GECI
Jurnal Generasi Ceria Indonesia
Volume: 3 | Nomor 2 | Nopember 2025 | E-ISSN : 2987-9264 | DOI: 10.47709/geci.v3i2.9768

Hubungan Persepsi Aktivitas HOTS dengan Persepsi Berpikir Kritis Mahasiswa PAI

Author:
Neli Hidayah[1]
Uray Ferry Haryanto[2]
Ahmad Fadhil Rizki[3]

Afiliation:
[1,2,3] Institut Agama Islam Miftahul Ulum Tanjungpinang

Corresponding email
neli.hidayah3008@gmail.com

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License

Abstrak:

Latar belakang: Penilaian pembelajaran Pendidikan Agama Islam (PAI) perlu membedakan tuntutan soal HOTS, pengalaman berpikir yang dilaporkan mahasiswa, dan kemampuan berpikir kritis yang ditunjukkan melalui kinerja. Penelitian ini bertujuan mendeskripsikan dua skor pelaporan diri dan menelaah hubungan yang dilaporkan di antara keduanya.

Metode penelitian: Penelitian kuantitatif deskriptif korelasional dilaksanakan pada April–Juni 2025 di IAI Miftahul Ulum Tanjungpinang. Dari populasi 233 mahasiswa PAI kelas pagi dan malam, penelitian melaporkan 80 responden yang dipilih secara purposif. Dua angket masing-masing terdiri atas 15 pernyataan Likert lima tingkat.

Hasil penelitian: Penghitungan berdasarkan total skor menghasilkan rerata total 56,53 untuk X dan 55,11 untuk Y, masing-masing pada rentang teoritis 15–75 jika respons lengkap dan penskoran sesuai. Koefisien yang dilaporkan sebagai korelasi Spearman ialah 0,774 dengan $p < 0,001$. Arah asosiasi tersebut konsisten dengan keterkaitan dua penilaian diri, tetapi belum memisahkan hubungan konstruk dari kemiripan isi dan metode pengukuran. Daftar butir, bukti validitas dan reliabilitas, rincian seleksi responden, data berpasangan, serta keluaran statistik belum tersedia untuk pemeriksaan ulang.

Kesimpulan: Kesimpulan dibatasi pada informasi skor yang dilaporkan; keabsahan inferensi hubungan masih memerlukan verifikasi empiris. Penelitian belum membuktikan kualitas soal dosen, kemampuan berpikir kritis aktual, atau pengaruh kausal penerapan HOTS.

Kata kunci: berpikir kritis; HOTS; korelasi Spearman; pelaporan diri; Pendidikan Agama Islam

Pendahuluan

Pembelajaran Pendidikan Agama Islam (PAI) di perguruan tinggi menuntut mahasiswa menelaah informasi, memeriksa alasan, dan merumuskan kesimpulan yang dapat dipertanggungjawabkan. Tugas Higher Order Thinking Skills (HOTS) dapat menyediakan kesempatan untuk melakukan analisis, evaluasi, dan kreasi. Akan tetapi, tuntutan kognitif tugas, pengalaman mahasiswa mengerjakannya, dan mutu penalaran dalam jawaban merupakan objek pengukuran yang berbeda. Pemisahan objek tersebut menentukan kesimpulan yang dapat ditarik dari penelitian tentang HOTS dan berpikir kritis.

Baharsyah et al. (2023) memeriksa penyelesaian dan penyusunan soal HOTS oleh mahasiswa PAI melalui soal, rancangan soal, angket, dan wawancara. Studi kasus tersebut menunjukkan kesulitan pada tuntutan evaluasi dan kreasi, sehingga memberi perhatian pada bukti pekerjaan mahasiswa. Ananda dan Maemonah (2022) membahas implementasi asesmen HOTS dalam pembelajaran PAI berbasis proyek di SMP melalui wawancara dan dokumentasi. Keduanya menghubungkan asesmen dengan praktik atau hasil pekerjaan, tetapi berbeda dalam jenjang, desain, dan unit analisis; hasilnya tidak dapat langsung digunakan untuk menyimpulkan kondisi semua mahasiswa PAI.

Di sisi konseptual, Loyens et al. (2023) menunjukkan pentingnya memperjelas definisi dan pengukuran berpikir tingkat tinggi, berpikir kritis, serta berpikir kritis analitis dalam lingkungan pembelajaran berbasis

Volume: 3 | Nomor 2 | Nopember 2025 | E-ISSN : 2987-9264 | DOI: 10.47709/geci.v3i2.9768

masalah dan proyek. Kedekatan konsep membuat hubungan antarukuran perlu ditafsirkan dengan hati-hati. Validasi konstruk juga tidak cukup diwakili oleh satu indeks konsistensi internal, sebagaimana ditekankan Flake et al. (2017). Dengan demikian, persoalan ilmiahnya bukan semata apakah dua skor berkorelasi, melainkan apakah keduanya mengandung informasi yang dapat dibedakan.

Penelitian ini menempatkan persepsi mahasiswa sebagai sumber informasi tentang bagaimana mereka menilai aktivitas berpikir dalam pembelajaran PAI dan penalaran dirinya. Informasi tersebut penting bagi evaluasi pembelajaran karena dapat memperlihatkan apakah dua penilaian mahasiswa bergerak searah. Namun, penilaian diri belum menunjukkan kesesuaian antara keyakinan dan kemampuan aktual. Korelasi kedua angket hanya dapat memberi tambahan informasi tentang hubungan persepsi apabila isi, reliabilitas, dan pemisahan ukurannya dapat dipertanggungjawabkan.

Celah yang menjadi fokus ialah perbedaan antara bukti tugas atau kinerja dalam penelitian terdahulu dan hubungan penilaian diri dalam konteks mahasiswa PAI yang diteliti. Kontribusi yang dituju berupa deskripsi keterkaitan persepsi dengan batas pengukuran yang jelas, bukan penemuan bahwa HOTS dan berpikir kritis berkaitan secara umum. Penelitian tidak mengklaim bahwa lokasi institusi merupakan kebaruan substantif, dan tidak menggunakan dugaan dominasi hafalan atau rendahnya kemampuan mahasiswa sebagai fakta awal tanpa bukti.

Dua pertanyaan diajukan: bagaimana skor angket yang dilaporkan untuk aktivitas HOTS dan berpikir kritis; serta bagaimana arah dan besaran hubungan yang dilaporkan di antara keduanya. Label X dan Y dipakai untuk menandai kelompok skor, bukan untuk menetapkan hubungan sebab-akibat. Hipotesis statistik mengenai asosiasi peringkat perlu dibedakan dari pertanyaan validitas konstruk; hasil korelasi yang signifikan tidak dengan sendirinya menjawab apakah kedua angket mengukur hal yang berbeda.

Studi Literatur

Tuntutan HOTS dan pengalaman mahasiswa

HOTS merujuk pada tuntutan berpikir yang melampaui pengingatan informasi, termasuk analisis, evaluasi, dan kreasi. Tuntutan tersebut perlu terlihat pada stimulus, pertanyaan, alasan yang diminta, dan kriteria penilaian. Label HOTS atau penggunaan kata kerja tertentu belum cukup untuk membuktikan kualitas soal. Pemeriksaan dokumen tugas dan jawaban mahasiswa, seperti yang menjadi perhatian Baharsyah et al. (2023) serta Ananda dan Maemonah (2022), perlu dibedakan dari pernyataan mahasiswa bahwa dirinya dapat melakukan aktivitas tersebut.

Dalam naskah ini, X dibatasi sebagai skor persepsi aktivitas HOTS karena uraian indikator yang tersedia banyak merujuk pada penguraian informasi, penilaian solusi, dan penyusunan pemahaman. X tidak langsung diperlakukan sebagai ukuran frekuensi pemberian atau mutu soal dosen. Pembatasan istilah ini memperjelas jangkauan interpretasi; ia tidak menggantikan pemeriksaan butir asli dan tidak mengesahkan konstruk baru.

Berpikir kritis dan pemisahan konstruk

Facione (1990) merumuskan berpikir kritis melalui interpretasi, analisis, evaluasi, inferensi, penjelasan, dan regulasi diri. Kerangka tersebut mengarahkan perhatian pada mutu alasan dan penggunaan bukti. Penelitian ini tidak menyatakan menggunakan tes standar Facione. Y merupakan angket dengan cakupan penjelasan sederhana, keterampilan dasar, penyimpulan, penjelasan lanjutan, serta strategi dan taktik. Persetujuan terhadap pernyataan kemampuan merupakan penilaian diri; kemampuan aktual memerlukan jawaban atau tindakan yang dinilai menggunakan kriteria.

Pembedaan konstruk menuntut alasan konseptual serta bukti empiris pada ukurannya. Rönkkö dan Cho (2022) membedakan validitas diskriminan dari sekadar penggunaan label yang berbeda atau korelasi skor yang belum memperhitungkan kesalahan pengukuran. Karena itu, korelasi 0,774 tidak dapat dipakai

GECI
Jurnal Generasi Ceria Indonesia
Volume: 3 | Nomor 2 | Nopember 2025 | E-ISSN : 2987-9264 | DOI: 10.47709/geci.v3i2.9768

sekaligus sebagai bukti adanya hubungan dua konstruk dan sebagai bukti bahwa keduanya terpisah. Kemiripan analisis, evaluasi, dan penalaran pada X serta Y harus diperiksa pada tingkat butir.

Metode Penelitian

Desain konteks dan responden

Penelitian menggunakan pendekatan kuantitatif deskriptif korelasional di IAI Miftahul Ulum Tanjungpinang pada April–Juni 2025. Populasi yang dilaporkan mencakup 233 mahasiswa PAI kelas pagi dan malam tahun akademik 2024/2025. Jumlah responden yang dilaporkan ialah 80 mahasiswa, dipilih melalui purposive sampling. Desain tidak mencakup perlakuan, kelompok pembanding, atau pengukuran awal–akhir, sehingga tidak digunakan untuk menguji perubahan kemampuan maupun pengaruh kausal.

Jumlah 80 setara dengan 34,33% populasi yang dilaporkan. Proporsi ini hanya mendeskripsikan cakupan jumlah, bukan membuktikan keterwakilan atau kecukupan sampel. Dasar penetapan ukuran sampel perlu berhubungan dengan tujuan inferensi dan proses penelitian yang sebenarnya (Lakens, 2022). Dalam informasi yang tersedia, alasan penetapan 80 responden, kriteria purposif yang diterapkan, jumlah diundang dan nonrespons, serta distribusi kelas, semester, dan mata kuliah belum terdokumentasi. Karena itu, penelitian tidak mengklaim sampel acak, perbandingan kelas pagi–malam, atau keterwakilan seluruh mahasiswa PAI.

Instrumen dan pembentukan skor

Data utama dilaporkan berasal dari dua angket, masing-masing 15 pernyataan. Alternatif respons ialah sangat setuju (5), setuju (4), netral (3), tidak setuju (2), dan sangat tidak setuju (1). Jika seluruh item lengkap dan arah skor telah ditetapkan dengan benar, rentang teoritis total per responden ialah 15–75 pada setiap angket. Jumlah maksimum bagi 80 responden ialah 6.000. Pernyataan negatif, jika ada dalam instrumen asli, memerlukan pembalikan skor sebelum penjumlahan; keberadaan dan prosedurnya belum dapat dikonfirmasi dari daftar butir.

Tabel 1. Objek pengukuran dan risiko tumpang tindih

Kelompok skor	Cakupan yang dilaporkan	Batas interpretasi
X Persepsi aktivitas HOTS	Penguraian informasi, pengenalan hubungan, penggunaan pertanyaan, penilaian solusi, dan penyusunan pemahaman.	Belum menjadi ukuran kualitas atau frekuensi soal dosen; beberapa indikator berdekatan dengan penilaian kemampuan diri.
Y Persepsi berpikir kritis	Penjelasan, pemeriksaan informasi, penyimpulan, penjelasan lanjutan, serta strategi dan taktik.	Belum menjadi ukuran kinerja penalaran melalui tes; potensi kemiripan dengan X perlu ditelaah pada setiap butir.

Tabel 1 merupakan pemetaan cakupan yang dilaporkan, bukan kisi-kisi lengkap atau hasil validasi. Daftar 30 butir beserta sumber, dimensi, dan arah penskorannya belum tersedia untuk pemeriksaan ulang. Oleh sebab itu, tidak disusun butir pengganti yang kemudian diklaim sebagai instrumen penelitian terdahulu. Wawancara, observasi, dan dokumentasi yang disebut sebagai teknik pendukung juga tidak dijadikan bukti triangulasi karena protokol serta hasilnya belum disertakan.

Pengujian instrumen disebut menggunakan IBM SPSS Statistics, tetapi hasil korelasi item–total, peserta uji coba, keputusan item, dan Cronbach's alpha masing-masing angket belum tersedia. Penyebutan perangkat lunak tidak membuktikan kelayakan instrumen. Cronbach's alpha berkaitan dengan konsistensi internal dan tidak menggantikan bukti validitas atau pemisahan konstruk (Tavakol & Dennick, 2011). Karena data per item belum tersedia, revisi ini tidak melaporkan hasil uji validitas, reliabilitas, atau validitas diskriminan yang baru.

GECI
Jurnal Generasi Ceria Indonesia
Volume: 3 | Nomor 2 | Nopember 2025 | E-ISSN : 2987-9264 | DOI: 10.47709/geci.v3i2.9768

Analisis dan status pemeriksaan ulang

Analisis deskriptif menggunakan total skor yang dilaporkan. Persentase terhadap maksimum dihitung dengan rumus total skor dibagi 6.000, dikalikan 100%; rerata total diperoleh dengan membagi total skor dengan 80. Semua hasil dibulatkan dua desimal. Penghitungan tersebut bersifat aritmetika dan mengasumsikan respons lengkap serta penskoran yang benar. Persentase tidak diperlakukan sebagai proporsi mahasiswa yang kompeten atau diberi kategori baik dan kurang tanpa dasar ambang yang jelas.

Analisis hubungan yang disebut dalam pelaporan akhir ialah korelasi Spearman. Spearman mengukur asosiasi monotonik berdasarkan peringkat, berbeda dari hubungan linear yang diringkas Pearson (Schober et al., 2018). Penggunaannya perlu mempertimbangkan pembentukan skor, independensi observasi, peringkat yang sama, nilai ekstrem, serta pola hubungan kedua skor. Distribusi tidak normal bukan satu-satunya pertimbangan dan tidak menjamin bahwa hubungan bersifat monotonik.

Nilai Kolmogorov–Smirnov $p = 0{,}048$ yang sebelumnya disebut berasal dari residual tidak digunakan untuk menetapkan distribusi X atau Y maupun membenarkan pemilihan Spearman. Model pembentuk residual dan keluaran uji tersebut tidak tersedia. Karena data berpasangan serta output asli belum disertakan, koefisien korelasi tidak dihitung ulang; pemeriksaan monotonisitas, interval kepercayaan, dan jumlah pasangan valid juga belum dapat dilakukan. Pelaporan 0,000 ditulis $p < 0{,}001$, sementara jenis uji satu atau dua sisi masih memerlukan konfirmasi.

Hasil

Deskripsi skor yang dapat dihitung dari rekap

Total skor yang dilaporkan ialah 4.522 untuk X dan 4.409 untuk Y. Tabel 2 menyajikan hasil penghitungan berdasarkan 80 responden, 15 item per angket, dan skor maksimum item sebesar lima. Hanya rasio dan rerata yang dapat diturunkan dari rekap tersebut; statistik sebaran serta distribusi antarresponden tidak dapat direkonstruksi dari total.

Tabel 2. Penghitungan deskriptif dari total skor yang dilaporkan

Kelompok skor	Total skor	Maksimum teoritis	Persentase maksimum	Rerata total
X Persepsi aktivitas HOTS	4.522	6.000	75,37%	56,53
Y Persepsi berpikir kritis	4.409	6.000	73,48%	55,11

Catatan: rerata total berada pada rentang teoritis 15–75 jika seluruh 15 item terisi dan arah skor benar. Sumber total ialah rekap dalam naskah, bukan data individual yang telah diaudit. Persentase maksimum tidak menggunakan titik nol kemampuan dan tidak menunjukkan persentase mahasiswa yang menguasai kompetensi.

Perbedaan kedua rerata tidak digunakan untuk menyatakan bahwa kemampuan HOTS lebih tinggi daripada berpikir kritis, karena kesetaraan isi dan metrik kedua angket belum dibuktikan. Rerata, simpangan baku, minimum dan maksimum empiris, serta sebaran kelas pagi dan malam memerlukan data individual. Uraian persetujuan pada butir tertentu tidak dipertahankan sebagai temuan dimensi karena frekuensi setiap kategori dan redaksi butir belum tersedia untuk pemeriksaan.

GECI
Jurnal Generasi Ceria Indonesia
Volume: 3 | Nomor 2 | Nopember 2025 | E-ISSN : 2987-9264 | DOI: 10.47709/geci.v3i2.9768

Pelaporan hubungan antarskor

Tabel 3. Ringkasan korelasi dalam pelaporan penelitian

Komponen	Nilai yang dilaporkan	Status pemeriksaan
Metode korelasi	Spearman	Perlu dikonfirmasi pada keluaran Correlations asli.
Koefisien	0,774	Belum dihitung ulang dari pasangan skor individual.
Signifikansi	$p < 0,001$	Penulisan ulang dari 0,000; jenis uji perlu dikonfirmasi.
Responden	80 mahasiswa	Jumlah responden yang dilaporkan; N pasangan valid belum dikonfirmasi.

Secara deskriptif, tanda positif pada koefisien menunjukkan kecenderungan searah kedua skor menurut pelaporan tersebut. Namun, keluaran yang dapat ditelusuri belum tersedia untuk memastikan nama prosedur, pasangan valid, nilai p, penanganan data hilang, dan pola monotonik. Oleh karena itu, hipotesis substantif tentang hubungan dua konstruk yang berbeda belum dinyatakan terkonfirmasi. Angka dalam Tabel 3 tidak merupakan hasil analisis ulang pada revisi ini.

Tidak disajikan koefisien determinasi sebagai proporsi kemampuan yang dijelaskan oleh HOTS. Kuadrat korelasi peringkat tidak otomatis sama dengan R^2 regresi pada skor asli, dan bukan ukuran pengaruh kausal. Klaim kontribusi 62,5% atau 65,2% tidak digunakan karena model dan output pendukungnya tidak tersedia.

Pembahasan

Makna hubungan persepsi bagi evaluasi pembelajaran PAI

Arah asosiasi yang dilaporkan dapat dibaca sebagai indikasi bahwa mahasiswa yang menilai aktivitas HOTS secara lebih positif juga cenderung menilai penalaran dirinya secara lebih positif. Ada dua penjelasan yang perlu dibedakan: pengalaman pembelajaran dapat berkaitan dengan penilaian diri, atau kemiripan butir dan cara menjawab dapat menghasilkan keterkaitan skor. Data yang tersedia belum dapat memilih di antara penjelasan tersebut. Karena itu, besarnya koefisien tidak diperlakukan sebagai bukti bahwa praktik dosen telah meningkatkan kemampuan mahasiswa.

Informasi persepsi dapat berguna untuk menentukan aspek pengalaman belajar yang perlu diperiksa lebih lanjut. Jika mahasiswa merasa sering melakukan analisis tetapi jawaban kasusnya tidak memuat alasan yang memadai, tindak lanjutnya berbeda dari kondisi ketika keduanya sejalan. Penelitian ini belum menguji kesesuaian persepsi dan kinerja tersebut. Kontribusi empiris yang dapat dipertahankan saat ini terbatas pada ringkasan skor serta asosiasi yang dilaporkan, dan belum mencakup penjelasan mekanisme pembelajaran atau ketepatan penilaian diri.

Perbandingan dengan penelitian terdahulu perlu menjaga perbedaan bukti. Baharsyah et al. (2023) menelaah pekerjaan mahasiswa, sedangkan Ananda dan Maemonah (2022) membahas praktik asesmen melalui dokumentasi dan wawancara. Respons angket yang positif dalam penelitian ini dapat berdampingan dengan kesulitan kinerja dalam konteks lain tanpa menjadi pertentangan langsung. Kajian terdahulu tersebut tidak digunakan untuk menggantikan bukti instrumen maupun mengesahkan koefisien yang belum diperiksa.

Tumpang tindih isi dan metode pengukuran

X dan Y sama-sama mencakup kegiatan analisis, evaluasi, serta penalaran. Ketika kedua angket berisi pernyataan kecakapan diri yang mirip, korelasi dapat mencerminkan kesamaan isi alih-alih hubungan antara paparan tugas dan kemampuan. Pembatasan judul menjadi persepsi membantu mencegah klaim berlebihan, tetapi tidak menyelesaikan masalah pengukuran. Pemeriksaan redaksi, struktur, dan pemisahan ukuran tetap diperlukan, sejalan dengan perhatian Loyens et al. (2023) dan Flake et al. (2017).

Volume: 3 | Nomor 2 | Nopember 2025 | E-ISSN : 2987-9264 | DOI: 10.47709/geci.v3i2.9768

Kedua skor juga diperoleh dari sumber yang sama dan menggunakan bentuk respons yang sama. Podsakoff et al. (2003) menjelaskan bahwa karakteristik metode dapat memengaruhi kovariasi pengukuran. Hal ini merupakan kemungkinan yang perlu diuji, bukan bukti bahwa seluruh korelasi disebabkan bias metode. Tanpa ukuran dari dokumen tugas atau kinerja, kontribusi metode dan hubungan substantif belum dapat dipisahkan.

Pengetahuan awal, literasi membaca, motivasi, pengalaman berdiskusi, dan dukungan dosen dapat berkaitan dengan kedua penilaian diri. Perbedaan paparan belajar antara kelas pagi dan malam juga belum dapat diperiksa. Faktor-faktor ini tidak dinyatakan telah dikendalikan. Desain korelasional dan sampling nonprobabilitas tidak menetapkan arah hubungan maupun menghilangkan penjelasan alternatif.

Kebutuhan verifikasi dan batas kontribusi

Pemeriksaan lanjutan terhadap instrumen perlu dimulai dari butir yang benar-benar digunakan, bukan penyusunan ulang setelah hasil diketahui. Redaksi, dimensi, sumber pengembangan, arah skor, dan aturan data hilang perlu dicocokkan dengan respons per item. Korelasi item–total dan konsistensi internal dapat membantu evaluasi, tetapi tidak otomatis membuktikan dua konstruk berbeda. Penilaian validitas diskriminan perlu mempertimbangkan struktur pengukuran dan ketidakpastiannya; korelasi total yang tinggi atau rendah saja tidak cukup (Rönkkö & Cho, 2022). Pemeriksaan tersebut belum dilakukan pada revisi ini karena datanya tidak disertakan.

Jika audit menunjukkan bahwa X lebih banyak mengukur kecakapan diri daripada pengalaman tugas, hasil tidak boleh dipertahankan sebagai hubungan penerapan soal dengan kemampuan. Mengubah, membuang, atau memindahkan item akan mengubah definisi skor dan memerlukan analisis ulang. Sampel 80 dengan 30 butir juga tidak otomatis memadai untuk setiap model faktor; kelayakan estimasi perlu diperiksa berdasarkan model dan karakter data, bukan diasumsikan dari rasio sederhana.

Batas utama penelitian mencakup belum tersedianya bukti instrumen, rincian pemilihan responden, data individual, serta keluaran analisis untuk pemeriksaan ulang. Pengakuan batas tersebut memperjelas status kesimpulan, tetapi tidak menggantikan perbaikan empiris. Penelitian belum menyediakan dasar yang cukup untuk menilai efektivitas soal HOTS, membandingkan kelas, atau menyimpulkan kemampuan aktual. Kepastian tentang hubungan dua konstruk baru dapat dinilai setelah masalah pengukuran dan keterlacakan statistik diselesaikan.

Kesimpulan

Rekap yang dilaporkan untuk 80 mahasiswa menghasilkan rerata total 56,53 pada angket X dan 55,11 pada angket Y, dengan rentang teoritis 15–75 apabila respons lengkap serta penskoran benar. Koefisien yang dilaporkan sebagai Spearman sebesar 0,774 dengan $p < 0,001$ memberi indikasi hubungan searah antara dua skor pelaporan diri. Namun, validitas instrumen, pemisahan konstruk, proses sampling, dan analisis berpasangan belum dapat diverifikasi. Temuan hubungan karenanya masih memerlukan konfirmasi dan belum menjadi bukti efektivitas pembelajaran atau kemampuan berpikir kritis aktual.

Implikasi penelitian ialah perlunya memisahkan penilaian pengalaman belajar dari pengukuran kemampuan sebelum menarik simpulan pedagogis. Bukti dokumen tugas dan tes penalaran akan memberi informasi yang berbeda dari angket persepsi. Penambahan kedua ukuran tersebut perlu dilaporkan sebagai tahap pengumpulan data baru apabila memang belum pernah dilakukan, dan tidak dapat dilekatkan secara retrospektif pada koefisien penelitian ini.

Daftar Pustaka

Ananda, W., & Maemonah. (2022). Implementasi asesmen kognitif berbasis HOTS materi PAI dengan pembelajaran berbasis proyek di sekolah menengah pertama. *Edukatif: Jurnal Ilmu Pendidikan, 4*(5), 6564–6575. https://doi.org/10.31004/edukatif.v4i5.3179

Volume: 3 | Nomor 2 | Nopember 2025 | E-ISSN : 2987-9264 | DOI: 10.47709/geci.v3i2.9768

Baharsyah, A. F., Fauzi, M., Sabarudin, & Suprapni, Y. (2023). Kemampuan mahasiswa dalam menyelesaikan soal-soal HOTS pada materi PAI sekolah/madrasah. *Jurnal Tawadhu, 7*(1), 1–11. https://doi.org/10.52802/twd.v7i1.519

Facione, P. A. (1990). *Critical thinking: A statement of expert consensus for purposes of educational assessment and instruction. Research findings and recommendations* (ED315423) [Laporan untuk American Philosophical Association]. https://eric.ed.gov/?id=ED315423

Flake, J. K., Pek, J., & Hehman, E. (2017). Construct validation in social and personality research: Current practice and recommendations. *Social Psychological and Personality Science, 8*(4), 370–378. https://doi.org/10.1177/1948550617693063

Lakens, D. (2022). Sample size justification. *Collabra: Psychology, 8*(1), Article 33267. https://doi.org/10.1525/collabra.33267

Loyens, S. M. M., van Meerten, J. E., Schaap, L., & Wijnia, L. (2023). Situating higher-order, critical, and critical-analytic thinking in problem- and project-based learning environments: A systematic review. *Educational Psychology Review, 35*, Article 39. https://doi.org/10.1007/s10648-023-09757-x

Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology, 88*(5), 879–903. https://doi.org/10.1037/0021-9010.88.5.879

Rönkkö, M., & Cho, E. (2022). An updated guideline for assessing discriminant validity. *Organizational Research Methods, 25*(1), 6–14. https://doi.org/10.1177/1094428120968614

Schober, P., Boer, C., & Schwarte, L. A. (2018). Correlation coefficients: Appropriate use and interpretation. *Anesthesia & Analgesia, 126*(5), 1763–1768. https://doi.org/10.1213/ANE.0000000000002864

Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International Journal of Medical Education, 2*, 53–55. https://doi.org/10.5116/ijme.4dfb.8dfd