

Interpretasi dan Pelaporan Hasil Analisis Statistik Berbasis Uji Perbedaan Dua Kelompok dalam Penelitian Pendidikan Agama Islam

Rahayu Suprapti¹, Leny Misdiyanti², Fajri Ismail³, Muhammad Win Afgani⁴

^{1,2,3}Universitas Islam Negeri Raden Fatah Palembang, Indonesia

E-Mail: rahayusuprapti.pns@gmail.com¹

Article History:

Received: 13 Juni 2026

Revised: 26 Juni 2026

Accepted: 28 Juni 2026

Keywords: analisis statistik, uji perbedaan, interpretasi statistik, pelaporan hasil penelitian, Pendidikan Agama Islam, metodologi penelitian.

Abstract: Interpretasi dan pelaporan hasil analisis statistik merupakan tahapan penting dalam penelitian kuantitatif karena menentukan validitas argumentasi ilmiah dan kualitas publikasi. Namun, berbagai penelitian Pendidikan Agama Islam (PAI) masih menunjukkan kelemahan dalam menafsirkan hasil uji perbedaan dua kelompok serta menyajikan temuan secara sistematis. Penelitian ini bertujuan menganalisis prinsip-prinsip interpretasi dan pelaporan hasil analisis statistik berbasis uji perbedaan dua kelompok dalam penelitian PAI. Penelitian menggunakan pendekatan kualitatif melalui studi literatur terhadap buku metodologi, pedoman analisis statistik, dan artikel ilmiah yang relevan. Hasil kajian menunjukkan bahwa interpretasi hasil analisis harus didasarkan pada signifikansi statistik yang dipadukan dengan makna substantif, sehingga temuan tidak hanya berorientasi pada nilai statistik tetapi juga pada implikasi pendidikan. Selain itu, pelaporan hasil analisis perlu disusun secara sistematis melalui penyajian statistik deskriptif, pengujian asumsi, hasil uji hipotesis, serta pembahasan yang terintegrasi dengan kerangka teori. Penelitian ini menegaskan bahwa pelaporan statistik yang akurat, transparan, dan kontekstual berkontribusi terhadap peningkatan kualitas metodologis, kredibilitas, dan reproduktibilitas penelitian Pendidikan Agama Islam.

PENDAHULUAN

Penelitian kuantitatif dalam bidang pendidikan, termasuk Pendidikan Agama Islam (PAI), memiliki peran strategis dalam menghasilkan temuan yang objektif dan berbasis data. Salah satu teknik analisis yang sering digunakan adalah uji perbedaan dua kelompok, seperti independent sample t-test, yang bertujuan untuk mengetahui perbedaan rata-rata antara kelompok eksperimen dan kontrol (Ghozali, 2021; Sugiyono, 2019). Penggunaan analisis statistik inferensial ini menjadi penting karena mampu memberikan dasar pengambilan keputusan yang rasional dalam mengevaluasi efektivitas pembelajaran (Siregar, 2017; Field, 2018). Sejalan dengan itu, Afgan (2023) menegaskan bahwa uji perbedaan dua kelompok merupakan bagian fundamental dari

statistik inferensial dalam penelitian pendidikan, yang tidak hanya berfungsi untuk menguji hipotesis, tetapi juga untuk memberikan dasar empiris dalam menentukan efektivitas perlakuan pembelajaran secara objektif dan terukur.

Dalam konteks pendidikan modern, pendekatan berbasis data juga didukung oleh perkembangan teknologi analisis statistik yang semakin canggih dan mudah diakses oleh peneliti (Johnson & Christensen, 2020). Selain itu, tuntutan terhadap penelitian berbasis evidence-based practice semakin menguat dalam meningkatkan kualitas pembelajaran, termasuk dalam PAI (Mulyasa, 2021; Darling-Hammond, 2020). Namun demikian, berbagai studi menunjukkan bahwa masih banyak peneliti yang mengalami kesulitan dalam memahami dan menginterpretasikan hasil analisis statistik secara tepat (Hidayat, 2020; Widhiarso, 2020), sehingga berpotensi menimbulkan kesalahan dalam penarikan kesimpulan penelitian. Dalam hal ini, Afgan (2023) juga menyoroti bahwa kelemahan utama peneliti pendidikan bukan hanya pada teknik perhitungan statistik, tetapi juga pada ketepatan dalam memilih dan memahami jenis uji inferensial yang sesuai dengan karakteristik data penelitian.

Permasalahan utama dalam penelitian kuantitatif tidak hanya terletak pada proses analisis data, tetapi juga pada interpretasi hasil statistik yang sering kali kurang tepat. Banyak peneliti yang hanya berfokus pada nilai signifikansi (p -value) tanpa mempertimbangkan makna praktis dari hasil yang diperoleh (Cumming, 2014). Padahal, interpretasi statistik yang baik harus mencakup pemahaman terhadap ukuran efek (effect size) serta relevansi hasil dengan konteks penelitian (Lakens, 2013; Gravetter & Wallnau, 2017). Kesalahan dalam interpretasi dapat menyebabkan kesimpulan yang bias dan menurunkan validitas penelitian (Rohman, 2021).

Selain itu, transparansi dan reproduksibilitas hasil penelitian juga menjadi isu penting dalam dunia akademik saat ini (Nosek *et al.*, 2022). Afgan (2023) memperkuat hal tersebut dengan menegaskan bahwa dalam penelitian pendidikan, khususnya PAI, interpretasi hasil uji perbedaan tidak boleh dilepaskan dari konteks pedagogis, karena hasil statistik harus mampu diterjemahkan menjadi dasar pengambilan keputusan pembelajaran yang bermakna. Oleh karena itu, peneliti dituntut untuk tidak hanya mampu melakukan analisis statistik, tetapi juga memahami cara menginterpretasikan hasil tersebut secara komprehensif dan ilmiah (Hidayat, 2020).

METODE PENELITIAN

Penelitian ini menggunakan pendekatan kualitatif dengan jenis penelitian studi literatur (*library research*), yang bertujuan untuk mengkaji secara mendalam konsep interpretasi dan pelaporan hasil analisis statistik berbasis uji perbedaan dua kelompok dalam penelitian Pendidikan Agama Islam (PAI). Pendekatan ini dipilih karena mampu memberikan pemahaman komprehensif melalui analisis terhadap berbagai sumber ilmiah yang relevan dan mutakhir (Sugiyono, 2019; Fraenkel *et al.*, 2019). Studi literatur memungkinkan peneliti untuk mengidentifikasi pola, konsep, serta kesenjangan penelitian yang berkaitan dengan praktik interpretasi dan pelaporan statistik dalam penelitian pendidikan (Johnson & Christensen, 2020).

Sumber data dalam penelitian ini terdiri dari data sekunder yang diperoleh dari berbagai referensi ilmiah, seperti artikel jurnal nasional terakreditasi (SINTA 2–3), jurnal internasional bereputasi (Scopus), serta buku teks metodologi penelitian dan statistik pendidikan. Kriteria pemilihan sumber didasarkan pada relevansi dengan topik penelitian, kredibilitas penerbit, serta keterbaruan publikasi, dengan prioritas pada sumber yang diterbitkan dalam lima hingga sepuluh tahun terakhir (Nosek *et al.*, 2022; Hidayat, 2020). Selain itu, artikel yang memiliki DOI menjadi pertimbangan utama untuk menjamin keabsahan dan kemudahan penelusuran sumber ilmiah.

Teknik pengumpulan data dilakukan melalui metode dokumentasi, yaitu dengan cara mengidentifikasi, mengumpulkan, dan mengklasifikasikan berbagai literatur yang berkaitan

dengan interpretasi dan pelaporan hasil analisis statistik, khususnya yang menggunakan uji perbedaan dua kelompok seperti uji *t* dua sampel independen. Proses pengumpulan data dilakukan secara sistematis melalui database jurnal ilmiah seperti Google Scholar, DOAJ, dan portal jurnal nasional terakreditasi (APA, 2020). Data yang terkumpul kemudian diseleksi dan dikategorikan berdasarkan tema, seperti konsep interpretasi statistik, kesalahan umum dalam analisis data, serta standar pelaporan hasil penelitian (Widhiarso, 2020).

Teknik analisis data dalam penelitian ini menggunakan analisis isi (*content analysis*), yaitu dengan menelaah secara mendalam isi dari berbagai sumber literatur untuk menemukan pola, hubungan, dan sintesis konsep yang relevan dengan tujuan penelitian (Krippendorff, 2018). Analisis dilakukan melalui beberapa tahap, yaitu reduksi data, penyajian data, dan penarikan kesimpulan. Reduksi data dilakukan dengan memilih informasi yang relevan, sedangkan penyajian data dilakukan dalam bentuk narasi deskriptif yang sistematis. Selanjutnya, penarikan kesimpulan dilakukan dengan mengintegrasikan berbagai temuan untuk menghasilkan pemahaman yang komprehensif mengenai interpretasi dan pelaporan hasil analisis statistik (Miles *et al.*, 2018). Teknik analisis data dalam penelitian ini menggunakan analisis isi (*content analysis*), yaitu dengan menelaah berbagai literatur terkait interpretasi dan pelaporan uji perbedaan dua kelompok. Dalam konteks statistik inferensial, Afgan (2023) menjelaskan bahwa pemahaman terhadap prosedur uji seperti *t*-test tidak hanya mencakup perhitungan, tetapi juga pemahaman asumsi, jenis data, serta tujuan komparasi antar kelompok dalam penelitian pendidikan.

Untuk menjamin keabsahan data, penelitian ini menggunakan teknik triangulasi sumber, yaitu dengan membandingkan informasi dari berbagai referensi yang berbeda untuk memperoleh kesimpulan yang lebih valid dan objektif (Sugiyono, 2019). Selain itu, peneliti juga melakukan pengecekan ulang terhadap konsistensi data dan kesesuaian antara teori dan praktik yang ditemukan dalam literatur. Dengan demikian, metode penelitian ini diharapkan mampu menghasilkan kajian yang mendalam, sistematis, dan dapat dipertanggungjawabkan secara ilmiah dalam bidang Pendidikan Agama Islam.

HASIL DAN PEMBAHASAN

Hasil

Hasil penelitian ini diperoleh melalui analisis literatur terhadap berbagai sumber ilmiah yang membahas praktik interpretasi dan pelaporan hasil uji perbedaan dua kelompok, khususnya *independent sample t-test* dalam penelitian Pendidikan Agama Islam (PAI). Temuan penelitian menunjukkan bahwa terdapat empat komponen utama dalam pelaporan hasil analisis statistik yang baik dan sistematis, yaitu: (1) statistik deskriptif, (2) uji prasyarat, (3) hasil uji hipotesis, dan (4) interpretasi hasil.

Tabel 1. Sintesis Temuan Literatur tentang Uji Perbedaan Dua Kelompok dalam Penelitian Pendidikan

Penulis	Komponen Analisis	Temuan Utama	Standar Pelaporan	Implikasi dalam Penelitian PAI
Field (2018)	Statistik deskriptif	Banyak penelitian belum lengkap menyajikan mean, SD, min-max	Harus menyajikan mean, SD, N lengkap per kelompok	Menentukan gambaran awal perbedaan hasil belajar PAI
Ghasemi & Zahediasl (2012)	Uji prasyarat	Uji normalitas & homogenitas sering tidak dilaporkan	Data harus normal & homogen ($p > 0,05$)	Menjamin validitas uji <i>t</i> dalam penelitian PAI

Penulis	Komponen Analisis	Temuan Utama	Standar Pelaporan	Implikasi dalam Penelitian PAI
Wasserstein & Lazar (2016)	Uji hipotesis	Fokus hanya pada p-value tanpa t dan df lengkap	Harus mencantumkan t, df, p-value	Menghindari kesimpulan statistik yang tidak lengkap
Lakens (2013)	<i>Effect size</i>	Sangat jarang digunakan dalam penelitian PAI	Wajib melaporkan Cohen's d atau η^2	Menilai kekuatan pengaruh pembelajaran
Cumming (2014)	Interpretasi	Terbatas pada "signifikan/tidak signifikan"	Harus mengaitkan konteks pendidikan	Menguatkan makna pembelajaran PAI

Berdasarkan tabel tersebut, dapat diketahui bahwa pelaporan hasil analisis statistik dalam penelitian Pendidikan Agama Islam masih mengalami beberapa kelemahan utama, terutama pada aspek kelengkapan pelaporan dan kedalaman interpretasi hasil.

Pertama, pada statistik deskriptif, ditemukan bahwa sebagian besar penelitian belum menyajikan data secara lengkap. Padahal, menurut Field (2018), statistik deskriptif merupakan dasar penting dalam memahami distribusi data sebelum dilakukan uji inferensial.

Kedua, pada uji prasyarat, banyak penelitian tidak melaporkan uji normalitas dan homogenitas secara eksplisit. Padahal Ghasemi & Zahediasl (2012) menegaskan bahwa asumsi normalitas merupakan syarat utama dalam penggunaan uji parametrik seperti t-test.

Ketiga, pada uji hipotesis, peneliti cenderung hanya melaporkan nilai signifikansi (p-value), tanpa menyertakan nilai t dan derajat kebebasan (df). Hal ini bertentangan dengan standar pelaporan statistik modern yang menekankan transparansi hasil analisis (Wasserstein & Lazar, 2016).

Keempat, penggunaan effect size masih sangat rendah, padahal Lakens (2013) menegaskan bahwa signifikansi statistik tanpa effect size tidak cukup untuk menilai makna praktis dalam penelitian pendidikan.

Kelima, pada aspek interpretasi hasil, banyak penelitian berhenti pada keputusan "signifikan atau tidak signifikan", tanpa mengaitkan hasil dengan konteks pembelajaran PAI secara substantif. Hal ini sejalan dengan kritik Cumming (2014) bahwa praktik NHST sering disalahgunakan tanpa interpretasi yang bermakna.

Berdasarkan sintesis literatur, ditemukan bahwa penggunaan uji perbedaan dua kelompok masih sering tidak disertai pemahaman konseptual yang kuat terhadap dasar statistiknya. Afgan (2023) menyatakan bahwa dalam banyak penelitian pendidikan, uji t sering digunakan secara prosedural tanpa mempertimbangkan kesesuaian desain penelitian dan tujuan analisis.

Hal ini terlihat dari masih banyaknya penelitian yang hanya berfokus pada output statistik tanpa memahami makna inferensial dari hasil perbandingan antar kelompok.

Pembahasan

Interpretasi Hasil Uji Perbedaan Dua Kelompok dalam Penelitian PAI

Interpretasi hasil uji perbedaan dua kelompok dalam penelitian Pendidikan Agama Islam (PAI) tidak cukup hanya berhenti pada pembacaan nilai signifikansi statistik (*p-value*), tetapi harus dipahami secara lebih mendalam dan komprehensif. Secara statistik, nilai $p < 0,05$ memang menunjukkan adanya perbedaan yang signifikan antara kelompok eksperimen dan kelompok kontrol, namun makna tersebut belum sepenuhnya menjelaskan dampak pendidikan yang sebenarnya. Menurut Wasserstein & Lazar (2016), penggunaan *p-value* secara tunggal tanpa interpretasi substantif dapat menyesatkan dalam pengambilan kesimpulan ilmiah. Oleh karena itu, interpretasi harus mencakup keterkaitan antara hasil statistik dengan tujuan pembelajaran PAI yang

menekankan pembentukan karakter, nilai spiritual, dan akhlak peserta didik. Dalam konteks ini, hasil uji statistik harus diterjemahkan sebagai perubahan kualitas pembelajaran, bukan hanya perubahan angka. Dengan demikian, interpretasi yang baik harus mampu menjembatani antara hasil kuantitatif dan makna edukatif yang lebih luas dalam pendidikan Islam berbasis nilai.

Dalam konteks Pendidikan Agama Islam, hasil uji perbedaan dua kelompok tidak hanya mencerminkan peningkatan aspek kognitif seperti pengetahuan siswa, tetapi juga mencakup dimensi afektif dan spiritual yang menjadi inti dari pendidikan Islam. Misalnya, ketika suatu model pembelajaran berbasis nilai-nilai Islam menunjukkan perbedaan signifikan secara statistik, maka interpretasi yang tepat adalah adanya peningkatan internalisasi nilai religius, bukan sekadar peningkatan skor hasil belajar. Hal ini sejalan dengan pendapat Cumming (2014) yang menegaskan bahwa interpretasi hasil statistik harus mempertimbangkan makna praktis (*practical significance*), bukan hanya signifikansi statistik. Dalam pembelajaran PAI, perubahan perilaku seperti meningkatnya kedisiplinan ibadah, sikap toleransi, dan akhlak mulia menjadi indikator penting yang harus dikaitkan dengan hasil analisis statistik. Oleh karena itu, interpretasi hasil penelitian PAI harus bersifat holistik, mengintegrasikan data kuantitatif dengan realitas pendidikan yang berbasis nilai-nilai keislaman dan karakter peserta didik.

Selain itu, *effect size* merupakan komponen penting dalam interpretasi hasil uji perbedaan dua kelompok yang sering diabaikan dalam penelitian Pendidikan Agama Islam. Menurut Lakens (2013), *effect size* seperti Cohen's *d* memberikan informasi mengenai besar kecilnya pengaruh suatu perlakuan, sehingga tidak hanya bergantung pada nilai signifikansi statistik. Sebagai contoh, nilai Cohen's *d* = 0,8 menunjukkan efek yang besar dan secara pedagogis memiliki dampak yang kuat terhadap hasil belajar siswa. Dalam konteks PAI, hal ini berarti bahwa suatu strategi pembelajaran tidak hanya efektif secara statistik, tetapi juga memiliki dampak nyata terhadap pembentukan karakter religius dan pemahaman nilai-nilai keislaman. Tanpa *effect size*, peneliti berisiko menyimpulkan hasil secara berlebihan hanya berdasarkan *p-value*. Oleh karena itu, interpretasi yang baik harus menggabungkan signifikansi statistik dan signifikansi praktis agar hasil penelitian lebih bermakna dan dapat diterapkan dalam praktik pembelajaran yang sesungguhnya.

Lebih jauh, interpretasi hasil uji perbedaan dua kelompok dalam penelitian PAI juga harus mempertimbangkan konteks sosial dan budaya peserta didik. Pendidikan Agama Islam tidak dapat dipisahkan dari nilai-nilai lokal, tradisi masyarakat, serta lingkungan sosial yang membentuk perilaku siswa. Menurut Field (2018), interpretasi data statistik yang baik harus selalu dikaitkan dengan konteks penelitian agar tidak terjadi generalisasi yang keliru. Afgan (2023) menegaskan bahwa interpretasi hasil uji perbedaan dua kelompok harus dipahami sebagai proses menghubungkan hasil statistik dengan konteks pendidikan, bukan sekadar keputusan signifikan atau tidak signifikan. Dalam konteks PAI, hal ini berarti bahwa hasil uji tidak hanya menunjukkan perbedaan nilai, tetapi juga mencerminkan efektivitas strategi pembelajaran dalam membentuk aspek kognitif, afektif, dan spiritual peserta didik.

Dalam konteks PAI, hasil analisis statistik harus dijelaskan dalam bentuk perubahan nyata seperti peningkatan sikap religius, penguatan akhlak, dan kemampuan bersosialisasi secara Islami. Dengan demikian, interpretasi tidak hanya bersifat numerik, tetapi juga kontekstual dan transformatif. Hal ini menjadikan hasil penelitian lebih relevan dengan kebutuhan pendidikan Islam yang berorientasi pada pembentukan manusia berkarakter dan berakhlak mulia dalam kehidupan sehari-hari.

Pelaporan Hasil Analisis Statistik dalam Penelitian PAI

Pelaporan hasil analisis statistik dalam penelitian Pendidikan Agama Islam (PAI) harus mengikuti standar ilmiah yang sistematis, transparan, dan dapat direplikasi. Menurut American Psychological Association (APA, 2020), pelaporan hasil penelitian kuantitatif tidak hanya menampilkan hasil akhir, tetapi juga seluruh proses analisis secara runtut mulai dari statistik deskriptif hingga interpretasi hasil. Hal ini penting karena transparansi pelaporan berpengaruh terhadap validitas dan kredibilitas penelitian. Selain itu, Afgan (2023) menyoroti bahwa pelaporan hasil uji statistik harus dilakukan secara sistematis mulai dari pemilihan uji, pemeriksaan asumsi, hingga interpretasi hasil. Ketidaktepatan dalam pelaporan dapat menyebabkan kesalahan dalam penarikan kesimpulan, meskipun hasil uji secara statistik terlihat signifikan.

Dalam konteks ini, peneliti PAI dituntut untuk tidak hanya menyajikan hasil secara ringkas, tetapi juga memberikan informasi yang lengkap agar pembaca dapat memahami proses pengambilan keputusan statistik secara utuh. Selain itu, standar pelaporan modern juga menekankan pentingnya keterbukaan data agar hasil penelitian dapat diuji ulang oleh peneliti lain. Dengan demikian, pelaporan statistik dalam penelitian PAI tidak hanya bersifat teknis, tetapi juga merupakan bagian dari etika ilmiah dalam penelitian pendidikan berbasis data.

Berdasarkan standar APA Style (2020), pelaporan hasil analisis statistik dalam penelitian kuantitatif harus mencakup empat komponen utama. Pertama, statistik deskriptif harus disajikan secara lengkap untuk masing-masing kelompok, termasuk nilai mean, standar deviasi, dan jumlah sampel. Kedua, uji prasyarat seperti uji normalitas dan homogenitas varians wajib dilaporkan untuk memastikan bahwa data memenuhi asumsi uji parametrik. Ketiga, hasil uji hipotesis harus ditulis secara lengkap, mencakup nilai t , derajat kebebasan (df), dan nilai signifikansi (p -value), misalnya $t(58) = 2.45$, $p = 0.018$. Keempat, peneliti juga harus menyertakan ukuran efek (*effect size*) seperti Cohen's d untuk menggambarkan kekuatan pengaruh secara praktis. Menurut Lakens (2013), pelaporan effect size sangat penting untuk melengkapi informasi signifikansi statistik agar hasil penelitian lebih bermakna secara pedagogis.

Selain itu, pelaporan hasil analisis statistik dalam penelitian PAI juga harus diintegrasikan dengan teori pendidikan Islam agar memiliki makna yang lebih kontekstual. Hasil statistik tidak boleh hanya disajikan sebagai angka, tetapi harus dihubungkan dengan konsep pendidikan karakter, pembelajaran berbasis nilai, dan internalisasi akhlak dalam proses pembelajaran. Dengan demikian, hasil penelitian dapat memberikan kontribusi nyata terhadap pengembangan praktik pembelajaran PAI di sekolah. Misalnya, jika terdapat perbedaan signifikan antara dua kelompok, maka peneliti perlu menjelaskan bagaimana perbedaan tersebut berdampak pada peningkatan sikap religius, kedisiplinan ibadah, atau akhlak peserta didik. Pendekatan ini menjadikan pelaporan statistik tidak hanya bersifat kuantitatif, tetapi juga memiliki nilai edukatif dan spiritual yang sesuai dengan tujuan pendidikan Islam.

Wasserstein & Lazar (2016) menegaskan bahwa pelaporan statistik modern harus meninggalkan praktik "p-value only" dan beralih pada pelaporan yang lebih komprehensif, transparan, serta berbasis interpretasi yang lebih luas. Hal ini menunjukkan bahwa pelaporan statistik tidak lagi cukup hanya dengan menyatakan signifikan atau tidak signifikan, tetapi harus disertai dengan informasi tambahan seperti effect size, interval kepercayaan, serta konteks penelitian. Dalam penelitian PAI, pendekatan ini sangat relevan karena memungkinkan peneliti untuk memberikan gambaran yang lebih utuh tentang dampak pembelajaran terhadap peserta didik. Dengan demikian, pelaporan hasil analisis statistik yang baik tidak hanya meningkatkan kualitas ilmiah penelitian, tetapi juga memperkuat kontribusinya terhadap pengembangan pendidikan Islam berbasis bukti (*evidence-based Islamic education*).

Sintesis dan Implikasi dalam Penelitian PAI

Secara keseluruhan, hasil penelitian ini menunjukkan bahwa kelemahan utama dalam penelitian Pendidikan Agama Islam (PAI) tidak hanya terletak pada aspek teknik analisis statistik, tetapi lebih dominan pada kemampuan peneliti dalam melakukan interpretasi dan pelaporan hasil statistik secara tepat, sistematis, dan bermakna. Banyak penelitian masih berhenti pada tahap pengujian hipotesis tanpa mengembangkan pemahaman yang lebih dalam terhadap makna substantif dari hasil tersebut dalam konteks pendidikan.

Hal ini menunjukkan bahwa literasi statistik peneliti PAI masih perlu ditingkatkan agar mampu mengintegrasikan antara hasil kuantitatif dan implikasi pedagogis secara lebih komprehensif. Kemampuan memahami, menafsirkan, dan mengkomunikasikan hasil statistik secara benar menjadi aspek penting dalam menghasilkan penelitian yang tidak hanya kuat secara metodologis, tetapi juga relevan secara praktis dalam dunia pendidikan.

Selain itu, standar pelaporan internasional seperti APA Style (2020) belum sepenuhnya diterapkan secara konsisten dalam penelitian-penelitian PAI. Kondisi ini berdampak pada rendahnya tingkat transparansi, akurasi pelaporan, serta replikabilitas penelitian. Akibatnya, hasil penelitian sering kali sulit diverifikasi atau dibandingkan dengan penelitian lain yang sejenis. Secara lebih luas, Afgan (2023) menekankan bahwa rendahnya literasi statistik di kalangan peneliti pendidikan menyebabkan kesalahan umum berupa penggunaan uji statistik yang tidak tepat serta interpretasi yang terlalu sederhana. Hal ini berdampak pada kualitas penelitian yang kurang optimal dalam memberikan kontribusi terhadap pengembangan praktik pembelajaran.

Di samping itu, penggunaan *effect size* masih relatif jarang ditemukan dalam penelitian PAI, padahal indikator ini sangat penting untuk menilai kekuatan pengaruh suatu intervensi pembelajaran secara praktis, tidak hanya berdasarkan signifikansi statistik semata. Ketidadaan pelaporan *effect size* menyebabkan interpretasi hasil menjadi kurang lengkap dan berpotensi menyesatkan dalam pengambilan kesimpulan.

Lebih lanjut, interpretasi hasil penelitian sering kali belum mengaitkan temuan statistik dengan konteks nilai-nilai pendidikan Islam, seperti pembentukan karakter, penguatan akhlak, dan internalisasi nilai religius. Padahal, aspek ini merupakan inti dari tujuan pendidikan Islam yang tidak dapat dipisahkan dari proses pembelajaran.

Oleh karena itu, peningkatan kualitas interpretasi dan pelaporan statistik menjadi sangat penting dalam memperkuat validitas ilmiah penelitian PAI. Upaya ini juga diharapkan mampu mendorong pengembangan *evidence-based education* yang lebih kuat, kontekstual, dan relevan dengan kebutuhan pendidikan Islam di era modern.

KESIMPULAN

Penelitian ini menemukan bahwa interpretasi hasil uji perbedaan dua kelompok dalam penelitian Pendidikan Agama Islam sering kali terbatas pada nilai *p-value* tanpa mempertimbangkan *effect size*, makna praktis, dan konteks pedagogis seperti pembentukan karakter religius. Pelaporan juga menunjukkan kelemahan dalam kelengkapan data deskriptif, uji prasyarat, serta integrasi dengan teori pendidikan Islam, sehingga menurunkan transparansi dan kredibilitas ilmiah.

Namun, keterbatasan penelitian ini terletak pada pendekatan studi literatur yang bergantung pada sumber sekunder, sehingga belum menguji aplikasi empiris di lapangan PAI. Saran untuk penelitian mendatang mencakup studi empiris dengan data primer dan analisis *mixed-methods* untuk validasi lebih kuat, serta implikasi praktis berupa pelatihan literasi statistik bagi peneliti PAI guna menghasilkan laporan yang lebih akurat dan berbasis bukti dalam pengembangan pembelajaran Islami.

DAFTAR PUSTAKA

- Afgan, M. W. (2023). *Pendekatan statistik inferensial dalam menguji perbedaan lebih dari dua kelompok pada data penelitian Pendidikan Agama Islam*. UIN Raden Fatah Palembang.
- American Psychological Association. (2020). *Publication manual of the American Psychological Association* (7th ed.). American Psychological Association.
- Arifin, Z. (2021). *Penelitian pendidikan: Metode dan paradigma baru*. Remaja Rosdakarya.
- Cumming, G. (2014). The new statistics: Why and how. *Psychological Science*, 25(1), 7–29. <https://doi.org/10.1177/0956797613504966>
- Darling-Hammond, L. (2020). Teaching and learning during the COVID-19 pandemic. Learning Policy Institute.
- Field, A. (2018). *Discovering statistics using IBM SPSS Statistics* (5th ed.). SAGE Publications. <https://doi.org/10.4135/9781526441140>
- Fraenkel, J. R., Wallen, N. E., & Hyun, H. H. (2019). *How to design and evaluate research in education* (10th ed.). McGraw-Hill Education.
- Ghasemi, A., & Zahediasl, S. (2012). Normality tests for statistical analysis: A guide for non-statisticians. *International Journal of Endocrinology and Metabolism*, 10(2), 486–489. <https://doi.org/10.5812/ijem.3505>
- Ghozali, I. (2021). *Aplikasi analisis multivariate dengan program IBM SPSS 26* (10th ed.). Badan Penerbit Universitas Diponegoro.
- Gravetter, F. J., & Wallnau, L. B. (2017). *Statistics for the behavioral sciences* (10th ed.). Cengage Learning.
- Hattie, J. (2017). *Visible learning for teachers: Maximizing impact on learning*. Routledge.
- Hidayat, A. (2020). Statistik dalam penelitian pendidikan: Teori dan praktik. *Jurnal Evaluasi Pendidikan*, 11(2), 101–115.
- Johnson, R. B., & Christensen, L. (2020). *Educational research: Quantitative, qualitative, and mixed approaches* (7th ed.). SAGE Publications.
- Krippendorff, K. (2018). *Content analysis: An introduction to its methodology* (4th ed.). SAGE Publications.
- Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for t-tests and ANOVAs. *Frontiers in Psychology*, 4, Article 863. <https://doi.org/10.3389/fpsyg.2013.00863>
- Maulana, R. (2023). Evaluasi pembelajaran Pendidikan Agama Islam berbasis data. *Jurnal Pendidikan Islam*, 9(1), 45–60.
- Miles, M. B., Huberman, A. M., & Saldaña, J. (2018). *Qualitative data analysis: A methods sourcebook* (4th ed.). SAGE Publications.
- Mulyasa, E. (2021). *Manajemen berbasis sekolah*. Remaja Rosdakarya.
- Nosek, B. A., et al. (2022). The preregistration revolution. *Proceedings of the National Academy of Sciences*, 119(16). <https://doi.org/10.1073/pnas.2122498119>
- Rahman, F. (2022). Analisis hasil belajar PAI dalam perspektif statistik pendidikan. *Jurnal Tarbiyah*, 14(2), 77–92.
- Rohman, M. (2021). Kesalahan interpretasi statistik dalam penelitian pendidikan. *Jurnal Penelitian Pendidikan*, 38(1), 33–44.
- Sari, D. P. (2022). Pembelajaran berbasis nilai dalam Pendidikan Agama Islam. *Jurnal Pendidikan Islam Indonesia*, 7(1), 12–25.
- Schunk, D. H. (2020). *Learning theories: An educational perspective*. Pearson.
- Siregar, S. (2017). *Statistik parametrik untuk penelitian kuantitatif*. Bumi Aksara.
- Sugiyono. (2019). *Metode penelitian pendidikan (Kuantitatif, kualitatif, R&D)*. Alfabeta.

Uno, H. B. (2020). *Teori motivasi dan pengukurannya*. Bumi Aksara.

Wasserstein, R. L., & Lazar, N. A. (2016). The ASA statement on p-values: Context, process, and purpose. *The American Statistician*, 70(2), 129–133.
<https://doi.org/10.1080/00031305.2016.1154108>

Widhiarso, W. (2020). Kesalahan umum dalam analisis data penelitian pendidikan. *Jurnal Psikologi*, 47(2), 150–162.