

MEREVOLUSI PENILAIAN LITERASI AL-QUR'AN MELALUI KECERDASAN BUATAN: MENINGKATKAN KUALITAS EVALUASI DAN HASIL PEMBELAJARAN

Abd. Madjid^{1*}, Akbar Nur Aziz², Siti Bahiroh³

^{1,2}Doktor Psikologi Pendidikan Islam, Universitas Muhammadiyah Yogyakarta, Indonesia

³Magister Ilmu Agama Islam, Universitas Muhammadiyah Yogyakarta, Indonesia

abdulmadjid@umy.ac.id

ABSTRAK

Abstrak. Integrasi Kecerdasan Buatan (AI) dalam penilaian pendidikan menawarkan peluang untuk meningkatkan kualitas evaluasi dan hasil belajar, namun penerapannya pada pendidikan literasi Al-Qur'an masih terbatas. Penelitian ini bertujuan menguji implementasi platform penilaian berbasis AI, Q-Rate (ngaji.ai), dalam mendukung kualitas penilaian dan kemampuan literasi Al-Qur'an siswa. Penelitian menggunakan desain *mixed methods* kuasi-eksperimental dengan pendekatan one-group pretest-posttest yang melibatkan 30 siswa SMK dan tiga guru Pendidikan Agama Islam di Yogyakarta. Data dianalisis menggunakan statistik deskriptif, uji paired-samples t-test, ukuran efek, dan analisis tematik. Hasil menunjukkan skor literasi Al-Qur'an meningkat dari 28,40 menjadi 41,75 (47,01%), dengan perbedaan yang signifikan ($t(29) = -8,75$; $p < 0,001$) dan ukuran efek yang besar (Cohen's $d = 1,20$). Guru juga memberikan penilaian positif terhadap objektivitas, efisiensi, kualitas umpan balik, dan dokumentasi penilaian. Temuan ini menunjukkan bahwa penerapan penilaian berbasis AI berpotensi mendukung peningkatan kualitas evaluasi dan hasil belajar literasi Al-Qur'an. Namun, mengingat penggunaan desain one-group pretest-posttest tanpa kelompok kontrol, temuan ini perlu ditafsirkan secara hati-hati dan memerlukan konfirmasi melalui penelitian dengan desain eksperimental yang lebih kuat.

Kata Kunci: Kecerdasan Buatan; Penilaian Literasi Al-Qur'an; Penilaian berbasis AI; Pendidikan Islam; Hasil Pembelajaran; Teknologi Pendidikan.

Abstrak: *The integration of Artificial Intelligence (AI) in educational assessment offers opportunities to improve the quality of evaluation and learning outcomes, but its application to Qur'anic literacy education is still limited. This study aims to test the implementation of an AI-based assessment platform, Q-Rate (ngaji.ai), in supporting the quality of assessment and literacy skills of students. The study used a quasi-experimental mixed methods design with a one-group pretest-posttest approach involving 30 vocational school students and three Islamic Religious Education teachers in Yogyakarta. Data were analyzed using descriptive statistics, paired-samples t-test, effect size, and thematic analysis. The results showed that the Qur'an literacy score increased from 28.40 to 41.75 (47.01%), with a significant difference ($t(29) = -8.75$; $p < 0.001$) and a large effect size (Cohen's $d = 1.20$). Teachers also provide positive assessments of objectivity, efficiency, quality of feedback, and assessment documentation. These findings show that the application of AI-based assessments has the potential to support improving the quality of Qur'anic literacy evaluation and learning outcomes. However, given the use of a one-group pretest-posttest design without a control group, these findings need to be interpreted with caution and require confirmation through research with a more robust experimental design.*

Keywords: *Artificial Intelligence; Qur'anic Literacy Assessment; AI-Based Assessment; Islamic Education; Learning Outcomes; Educational Technology.*



Article History:

Received: 02-07-2026

Revised : 14-07-2026

Accepted: 21-07-2026

Online : 01-08-2026



*This is an open access article under the
CC-BY-SA license*

A. LATAR BELAKANG

Kecerdasan Buatan (AI) telah muncul sebagai salah satu teknologi paling transformatif dalam pendidikan kontemporer, membentuk kembali praktik instruksional, lingkungan belajar, dan sistem penilaian pendidikan. Kemajuan terbaru dalam pembelajaran mesin, pemrosesan bahasa alami, pengenalan suara, dan analitik pembelajaran telah memungkinkan pengembangan sistem pendidikan cerdas yang mampu memberikan instruksi adaptif, umpan balik yang dipersonalisasi, dan proses evaluasi otomatis (Chen et al., 2022). Karena institusi pendidikan semakin merangkul transformasi digital, penilaian bertenaga AI telah menjadi mekanisme strategis untuk meningkatkan kualitas pembelajaran, efisiensi penilaian, dan pengambilan keputusan berbasis bukti (George & Wooden, 2023).

Penilaian merupakan komponen mendasar dari kualitas pendidikan karena memberikan informasi mengenai kemajuan belajar siswa, efektivitas instruksional, dan pencapaian tujuan pembelajaran. Metode penilaian tradisional mengalami keterbatasan yang terkait dengan subjektivitas, inkonsistensi, umpan balik yang tertunda, dan beban administratif, terutama dalam konteks yang melibatkan evaluasi berbasis kinerja (Bennett, 2015). Menanggapi tantangan ini, sistem penilaian berbasis AI telah mendapatkan perhatian yang semakin meningkat karena kemampuannya untuk mengotomatiskan proses penilaian, menghasilkan umpan balik waktu nyata (Chen et al., 2022), dan menghasilkan analitik pembelajaran terperinci yang mendukung peningkatan pembelajaran berkelanjutan (Grosbeck et al., 2024).

Perhatian yang relatif terbatas telah dicurahkan pada pendidikan agama, khususnya pada penilaian literasi Al-Qur'an (Hikmah, 2022; Shodiq & Makrufi, 2026). Penilaian literasi Al-Qur'an berbeda secara substansial dari penilaian akademik konvensional karena melibatkan evaluasi beberapa dimensi yang dinilai melalui ujian lisan dan observasi guru. Meskipun pendekatan tersebut mempertahankan keaslian pedagogis, mereka sering menghadapi tantangan yang terkait dengan subjektivitas evaluator, inkonsistensi penilaian, dokumentasi terbatas, dan umpan balik individual yang tidak memadai (Lin et al., 2022). Akibatnya, praktik penilaian konvensional mungkin tidak cukup mendukung pemantauan sistematis kemajuan siswa atau memfasilitasi peningkatan instruksional berbasis data.

Sejalan dengan meningkatnya transformasi digital dalam pendidikan, UNESCO (2023) menekankan bahwa pemanfaatan AI dalam asesmen perlu diarahkan untuk menghasilkan evaluasi yang lebih adaptif, transparan, dan berpusat pada peserta didik. Namun, implementasi rekomendasi tersebut masih didominasi pada mata pelajaran umum, sementara bukti empiris mengenai penerapannya dalam evaluasi literasi Al-Qur'an masih sangat terbatas. Kondisi ini menunjukkan adanya kebutuhan mendesak untuk mengembangkan dan mengevaluasi sistem asesmen berbasis AI yang sesuai dengan karakteristik pembelajaran agama.

AI menawarkan potensi yang cukup besar untuk mengatasi keterbatasan ini. Melalui analisis ucapan otomatis, pengenalan pola, dan sistem umpan balik cerdas, AI dapat menganalisis kinerja pembacaan, mengidentifikasi kesalahan pengucapan, mengevaluasi aplikasi tajwid, dan memberikan umpan balik korektif langsung kepada pelajar (Sun, 2023). Studi sebelumnya secara menunjukkan umpan balik tepat waktu menjadi faktor yang mempengaruhi pencapaian pembelajaran dan perolehan keterampilan (Hattie & Timperley, 2007). AI telah menunjukkan peluang yang menjanjikan untuk mengintegrasikan teknologi digital ke dalam pendidikan Islam.

Penelitian terdahulu secara konsisten menunjukkan bahwa AI berpotensi meningkatkan objektivitas penilaian, mempercepat penyampaian umpan balik, serta mendukung personalisasi pembelajaran pada berbagai bidang, seperti pembelajaran bahasa, matematika, dan sains (Chen et al., 2022; George & Wooden, 2023a). Namun, dalam konteks pendidikan Islam, pemanfaatan AI masih didominasi sebagai media pembelajaran Al-Qur'an, aplikasi tajwid, atau sistem tutor digital, sementara penerapannya sebagai instrumen asesmen masih sangat terbatas (Madjid & Shodiq, 2026; Salsabila & Atikoh, 2024). Akibatnya, bukti empiris mengenai kemampuan AI dalam meningkatkan kualitas evaluasi sekaligus hasil belajar literasi Al-Qur'an masih minim. Selain itu, penelitian AI dalam pendidikan masih didominasi oleh konteks pembelajaran umum, sehingga kajian mengenai asesmen berbasis AI pada pendidikan Islam belum berkembang secara memadai.

Berangkat dari kesenjangan tersebut, penelitian ini menawarkan kebaruan dengan memposisikan AI sebagai sistem asesmen digital untuk mengevaluasi literasi Al-Qur'an, bukan sekadar sebagai media pembelajaran. Penelitian ini tidak hanya menguji perubahan hasil belajar siswa, tetapi juga menganalisis pengaruh asesmen berbasis AI terhadap kualitas evaluasi melalui aspek objektivitas, efisiensi, kualitas umpan balik, dan dokumentasi penilaian. Oleh karena itu, penelitian ini bertujuan menguji efektivitas platform penilaian berbasis AI dalam meningkatkan hasil belajar literasi Al-Qur'an sekaligus mengevaluasi kontribusinya terhadap kualitas asesmen. Temuan penelitian diharapkan dapat memperluas literatur Artificial Intelligence in Education dengan menghadirkan bukti empiris mengenai penerapan AI sebagai sistem asesmen dalam pendidikan Islam.

B. METODE

1. Desain Penelitian dan Partisipan

Penelitian ini menggunakan desain mixed methods kuasi-eksperimental dengan pendekatan *one-group* pretest-posttest yang bertujuan mengevaluasi implementasi awal platform AI dalam konteks pembelajaran yang berlangsung secara alami di sekolah, sehingga pembentukan kelompok kontrol tidak memungkinkan tanpa mengganggu proses pembelajaran reguler. Meskipun desain ini memiliki keterbatasan dalam validitas internal,

pendekatan ini sesuai untuk mengevaluasi kelayakan awal dan efektivitas praktis suatu inovasi pendidikan sebelum dilakukan pengujian menggunakan desain eksperimental yang lebih ketat. Komponen kuantitatif mengukur perubahan kemampuan literasi Al-Qur'an sebelum dan sesudah intervensi, sedangkan komponen kualitatif mengeksplorasi persepsi guru terhadap implementasi sistem AI (Cresswell, 2016).

Penelitian dilaksanakan di SMK Negeri 1 Godean, Yogyakarta, dengan melibatkan 30 siswa dan tiga guru Pendidikan Agama Islam yang dipilih melalui purposive sampling berdasarkan partisipasi aktif dalam pembelajaran BTQ, penyelesaian pretest dan posttest, serta kesediaan mengikuti seluruh rangkaian penelitian (Schoonenboom, 2024). Untuk meminimalkan ancaman validitas internal, seluruh peserta memperoleh intervensi yang sama, menggunakan materi BTQ yang seragam, mengikuti jadwal pembelajaran yang identik, dan dinilai menggunakan rubrik serta prosedur penilaian yang konsisten sepanjang penelitian. Selain itu, triangulasi antara hasil kuantitatif dan wawancara guru digunakan untuk meningkatkan kredibilitas temuan.

2. Intervensi Penilaian Berbasis AI

Intervensi menggunakan Q-Rate (ngaji.ai), platform penilaian berbasis AI yang menganalisis akurasi makhraj, penerapan tajwid, kefasihan membaca, serta memberikan umpan balik otomatis mengenai kesalahan dan rekomendasi perbaikan. Intervensi dilaksanakan selama enam minggu dengan frekuensi dua sesi setiap minggu yang terintegrasi dalam kegiatan BTQ. Setiap sesi berlangsung sekitar 30–40 menit, diawali dengan pembacaan ayat Al-Qur'an sesuai materi BTQ yang sedang dipelajari, kemudian siswa mengunggah rekaman bacaan melalui platform Q-Rate (ngaji.ai). Sistem AI menganalisis akurasi makhraj, penerapan tajwid, kefasihan membaca, dan kesalahan pelafalan, kemudian menghasilkan umpan balik otomatis beserta rekomendasi perbaikan. Guru mendampingi proses penggunaan platform, membantu interpretasi hasil penilaian, serta memberikan penguatan terhadap umpan balik yang dihasilkan AI. Minggu pertama digunakan untuk pretest, minggu kedua pelatihan guru, minggu ketiga hingga kelima implementasi intervensi, dan minggu keenam pelaksanaan posttest, pengisian kuesioner, serta wawancara.

Kinerja literasi Al-Qur'an diukur menggunakan rubrik sepuluh indikator dengan skala Likert lima poin. Persepsi guru mengenai kualitas evaluasi diukur menggunakan kuesioner 20 butir yang mencakup objektivitas, akurasi, efisiensi, kualitas umpan balik, dan dokumentasi penilaian (Campbell et al., 2020; Cohen, 1988), kemudian diperdalam melalui wawancara semi-terstruktur (Redecker & Punie, 2017). Validitas isi instrumen ditentukan melalui penilaian tiga ahli menggunakan Content Validity Index ($CVI \geq 0,80$) (Polit & Beck, 2021), sedangkan reliabilitas diuji

menggunakan Cronbach's Alpha dengan batas penerimaan $\geq 0,70$ (Field, 2022).

3. Analisis Data dan Etika Penelitian

Data kuantitatif dianalisis menggunakan IBM SPSS Statistics 29 melalui uji normalitas Shapiro–Wilk (Shapiro & Wilk, 1965), statistik deskriptif, paired-samples t-test, dan ukuran efek Cohen's d (Kraft, 2020). Data kualitatif dianalisis menggunakan Reflexive Thematic Analysis (Braun & Clarke, 2022). Kemudian diintegrasikan dengan temuan kuantitatif melalui triangulasi metodologis. Penelitian ini memperoleh persetujuan dari Komite Etik Riset Universitas Muhammadiyah Yogyakarta, dengan seluruh peserta memberikan persetujuan tertulis dan dijamin kerahasiaan serta anonimitas data penelitian.

C. HASIL DAN PEMBAHASAN

1. Perubahan Kinerja Literasi Al-Qur'an Siswa

Hasil statistik deskriptif menunjukkan peningkatan substansial dalam prestasi siswa setelah intervensi enam minggu, seperti terlihat pada Tabel 1.

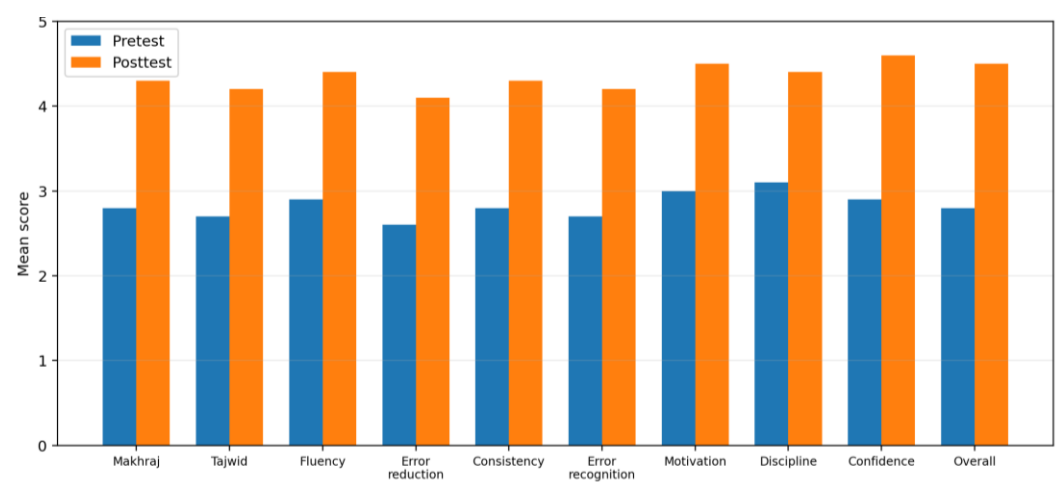
Tabel 1. Statistik Deskriptif Nilai Literasi Al-Qur'an Siswa (n = 30)

Variabel	Rata-rata Pretest	Rata-rata Posttest	Perbedaan Rata-Rata	SD	Improvement (%)
Skor Literasi Al-Qur'an	28.40	41.75	13.35	4.20	47.01

Data pada Tabel 1 menunjukkan adanya peningkatan skor literasi Al-Qur'an setelah intervensi selama enam minggu. Peningkatan terjadi pada kemampuan membaca dan indikator literasi yang diukur. Semua indikator mengalami peningkatan nilai setelah intervensi, dengan peningkatan terbesar pada kepercayaan diri membaca dan perkembangan literasi secara keseluruhan.

2. Peningkatan Kinerja di Seluruh Indikator Penilaian

Guna memeriksa dimensi literasi Al-Qur'an mana yang paling diuntungkan dari intervensi, nilai pretest dan posttest dianalisis di sepuluh indikator penilaian, seperti terlihat pada Gambar 1.



Gambar 1. Peningkatan pra-pasca tes

Untuk mendapatkan pemahaman yang lebih komprehensif tentang peningkatan pembelajaran, setiap indikator literasi Al-Qur'an dianalisis secara terpisah, seperti terlihat pada Tabel 2.

Tabel 2. Peningkatan Kinerja Literasi Al-Qur'an di Seluruh Indikator Penilaian

Indikator Penilaian	Rata-rata Pretest	Rata-rata Posttest	Peningkatan
Akurasi Makhraj	2.80	4.30	+1.50
Penerapan Tajwid	2.70	4.20	+1.50
Kelancaran Membaca	2.90	4.40	+1.50
Pengurangan Kesalahan	2.60	4.10	+1.50
Konsistensi Membaca	2.80	4.30	+1.50
Pengenalan Kesalahan	2.70	4.20	+1.50
Motivasi Belajar	3.00	4.50	+1.50
Disiplin Belajar	3.10	4.40	+1.30
Kepercayaan Diri dalam Membaca	2.90	4.60	+1.70
Perkembangan Secara Keseluruhan	2.80	4.50	+1.70

Hasilnya menunjukkan peningkatan di semua dimensi literasi Al-Qur'an yang terukur. Peningkatan substansial dalam akurasi makhraj, aplikasi tajwid, dan kefasihan membaca lebih lanjut menunjukkan bahwa umpan balik yang dihasilkan AI memfasilitasi pembelajaran yang ditargetkan dan koreksi kesalahan.

3. Pengaruh Penilaian Berbasis AI terhadap Hasil Pembelajaran

Sebelum melakukan analisis inferensial, normalitas data dinilai menggunakan uji Shapiro-Wilk. Uji-t sampel berpasangan kemudian dilakukan untuk memeriksa perbedaan antara skor pretest dan posttest, seperti terlihat pada Tabel 3.

Tabel 3. Hasil Uji t Sampel Berpasangan

Variabel	Rata-rata Pretest	Rata-rata Posttest	t-value	p-value
Skor Literasi Al-Qur'an	28.40	41.75	-8.75	.000

Analisis mengungkapkan perbedaan yang signifikan secara statistik antara skor pretest dan posttest, $t(29) = -8,75$, $p < 0,001$. Oleh karena itu, null hypothesis ditolak. Temuan penelitian ini mengindikasikan bahwa penerapan AI berkaitan dengan peningkatan kinerja literasi Al-Qur'an siswa.

4. Besarnya Efek Intervensi

Guna menentukan signifikansi praktis intervensi, ukuran efek d Cohen dihitung, seperti terlihat pada Tabel 4.

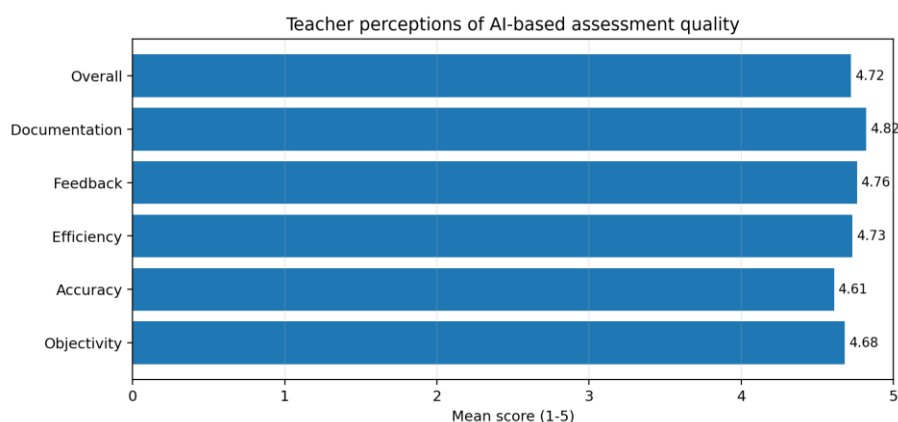
Tabel 4. Analisis Ukuran Efek

Variabel	Cohen's d	Interpretasi
Skor Literasi Al-Qur'an	1.20	Pengaruh Besar

Nilai ukuran efek 1,20 melebihi ambang batas untuk efek besar, temuan ini memberikan indikasi adanya manfaat pendidikan yang substansial. Ukuran efek yang besar sangat penting karena menunjukkan bahwa perbaikan yang diamati bermakna secara pendidikan dan kemungkinan akan direplikasi dalam konteks pembelajaran yang serupa (George & Wooden, 2023).

5. Persepsi Guru Terhadap Kualitas Evaluasi

Berikut persepsi guru terhadap kualitas evaluasi, seperti terlihat pada Gambar 4.

**Gambar 4.** Persepsi guru tentang kualitas penilaian berbasis AI

Untuk menilai pengaruh AI terhadap kualitas evaluasi, guru mengisi kuesioner evaluasi pasca-intervensi, seperti terlihat pada Tabel 5.

Tabel 5. Persepsi Guru tentang Kualitas Penilaian Berbasis AI

Dimensi	Nilai Rata-rata (M)
Objektivitas Penilaian	4.68
Akurasi Penilaian	4.61
Efisiensi Penilaian	4.73
Kualitas Umpan Balik	4.76
Dokumentasi Penilaian	4.82
Kualitas Evaluasi Secara Keseluruhan	4.72

Guru melaporkan persepsi yang sangat positif di semua dimensi kualitas evaluasi. Hal ini menunjukkan bahwa sistem AI secara efektif mendukung praktik penilaian sistematis dan peningkatan penyampaian umpan balik (Mattar et al., 2022).

6. Temuan Kualitatif

Analisis tematik wawancara guru mengidentifikasi tiga tema utama:

- a. Tema 1: Peningkatan Objektivitas Penilaian
Guru melaporkan bahwa evaluasi berbantuan AI mengurangi subjektivitas dalam penilaian dan mempromosikan konsistensi yang lebih besar dalam keputusan penilaian. "*Sistem ini mengevaluasi siswa menggunakan kriteria yang sama setiap saat, membuat proses penilaian jauh lebih adil.*"
- b. Tema 2: Kualitas Umpan Balik yang Ditingkatkan
Peserta menyoroti nilai umpan balik langsung dalam membantu siswa mengidentifikasi dan memperbaiki kesalahan pembacaan. "*Siswa menerima umpan balik langsung setelah pembacaan, memungkinkan mereka untuk berkembang lebih cepat.*"
- c. Tema 3: Peningkatan Efisiensi Penilaian
Guru menekankan bahwa platform AI mengurangi beban kerja penilaian dan memfasilitasi dokumentasi kemajuan siswa. "*Platform ini menghemat banyak waktu dan menyediakan catatan komprehensif tentang perkembangan siswa.*" Secara kolektif, temuan ini menunjukkan bahwa sistem penilaian berbasis AI berkontribusi positif terhadap kualitas evaluasi dan hasil pembelajaran, mendukung integrasinya ke dalam praktik pendidikan Islam kontemporer.

7. Pembahasan

- a. Penilaian Berbasis AI sebagai Katalis untuk Meningkatkan Hasil Pembelajaran Literasi Al-Qur'an
Temuan penelitian ini menunjukkan bahwa penerapan sistem penilaian berbasis Kecerdasan Buatan (AI) secara signifikan meningkatkan kinerja literasi Al-Qur'an siswa. Peningkatan substansial dalam skor rata-rata dari 28,40 menjadi 41,75, disertai dengan ukuran efek yang besar (Cohen's $d = 1,20$), menunjukkan bahwa penilaian yang dibantu AI dapat menghasilkan peningkatan

pendidikan yang berarti. Temuan ini mendukung penelitian sebelumnya yang menunjukkan bahwa teknologi AI berkontribusi pada peningkatan hasil pembelajaran dengan memberikan dukungan adaptif, pengalaman belajar yang dipersonalisasi, dan mekanisme umpan balik langsung (Chen et al., 2022).

Dari perspektif ilmu pembelajaran, efektivitas penilaian berbasis AI dapat dijelaskan melalui konsep penilaian formatif. Black dan (Wiliam, 2017) berpendapat bahwa penilaian menjadi paling efektif ketika berfungsi tidak hanya sebagai alat pengukuran tetapi sebagai mekanisme untuk mendukung pembelajaran. Dalam penelitian ini, sistem AI terus menganalisis kinerja pembacaan siswa dan memberikan umpan balik langsung mengenai akurasi pengucapan, kepatuhan tajwid, dan kefasihan membaca. Umpan balik semacam itu memungkinkan peserta didik untuk mengidentifikasi kelemahan dan melakukan koreksi segera, sehingga mempercepat peningkatan pembelajaran.

Peningkatan signifikan yang diamati di semua indikator penilaian menunjukkan bahwa evaluasi berbasis AI tidak hanya memengaruhi hasil pembelajaran yang terisolasi tetapi berkontribusi pada peningkatan kompetensi literasi Al-Qur'an yang komprehensif. Temuan ini konsisten dengan Teori Umpan Balik Hattie & Timperley (2017) yang mengemukakan bahwa umpan balik yang tepat waktu, spesifik, dan dapat ditindaklanjuti merupakan salah satu penentu terkuat dari pencapaian pembelajaran. Melalui umpan balik yang dihasilkan AI, siswa menerima informasi berkelanjutan tentang kinerja mereka, mengurangi ketidakpastian dan mempromosikan strategi pembelajaran yang lebih efektif.

Peningkatan kemampuan literasi Al-Qur'an pada penelitian ini kemungkinan terjadi karena sistem AI menyediakan siklus umpan balik yang lebih cepat, spesifik, dan berulang dibandingkan evaluasi konvensional. Dalam pembelajaran keterampilan seperti membaca Al-Qur'an, kemampuan memperbaiki kesalahan secara langsung memungkinkan peserta didik melakukan proses error correction secara berkelanjutan sehingga kesalahan pengucapan tidak terus berulang. Mekanisme ini sejalan dengan teori deliberate practice yang menekankan bahwa latihan yang disertai umpan balik segera akan menghasilkan peningkatan keterampilan yang lebih efektif dibandingkan latihan tanpa umpan balik yang memadai (Ericsson et al., 1993). Dengan demikian, peningkatan yang diamati pada penelitian ini secara teoritis dapat dijelaskan melalui kombinasi umpan balik adaptif, latihan berulang, dan pemantauan kemajuan yang difasilitasi oleh sistem AI.

b. Meningkatkan Dimensi Teknis dan Afektif Literasi Al-Qur'an

Temuan penting dari penelitian ini adalah bahwa perbaikan terjadi tidak hanya dalam dimensi teknis literasi Al-Qur'an tetapi juga dalam aspek afektif seperti motivasi, kepercayaan diri, dan disiplin belajar. Sementara peningkatan substansial diamati dalam akurasi makhrāj, penerapan tajwid, dan kefasihan membaca, peningkatan terbesar terjadi dalam kepercayaan membaca dan pengembangan literasi secara keseluruhan.

Temuan ini sangat signifikan karena pembelajaran literasi Al-Qur'an melibatkan proses kognitif dan afektif. Pembacaan yang efektif membutuhkan penguasaan teknis pengucapan dan tajwid, sedangkan pembelajaran yang sukses juga bergantung pada motivasi, kepercayaan diri, dan kemauan peserta didik untuk terlibat dalam latihan berulang (Hikmah, 2022). Peningkatan kepercayaan diri siswa menunjukkan bahwa penilaian berbantuan AI menciptakan lingkungan evaluasi yang lebih mendukung dan tidak terlalu mengintimidasi.

Dari perspektif teori *Self-Regulated Learning* (SRL), umpan balik yang didukung teknologi dapat memperkuat kemampuan peserta didik untuk memantau, mengevaluasi, dan mengatur proses pembelajaran mereka sendiri (Zimmerman, 2008). Platform AI memungkinkan siswa untuk berulang kali meninjau kinerja mereka, mengenali kesalahan secara mandiri, dan memantau kemajuan dari waktu ke waktu. Kesempatan untuk pemantauan diri seperti itu dapat menjelaskan peningkatan motivasi dan kepercayaan diri yang diamati.

Temuan ini sejalan dengan penelitian sebelumnya yang menunjukkan bahwa lingkungan belajar digital dapat meningkatkan keterlibatan dan otonomi belajar siswa dengan memberikan informasi kinerja yang transparan dan dukungan pembelajaran individual (Demartini et al., 2024). Akibatnya, penilaian berbasis AI tampaknya mampu mendukung dimensi teknis dan psikologis dari pengembangan literasi Al-Qur'an.

c. AI dan Peningkatan Kualitas Evaluasi

Selain meningkatkan hasil belajar siswa, penelitian ini menunjukkan bahwa AI secara signifikan meningkatkan kualitas evaluasi. Guru melaporkan tingkat kepuasan yang tinggi mengenai objektivitas, akurasi, efisiensi, kualitas umpan balik, dan dokumentasi penilaian. Temuan ini mendukung penelitian sebelumnya yang menekankan potensi sistem penilaian digital untuk meningkatkan keandalan dan konsistensi dalam evaluasi pendidikan (Jurane-Bremene, 2023).

Penilaian literasi Al-Qur'an tradisional sering kali sangat bergantung pada observasi guru dan ujian lisan. Meskipun metode tersebut memberikan wawasan pedagogis yang berharga, metode tersebut dapat dipengaruhi oleh penilaian subjektif, kelelahan evaluator, dan

inkonsistensi di seluruh sesi penilaian (Grosseck et al., 2024). Integrasi AI mengurangi keterbatasan ini dengan menerapkan kriteria evaluasi standar dan menghasilkan analisis kinerja yang konsisten.

Peringkat tinggi untuk dokumentasi penilaian menunjukkan kontribusi penting lainnya dari penilaian berbasis AI. Platform digital secara otomatis menyimpan data kinerja, memungkinkan pendidik untuk melacak kemajuan siswa secara longitudinal. Kemampuan ini mengatasi tantangan lama dalam pendidikan literasi Al-Qur'an, di mana catatan penilaian sering terfragmentasi atau tidak cukup didokumentasikan. Konsisten dengan Bennett (2015), sistem penilaian yang ditingkatkan teknologi memfasilitasi pengambilan keputusan pendidikan berbasis bukti dengan mengubah data penilaian menjadi informasi yang dapat ditindaklanjuti.

Selain itu, peringkat yang kuat untuk kualitas umpan balik menunjukkan kapasitas sistem AI untuk mendukung praktik penilaian formatif. Tidak seperti pendekatan penilaian konvensional yang sering memberikan umpan balik yang tertunda, evaluasi berbantuan AI menawarkan respons langsung yang dapat digunakan pelajar untuk meningkatkan kinerja. Temuan ini memperkuat penelitian sebelumnya yang menekankan bahwa umpan balik yang cepat sangat penting untuk akuisisi keterampilan dan pengembangan kinerja yang efektif (Hattie & Timperley, 2007).

d. **Menafsirkan Temuan Melalui Model Penerimaan Teknologi (TAM)**

Persepsi positif yang dilaporkan oleh guru dapat ditafsirkan melalui Model Penerimaan Teknologi (TAM) yang diusulkan oleh Davis (1989). Menurut TAM, adopsi teknologi terutama dipengaruhi oleh kegunaan yang dirasakan dan kemudahan penggunaan yang dirasakan. Skor tinggi yang dilaporkan untuk efisiensi, akurasi, dan kualitas umpan balik menunjukkan bahwa guru menganggap sistem AI sangat berguna dalam mendukung kegiatan penilaian. Peserta menunjukkan bahwa teknologi ini mengurangi beban kerja penilaian, meningkatkan konsistensi, dan memfasilitasi manajemen data. Manfaat ini sangat sesuai dengan konsep kegunaan yang dirasakan yang dijelaskan dalam literatur TAM.

Demikian pula, guru melaporkan bahwa platform menyederhanakan prosedur penilaian dan memungkinkan praktik evaluasi yang lebih sistematis. Temuan tersebut menunjukkan persepsi positif mengenai kemudahan penggunaan, yang semakin mendorong penerimaan dan adopsi teknologi (Alshammari & Alkhwalidi, 2025). Oleh karena itu, keberhasilan implementasi penilaian berbasis AI dalam penelitian ini memberikan dukungan empiris bagi TAM dalam konteks pendidikan Islam.

Temuan ini juga menunjukkan bahwa ketika teknologi pendidikan memberikan bukti awal dapat meningkatkan efektivitas instruksional dan mengurangi beban administrasi, guru menjadi lebih bersedia untuk mengintegrasikannya ke dalam praktik pendidikan rutin. Wawasan ini sangat penting bagi pembuat kebijakan yang ingin mempromosikan transformasi digital di lembaga pendidikan agama.

e. Kontribusi Kecerdasan Buatan dalam Pendidikan dan Pendidikan Islam

Studi ini berkontribusi pada dua domain ilmiah penting: Kecerdasan Buatan dalam Pendidikan dan penelitian Pendidikan Islam. Pertama, dalam literatur AIEd, studi sebelumnya terutama berfokus pada matematika, sains, pembelajaran bahasa, dan konteks pendidikan tinggi (Chen et al., 2022). Relatif sedikit penelitian yang telah meneliti aplikasi AI di lingkungan belajar berbasis agama atau agama. Dengan menunjukkan efektivitas penilaian berbantuan AI dalam pendidikan literasi Al-Qur'an, penelitian ini memperluas cakupan penelitian AIEd ke dalam domain yang sebelumnya kurang dieksplorasi. Kedua, dalam beasiswa pendidikan Islam, penelitian yang ada sebagian besar menekankan metode pengajaran, pengembangan kurikulum, dan media pembelajaran daripada inovasi penilaian (Mubarok et al., 2025). Temuan ini menunjukkan bahwa inovasi teknologi dapat memainkan peran transformatif tidak hanya dalam pengajaran dan pembelajaran tetapi juga dalam praktik penilaian. Akibatnya, AI harus dipandang tidak hanya sebagai alat teknologi tetapi sebagai sumber daya pendidikan strategis yang mampu meningkatkan kualitas dan akuntabilitas pendidikan dalam institusi Islam. Studi ini juga menyoroti bagaimana transformasi digital dapat diterapkan tanpa mengorbankan nilai-nilai agama dan pedagogis. Alih-alih menggantikan guru, AI berfungsi sebagai teknologi bantu yang memperkuat kapasitas guru untuk melakukan praktik penilaian yang objektif, efisien, dan berdasarkan data.

Penelitian ini memiliki beberapa keterbatasan yang perlu dipertimbangkan dalam menginterpretasikan temuan. Pertama, penggunaan desain one-group pretest-posttest tanpa kelompok kontrol membatasi kemampuan penelitian dalam mengatribusikan seluruh peningkatan hasil belajar secara eksklusif kepada intervensi AI karena masih dimungkinkan adanya pengaruh faktor lain, seperti pengalaman belajar selama penelitian atau efek pengukuran berulang. Kedua, penelitian hanya melibatkan 30 siswa dari satu sekolah sehingga generalisasi temuan ke populasi yang lebih luas perlu dilakukan secara hati-hati. Ketiga, penelitian dilakukan dalam durasi enam minggu sehingga belum mampu mengevaluasi keberlanjutan dampak penggunaan AI terhadap retensi kemampuan literasi Al-Qur'an. Oleh karena itu, penelitian selanjutnya perlu

menggunakan desain eksperimen dengan kelompok kontrol, melibatkan sampel yang lebih besar dan beragam, serta mengevaluasi implementasi jangka panjang pada berbagai jenjang pendidikan Islam.

D. SIMPULAN DAN SARAN

Penelitian ini menunjukkan bahwa implementasi sistem penilaian berbasis Kecerdasan Buatan (AI) berkaitan dengan peningkatan kualitas evaluasi dan hasil belajar literasi Al-Qur'an siswa pada konteks penelitian ini. Penerapan AI diikuti oleh peningkatan kemampuan literasi Al-Qur'an serta persepsi positif guru terhadap objektivitas, akurasi, efisiensi, kualitas umpan balik, dan dokumentasi penilaian. Temuan ini memberikan bukti awal mengenai potensi penerapan AI dalam kajian *Artificial Intelligence in Education* (AIEd), khususnya sebagai sistem asesmen digital pada pendidikan Islam yang berfungsi mendukung, bukan menggantikan, peran guru dalam proses evaluasi. Secara praktis, hasil penelitian mengindikasikan bahwa sistem penilaian berbasis AI berpotensi mendukung praktik evaluasi yang lebih objektif, sistematis, dan berbasis data di lembaga pendidikan Islam. Namun, mengingat penelitian ini menggunakan desain one-group pretest-posttest, melibatkan sampel yang relatif kecil, dan dilakukan pada satu institusi, temuan perlu ditafsirkan secara hati-hati. Penelitian selanjutnya disarankan menggunakan desain eksperimen yang lebih kuat, seperti pretest-posttest control group, melibatkan sampel yang lebih besar, serta menguji implementasi sistem asesmen berbasis AI pada berbagai jenjang dan institusi pendidikan agar diperoleh bukti empiris yang lebih kuat dan memiliki daya generalisasi yang lebih baik.

UCAPAN TERIMA KASIH

Kami mengucapkan terima kasih kepada SMK Negeri 1 Godean yang telah menjadi mitra dalam kegiatan pengabdian kepada masyarakat ini. Kami juga menyampaikan apresiasi kepada Lembaga Pengabdian Masyarakat Universitas Muhammadiyah Yogyakarta yang telah memberikan dukungan dalam pelaksanaan program ini.

DAFTAR RUJUKAN

- Alshammari, S. H., & Alkhwalid, A. F. (2025). An integrated approach using social support theory and technology acceptance model to investigate the sustainable use of digital learning technologies. *Scientific Reports*, 15(1), 342. <https://doi.org/10.1038/s41598-024-83450-z>
- Bennett, R. E. (2015). The changing nature of educational assessment. *Review of Research in Education*, 39(1), 370–407. <https://doi.org/10.3102/0091732X14554179>
- Braun, V., & Clarke, V. (2022). Thematic analysis: A practical guide. *SAGE*, 1(33), 46–50. <https://doi.org/10.53841/bpsqmp.2022.1.33.46>
- Campbell, S., Greenwood, M., Prior, S., Shearer, T., Walkem, K., Young, S., Bywaters, D., & Walker, K. (2020). Purposive sampling: Complex or simple? Research

- case examples. *Journal of Research in Nursing*, 25(8), 652–661. <https://doi.org/10.1177/1744987120927206>
- Chen, X., Zou, D., Xie, H., Cheng, G., & Liu, C. (2022). Two decades of artificial intelligence in education: Contributors, collaborations, research topics, challenges, and future directions. *Educational Technology & Society*, 25(1), 28–47. <https://doi.org/10.1007/s10639-022-11209-y>
- Cohen, J. (1988). Statistical power analysis for the behavioral sciences (2nd ed.). Lawrence Erlbaum Associates, 84(408), 1096. <https://doi.org/10.2307/2290095>
- Cresswell, J. W. (2016). Revisiting Mixed Methods and Advancing Scientific Practices. In *The Oxford Handbook of Multimethod and Mixed Methods Research Inquiry* (pp. 57–71). Oxford Academic.
- Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 13(3), 319–340.
- Demartini, C. G., Sciascia, L., Bosso, A., & Manuri, F. (2024). Artificial intelligence bringing improvements to adaptive learning in education: A case study. *Sustainability*, 16(3), 1347. <https://doi.org/10.3390/su16031347>
- Field, A. (2022). *Discovering statistics using IBM SPSS Statistics (6th ed.)*. SAGE. <https://doi.org/10.3389/fpsy.2019.00864>
- George, B., & Wooden, O. (2023a). administrative sciences Managing the Strategic Transformation of Higher Education through Artificial Intelligence. *Administrative Sciences*, 13(196), 1–20. <https://doi.org/10.3390/admsci13090196> Received:
- George, B., & Wooden, O. (2023). Managing the strategic transformation of higher education through artificial intelligence. *Administrative Sciences*, 13(9), 196. <https://doi.org/10.3390/admsci13090196>
- Giannini, S. (2023). *Generative AI and the future of education*. UNESCO. <https://doi.org/10.54675/HOXG8740>
- Grossec, G., Bran, R. A., & Țîru, L. G. (2023). Digital assessment: A survey of Romanian higher education teachers' practices and needs. *Education Sciences*, 14(1), 32. <https://doi.org/10.3390/educsci14010032>
- Hattie, J., & Timperley, H. (2017). The power of feedback. *Review of Educational Research*, 77(1), 81–112. <https://doi.org/10.3102/003465430298487>
- Hikmah, N. (2022). Implementation of learning to read writing the Quran (BTQ) Qutab system and their relevance in early children's education. *Edukasi Islami: Jurnal Pendidikan Islam*, 11(1), 247–266. <https://doi.org/10.30868/ei.v11i01.2252>
- Jurāne-Brēmane, A. (2023). Digital assessment in technology-enriched education: Thematic review. *Education Sciences*, 13(5), 522. <https://doi.org/10.3390/educsci13050522>
- Kraft, M. A. (2020). Interpreting effect sizes of education interventions. *Educational Researcher*, 49(4), 241–253. <https://doi.org/10.3102/0013189X20912798>
- Lin, V., Yeh, H.-C., & Chen, N.-S. (2022). A systematic review on oral interactions in robot-assisted language learning. *Electronics*, 11(2), 290. <https://doi.org/10.3390/electronics11020290>
- Madjid, A., & Shodiq, S. F. (2026). Enhancing religious character through a technology-based learning model in boarding school: Examining the influence of wellbeing, critical thinking skills, and social skills as moderator variables. *Frontiers in Education*, 11, 1758690. <https://doi.org/10.3389/feduc.2026.1758690>
- Mattar, J., Ramos, D. K., & Lucas, M. (2022). DigComp-based digital competence assessment tools: Literature review and instrument analysis. *Education and Information Technologies*, 27, 10843–10867. <https://doi.org/10.1007/s10639-022-11034-3>
- Mubarok, A., Munawaroh, Saepudin, Mifatahudin, Sudrajat, S., & Rahmatullah, M.

- A. (2025). Efforts to improve the quality of Islamic religious education through the contextual teaching and learning (CTL) model in the Qur'an reading and writing (BTQ) material. *Dirasah: Jurnal Pendidikan Islam*, 6(2). <https://doi.org/10.29210/810667700>
- Polit, D. F., & Beck, C. T. (2021). *Nursing research: Generating and assessing evidence for nursing practice (11th ed.)*. Wolters Kluwer. <https://doi.org/10.1093/eurjcn/zvab038>
- Redecker, C., & Punie, Y. (2017). *European framework for the digital competence of educators: DigCompEdu*. Publications Office of the European Union. <https://doi.org/10.2760/159770>
- Salsabila, I. A., & Atikoh, N. (2024). Implementasi aplikasi “ngaji.ai” dalam pembelajaran Al-Qur'an. *Sosaintek: Jurnal Ilmu Sosial Sains Dan Teknologi*, 1(4), 265–272. <https://doi.org/10.33367/sosaintek.v1i4.7318>
- Schoonenboom, J. (2024). Design patterns in mixed methods research. *Journal of Mixed Methods Research*, 18(3), 383–392. <https://doi.org/10.1177/15586898241257284>
- Shapiro, S. S., & Wilk, M. B. (1965). An analysis of variance test for normality. *Biometrika*, 52(3–4), 591–611. <https://doi.org/10.2307/2333709>
- Shodiq, S. F., & Makrufi, A. D. (2026). Empathy as a moderator among social competence, peer support, and adolescent responsibility in Indonesian adolescents: An empirical study. *International Journal of Adolescence and Youth*, 31(1), 2657347. <https://doi.org/10.1080/02673843.2026.2657347>
- Sun, W. (2023). The impact of automatic speech recognition technology on second language pronunciation and speaking skills of EFL learners: A mixed methods investigation. *Frontiers in psychology*, 14, 1210187. <https://doi.org/10.3389/fpsyg.2023.1210187>
- Wiliam, D. (2017). Assessment for learning: meeting the challenge of implementation. *Assessment in Education: Principles, Policy & Practice*, 25(6), 682–685. <https://doi.org/10.1080/0969594x.2017.1401526>
- Zimmerman, B. J. (2008). Investigating self-regulation and motivation: Historical background, methodological developments, and future prospects. *American Educational Research Journal*, 45(1), 166–183. <https://doi.org/10.3102/0002831207312909>