya yang dipublikasikan pada rentang tahun 2020–2025. Literatur yang dijadikan sampel dipilih karena memiliki keterkaitan langsung dengan konsep evaluasi pembelajaran, kualitas instrumen penilaian, taksonomi Bloom revisi, serta asesmen berbasis Higher Order Thinking Skills (HOTS).

Teknik pengambilan sampel menggunakan purposive sampling, yaitu pemilihan sumber literatur secara sengaja berdasarkan kriteria tertentu yang telah ditetapkan. Adapun kriteria yang digunakan meliputi: (1) sumber membahas evaluasi pembelajaran atau asesmen kognitif dalam PAI, (2) memiliki landasan teori yang kuat dan relevan, (3) berasal dari publikasi ilmiah yang dapat dipertanggungjawabkan secara akademik, serta (4) diterbitkan dalam rentang waktu 2020–2025. Pendekatan ini digunakan agar data yang diperoleh bersifat fokus, relevan, dan mendukung tujuan analisis isi dalam penelitian kepustakaan ini.

2.3. Instrumen

Instrumen penelitian dalam studi kepustakaan ini disusun dalam bentuk tabel analisis literatur dan matriks sintesis penelitian terdahulu untuk memperoleh data konseptual mengenai optimalisasi evaluasi ranah kognitif dalam pembelajaran Pendidikan Agama Islam (PAI). Instrumen analisis literatur digunakan untuk mengidentifikasi substansi teoritis yang berkaitan dengan konsep evaluasi pembelajaran, taksonomi Bloom revisi, asesmen HOTS, validitas instrumen, reliabilitas, daya pembeda soal, tingkat kesukaran, serta implementasi evaluasi kognitif dalam pembelajaran PAI. Analisis dilakukan terhadap buku dan artikel ilmiah dengan indikator meliputi identitas sumber, fokus kajian, teori utama, pendekatan penelitian, teknik analisis data, hasil penelitian, serta relevansi terhadap penelitian yang dilakukan. Instrumen dalam penelitian kepustakaan berfungsi sebagai alat sistematis untuk menyeleksi dan menginterpretasikan data tekstual agar menghasilkan sintesis ilmiah yang objektif. Perspektif tersebut diperkuat oleh Hamzah dan Mudlofir (Hamzah & Mudlofir, 2025) yang menjelaskan bahwa penelitian berbasis literatur memerlukan instrumen kategorisasi data agar analisis konseptual lebih terarah dan mendalam. Penggunaan instrumen tersebut memungkinkan peneliti memperoleh pemetaan komprehensif mengenai kecenderungan penelitian evaluasi ranah kognitif dalam pembelajaran PAI sekaligus mengidentifikasi kesenjangan penelitian secara sistematis.

Tabel Matriks Penelitian Terdahulu.

No	Peneliti	Tahun	Judul Penelitian	Metode	Fokus Kajian
1	Amalia	2023	Analisis Kualitas Instrumen Asesmen Ranah Kognitif Pendidikan Agama Islam	Kualitatif	Kualitas instrumen asesmen

2	Zakiyyah et al.	2024	Analisis Instrumen Tes Kognitif Mata Pelajaran PAI di Sekolah Dasar	Deskriptif kualitatif	Instrumen tes kognitif PAI
3	Supriadi et al.	2025	Pengembangan Instrumen Penilaian PAI dengan Pendekatan HOTS	Research and Development (R&D)	Pengembangan instrumen HOTS
4	Rahmawati	2021	Analisis Tingkat Kognitif Soal PAT PAI Berbasis AKM	Analisis isi	Taksonomi Bloom pada soal PAI
5	Diki et al.	2022	Pengembangan Instrumen Tes Kemampuan Berpikir Analisis	Pengembangan instrumen	Kemampuan berpikir analitis

2.4. Prosedur Pengumpulan Data

Prosedur pengumpulan data dalam penelitian ini dilakukan melalui teknik studi literatur (*library research*) dan dokumentasi terhadap berbagai sumber ilmiah yang relevan dengan fokus penelitian mengenai optimalisasi evaluasi ranah kognitif dalam pembelajaran Pendidikan Agama Islam (PAI). Pendekatan studi literatur digunakan karena penelitian ini bertujuan untuk mengidentifikasi, menelaah, dan mensintesis berbagai hasil penelitian terdahulu guna memperoleh pemahaman konseptual yang mendalam mengenai kualitas instrumen evaluasi pembelajaran PAI. Studi literatur memungkinkan peneliti mengembangkan konstruksi pengetahuan secara sistematis berdasarkan sumber akademik yang valid dan relevan.

Tahap awal pengumpulan data dilakukan dengan menentukan fokus kajian dan kata kunci penelitian, seperti “evaluasi pembelajaran PAI”, “instrumen HOTS”, “evaluasi kognitif”, “validitas dan reliabilitas instrumen”, serta “asesmen autentik”. Kata kunci tersebut digunakan untuk menelusuri berbagai sumber ilmiah melalui database akademik seperti Google Scholar, Garuda, DOAJ, dan jurnal nasional terakreditasi. Pemilihan sumber dilakukan secara selektif dengan mempertimbangkan relevansi tema, kualitas publikasi, tahun terbit, serta keterkaitan substansi dengan fokus penelitian agar data yang diperoleh benar-benar mendukung tujuan penelitian secara akademik.

Selanjutnya, peneliti melakukan proses identifikasi dan seleksi sumber data dengan menetapkan kriteria inklusi dan eksklusi. Kriteria inklusi meliputi artikel ilmiah, jurnal nasional maupun internasional, prosiding, dan buku akademik yang membahas evaluasi pembelajaran PAI, instrumen evaluasi kognitif, HOTS, validitas, reliabilitas, dan asesmen autentik. Sementara itu, sumber yang tidak memiliki relevansi langsung dengan topik penelitian atau tidak memenuhi standar akademik dikeluarkan dari proses analisis. Tahap seleksi ini bertujuan untuk memastikan bahwa data yang digunakan memiliki validitas ilmiah, kredibilitas sumber, dan dapat dipertanggungjawabkan secara akademik.

Setelah sumber data terpilih, peneliti melakukan dokumentasi data dengan cara membaca secara mendalam, mencatat informasi penting, mengelompokkan temuan, serta mengidentifikasi konsep-konsep utama yang berkaitan dengan kualitas evaluasi ranah kognitif dalam pembelajaran PAI. Proses dokumentasi dilakukan secara sistematis agar data yang diperoleh tersusun secara terstruktur dan memudahkan proses analisis. Data yang telah terkumpul kemudian diklasifikasikan berdasarkan tema-tema tertentu, seperti dominasi LOTS, pengembangan HOTS, validitas instrumen, reliabilitas soal, analisis butir soal, dan pendekatan evaluasi autentik.

2.5. Analisis data

Analisis data dalam penelitian ini menggunakan teknik analisis isi (*content analysis*) yang bertujuan untuk memahami makna, pola, dan kecenderungan informasi dari berbagai sumber literatur secara sistematis dan mendalam. Teknik *content analysis* digunakan karena penelitian ini berfokus pada interpretasi konseptual terhadap berbagai temuan penelitian terdahulu mengenai evaluasi pembelajaran PAI. Analisis isi memungkinkan peneliti mengidentifikasi pola-pola tertentu dalam data tekstual sehingga menghasilkan interpretasi yang valid dan kontekstual terhadap fenomena yang diteliti.

Tahap pertama analisis data dilakukan melalui reduksi data. Pada tahap ini, peneliti menyeleksi, menyederhanakan, dan memfokuskan data yang relevan dengan tujuan penelitian. Informasi yang tidak berkaitan langsung dengan evaluasi ranah kognitif dalam pembelajaran PAI dieliminasi, sedangkan data yang memiliki keterkaitan substantif dipertahankan untuk dianalisis lebih lanjut. Reduksi data bertujuan untuk mempermudah peneliti menemukan fokus kajian, pola utama, serta hubungan antarvariabel dalam berbagai sumber literatur yang dianalisis.

Tahap kedua yaitu kategorisasi data. Data yang telah direduksi kemudian dikelompokkan ke dalam beberapa tema utama berdasarkan kesamaan konsep dan fokus pembahasan. Kategori tersebut meliputi dominasi soal LOTS, implementasi HOTS, kualitas instrumen evaluasi, validitas dan reliabilitas soal, analisis butir soal, serta pendekatan evaluasi autentik. Proses kategorisasi dilakukan agar peneliti lebih mudah memahami keterkaitan antarhasil penelitian dan menemukan kecenderungan umum yang muncul dalam berbagai literatur.

Tahap ketiga adalah interpretasi data. Pada tahap ini, peneliti menafsirkan berbagai temuan penelitian dengan menghubungkan hasil analisis literatur terhadap teori evaluasi pembelajaran, taksonomi Bloom revisi, serta konsep asesmen autentik. Interpretasi dilakukan secara kritis dan mendalam agar diperoleh pemahaman yang komprehensif mengenai kondisi evaluasi ranah kognitif dalam pembelajaran PAI serta berbagai faktor yang memengaruhi kualitas instrumen evaluasi.

Tahap terakhir yaitu penarikan kesimpulan. Kesimpulan disusun berdasarkan hasil sintesis seluruh temuan penelitian yang telah dianalisis sehingga menghasilkan konstruksi konseptual mengenai optimalisasi evaluasi kognitif dalam pembelajaran PAI. Proses penarikan kesimpulan dilakukan secara sistematis agar hasil penelitian mampu memberikan kontribusi teoritis maupun praktis dalam pengembangan evaluasi pembelajaran yang lebih autentik, kontekstual, dan berbasis HOTS.

3. Hasil

3.1 Karakteristik Penelitian Evaluasi Ranah Kognitif PAI

Hasil analisis terhadap berbagai literatur menunjukkan bahwa penelitian mengenai evaluasi ranah kognitif dalam pembelajaran Pendidikan Agama Islam (PAI) mengalami perkembangan yang cukup signifikan dalam lima tahun terakhir. Perkembangan tersebut terlihat dari perubahan orientasi penelitian yang sebelumnya lebih menitikberatkan pada pencapaian hasil belajar peserta didik menuju kajian yang lebih mendalam mengenai kualitas instrumen evaluasi sebagai fondasi utama dalam proses penilaian pembelajaran. Perubahan arah penelitian tersebut menunjukkan bahwa evaluasi pembelajaran tidak lagi dipahami sekadar sebagai alat pengukuran hasil belajar, tetapi juga sebagai instrumen strategis untuk menilai kualitas proses berpikir peserta didik secara komprehensif.

Kecenderungan tersebut terlihat dalam penelitian (Zakkiyah et al., 2024), (Rahmawati, 2021), dan (Amalia, 2023) yang sama-sama menempatkan instrumen evaluasi sebagai fokus utama kajian penelitian. Meskipun memiliki objek dan pendekatan yang berbeda, ketiga penelitian tersebut menunjukkan pola yang relatif serupa, yaitu adanya perhatian terhadap kualitas konstruksi soal, distribusi level kognitif, serta kemampuan instrumen dalam mengukur kemampuan berpikir

peserta didik. Sintesis dari beberapa penelitian tersebut menunjukkan bahwa kualitas evaluasi pembelajaran PAI sangat dipengaruhi oleh kualitas instrumen penilaian yang digunakan guru.

Selain itu, perkembangan penelitian evaluasi PAI juga menunjukkan adanya pergeseran paradigma dari evaluasi konvensional menuju evaluasi berbasis Higher Order Thinking Skills (HOTS). Penelitian (Supriadi et al., 2025) menegaskan bahwa pengembangan instrumen berbasis HOTS mampu meningkatkan kualitas pembelajaran PAI karena peserta didik tidak hanya dituntut memahami materi secara tekstual, tetapi juga mampu melakukan analisis, evaluasi, dan pemecahan masalah kontekstual berdasarkan nilai-nilai keislaman. Temuan tersebut memperlihatkan bahwa instrumen evaluasi tidak hanya berfungsi sebagai alat ukur hasil belajar, tetapi juga berperan dalam membentuk kemampuan berpikir kritis peserta didik.

Secara konseptual, hubungan antara evaluasi ranah kognitif, kualitas instrumen, dan HOTS menunjukkan keterkaitan yang sangat erat. Evaluasi ranah kognitif membutuhkan instrumen berkualitas agar hasil pengukuran valid dan reliabel, sedangkan pengembangan HOTS menjadi indikator bahwa evaluasi telah mampu mengukur kemampuan berpikir tingkat tinggi peserta didik. Dengan demikian, kualitas instrumen penilaian menjadi variabel utama yang menentukan efektivitas evaluasi pembelajaran PAI.

Temuan tersebut menunjukkan bahwa penelitian evaluasi pembelajaran PAI mulai bergerak menuju pendekatan evaluasi yang lebih autentik dan kontekstual. Namun demikian, sebagian besar penelitian masih berfokus pada aspek teknis analisis instrumen dan belum banyak mengintegrasikan hubungan antara kualitas instrumen dengan optimalisasi evaluasi pembelajaran secara menyeluruh. Kondisi tersebut menunjukkan adanya ruang pengembangan penelitian yang masih terbuka dalam kajian evaluasi ranah kognitif PAI.

3.2 Dominasi LOTS dalam Instrumen Evaluasi PAI

Salah satu pola yang paling sering muncul dalam berbagai penelitian adalah dominasi soal berbasis Lower Order Thinking Skills (LOTS) pada evaluasi pembelajaran PAI. Penelitian (Zakkiyah et al., 2024) dan (Rahmawati, 2021) memperlihatkan bahwa sebagian besar soal masih berada pada level kognitif C1 dan C2 berdasarkan Taksonomi Bloom revisi. Artinya, evaluasi lebih banyak mengukur kemampuan mengingat dan memahami dibanding kemampuan berpikir tingkat tinggi.

Fenomena tersebut cukup mudah ditemukan dalam praktik pembelajaran PAI. Soal evaluasi umumnya masih berorientasi pada hafalan materi, definisi, atau penyebutan dalil keagamaan. Peserta didik memang dapat memperoleh nilai tinggi, tetapi belum tentu memiliki kemampuan menganalisis atau mengaitkan materi agama dengan persoalan kehidupan nyata. Dalam konteks pendidikan abad ke-21, kondisi semacam ini tentu menjadi tantangan tersendiri.

Beberapa penelitian menunjukkan bahwa keterbatasan implementasi HOTS tidak hanya disebabkan oleh bentuk soal, tetapi juga dipengaruhi oleh pemahaman guru terhadap penyusunan instrumen evaluasi. (Supriadi et al., 2025) menjelaskan bahwa pengembangan instrumen HOTS membutuhkan kemampuan pedagogis dan penguasaan indikator berpikir tingkat tinggi yang tidak sederhana. Karena itu, banyak guru masih merasa lebih mudah menggunakan soal objektif konvensional dibanding menyusun soal yang menuntut analisis dan refleksi.

Pandangan (Zahra & At-Taqiyyah, 2024) memperkuat hal tersebut. Mereka menjelaskan bahwa evaluasi berbasis HOTS seharusnya mampu mendorong peserta didik berpikir reflektif serta memecahkan persoalan secara kontekstual. Dengan kata lain, HOTS bukan hanya soal tingkat kesulitan, melainkan berkaitan dengan bagaimana peserta didik menggunakan pengetahuan keagamaan untuk memahami realitas sosial di sekitarnya.

Jika kondisi ini terus berlangsung, evaluasi pembelajaran PAI berisiko hanya menjadi alat administratif yang mengukur hafalan materi. Padahal, pembelajaran agama idealnya juga membentuk kemampuan berpikir kritis dan kesadaran reflektif peserta didik. Oleh sebab itu,

dominasi LOTS dalam instrumen evaluasi menunjukkan perlunya transformasi penilaian menuju pendekatan yang lebih kontekstual dan berorientasi pada pengembangan kemampuan berpikir tingkat tinggi.

3.3 Kualitas Instrumen Penilaian PAI

Kualitas instrumen penilaian menjadi salah satu persoalan yang paling banyak dibahas dalam penelitian evaluasi pembelajaran PAI. Dari berbagai literatur yang dianalisis, terlihat bahwa validitas, reliabilitas, tingkat kesukaran, dan daya pembeda soal merupakan aspek utama yang menentukan kualitas suatu instrumen evaluasi.

(Amalia, 2023) menemukan bahwa instrumen asesmen ranah kognitif PAI di perguruan tinggi belum memiliki keseimbangan antara indikator HOTS dan LOTS. Beberapa instrumen bahkan belum melalui pengujian empiris secara memadai. Situasi ini menunjukkan bahwa proses penyusunan instrumen evaluasi sering kali lebih berfokus pada penyelesaian administratif dibanding memastikan kualitas pengukuran.

Hal serupa juga terlihat pada penelitian (Diki et al., 2022). Penelitian tersebut menegaskan pentingnya pengembangan instrumen berbasis kemampuan berpikir analitis agar evaluasi pembelajaran mampu menghasilkan pengukuran yang lebih akurat. Instrumen yang baik tidak hanya menyajikan soal, tetapi juga harus mampu membedakan tingkat kemampuan peserta didik secara objektif.

Dalam praktik pembelajaran, penggunaan tes objektif memang dianggap lebih praktis dan efisien. Namun, pendekatan ini sering kali kurang mampu menggambarkan kemampuan autentik peserta didik secara menyeluruh. Akibatnya, hasil evaluasi belum sepenuhnya mencerminkan kemampuan berpikir peserta didik yang sebenarnya. (Rosyidi, 2020) bahkan menegaskan bahwa instrumen evaluasi yang tidak memenuhi standar kualitas dapat menghasilkan bias pengukuran dan memengaruhi efektivitas pembelajaran.

Di sisi lain, hubungan antara validitas, reliabilitas, tingkat kesukaran, dan daya pembeda sebenarnya saling berkaitan. Instrumen yang valid tetapi tidak reliabel tetap akan menghasilkan pengukuran yang lemah. Begitu pula soal yang terlalu mudah atau terlalu sulit tidak akan mampu memetakan kemampuan peserta didik secara proporsional. Karena itu, kualitas instrumen penilaian tidak dapat dipahami secara parsial.

Persoalan tersebut memperlihatkan bahwa optimalisasi evaluasi ranah kognitif tidak cukup hanya dengan menambah jumlah soal HOTS. Yang jauh lebih penting adalah bagaimana instrumen disusun secara sistematis agar benar-benar mampu memberikan gambaran autentik mengenai kemampuan peserta didik dalam pembelajaran PAI.

3.4 Kesenjangan Penelitian Evaluasi Kognitif PAI

Jika diperhatikan lebih mendalam, sebagian besar penelitian evaluasi pembelajaran PAI masih bergerak pada ranah analisis teknis instrumen. Penelitian (Zakkiyah et al., 2024) dan (Rahmawati, 2021) misalnya, lebih banyak membahas distribusi level kognitif soal berdasarkan Taksonomi Bloom revisi. Sementara itu, (Amalia, 2023) dan (Diki et al., 2022) cenderung fokus pada kualitas teknis instrumen seperti validitas dan kemampuan analitis soal.

Meskipun kajian tersebut penting, ada kecenderungan bahwa penelitian evaluasi PAI masih dilakukan secara parsial. Banyak penelitian berhenti pada identifikasi kelemahan soal tanpa menghubungkannya dengan efektivitas evaluasi pembelajaran secara lebih luas. Akibatnya, hubungan antara kualitas instrumen dan optimalisasi evaluasi ranah kognitif belum banyak dibahas secara mendalam.

Padaahal, evaluasi pembelajaran tidak hanya berkaitan dengan baik atau buruknya soal. Evaluasi juga menentukan bagaimana guru memahami perkembangan kemampuan peserta didik dan

mengambil keputusan pembelajaran. (Nuryani et al., 2024) menjelaskan bahwa evaluasi yang baik harus mampu memberikan gambaran autentik mengenai tingkat pemahaman peserta didik terhadap materi keagamaan. Pendapat tersebut diperkuat (Harsoyo et al., 2025) yang menegaskan bahwa kualitas instrumen akan memengaruhi ketepatan pengambilan keputusan pendidikan.

Dari berbagai penelitian yang dianalisis, terlihat bahwa belum banyak kajian yang secara khusus mengintegrasikan kualitas instrumen, implementasi HOTS, dan efektivitas evaluasi pembelajaran dalam satu konstruksi konseptual yang utuh. Di sinilah letak kesenjangan penelitian yang masih terbuka.

Berdasarkan kondisi tersebut, penelitian ini mencoba menempatkan kualitas instrumen penilaian sebagai faktor utama dalam optimalisasi evaluasi ranah kognitif pembelajaran PAI. Dengan demikian, evaluasi tidak hanya dipahami sebagai proses pengukuran hasil belajar, tetapi juga sebagai upaya membangun kualitas pembelajaran yang lebih reflektif, kontekstual, dan bermakna.

3.5 Visualisasi Pola Temuan Literatur

Pola temuan literatur mengenai evaluasi pembelajaran Pendidikan Agama Islam (PAI) menunjukkan bahwa proses evaluasi ranah kognitif masih menghadapi berbagai tantangan, baik dari aspek penyusunan instrumen maupun pendekatan penilaian yang digunakan. Berdasarkan hasil analisis berbagai penelitian, sebagian besar evaluasi pembelajaran PAI masih didominasi oleh pendekatan konvensional yang berorientasi pada penguasaan materi secara teoritis dan hafalan. Instrumen penilaian lebih banyak mengukur kemampuan berpikir tingkat rendah atau *Lower Order Thinking Skills* (LOTS), khususnya pada level mengingat (C1) dan memahami (C2). Kondisi ini memperlihatkan bahwa evaluasi pembelajaran belum sepenuhnya diarahkan untuk mengembangkan kemampuan berpikir tingkat tinggi peserta didik (Maulina, n.d.).

Dominasi soal LOTS menyebabkan peserta didik cenderung hanya berfokus pada penguasaan informasi dan hafalan konsep keagamaan tanpa didorong untuk melakukan analisis, evaluasi, maupun pemecahan masalah secara kritis. Padahal, pembelajaran PAI pada hakikatnya tidak hanya menekankan penguasaan pengetahuan agama, tetapi juga pengembangan kemampuan berpikir reflektif dan penerapan nilai-nilai Islam dalam kehidupan sehari-hari. Oleh karena itu, evaluasi yang masih berorientasi pada LOTS dinilai kurang relevan dengan kebutuhan pembelajaran abad ke-21 yang menuntut peserta didik memiliki kemampuan berpikir kritis, kreatif, kolaboratif, dan komunikatif (Hilmiyati et al., 2025).

Selain persoalan level kognitif, penggunaan instrumen berbasis *Higher Order Thinking Skills* (HOTS) dalam pembelajaran PAI juga masih relatif terbatas. Beberapa penelitian menunjukkan bahwa guru masih mengalami kesulitan dalam menyusun soal HOTS karena kurangnya pemahaman mengenai karakteristik soal berbasis analisis, evaluasi, dan kreasi. Di samping itu, sebagian guru masih beranggapan bahwa mata pelajaran PAI lebih mudah diukur melalui soal hafalan dan tes objektif. Akibatnya, bentuk evaluasi yang digunakan cenderung monoton dan belum mampu menggambarkan kemampuan autentik peserta didik secara menyeluruh (Efendi et al., 2025).

Keterbatasan penggunaan instrumen HOTS juga berdampak pada rendahnya kemampuan peserta didik dalam menghubungkan materi PAI dengan konteks kehidupan nyata. Evaluasi pembelajaran seharusnya mampu menjadi sarana untuk melatih peserta didik berpikir kritis terhadap fenomena sosial, etika, dan nilai-nilai keislaman yang berkembang di masyarakat. Dengan demikian, pengembangan instrumen evaluasi berbasis HOTS menjadi penting agar pembelajaran PAI tidak hanya bersifat teoritis, tetapi juga kontekstual dan aplikatif (Maulina, n.d.).

Di sisi lain, kualitas instrumen evaluasi juga menjadi persoalan yang cukup menonjol dalam berbagai penelitian. Sejumlah studi menunjukkan bahwa validitas dan reliabilitas instrumen belum seluruhnya diuji secara empiris sebelum digunakan dalam proses evaluasi. Validitas instrumen sangat penting untuk memastikan bahwa soal benar-benar mengukur kompetensi yang ingin dicapai, sedangkan reliabilitas berkaitan dengan konsistensi hasil pengukuran. Ketika instrumen

tidak memenuhi kedua aspek tersebut, hasil evaluasi berpotensi bias dan kurang akurat dalam menggambarkan kemampuan peserta didik yang sebenarnya (Elviana, 2020).

Selain validitas dan reliabilitas, analisis instrumen seperti tingkat kesukaran soal, daya pembeda, dan efektivitas pengecoh juga masih belum dilakukan secara komprehensif. Padahal, analisis butir soal sangat penting untuk mengetahui kualitas instrumen evaluasi dan memperbaiki kelemahan soal sebelum digunakan kembali. Kurangnya analisis instrumen menyebabkan evaluasi pembelajaran belum berjalan secara optimal dan berpotensi menghasilkan keputusan penilaian yang kurang objektif (Elviana, 2020)

Tabel 1 Visualisasi Pola Temuan Literatur

No.	Aspek Temuan	Kecenderungan Temuan Literatur	Implikasi
1	Level Kognitif Soal	Dominasi LOTS (C1-C2)	Kemampuan analitis peserta didik belum terukur optimal
2	Instrumen HOTS	Masih terbatas digunakan	Evaluasi belum kontekstual
3	Validitas Instrumen	Sebagian belum diuji empiris	Hasil evaluasi berpotensi bias
4	Reliabilitas Soal	Belum konsisten	Akurasi pengukuran rendah
5	Analisis Instrumen	Belum komprehensif	Optimalisasi evaluasi belum maksimal
6	Pendekatan Evaluasi	Dominan tes objektif	Kurang menggambarkan kemampuan autentik

Berdasarkan grafik visualisasi pola temuan literatur di atas, terlihat bahwa dominasi penggunaan soal LOTS dan pendekatan tes objektif menjadi kecenderungan paling tinggi dalam evaluasi pembelajaran PAI. Temuan tersebut menunjukkan bahwa evaluasi masih lebih menekankan aspek kognitif dasar dibanding kemampuan berpikir tingkat tinggi. Sementara itu, keterbatasan penggunaan instrumen HOTS memperlihatkan bahwa inovasi evaluasi pembelajaran belum berkembang secara optimal (Maulina, n.d.).

Grafik juga menunjukkan bahwa aspek validitas, reliabilitas, dan analisis instrumen masih menjadi kelemahan dalam pelaksanaan evaluasi pembelajaran PAI. Hal tersebut mengindikasikan bahwa sebagian instrumen evaluasi belum memenuhi standar kualitas pengukuran pendidikan. Akibatnya, hasil evaluasi berpotensi kurang objektif dan belum mampu memberikan gambaran menyeluruh mengenai perkembangan kemampuan peserta didik (Elviana, 2020).

Menariknya, beberapa penelitian memperlihatkan bahwa ketika instrumen evaluasi dirancang secara sistematis, berbasis HOTS, dan didukung pengujian kualitas soal yang baik, hasil evaluasi menjadi lebih autentik dan kontekstual. Evaluasi tidak lagi hanya mengukur kemampuan hafalan, tetapi juga mampu menilai kemampuan analitis, reflektif, dan pemecahan masalah peserta didik dalam kehidupan nyata. Oleh karena itu, optimalisasi evaluasi ranah kognitif dalam pembelajaran PAI memerlukan integrasi antara pengembangan instrumen HOTS, pengujian validitas dan reliabilitas soal, serta penerapan evaluasi autentik agar proses pembelajaran dapat berkembang secara lebih efektif dan berkelanjutan (Efendi et al., 2025).

4. Pembahasan

4.1 Dominasi LOTS dan Tantangan Evaluasi Pembelajaran PAI

Temuan penelitian menunjukkan bahwa instrumen evaluasi pembelajaran PAI masih didominasi oleh soal berbasis Lower Order Thinking Skills (LOTS), khususnya pada level C1 dan C2 berdasarkan Taksonomi Bloom revisi. Kondisi tersebut menandakan bahwa evaluasi pembelajaran masih berorientasi pada kemampuan mengingat dan memahami materi, sedangkan kemampuan berpikir kritis, analitis, dan reflektif peserta didik belum sepenuhnya terakomodasi dalam proses penilaian.

Jika dikaitkan dengan teori Taksonomi Bloom revisi Anderson dan Krathwohl, kemampuan berpikir tingkat rendah memang menjadi fondasi awal pembelajaran kognitif. Akan tetapi, pembelajaran abad ke-21 menuntut peserta didik tidak berhenti pada tahap mengingat informasi, melainkan mampu menganalisis, mengevaluasi, dan menciptakan gagasan baru berdasarkan pengetahuan yang dimiliki. Dalam konteks pembelajaran PAI, kemampuan tersebut penting karena materi keagamaan tidak hanya dipahami secara tekstual, tetapi juga harus mampu diimplementasikan dalam kehidupan sosial secara kontekstual.

Fenomena dominasi LOTS dalam evaluasi PAI memperlihatkan adanya kesenjangan antara tujuan pembelajaran dan praktik evaluasi di lapangan. Pada banyak kasus, pembelajaran PAI sebenarnya telah mengarahkan peserta didik untuk memahami nilai-nilai Islam secara mendalam, tetapi evaluasi yang digunakan masih bersifat hafalan. Akibatnya, proses evaluasi belum mampu menggambarkan kemampuan berpikir peserta didik secara utuh.

Penelitian (Zakkiyah et al., 2024) dan (Rahmawati, 2021) memperlihatkan bahwa distribusi soal evaluasi PAI masih didominasi level kognitif rendah. Temuan tersebut memperkuat argumentasi bahwa implementasi HOTS dalam evaluasi pembelajaran PAI masih menghadapi tantangan pedagogis. Guru cenderung menggunakan bentuk evaluasi yang praktis dan mudah disusun dibanding merancang instrumen yang menuntut kemampuan berpikir tingkat tinggi.

Dari berbagai penelitian yang dianalisis, terlihat bahwa dominasi LOTS bukan hanya persoalan teknis penyusunan soal, tetapi juga berkaitan dengan paradigma evaluasi pembelajaran yang masih konvensional. Evaluasi masih dipahami sebagai alat mengukur hasil akhir pembelajaran, bukan sebagai sarana membangun kemampuan berpikir peserta didik. Oleh sebab itu, optimalisasi evaluasi ranah kognitif memerlukan perubahan paradigma evaluasi menuju pendekatan yang lebih reflektif, kontekstual, dan berbasis HOTS.

4.2 Kualitas Instrumen dan Akurasi Evaluasi Pembelajaran

Hasil penelitian juga menunjukkan bahwa kualitas instrumen penilaian memiliki hubungan yang sangat erat dengan akurasi evaluasi pembelajaran. Validitas, reliabilitas, tingkat kesukaran, dan daya pembeda menjadi indikator utama dalam menentukan apakah suatu instrumen mampu mengukur kemampuan peserta didik secara tepat.

Dalam teori evaluasi pendidikan, instrumen yang valid berarti mampu mengukur kompetensi yang seharusnya diukur, sedangkan reliabilitas berkaitan dengan konsistensi hasil pengukuran. Ketika instrumen tidak memenuhi kedua aspek tersebut, hasil evaluasi berpotensi menghasilkan bias dan tidak mencerminkan kemampuan peserta didik secara objektif. Temuan (Amalia, 2023) menunjukkan bahwa instrumen evaluasi PAI masih belum memiliki keseimbangan indikator HOTS dan LOTS secara proporsional. Hal ini menunjukkan bahwa kualitas evaluasi pembelajaran belum sepenuhnya didukung oleh kualitas instrumen yang memadai.

Kondisi tersebut memiliki implikasi yang cukup luas dalam proses pembelajaran. Evaluasi yang tidak akurat dapat memengaruhi pengambilan keputusan guru, baik dalam menentukan strategi pembelajaran, tindak lanjut evaluasi, maupun pemetaan kemampuan peserta didik. Dalam praktik pembelajaran PAI, instrumen yang kurang berkualitas sering kali menghasilkan penilaian yang

hanya menekankan aspek hafalan, padahal tujuan pembelajaran agama juga mencakup pembentukan kemampuan berpikir reflektif dan pemahaman kontekstual.

Penelitian (Diki et al., 2022) memperlihatkan bahwa pengembangan instrumen berbasis kemampuan berpikir analitis mampu meningkatkan ketepatan evaluasi pembelajaran. Temuan tersebut menunjukkan bahwa kualitas instrumen tidak dapat dipisahkan dari kualitas hasil evaluasi. Semakin baik kualitas instrumen, maka semakin akurat pula informasi yang diperoleh mengenai perkembangan kemampuan peserta didik.

Dari hasil sintesis penelitian dapat dipahami bahwa kualitas instrumen penilaian bukan sekadar aspek administratif dalam evaluasi pembelajaran. Instrumen justru menjadi fondasi utama dalam menentukan efektivitas evaluasi. Oleh sebab itu, optimalisasi evaluasi ranah kognitif dalam pembelajaran PAI harus dimulai dari pengembangan instrumen yang valid, reliabel, dan mampu mengukur kemampuan berpikir tingkat tinggi secara komprehensif.

4.3 Implementasi HOTS dalam Evaluasi Pembelajaran PAI

Perkembangan penelitian evaluasi pembelajaran PAI menunjukkan bahwa implementasi Higher Order Thinking Skills (HOTS) mulai menjadi perhatian penting dalam pengembangan instrumen penilaian. Hal tersebut sejalan dengan tuntutan pendidikan modern yang menempatkan kemampuan berpikir kritis, kreatif, dan problem solving sebagai kompetensi utama peserta didik.

Namun demikian, hasil penelitian memperlihatkan bahwa implementasi HOTS dalam evaluasi PAI masih belum optimal. Banyak instrumen evaluasi yang secara formal mengklaim berbasis HOTS, tetapi pada praktiknya masih berada pada level memahami dan mengingat. Situasi tersebut menunjukkan bahwa pemahaman terhadap konsep HOTS belum sepenuhnya diikuti kemampuan teknis dalam menyusun instrumen evaluasi.

Menurut (Zahra & At-Taqiyyah, 2024), evaluasi berbasis HOTS seharusnya mampu mengarahkan peserta didik pada kemampuan berpikir reflektif dan pemecahan masalah kontekstual. Dalam pembelajaran PAI, kemampuan tersebut sangat penting karena peserta didik tidak hanya dituntut memahami ajaran agama secara teoritis, tetapi juga mampu mengimplementasikan nilai-nilai Islam dalam kehidupan sehari-hari.

Jika dikaji lebih mendalam, implementasi HOTS dalam evaluasi PAI sebenarnya memiliki hubungan erat dengan pendekatan evaluasi autentik. Evaluasi autentik memungkinkan peserta didik menunjukkan kemampuan berpikir melalui analisis kasus, pemecahan masalah, maupun interpretasi nilai-nilai keagamaan dalam konteks sosial. Oleh sebab itu, instrumen HOTS tidak cukup hanya diwujudkan melalui soal sulit, tetapi harus mampu menghadirkan situasi yang menuntut peserta didik berpikir secara mendalam.

Temuan ini memperlihatkan bahwa pengembangan evaluasi berbasis HOTS memerlukan kesiapan pedagogis guru, penguasaan konstruksi soal, serta perubahan paradigma evaluasi pembelajaran. Evaluasi tidak lagi dipahami sekadar alat pengukur hasil belajar, tetapi menjadi bagian dari proses pembentukan kemampuan berpikir peserta didik secara berkelanjutan.

4.4 Optimalisasi Evaluasi Ranah Kognitif dalam Pembelajaran PAI

Berdasarkan keseluruhan hasil penelitian, optimalisasi evaluasi ranah kognitif dalam pembelajaran PAI tidak dapat dilakukan hanya dengan memperbaiki bentuk soal semata. Optimalisasi evaluasi memerlukan integrasi antara kualitas instrumen penilaian, implementasi HOTS, dan pendekatan evaluasi autentik dalam proses pembelajaran.

Selama ini, sebagian penelitian lebih banyak berfokus pada analisis teknis instrumen seperti validitas, reliabilitas, dan tingkat kesukaran soal. Meskipun aspek tersebut penting, evaluasi pembelajaran sebenarnya memiliki dimensi yang lebih luas. Evaluasi bukan hanya mengukur

capaian belajar, tetapi juga menentukan bagaimana guru memahami perkembangan kemampuan berpikir peserta didik.

Dari berbagai penelitian yang dianalisis terlihat bahwa kualitas instrumen penilaian memiliki posisi strategis dalam menentukan efektivitas evaluasi pembelajaran. Instrumen yang baik akan menghasilkan informasi yang lebih autentik mengenai kemampuan peserta didik, sedangkan instrumen yang lemah berpotensi menghasilkan bias pengukuran. Oleh karena itu, kualitas instrumen perlu dipahami sebagai bagian integral dari optimalisasi evaluasi pembelajaran PAI.

Novelty penelitian ini terletak pada upaya mengintegrasikan analisis kualitas instrumen dengan optimalisasi evaluasi ranah kognitif dalam pembelajaran PAI. Penelitian sebelumnya cenderung membahas kedua aspek tersebut secara terpisah. Sementara itu, penelitian ini menunjukkan bahwa kualitas instrumen, implementasi HOTS, dan efektivitas evaluasi pembelajaran merupakan satu kesatuan yang saling berkaitan.

Berdasarkan sintesis konseptual yang diperoleh, optimalisasi evaluasi ranah kognitif dalam pembelajaran PAI dapat dibangun melalui tiga komponen utama, yaitu pengembangan instrumen berbasis HOTS, pengujian kualitas instrumen secara sistematis, dan penerapan evaluasi autentik dalam pembelajaran. Ketiga komponen tersebut saling mendukung dalam menghasilkan evaluasi yang lebih reflektif, kontekstual, dan mampu menggambarkan kemampuan peserta didik secara komprehensif.

Dengan demikian, evaluasi pembelajaran PAI tidak lagi dipahami hanya sebagai proses pemberian nilai, tetapi sebagai bagian dari upaya membangun kualitas pembelajaran dan pengembangan kemampuan berpikir peserta didik secara berkelanjutan.

5. Kesimpulan

Hasil penelitian ini menegaskan bahwa optimalisasi evaluasi ranah kognitif dalam pembelajaran Pendidikan Agama Islam (PAI) tidak dapat dipisahkan dari kualitas instrumen penilaian yang digunakan. Analisis berbagai literatur mengungkap satu pola yang berulang: instrumen evaluasi PAI masih terlalu berat ke arah pengukuran kemampuan berpikir tingkat rendah atau *Lower Order Thinking Skills* (LOTS), sehingga kemampuan berpikir kritis, analitis, dan reflektif peserta didik tidak terukur secara memadai apalagi dikembangkan. Ini bukan sekadar persoalan teknis penyusunan soal. Kualitas instrumen penilaian adalah faktor penentu apakah evaluasi pembelajaran benar-benar mencerminkan kemampuan peserta didik atau hanya mengukur daya ingat jangka pendek mereka. Penelitian ini juga menemukan bahwa validitas, reliabilitas, tingkat kesukaran soal, dan penerapan *Higher Order Thinking Skills* (HOTS) merupakan komponen yang tidak bisa ditawar dalam membangun evaluasi kognitif yang komprehensif dan autentik. Dengan kata lain, optimalisasi evaluasi pembelajaran PAI tidak cukup diupayakan melalui peningkatan hasil belajar semata. Ia menuntut pengembangan instrumen evaluasi yang sistematis, adaptif, dan benar-benar kontekstual terhadap kebutuhan pembelajaran abad ke-21 bukan instrumen yang diadaptasi seadanya dari generasi kurikulum sebelumnya.

Kontribusi teoretis penelitian ini terletak pada penguatan kajian evaluasi ranah kognitif dalam pembelajaran PAI melalui sintesis konseptual mengenai kualitas instrumen penilaian berbasis HOTS dan analisis kualitas soal secara terintegrasi. Selama ini, dua hal itu sering dikaji secara terpisah penelitian ini berusaha menyatukannya dalam satu kerangka yang lebih koheren. Secara praktis, temuan penelitian ini membawa implikasi langsung bagi guru, lembaga pendidikan, dan pengembang kurikulum: sudah saatnya instrumen evaluasi disusun bukan sekadar memenuhi syarat administratif, melainkan benar-benar dirancang untuk mengukur kemampuan berpikir tingkat tinggi peserta didik secara autentik. Implementasi evaluasi berbasis HOTS berpotensi mendorong transformasi pembelajaran PAI dari pendekatan yang bertumpu pada hafalan menuju pembelajaran yang lebih reflektif dan kontekstual perubahan yang dibutuhkan, tetapi tidak akan

terjadi dengan sendirinya tanpa instrumen yang mendukungnya. Penelitian ini pun membuka jalur bagi studi lanjutan yang lebih empiris, misalnya analisis kualitas instrumen evaluasi PAI berbasis teknologi digital, pemodelan Rasch, maupun asesmen autentik yang mempertimbangkan konteks sosial-keagamaan peserta didik secara lebih spesifik. Pada akhirnya, optimalisasi evaluasi ranah kognitif perlu diperlakukan sebagai bagian yang tidak terpisahkan dari upaya peningkatan mutu pembelajaran Pendidikan Agama Islam secara berkelanjutan bukan sebagai agenda sampingan yang dikerjakan jika ada waktu.

Referensi

- Adzillah, S. N. I., & Hamami, T. (2025). Pengembangan Asesmen Kognitif dan Psikomotorik Agama Siswa. *El-Ihsan: Journal of Contemporary Islamic Education*, 1(2), 90–109.
- Amalia, J. (2023). ANALISIS KUALITAS INSTRUMEN ASESMEN RANAH KOGNITIF PENDIDIKAN AGAMA ISLAM DI UNIVERSITAS NEGERI YOGYAKARTA. *Dar El-Ilmi : Jurnal Studi Keagamaan, Pendidikan Dan Humaniora*, 10(1), 1–22. <https://doi.org/10.52166/darelilmi.v10i1.3787>
- Astuti, F. (2021). Analisis Ranah Kognitif Taksonomi Bloom Revisi pada Soal Ujian Sekolah Bahasa Jawa. *Piwulang : Jurnal Pendidikan Bahasa Jawa*, 9(1), Article 1. <https://doi.org/10.15294/piwulang.v9i1.47031>
- Diki, D. M., Farida, I., & Hardiansyah, D. (2022). Pengembangan Instrumen Tes Kemampuan Berpikir Analisis Pada Materi Akhlak Terpuji Dalam Pergaulan Remaja. *Al-Hikmah: Jurnal Agama Dan Ilmu Pengetahuan*, 19(1), 95–108. [https://doi.org/10.25299/al-hikmah:jaip.2022.vol19\(1\).9094](https://doi.org/10.25299/al-hikmah:jaip.2022.vol19(1).9094)
- Efendi, R. A., Rudi, H., & Prasetya, I. (2025). Pengembangan Instrumen Evaluasi Pembelajaran Pendidikan Agama Islam Berbasis Multimedia dengan Menggunakan Wondershare Quiz Creator di SMA se-Kabupaten Tapanuli Tengah. *Jurnal Mudarrisuna: Media Kajian Pendidikan Agama Islam*, 14(4). <https://jurnal.ar-raniry.ac.id/index.php/mudarrisuna/article/view/24099>
- Elviana. (2020). Analisis Butir Soal Evaluasi Pembelajaran PAI Menggunakan Program Anates. *Jurnal Mudarrisuna: Media Kajian Pendidikan Agama Islam*, 10(2), 195–214.
- Erviana, R., Fadilla, A., Sutopo, S., & Musrikah, M. (2025). Analisis Butir Soal Sumatif Akhir Semester (SAS) Mata Pelajaran Matematika MTs Maftahul Ulum Karangsono. *AL KHAWARIZMI: Jurnal Pendidikan Matematika*, 5(2). <https://doi.org/10.46368/kjpm.v5i2.3529>
- Fitriana, R., & Mustofa, T. A. (2023). Teknik evaluasi kognitif dalam pembelajaran pendidikan agama islam di SMK Negeri 6 Surakarta. *At Tuots: Jurnal Pendidikan Islam*, 976–983. <https://doi.org/10.51468/jpi.v5i2.312>

- Hamzah, A., & Mudlofir, A. (2025). Menulik Tantangan Guru Pendidikan Agama Islam Dalam Pengembangan Kuriukulum Berbasis Teknologi Pada Era Society 5.0. *Jurnal Ilmiah Profesi Pendidikan*, 10(3), 2659–2669. <https://doi.org/10.29303/jipp.v10i3.3485>
- Harsoyo, F. H., Ramadhan, N., Mutiah, S. H., Akil, & Aziz, A. (2025). ANALISIS TEKNIK PENGOLAHAN HASIL EVALUASI PEMBELAJARAN PENDIDIKAN AGAMA ISLAM (SKOR TES, SKOR TOTAL, DAN KONVERSI NILAI) BERBASIS BUDAYA LOKAL. *Edukreatif: Jurnal Kreativitas dalam Pendidikan*, 6(3). <https://ijurnal.com/1/index.php/jkp/article/view/811>
- Hilmiyati, F., Pratiwi, H., & Fajriah, N. (2025). Implementasi Evaluasi Pembelajaran PAI pada Kurikulum Merdeka Belajar di Sekolah Penggerak Provinsi Banten. *Jurnal Mudarrisuna: Media Kajian Pendidikan Agama Islam*, 15(3). <https://jurnal.ar-raniry.ac.id/index.php/mudarrisuna/article/view/23246>
- Maulina, W. (n.d.). *Pengembangan Instrumen Asesmen Berbasis Hots pada Mata Pelajaran Pendidikan Agama Islam dan Budi Pekerti di SMAN 6 Padang*.
- Meiliani, R., Supriyanto, T., Astuti, T., & Yulianto, S. (2023). ANALISIS INSTRUMEN PENILAIAN AUTENTIK PADA BUKU SISWA KELAS V TEMA 6 EDISI REVISI 2017. *Jurnal Binagogik*, 10(1), 89–97. <https://doi.org/10.61290/pgsd.v10i1.15>
- Muti'ah, N., & Maemonah, M. (2022). Analisi Tes Butir Soal Uraian Pada Pembelajaran PAI Dengan Menggunakan Taksonomi Bloom Di Sd N Gedongtengen, Yogyakarta. *Raudhah Proud To Be Professionals : Jurnal Tarbiyah Islamiyah*, 7(1), 192–206. <https://doi.org/10.48094/raudhah.v7i1.533>
- Nuryani, A., Ropikoh, R., & Aziz, A. (2024). Evaluasi Hasil Belajar dalam Pembelajaran Pendidikan Agama Islam. *Jurnal Kreativitas Mahasiswa*, 2(1), 83–97.
- Pudjiati, I., & Madani, F. (n.d.). Asesmen Autentik Analisis Butir Soal Dengan Rasch Model Di Sekolah Dasar: Literature Review. *Mutiara : Jurnal Penelitian Dan Karya Ilmiah*, 1(4), 01. <https://doi.org/10.59059/MUTIARA.V1I4.325>
- Rahmawati, N. (2021). Analisis Tingkat Kognitif Soal Penilaian Akhir Tahun (Pat) Pendidikan Agama Islam Kelas X Berbasis Akm Berdasarkan Taksonomi Bloom DI Sma Antartika Sidoarjo. *Al Ta'dib : Jurnal Ilmu Pendidikan*. https://www.academia.edu/91612537/Analisis_Tingkat_Kognitif_Soal_Penilaian_Akhir_Tahun_Pat_Pendidikan_Agama_Islam_Kelas_X_Berbasis_Akm_Berdasarkan_Taksonomi_Bloom_DI_Sma_Antartika_Sidoarjo

- Ramadhan, M. R., & Subando, J. (2025). ANALISIS KUALITAS BUTIR SOAL FIQIH DAN KEMAMPUAN SISWA DI MADRASAH ALIYAH NEGERI 1 SURAKARTA. *Al'ulum Jurnal Pendidikan Islam*, 126–138. <https://doi.org/10.54090/alulum.698>
- Rizqillah, S., Ayu, D. K., Sulaiman, A., & Irma, C. N. (2023). Analisis Butir Soal PTS Bahasa Indonesia Kelas IX SMP Muhammadiyah Boarding School Bumiayu Menggunakan Model Evaluasi CIPP. *Diskursus: Jurnal Pendidikan Bahasa Indonesia*, 6(2), 176–193. <https://doi.org/10.30998/diskursus.v6i2.19705>
- Rosyidi, D. (2020). Teknik dan Instrumen Asesmen Ranah Kognitif. *TASYRI': JURNAL TARBIYAH-SYARI'AH ISLAMIYAH*, 27(1), 1–13.
- Sugiyono. (2022). *Metode Penelitian Kualitatif* (Edisi Revisi). Alfabeta.
- Suliswiyadi, S. (2020). Hierarki Ranah Pembelajaran Afektif Pendidikan Agama Islam dalam Perspektif Taksonomi Qur'ani. *Jurnal Tarbiyatuna*, 11(1), 61–76. <https://doi.org/10.31603/tarbiyatuna.v11i1.3451>
- Supriadi, Sahri, & Hasmy, A. (2025). PENGEMBANGAN INSTRUMEN PENILAIAN PEMBELAJARAN PAI DENGAN PENDEKATAN HIGHER ORDER THINKING SKILLS (HOTS) PADA SISWA TINGKAT DASAR. *Jurnal Pendidikan dan Pengajaran*, 6(3). <https://ijurnal.com/1/index.php/jpp/article/view/911>
- Yanti, N., Ikbali, M., Siskawati, Z., Yanti, N., Ikbali, M., & Siskawati, Z. (2022). Pelaksanaan Penilaian Ranah Kognitif pada Mata Pelajaran Pendidikan Agama Islam di SMK Negeri 1 Mandau. *Al-Mutharahah: Jurnal Penelitian Dan Kajian Sosial Keagamaan*, 19(2), 171–183. <https://doi.org/10.46781/al-mutharahah.v19i2.527>
- Yusuf, A. M. (2020). *Metode Penelitian: Kuantitatif, Kualitatif, dan Penelitian Gabungan* (Edisi Revisi). Kencana.
- Zahra, A. A., & At-Taqiyyah, A. K. (2024). Mengukur Kemampuan Kognitif Tingkat Tinggi dengan Pertanyaan Pilihan Ganda. *Jurnal Pelita Nusantara*, 1(4), 462–470. <https://doi.org/10.59996/jurnalpelitanusantara.v1i4.336>
- Zakkiyah, F., Armedi, R., & Zuhriyah, I. A. (2024). Analisis Instrumen Tes Kognitif Mata Pelajaran Pendidikan Agama Islam di Sekolah Dasar. *Jurnal Evaluasi Pendidikan*, 15(2), 63–75. <https://doi.org/10.21009/jep.v15i2.50143>