

DOI: <https://doi.org/10.37850/cendekia>.
<https://journal.stitaf.ac.id/index.php/cendekia>.

LITERASI PENILAIAN GURU ITU PENTING: STUDI KORELASIONAL TENTANG IMPLEMENTASI PENILAIAN FORMATIF DALAM PENDIDIKAN AGAMA ISLAM

Zuhriyyah Hidayati^{*1}, Sita Isna Malyuna,² Badrut Tamami³, Khofiatul Aisyah⁴

^{1*}Prodi PAI, Universitas Billfath, Indonesia

²Prodi PGSD Universitas Ronggolawe Tuban, Indonesia

³Prodi PAI Universitas Muhammadiyah Jember, Indonesia

⁴Prodi PAI, Universitas Billfath, Indonesia

e-mail: 1zuhriyyahhidayati@gmail.com, sitaisna93@gmail.com, badruttamami@unmuhjember.ac.id

Received 20 June 2024; Received in revised form 25 September 2024; Accepted 30 October 2024

Abstrak

Penelitian ini bertujuan menganalisis hubungan antara literasi penilaian formatif dengan praktik penilaian formatif guru PAI di Kabupaten Lamongan. Penelitian menggunakan pendekatan kuantitatif dengan desain korelasional. Sebanyak 267 guru PAI dipilih sebagai responden melalui teknik purposive sampling. Pengumpulan data dilakukan menggunakan dua instrumen, yaitu literasi penilaian formatif (LPF)/ (*Teacher Formative Assessment Literacy Scale [TFALS]*) yang terdiri atas 12 butir pernyataan dan praktik penilaian formatif (PPF)/ (*Teacher Formative Assessment Practice Scale [TFAPS]*) yang terdiri atas 10 butir pernyataan dengan skala Likert enam poin. Data dianalisis menggunakan statistik deskriptif, uji reliabilitas, uji normalitas Shapiro-Wilk, dan uji korelasi Spearman's rho menggunakan perangkat lunak JASP versi 0.19.3. Hasil penelitian menunjukkan bahwa kedua instrumen memiliki reliabilitas yang sangat tinggi (Cronbach's Alpha LPF = 0,919; PPF = 0,912). Uji normalitas menunjukkan bahwa data tidak berdistribusi normal ($p < 0,001$), sehingga analisis hubungan dilakukan menggunakan korelasi Spearman. Hasil analisis menunjukkan adanya hubungan positif yang sangat kuat antara literasi penilaian formatif dan praktik penilaian formatif guru ($\rho = 0,863$; $p < 0,001$; 95% CI = 0,816–0,902). Temuan ini menunjukkan bahwa guru dengan tingkat literasi penilaian formatif yang lebih tinggi cenderung menerapkan praktik penilaian formatif secara lebih baik dalam pembelajaran. Oleh karena itu, penguatan literasi penilaian formatif perlu menjadi fokus dalam program pendidikan dan pengembangan profesional guru untuk mendukung implementasi asesmen yang berorientasi pada peningkatan kualitas pembelajaran.

Kata kunci: Literasi Penilaian Formatif; Praktik Penilaian Formatif; Guru Pendidikan Agama Islam; Asesmen Formatif; Korelasi.

Abstract

This study aims to analyze the relationship between formative assessment literacy and the practice of formative assessment of PAI teachers in Lamongan Regency. The research uses a quantitative approach with a correlational design. A total of 267 PAI teachers were selected as respondents through purposive sampling techniques. Data collection was carried out using two instruments, namely Teacher Formative Assessment Literacy Scale (TFALS) which consists of 12 statements and formative assessment practices (PPF) / (Teacher Formative Assessment Practice Scale [TFAPS]) which consists of 10 statements with a six-point Likert scale. Data were analyzed using descriptive statistics, reliability test, Shapiro-Wilk normality test, and Spearman's rho correlation test using JASP software version 0.19.3. The results

showed that both instruments had very high reliability (Cronbach's Alpha LPF = 0.919; PPF = 0.912). The normality test showed that the data was not normally distributed ($p < 0.001$), so the relationship analysis was performed using Spearman correlation. The results of the analysis showed a very strong positive relationship between formative assessment literacy and teacher formative assessment practice ($\rho = 0.863$; $p < 0.001$; 95% CI = 0.816–0.902). These findings show that teachers with higher levels of formative assessment literacy tend to better implement formative assessment practices in learning. Therefore, strengthening formative assessment literacy needs to be a focus in teacher education and professional development programs to support the implementation of assessments that are oriented towards improving the quality of learning.

Keywords: *Formative Assessment Literacy; Formative Assessment Practices; Islamic Religious Education Teacher; Formative Assessments; Correlation.*

PENDAHULUAN

Dalam lanskap pendidikan abad ke-21 yang berkembang, asesmen formatif telah muncul sebagai pilar sentral dan tren dalam mempromosikan pembelajaran peserta didik dan meningkatkan kualitas pembelajaran (Sortwell et al., 2024; Sudakova et al., 2022). Tidak seperti asesmen sumatif yang berfungsi terutama untuk mengevaluasi hasil peserta didik pada akhir periode pembelajaran, asesmen formatif tertanam dalam proses pembelajaran berkelanjutan untuk memberikan umpan balik yang meningkatkan pengajaran dan pembelajaran (Winarno & Azies, 2024). Asesmen formatif juga sangat penting untuk membantu guru mengidentifikasi kebutuhan belajar siswa secara *real time* (Zhang, Yilmaz, Ustun, & Karaoglan Yilmaz, 2023) dan menyesuaikan strategi pembelajaran peserta didik karena asesmen ini dimaksudkan untuk memberikan informasi tentang keputusan instruksional pembelajaran (van der Steen et al., 2022). Keberhasilan asesmen formatif sangat bergantung pada bagaimana guru mengintegrasikan praktik-praktik asesmen tersebut ke dalam proses pembelajaran sehari-hari. Dengan demikian, kualitas praktik penilaian formatif menjadi faktor penting dalam menentukan efektivitas pembelajaran.

Meskipun asesmen formatif telah diakui sebagai praktik pembelajaran yang efektif, implementasinya di kelas masih menunjukkan berbagai variasi. Praktik penilaian formatif mencakup kegiatan seperti mengklarifikasi tujuan pembelajaran, menggali bukti belajar siswa, memberikan umpan balik yang konstruktif, memfasilitasi penilaian teman sebaya maupun penilaian diri, serta menggunakan hasil asesmen untuk menyesuaikan pembelajaran (Black & Wiliam, 2009). Praktik-praktik tersebut menuntut kompetensi guru yang memadai agar asesmen benar-benar berfungsi untuk meningkatkan pembelajaran. Dalam konteks pendidikan Indonesia saat ini, asesmen formatif telah mendapatkan penekanan baru melalui penerapan Kurikulum Merdeka berbasis pendekatan pembelajaran mendalam. Kurikulum ini memposisikan asesmen tidak hanya sebagai alat asesmen, tetapi juga sebagai media untuk pertumbuhan pembelajaran dan pengajaran reflektif. Kementerian Pendidikan menyoroti pentingnya asesmen



formatif sebagai bagian dari kerangka asesmen pembelajaran untuk menumbuhkan kemandirian, kreativitas, dan karakter siswa (Aditomo, 2024; Anggraena et al., 2021). Implementasi Kurikulum Merdeka menuntut guru tidak hanya memahami konsep asesmen formatif, tetapi juga mampu menerapkannya secara konsisten dalam praktik pembelajaran sehari-hari.

Berbagai penelitian menunjukkan bahwa praktik penilaian formatif guru masih belum optimal. Guru sering kali masih berfokus pada pemberian nilai daripada penggunaan informasi hasil asesmen untuk memperbaiki proses pembelajaran. Umpan balik yang diberikan cenderung bersifat evaluatif, belum mendorong refleksi belajar siswa, sementara kegiatan penilaian diri dan penilaian teman sebaya masih jarang diterapkan (Worabu, 2024; Yusof & Abdul Rahim, 2024). Bahkan ketika guru telah mengenal berbagai strategi asesmen formatif, penerapannya sering kali tidak dilakukan secara sistematis sehingga manfaat asesmen formatif belum sepenuhnya dirasakan oleh peserta didik (Popham, 2009).

Salah satu faktor yang diduga berkontribusi terhadap belum optimalnya praktik penilaian formatif adalah tingkat literasi penilaian formatif guru, terutama jika mengingat literasi asesmen formatif telah diintegrasikan ke dalam berbagai program pelatihan formal dan pengembangan profesi, seperti Program Pendidikan Profesi Guru (PPG), kurikulum pendidikan guru, dan inisiatif pengembangan profesi berkelanjutan seperti PKG/PKB (Li et al., 2023). Program-program ini bertujuan untuk mempersiapkan guru dengan kompetensi yang diperlukan untuk melakukan asesmen secara bermakna. Namun, penerjemahan teori ke dalam praktik tetap menjadi bidang yang menjadi perhatian, terutama di kalangan guru khusus mata pelajaran seperti guru PAI (Pendidikan Agama Islam).

Meskipun sejumlah penelitian telah mengkaji literasi asesmen formatif guru maupun praktik asesmen formatif secara terpisah, penelitian yang secara empiris menguji hubungan antara literasi penilaian formatif dan praktik penilaian formatif pada guru Pendidikan Agama Islam masih sangat terbatas. Sebagian besar penelitian sebelumnya dilakukan pada guru sekolah umum atau mahasiswa pendidikan, sehingga karakteristik guru PAI sebagai pendidik yang mengintegrasikan kompetensi pedagogik dengan nilai-nilai keislaman belum banyak mendapat perhatian. Oleh karena itu, masih diperlukan bukti empiris mengenai apakah tingkat literasi penilaian formatif benar-benar berkaitan dengan kualitas praktik penilaian formatif guru PAI.

Oleh karena itu, penelitian ini bertujuan untuk menganalisis hubungan antara literasi penilaian formatif dan praktik penilaian formatif guru Pendidikan Agama Islam di Kabupaten Lamongan. Literasi penilaian formatif diukur melalui dimensi konseptual, praktis, dan sosial-emosional, sedangkan praktik penilaian formatif mencakup berbagai strategi asesmen yang diterapkan guru selama proses pembelajaran. Hasil penelitian ini diharapkan dapat memberikan bukti empiris mengenai pentingnya literasi penilaian formatif sebagai dasar penguatan praktik

asesmen guru PAI, sekaligus menjadi rekomendasi bagi program pendidikan guru, pengembangan profesional berkelanjutan, dan perumusan kebijakan peningkatan kualitas pembelajaran.

METODE

Penelitian ini menggunakan pendekatan kuantitatif dengan desain korelasional untuk menganalisis hubungan antara literasi penilaian formatif dan praktik penilaian formatif guru Pendidikan Agama Islam (PAI) di Kabupaten Lamongan. Dalam penelitian ini, literasi penilaian formatif (LPF) berperan sebagai variabel independen (X), sedangkan praktik penilaian formatif (PPF) berperan sebagai variabel dependen (Y). Desain korelasional dipilih karena penelitian ini bertujuan mengidentifikasi arah dan kekuatan hubungan antara kedua variabel tanpa melakukan manipulasi terhadap variabel penelitian.

Populasi penelitian adalah seluruh guru Pendidikan Agama Islam yang mengajar pada jenjang sekolah dasar, sekolah menengah pertama, sekolah menengah atas, dan sekolah menengah kejuruan di Kabupaten Lamongan. Sampel penelitian berjumlah 267 guru PAI yang dipilih menggunakan teknik purposive sampling. Pengambilan sampel bertujuan dipilih karena penelitian ini bertujuan untuk menyelidiki literasi penilaian formatif di kalangan guru yang memiliki pengalaman profesional yang cukup terkait dengan implementasi penilaian formatif. Oleh karena itu, hanya guru yang aktif mengajar Pendidikan Agama Islam dan telah berpartisipasi dalam program pengembangan profesional seperti PPG atau PKB yang dianggap memenuhi syarat. Pengambilan sampel probabilitas tidak digunakan karena tidak semua guru PAI memenuhi kriteria inklusi ini, dan penelitian ini menekankan partisipan yang kaya informasi daripada keterwakilan statistik. Kriteria responden meliputi: (1) aktif mengajar mata pelajaran Pendidikan Agama Islam pada tahun pelajaran berjalan; dan (2) telah mengikuti program pengembangan profesi guru, seperti Pendidikan Profesi Guru (PPG) atau Pengembangan Keprofesian Berkelanjutan (PKB).

Data penelitian dikumpulkan menggunakan dua instrumen berupa angket yang diadaptasi dari *Teacher Formative Assessment Literacy Scale* (TFALS) dan *Teacher Formative Assessment Practice Scale* (TFAPS) yang dikembangkan oleh Yan dan Pastore (2022). Variabel literasi penilaian formatif (LPF)/ (*Teacher Formative Assessment Literacy Scale [TFALS]*) yang terdiri atas 12 butir pernyataan yang merepresentasikan dua dimensi, yaitu konseptual dan sosial-emosional. Sementara itu, variabel praktik penilaian formatif (PPF)/ (*Teacher Formative Assessment Practice Scale [TFAPS]*) yang menggambarkan praktik guru dalam menerapkan asesmen formatif selama proses pembelajaran, meliputi kegiatan mengklarifikasi tujuan pembelajaran, memperoleh bukti belajar peserta didik, memberikan umpan balik yang konstruktif, memfasilitasi penilaian diri dan penilaian teman sebaya, serta memanfaatkan informasi hasil asesmen untuk



menyesuaikan strategi pembelajaran. Seluruh butir pernyataan menggunakan skala Likert enam tingkat, mulai dari 1 (sangat tidak setuju) hingga 6 (sangat setuju).

Pengumpulan data dilakukan secara daring menggunakan Google Form yang disebarluaskan melalui forum Musyawarah Guru Mata Pelajaran (MGMP) Pendidikan Agama Islam Kabupaten Lamongan, grup WhatsApp guru PAI, serta media komunikasi resmi lainnya. Pengumpulan data berlangsung selama tiga minggu. Sebelum mengisi kuesioner, seluruh responden memperoleh penjelasan mengenai tujuan penelitian, kerahasiaan data, serta hak responden untuk berpartisipasi secara sukarela sebagai bentuk penerapan prinsip etika penelitian.

Sebagaimana disebutkan sebelumnya, kedua instrumen merupakan adaptasi dari *Teacher Formative Assessment Literacy Scale* (TFALS) dan *Teacher Formative Assessment Practice Scale* (TFAPS) yang dikembangkan oleh Yan dan Pastore (2022). Proses adaptasi meliputi *forward translation*, *backward translation*, *expert judgment*, dan *pilot testing* untuk memastikan kesesuaian bahasa, makna, dan konteks budaya Indonesia. Setelah proses adaptasi, kualitas psikometrik instrumen dievaluasi kembali pada sampel penelitian. Reliabilitas internal dianalisis menggunakan Cronbach's Alpha dan McDonald's Omega, sedangkan validitas konstruk dievaluasi melalui item-rest correlation dan standardized factor loading. Instrumen dinyatakan memenuhi kriteria apabila nilai Cronbach's Alpha dan McDonald's Omega $> 0,70$, nilai *item-rest correlation* $> 0,30$, serta *standardized factor loading* $> 0,50$.

Analisis data dilakukan menggunakan aplikasi JASP versi 0.19.3. Tahap awal analisis dilakukan dengan statistik deskriptif untuk menggambarkan karakteristik responden dan distribusi skor pada masing-masing variabel melalui nilai rata-rata, simpangan baku, frekuensi, persentase, nilai minimum, dan nilai maksimum. Sebelum pengujian hipotesis, dilakukan uji normalitas menggunakan Shapiro-Wilk untuk menentukan teknik analisis korelasi yang sesuai. Apabila data memenuhi asumsi normalitas, hubungan antarvariabel dianalisis menggunakan korelasi Pearson Product Moment. Sebaliknya, apabila data tidak memenuhi asumsi normalitas, hubungan antarvariabel dianalisis menggunakan korelasi Spearman's rho. Analisis korelasi dilakukan terhadap skor total literasi penilaian formatif dengan skor praktik penilaian formatif, serta terhadap masing-masing dimensi literasi penilaian formatif (konseptual, praktis, dan sosial-emosional) dengan praktik penilaian formatif guru. Seluruh pengujian statistik dilakukan pada taraf signifikansi $\alpha = 0,05$. Kekuatan hubungan antarvariabel diinterpretasikan berdasarkan klasifikasi koefisien korelasi menurut Akoglu (2018) sedangkan arah hubungan ditentukan berdasarkan tanda positif atau negatif dari koefisien korelasi yang diperoleh.

HASIL DAN PEMBAHASAN

Statistik Deskriptif

Analisis statistik deskriptif dilakukan untuk memberikan gambaran umum mengenai karakteristik data pada kedua variabel penelitian, yaitu literasi penilaian formatif (LPF)/ (*Teacher Formative Assessment Literacy Scale* [TFALS]) dan praktik penilaian formatif (PPF)/ (*Teacher Formative Assessment Practice Scale* [TFAPS]). Analisis ini mencakup jumlah responden, nilai rata-rata (*mean*), simpangan baku (*standard deviation*), serta rentang skor yang diperoleh responden. Informasi tersebut diperlukan sebagai langkah awal untuk memahami kecenderungan data sebelum dilakukan analisis inferensial.

Sebanyak 267 guru Pendidikan Agama Islam (PAI) yang mengajar pada berbagai jenjang pendidikan di Kabupaten Lamongan berpartisipasi dalam penelitian ini. Seluruh data yang diperoleh dinyatakan lengkap dan dapat dianalisis, sehingga tidak ditemukan data yang hilang (*missing value*).

Tabel 1. Statistik Deskriptif Variabel Penelitian

Variabel	N	Mean	SD	Minimum	Maksimum
Literasi Penilaian Formatif (LPF)	267	63,67	4,86	50	72
Praktik Penilaian Formatif (PPF)	267	53,14	4,32	41	60

Berdasarkan Tabel 1, skor literasi penilaian formatif guru menunjukkan nilai rata-rata sebesar 63,67 dengan simpangan baku 4,86. Nilai tersebut diperoleh dari rentang skor 50 hingga 72, yang menunjukkan bahwa seluruh responden memiliki skor yang berada pada kisaran menengah hingga tinggi. Simpangan baku yang relatif kecil mengindikasikan bahwa variasi skor antarresponden tidak terlalu besar, sehingga tingkat literasi penilaian formatif guru cenderung memperlihatkan pola yang relatif seragam.

Pada variabel praktik penilaian formatif, diperoleh nilai rata-rata sebesar 53,14 dengan simpangan baku 4,32. Skor responden berada pada rentang 41 hingga 60, yang menunjukkan bahwa variasi praktik penilaian formatif di antara guru juga relatif terbatas. Sebaran skor yang tidak terlalu lebar mengindikasikan bahwa sebagian besar responden memberikan jawaban yang berada pada kisaran nilai yang berdekatan, sehingga karakteristik data pada variabel ini juga menunjukkan tingkat homogenitas yang cukup baik.

Jika dibandingkan dengan skor maksimum teoritis masing-masing instrumen, rata-rata skor LPF mencapai sekitar 88,4% dari skor maksimum, sedangkan rata-rata skor PPF mencapai sekitar 88,6%. Temuan tersebut menunjukkan bahwa secara umum responden memperoleh skor yang mendekati batas maksimum instrumen pada kedua variabel penelitian. Dengan kata lain, sebagian besar guru Pendidikan Agama Islam melaporkan memiliki tingkat literasi penilaian formatif yang baik serta praktik penilaian formatif yang juga cenderung tinggi dalam pelaksanaan pembelajaran.



Meskipun demikian, hasil statistik deskriptif hanya memberikan informasi mengenai kecenderungan umum data dan belum dapat menjelaskan ada atau tidaknya hubungan antara literasi penilaian formatif dengan praktik penilaian formatif. Oleh karena itu, analisis selanjutnya difokuskan pada pengujian asumsi normalitas sebagai dasar penentuan teknik analisis inferensial yang sesuai, sebelum dilakukan pengujian hubungan antarvariabel menggunakan analisis korelasi.

Validitas Instrumen

Sebelum dilakukan pengujian hipotesis, instrumen penelitian dievaluasi kembali untuk memastikan bahwa instrumen hasil adaptasi memiliki kualitas psikometrik yang memadai pada konteks guru Pendidikan Agama Islam di Indonesia. Evaluasi dilakukan terhadap dua instrumen, yaitu Teacher Formative Assessment Literacy Scale (TFALS) dan Teacher Formative Assessment Practice Scale (TFAPS). Analisis mencakup pengujian reliabilitas internal menggunakan Cronbach's Alpha dan McDonald's Omega, serta pengujian validitas konstruk menggunakan corrected item-total correlation (item-rest correlation) dan standardized factor loading.

Hasil analisis menunjukkan bahwa kedua instrumen memiliki reliabilitas internal yang sangat baik. Instrumen TFALS memperoleh nilai Cronbach's Alpha sebesar 0,919 dan McDonald's Omega sebesar 0,919, sedangkan instrumen TFAPS memperoleh nilai Cronbach's Alpha sebesar 0,912 dan McDonald's Omega sebesar 0,913. Seluruh nilai tersebut melampaui batas minimum 0,70 sehingga menunjukkan konsistensi internal yang tinggi.

Validitas konstruk juga menunjukkan hasil yang memuaskan. Pada instrumen TFALS, nilai corrected item-total correlation berada pada rentang 0,571–0,767, sedangkan standardized factor loading berada pada rentang 0,594–0,794. Pada instrumen TFAPS, nilai corrected item-total correlation berada pada rentang 0,555–0,761, dengan standardized factor loading berkisar antara 0,582–0,803. Seluruh nilai tersebut berada di atas kriteria yang direkomendasikan (corrected item-total correlation > 0,30 dan standardized factor loading > 0,50), sehingga seluruh butir instrumen dinyatakan valid dan tidak terdapat item yang perlu dieliminasi.

Berdasarkan hasil evaluasi psikometrik tersebut, dapat disimpulkan bahwa instrumen TFALS dan TFAPS memiliki validitas konstruk dan reliabilitas internal yang sangat baik setelah proses adaptasi ke dalam konteks Indonesia. Dengan demikian, seluruh butir instrumen layak digunakan sebagai alat ukur dalam analisis hubungan antara literasi penilaian formatif dan praktik penilaian formatif guru Pendidikan Agama Islam.

Tabel 2. Hasil Evaluasi Psikometrik Instrumen

Aspek	TFALS	TFAPS	Kriteria	Interpretasi
Jumlah item	12	10	-	Seluruh item

Cronbach's Alpha	0,919	0,912	> 0,70	dianalisis Sangat reliabel
McDonald's Omega	0,919	0,913	> 0,70	Sangat reliabel
Corrected Item-Total Correlation	0,571–0,767	0,555–0,761	> 0,30	Seluruh item valid
Standardized Factor Loading	0,594–0,794	0,582–0,803	> 0,50	Seluruh item valid
Keputusan	Seluruh item dipertahankan	Seluruh item dipertahankan	-	Layak digunakan

Sumber: Hasil analisis menggunakan JASP versi 0.19.3.

Reliabilitas Instrumen

Sebelum dilakukan analisis hubungan antarvariabel, terlebih dahulu dilakukan pengujian reliabilitas terhadap kedua instrumen penelitian, yaitu literasi penilaian formatif (LPF)/ (*Teacher Formative Assessment Literacy Scale* [TFALS]) dan praktik penilaian formatif (PPF)/ (*Teacher Formative Assessment Practice Scale* [TFAPS]). Pengujian reliabilitas bertujuan untuk memastikan bahwa instrumen yang digunakan memiliki tingkat konsistensi internal yang memadai sehingga mampu menghasilkan pengukuran yang stabil dan dapat dipercaya. Dalam penelitian ini, reliabilitas diestimasi menggunakan dua koefisien, yaitu Cronbach's Alpha (α) dan McDonald's Omega (ω). Penggunaan kedua indeks tersebut dimaksudkan untuk memberikan estimasi reliabilitas yang lebih komprehensif, mengingat McDonald's Omega dipandang lebih robust terhadap variasi kontribusi masing-masing butir dibandingkan Cronbach's Alpha.

Tabel 3. Reliabilitas Instrumen Penelitian

Instrumen	Jumlah Item	McDonald's ω	Cronbach's α	Interpretasi
LPF	12	0,919	0,919	Sangat reliabel
PPF	10	0,913	0,912	Sangat reliabel

Berdasarkan Tabel 2, instrumen LPF memperoleh nilai McDonald's Omega sebesar 0,919 dan Cronbach's Alpha sebesar 0,919. Kedua koefisien tersebut berada jauh di atas batas minimum reliabilitas yang umum direkomendasikan, yaitu 0,70. Hasil ini menunjukkan bahwa dua belas butir pada instrumen LPF memiliki tingkat konsistensi internal yang sangat tinggi, sehingga mampu mengukur konstruk literasi penilaian formatif secara konsisten pada seluruh responden. Skor yang cenderung tinggi dan homogen ini perlu dicermati sebagai kemungkinan indikasi ceiling effect atau bias desirabilitas sosial pada data self-report, yang akan didiskusikan lebih lanjut pada bagian Pembahasan.

Hasil yang serupa juga ditunjukkan oleh instrumen PPF. Instrumen yang terdiri atas sepuluh butir ini memperoleh nilai McDonald's Omega sebesar 0,913 dan Cronbach's Alpha sebesar 0,912. Nilai tersebut mengindikasikan bahwa seluruh butir penyusun PPF memiliki keterkaitan yang kuat dalam mengukur



konstruk praktik penilaian formatif. Dengan demikian, instrumen ini juga memiliki tingkat reliabilitas yang sangat baik dan layak digunakan sebagai alat ukur dalam penelitian ini.

Secara keseluruhan, kedua instrumen penelitian menunjukkan koefisien reliabilitas di atas 0,90, yang mengindikasikan tingkat konsistensi internal yang sangat tinggi. Temuan ini memberikan keyakinan bahwa data yang diperoleh melalui LPF dan PPF memiliki tingkat keandalan yang memadai sehingga dapat digunakan sebagai dasar untuk melakukan analisis inferensial pada tahap berikutnya.

Uji Normalitas

Sebelum dilakukan pengujian hipotesis, terlebih dahulu dilakukan uji normalitas untuk mengetahui apakah distribusi data pada masing-masing variabel memenuhi asumsi statistik parametrik. Pengujian ini penting karena hasil uji normalitas menjadi dasar dalam menentukan teknik analisis yang paling sesuai untuk menguji hubungan antarvariabel. Dalam penelitian ini, normalitas data diuji menggunakan Shapiro-Wilk Test, mengingat uji tersebut memiliki sensitivitas yang baik dalam mendeteksi penyimpangan distribusi data dan direkomendasikan untuk berbagai ukuran sampel.

Hasil uji normalitas menunjukkan bahwa variabel Literasi Penilaian Formatif Guru (LPF)) memperoleh nilai Shapiro-Wilk sebesar 0,888 dengan nilai signifikansi (p) < 0,001, sedangkan variabel Praktik Penilaian Formatif Guru (PPF) memperoleh nilai Shapiro-Wilk sebesar 0,886 dengan nilai signifikansi (p) < 0,001. Nilai signifikansi kedua variabel berada di bawah taraf signifikansi yang ditetapkan ($\alpha = 0,05$), sehingga hipotesis nol yang menyatakan bahwa data berdistribusi normal tidak dapat diterima.

Hasil tersebut mengindikasikan bahwa distribusi skor pada kedua variabel penelitian tidak mengikuti distribusi normal. Dengan demikian, asumsi normalitas sebagai salah satu prasyarat penggunaan teknik analisis parametrik tidak terpenuhi. Oleh karena itu, pengujian hubungan antara literasi penilaian formatif dan praktik penilaian formatif guru dilakukan menggunakan korelasi Spearman's rho, yaitu teknik analisis nonparametrik yang tidak mensyaratkan distribusi data normal dan sesuai digunakan untuk mengukur keeratan hubungan antarvariabel yang diukur menggunakan skala Likert atau data yang tidak memenuhi asumsi normalitas (Field, 2018a).

Temuan ini sekaligus menunjukkan bahwa pemilihan teknik analisis dalam penelitian dilakukan berdasarkan karakteristik empiris data, sehingga hasil pengujian hubungan yang diperoleh diharapkan memiliki tingkat ketepatan dan validitas yang lebih baik.

Hubungan antara Literasi Penilaian Formatif dan Praktik Penilaian Formatif Guru

Hubungan antara literasi penilaian formatif guru literasi penilaian formatif (LPF)/ (*Teacher Formative Assessment Literacy Scale* [TFALS]) dan praktik penilaian formatif (PPF)/ (*Teacher Formative Assessment Practice Scale* [TFAPS]) dianalisis menggunakan uji korelasi Spearman's rho. Pemilihan uji nonparametrik ini didasarkan pada hasil uji normalitas yang menunjukkan bahwa distribusi kedua variabel tidak memenuhi asumsi normalitas ($p < 0,001$). Hasil analisis korelasi disajikan pada Tabel 4.

Tabel 4. Hasil Uji Korelasi Spearman antara Literasi Penilaian Formatif dan Praktik Penilaian Formatif Guru

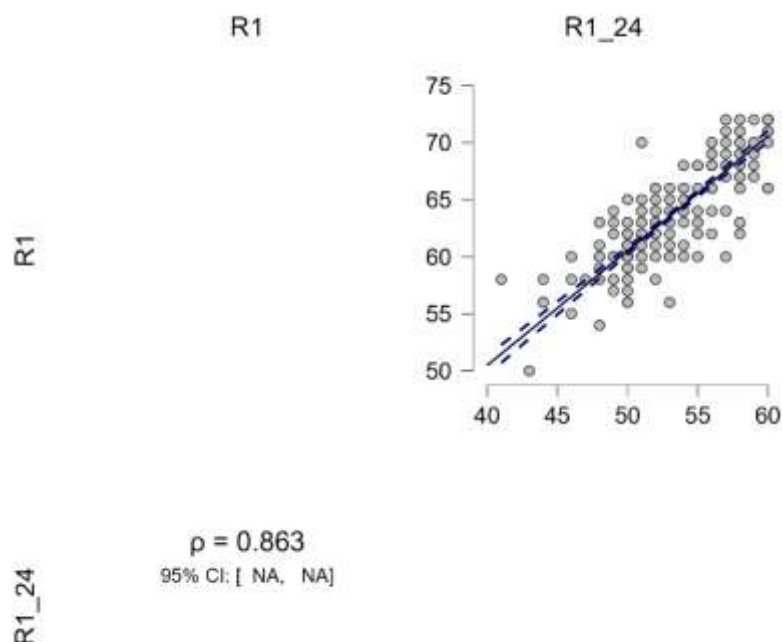
Variabel	n	ρ Spearman	95% CI	p	Interpretasi
LPF-PPF	267	0,863	0,816– 0,902	<0,001	Hubungan positif sangat kuat

Hasil analisis menunjukkan bahwa terdapat hubungan positif yang sangat kuat antara literasi penilaian formatif guru dan praktik penilaian formatif yang diterapkan dalam pembelajaran ($\rho = 0,863$; $p < 0,001$). Nilai signifikansi yang berada di bawah taraf kesalahan 5% menunjukkan bahwa hubungan tersebut signifikan secara statistik. Dengan demikian, hipotesis nol (H_0) yang menyatakan tidak terdapat hubungan antara kedua variabel ditolak, sedangkan hipotesis penelitian (H_1) diterima.

Besarnya koefisien korelasi menunjukkan bahwa peningkatan tingkat literasi penilaian formatif guru cenderung diikuti oleh peningkatan praktik penilaian formatif yang dilakukan di kelas. Berdasarkan klasifikasi kekuatan koefisien korelasi menurut Akoglu (2018), nilai koefisien sebesar 0,863 termasuk dalam kategori *strong correlation* (0,70–0,90). Hal ini mengindikasikan bahwa hubungan antara kedua variabel bukan hanya signifikan secara statistik, tetapi juga memiliki kekuatan hubungan yang tinggi.

Selain itu, interval kepercayaan 95% berada pada rentang 0,816 hingga 0,902, yang seluruhnya bernilai positif dan relatif sempit. Rentang interval tersebut menunjukkan bahwa estimasi koefisien korelasi memiliki tingkat presisi yang baik, sehingga memberikan keyakinan bahwa hubungan positif yang ditemukan pada sampel juga merepresentasikan hubungan yang kuat pada populasi guru Pendidikan Agama Islam di Kabupaten Lamongan.

Untuk memberikan gambaran visual mengenai hubungan kedua variabel, disajikan scatterplot pada gambar 1.



Gambar 1. Scatterplot Hubungan antara Literasi Penilaian Formatif (LPF) dan Praktik Penilaian Formatif (PPF)

Sebagaimana ditunjukkan pada Gambar 1, titik-titik data membentuk pola yang cenderung mengikuti garis diagonal dengan arah positif. Sebagian besar responden yang memiliki skor literasi penilaian formatif tinggi juga menunjukkan skor praktik penilaian formatif yang tinggi. Sebaliknya, responden dengan skor literasi yang relatif rendah cenderung memiliki skor praktik yang lebih rendah. Walaupun terdapat beberapa titik yang menyebar di sekitar garis kecenderungan, pola keseluruhan menunjukkan hubungan linear positif yang konsisten.

Visualisasi tersebut memperkuat hasil analisis statistik bahwa terdapat keterkaitan yang erat antara kemampuan guru dalam memahami konsep asesmen formatif dengan implementasinya dalam praktik pembelajaran. Dengan kata lain, guru yang memiliki literasi penilaian formatif yang lebih baik cenderung lebih mampu merancang, melaksanakan, memanfaatkan, dan menindaklanjuti asesmen formatif sebagai bagian dari proses pembelajaran yang berorientasi pada peningkatan kualitas belajar peserta didik.

Secara keseluruhan, temuan penelitian ini menunjukkan bahwa literasi penilaian formatif merupakan aspek yang berkaitan erat dengan praktik penilaian formatif guru. Semakin tinggi tingkat literasi yang dimiliki guru, semakin tinggi pula kecenderungan mereka untuk mengimplementasikan praktik penilaian formatif secara efektif dalam kegiatan pembelajaran.

Pembahasan

Hasil penelitian menunjukkan bahwa terdapat hubungan positif yang sangat kuat antara literasi penilaian formatif dan praktik penilaian formatif guru Pendidikan Agama Islam ($p = 0,863$; $p < 0,001$). Temuan ini mengindikasikan

bahwa guru yang memiliki pemahaman lebih baik mengenai konsep, tujuan, strategi, dan penggunaan hasil asesmen formatif cenderung lebih mampu mengimplementasikan praktik asesmen formatif secara konsisten dalam pembelajaran. Hasil penelitian ini sejalan dengan penelitian Yan dan Pastore (2022) yang menunjukkan bahwa literasi penilaian formatif merupakan prediktor penting praktik asesmen guru. Temuan serupa juga dilaporkan oleh DeLuca et al. (2015); Leontjev & deBoer (2020), dan Black dan Wiliam (2018) yang menegaskan bahwa kualitas praktik asesmen formatif sangat dipengaruhi oleh kemampuan guru dalam memahami prinsip-prinsip asesmen serta menggunakan informasi hasil asesmen untuk memperbaiki pembelajaran. Dengan demikian, literasi asesmen bukan hanya berupa penguasaan konsep, tetapi menjadi landasan utama dalam pengambilan keputusan pedagogis di kelas.

Temuan penelitian ini memperlihatkan koefisien korelasi yang relatif lebih tinggi dibandingkan sebagian besar penelitian sebelumnya. Salah satu kemungkinan penyebabnya adalah karakteristik responden penelitian yang merupakan guru Pendidikan Agama Islam yang telah banyak terpapar berbagai program pengembangan profesional, seperti Pendidikan Profesi Guru (PPG), Pengembangan Keprofesian Berkelanjutan (PKB), maupun implementasi Kurikulum Merdeka yang menempatkan asesmen formatif sebagai bagian integral dari proses pembelajaran. Kondisi tersebut memungkinkan guru tidak hanya memahami konsep asesmen formatif secara teoretis, tetapi juga mulai mengintegrasikannya dalam praktik pembelajaran sehari-hari. Temuan ini mendukung pandangan Andrade & Brookhart (2020), Brookhart (2011), dan Hattie & Timperley (2007), yang menegaskan bahwa peningkatan kompetensi asesmen guru melalui pelatihan berkelanjutan berkontribusi terhadap meningkatnya kualitas praktik pembelajaran.

Meskipun demikian, hubungan yang sangat kuat antara kedua variabel tidak serta-merta menunjukkan adanya hubungan sebab-akibat. Desain penelitian yang menggunakan pendekatan cross-sectional hanya mampu menjelaskan keeratan hubungan antarvariabel pada satu waktu pengamatan sehingga tidak dapat memastikan bahwa peningkatan literasi penilaian formatif secara langsung menyebabkan peningkatan praktik penilaian formatif. Selain itu, kedua variabel diukur menggunakan instrumen *self-report*, sehingga kemungkinan adanya *social desirability bias* maupun *common method bias* tetap perlu dipertimbangkan dalam menginterpretasikan hasil penelitian. Oleh karena itu, penelitian ini perlu dipahami sebagai bukti empiris mengenai keterkaitan yang kuat antara kedua konstruk, bukan sebagai bukti hubungan kausal. Analisis ini sejalan dengan pandangan Creswell & Creswell (2022) dan Field (2018b) mengenai keterbatasan penelitian korelasional berbasis survei.

Dalam konteks Pendidikan Agama Islam, temuan penelitian ini memiliki implikasi yang lebih luas daripada sekadar peningkatan kemampuan teknis guru



dalam menyusun instrumen penilaian. Pembelajaran PAI bertujuan mengembangkan aspek kognitif, afektif, dan psikomotorik secara terpadu sehingga memerlukan asesmen formatif yang mampu memberikan informasi autentik mengenai perkembangan peserta didik. Guru yang memiliki literasi asesmen tinggi akan lebih mampu memanfaatkan hasil asesmen untuk memberikan umpan balik yang konstruktif, melakukan diferensiasi pembelajaran, serta membimbing perkembangan karakter dan spiritual peserta didik secara berkelanjutan. Oleh karena itu, penguatan literasi asesmen formatif perlu menjadi bagian penting dalam pengembangan profesional guru PAI. Temuan ini mendukung pandangan peneliti sebelumnya mengenai pentingnya asesmen formatif sebagai instrumen peningkatan kualitas pembelajaran.

Secara akademik, penelitian ini memberikan kontribusi empiris terhadap pengembangan kajian literasi asesmen formatif di Indonesia, khususnya pada konteks guru Pendidikan Agama Islam yang hingga saat ini masih relatif jarang diteliti. Berbeda dengan penelitian sebelumnya yang umumnya dilakukan pada guru mata pelajaran umum atau konteks negara lain, penelitian ini menunjukkan bahwa hubungan antara literasi penilaian formatif dan praktik penilaian formatif juga ditemukan secara sangat kuat pada guru PAI di Indonesia. Temuan ini memperluas bukti empiris mengenai relevansi model TFALS dan TFAPS pada konteks budaya yang berbeda sekaligus memberikan dasar bagi penelitian selanjutnya untuk mengembangkan model struktural yang menjelaskan faktor-faktor lain yang memengaruhi praktik asesmen formatif guru. Dengan demikian, penelitian ini tidak hanya mengonfirmasi temuan penelitian terdahulu, tetapi juga memberikan kontribusi kontekstual terhadap pengembangan literasi asesmen pada pendidikan Islam.

KESIMPULAN DAN SARAN

Penelitian ini menunjukkan bahwa terdapat hubungan positif yang sangat kuat antara literasi penilaian formatif dan praktik penilaian formatif guru Pendidikan Agama Islam di Kabupaten Lamongan. Guru yang memiliki tingkat literasi penilaian formatif yang lebih tinggi cenderung lebih mampu mengimplementasikan praktik penilaian formatif secara efektif dalam proses pembelajaran. Temuan ini menegaskan bahwa penguatan literasi asesmen merupakan salah satu prasyarat penting dalam mendukung implementasi asesmen formatif dan kebijakan Kurikulum Merdeka pada mata pelajaran Pendidikan Agama Islam. Selain memberikan bukti empiris mengenai hubungan kedua variabel, penelitian ini juga memperkuat penggunaan instrumen TFALS dan TFAPS yang telah diadaptasi pada konteks guru PAI di Indonesia.

Penelitian ini memiliki beberapa keterbatasan. Pertama, penggunaan desain cross-sectional belum memungkinkan peneliti menjelaskan hubungan kausal antara literasi penilaian formatif dan praktik penilaian formatif. Kedua, penelitian

hanya dilakukan pada guru Pendidikan Agama Islam di Kabupaten Lamongan sehingga generalisasi hasil ke wilayah lain masih terbatas. Ketiga, data dikumpulkan menggunakan kuesioner sehingga masih berpotensi dipengaruhi oleh self-report bias. Oleh karena itu, penelitian selanjutnya disarankan menggunakan desain longitudinal atau mixed methods, melibatkan responden dari berbagai daerah di Indonesia, serta mengombinasikan data survei dengan observasi praktik pembelajaran atau wawancara agar diperoleh gambaran yang lebih komprehensif mengenai implementasi asesmen formatif.

REFERENSI

- Aditomo, A. (2024). *Panduan Pembelajaran dan Asesmen Pendidikan Anak Usia Dini, Pendidikan Dasar, dan Pendidikan Menengah Edisi Revisi Tahun 2024*. 1–72.
- Akoglu, H. (2018). *Turkish Journal of Emergency Medicine User ' s guide to correlation coefficients*. 18(August), 91–93. <https://doi.org/10.1016/j.tjem.2018.08.001>
- Andrade, H. L., & Brookhart, S. M. (2020). Classroom assessment as the co-regulation of learning. *Assessment in Education: Principles, Policy and Practice*, 27(4), 350–372. <https://doi.org/10.1080/0969594X.2019.1571992>
- Anggraena, Y., Felicia, N., G, D. E., Pratiwi, I., Utama, B., Alhapip, L., & Widiawati, D. (2021). Kajian Akademik Kurikulum Untuk Pemulihan Pembelajaran. In *Pusat Kurikulum dan Pembelajaran Badan Standar, Kurikulum, dan Asesmen Pendidikan Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi*.
- Black, P. (1998). Learning, league tables and national assessment: Opportunity lost or hope deferred? *Oxford Review of Education*, 24(1), 57 – 68. <https://doi.org/10.1080/0305498980240105>
- Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educ Asses Eval Acc*, December 2008, 5–31. <https://doi.org/10.1007/s11092-008-9068-5>
- Brookhart, S. M. (2011). Educational Assessment Knowledge and Skills for Teachers. *Educational Measurement: Issues and Practice*, 30(1), 3–12. <https://doi.org/10.1111/j.1745-3992.2010.00195.x>
- Creswell, J. W., & Creswell, J. D. (2022). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*. SAGE Publications. <https://books.google.co.id/books?id=Pr2VEAAQBAJ>
- DeLuca, C., LaPointe-McEwan, D., & Luhanga, U. (2015). Teacher assessment literacy: a review of international standards and measures. *Educational Assessment, Evaluation and Accountability*, 28(3), 251–272. <https://doi.org/10.1007/s11092-015-9233-6>
- Field, A. (2018). *Discovering Statistics Using IBM SPSS Statistics (5th ed.)*. Sage Publications.
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational*



- Research*, 77(1), 81–112. <https://doi.org/10.3102/003465430298487>
- Heritage, M. (2010). *Formative Assessment: Making it Happen in the Classroom*. Corwin Press. <https://doi.org/https://doi.org/10.4135/9781452219493>
- İZCİ, K., İLTER, İ., & IZGAR, G. (2024). Adapting the teacher formative assessment literacy scale into Turkish: Validation and reliability study. *International Journal of Assessment Tools in Education*, 11(1), 67–87. <https://doi.org/10.21449/ijate.1343373>
- Leontjev, D., & deBoer, M. (2020). Conceptualising assessment and learning in the CLIL context. An introduction. In *Assessment and Learning in Content and Language Integrated Learning (CLIL) Classrooms: Approaches and Conceptualisations*. Springer International Publishing. https://doi.org/10.1007/978-3-030-54128-6_1
- Li, Z., Yan, Z., Chan, K. K. Y., Zhan, Y., & Guo, W. Y. (2023). The role of a professional development program in improving primary teachers' formative assessment literacy. *Teacher Development*, 27(4), 447–467. <https://doi.org/10.1080/13664530.2023.2223595>
- Popham, W. J. (2009). Assessment literacy for teachers: Faddish or fundamental? *Theory into Practice*, 48(1), 4–11. <https://doi.org/10.1080/00405840802577536>
- Rohmawati, D., Nisa, R., & Hasyim, H. (2023). Gaya Mengajar Guru dalam Meningkatkan Hasil Belajar Kognitif pada Mata Pelajaran Fiqih Kelas II Madrasah Ibtidaiyah. *Ibtida'*, 4(01), 78-89. <https://doi.org/10.37850/ibtida'.v4i01.478>
- Sortwell, A., Trimble, K., Ferraz, R., Geelan, D. R., Hine, G., Ramirez-Campillo, R., Carter-Thuiller, B., Gkintoni, E., & Xuan, Q. (2024). A Systematic Review of Meta-Analyses on the Impact of Formative Assessment on K-12 Students' Learning: Toward Sustainable Quality Education. *Sustainability*, 16(17), 7826. <https://doi.org/10.3390/su16177826>
- Sudakova, N. E., Savina, T. N., Masalimova, A. R., Mikhaylovsky, M. N., Karandeeva, L. G., & Zhdanov, S. P. (2022). Online Formative Assessment in Higher Education: Bibliometric Analysis. *Education Sciences*, 12(3). <https://doi.org/10.3390/educsci12030209>
- van der Steen, J., van Schilt-Mol, T., van der Vleuten, C., & Joosten-ten Brinke, D. (2022). Supporting Teachers in Improving Formative Decision-Making: Design Principles for Formative Assessment Plans. *Frontiers in Education*, 7(June), 1–9. <https://doi.org/10.3389/feduc.2022.925352>
- Wiliam, P. B. and D. (2018). Classroom assessment and pedagogy. *Assessment in Education: Principles, Policy & Practice*, 25(6). <https://doi.org/https://doi.org/10.1080/0969594X.2018.1441807>
- Winarno, S., & Azies, H. Al. (2024). The Effectiveness of Continuous Formative Assessment in Hybrid Learning Models: An Empirical Analysis in Higher

- Education Institutions. *International Journal of Pedagogy and Teacher Education*, 8(1), 1. <https://doi.org/10.20961/ijpte.v8i1.89693>
- Worabu, R. D. (2024). Assessment Literacy: Secondary School Teachers' Conception of Formative Assessment. *Studies in English Language Teaching*, 12(2), p1. <https://doi.org/10.22158/selt.v12n2p1>
- Yan, Z., & Pastore, S. (2022). Are teachers literate in formative assessment? The development and validation of the Teacher Formative Assessment Literacy Scale. *Studies in Educational Evaluation*, 74(June), 101183. <https://doi.org/10.1016/j.stueduc.2022.101183>
- Yusof, I. J., & Abdul Rahim, F. E. (2024). Kappa Fleiss Analysis: Evidence Of Content Validity for Formative Assessment Literacy Test for Teachers of Fundamental Subjects in Vocational Special School/ Analisis Fleiss Kappa: Evidens Kesahan Kandungan Ujian Literasi Pentaksiran Formatif Guru-Guru Su. *Sains Humanika*, 16(2), 9–16. <https://doi.org/10.11113/sh.v16n2.2009>
- ZHANG, K., YILMAZ, R., USTUN, A. B., & KARAOĞLAN YILMAZ, F. G. (2023). Learning analytics in formative assessment: A systematic literature review. *Eğitimde ve Psikolojide Ölçme ve Değerlendirme Dergisi*, 14(Özel Sayı), 359–381. <https://doi.org/10.21031/epod.1272054>

