

PENGEMBANGAN INSTRUMEN PENILAIAN FORMATIF AUTENTIK
KETERAMPILAN MEMBACA INFERENSIAL BAHASA ARAB BERBASIS
APLIKASI WAYGROUND DI PONDOK PESANTREN SYAIKH
JAMILURRAHMAN BANTUL



Oleh : Zailani Novian

NIM : 24204021011

TESIS

Diajukan Kepada Program Magister (S2)
Fakultas Ilmu Tarbiyah dan Keguruan UIN Sunan Kalijaga
Untuk Memenuhi Salah Satu Syarat guna Memperoleh
Gelar Magister Pendidikan (M.Pd)

Program Studi Pendidikan Bahasa Arab

YOGYAKARTA

2026

PERNYATAAN KEASLIAN

Yang bertanda tangan di bawah ini:

Nama : Zailani Novian
NIM : 24204021011
Jenjang : Magister (S2)
Program Studi : Pendidikan Bahasa Arab

menyatakan bahwa naskah tesis ini secara keseluruhan adalah hasil penelitian/karya saya sendiri, kecuali pada bagian-bagian yang dirujuk sumbernya.

Yogyakarta, 6 Juni 2026

Saya yang menyatakan,



Zailani Novian, S.Pd

NIM : 24204021011

PERNYATAAN BEBAS PLAGIASI

Yang bertanda tangan di bawah ini:

Nama : Zailani Novian
NIM : 24204021011
Jenjang : Magister (S2)
Program Studi : Pendidikan Bahasa Arab

menyatakan bahwa naskah tesis ini secara keseluruhan benar-benar bebas dari plagiasi. Jika di kemudian hari terbukti melakukan plagiasi, maka saya siap ditindak sesuai ketentuan hukum yang berlaku.

Yogyakarta, 6 Juni 2026

Saya yang menyatakan,



Zailani Novian, S.Pd

NIM : 24204021011

MOTTO

“The mind passes from understanding less to understanding more.”¹

Mortimer J. Adler dan Charles Van Doren

لَا خَيْرَ فِي عِلْمٍ لَا فَهْمَ فِيهِ، وَلَا خَيْرَ فِي قِرَاءَةٍ لَا تَدَبُّرَ فِيهَا

“Tidak ada kebaikan pada ilmu yang tidak disertai pemahaman, dan tidak ada kebaikan pada bacaan yang tidak disertai perenungan.”

Ali bin Abi Thalib

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

¹ Charles Van Doren and Mortimer J Adler, *How to Read a Book: The Classic Guide to Intelligent Reading* (Blackstone Audio, 2012).

ABSTRAK

Zailani Novian, Pengembangan Instrumen Penilaian Formatif Autentik Keterampilan Membaca Inferensial Bahasa Arab Berbasis Aplikasi *Wayground* Di Pondok Pesantren Syaikh Jamilurrahman Bantul. **Tesis: Yogyakarta, Program Magister Pendidikan Bahasa Arab, Fakultas Ilmu Tarbiyah dan Keguruan, UIN Sunan Kalijaga, 2026.**

Penelitian ini dilatarbelakangi oleh penilaian keterampilan membaca bahasa Arab di Pondok Pesantren Syaikh Jamilurrahman Bantul yang masih cenderung berorientasi pada pemahaman literal, dilaksanakan secara konvensional, dan belum dimanfaatkan secara optimal sebagai sarana pemberian umpan balik. Penelitian ini bertujuan untuk: (1) menghasilkan instrumen penilaian formatif autentik berbasis aplikasi *Wayground* yang sesuai dengan karakteristik keterampilan membaca inferensial bahasa Arab dan konteks pesantren; (2) mendeskripsikan prosedur pengembangannya; serta (3) mengetahui kelayakan dan keterterapan instrumen yang dikembangkan.

Penelitian ini menggunakan metode penelitian dan pengembangan dengan model 4-D yang meliputi tahap *define*, *design*, *develop*, dan *disseminate*. Subjek uji coba terdiri atas 19 santri *mustawa awwal*. Data dikumpulkan melalui observasi, wawancara, dokumentasi, lembar validasi ahli, angket respons pengguna, lembar keterlaksanaan, dan tes membaca inferensial. Produk divalidasi oleh ahli materi bahasa Arab, guru/praktisi, dan ahli media. Data dianalisis secara deskriptif kuantitatif dan kualitatif, sedangkan karakteristik butir dianalisis menggunakan model Rasch yang mencakup unidimensionalitas, kesesuaian butir, reliabilitas, dan tingkat kesukaran.

Hasil penelitian menunjukkan bahwa (1) instrumen yang dikembangkan berbasis aplikasi *Wayground* memiliki karakteristik formatif, autentik, inferensial, digital, dan kontekstual pesantren. Produk terdiri atas indikator membaca inferensial, kisi-kisi, butir soal, pedoman penskoran, serta implementasi penilaian melalui aplikasi *Wayground*. (2) Hasil penilaian para ahli menunjukkan bahwa instrumen layak digunakan dengan skor rata-rata 84% dengan kategori sangat

Baik, kemudian melalui *expert judgment* CVR/CVI hasilnya Instrumen Valid. Analisis Rasch menunjukkan bahwa varians yang dapat dijelaskan oleh ukuran sebesar 32,0%, sehingga memenuhi batas minimal unidimensionalitas. Nilai Alpha Cronbach sebesar 0,77 menunjukkan konsistensi internal yang memadai. (3) Instrumen dapat diterapkan dalam empat kali pertemuan uji coba dengan tingkat keterlaksanaan pembelajaran 100 % dan membantu penyajian soal, pengumpulan respons, serta pengelolaan hasil penilaian. Dengan demikian, instrumen penilaian formatif autentik berbasis aplikasi *Wayground* layak digunakan sebagai produk awal secara terbatas untuk menilai keterampilan membaca inferensial bahasa Arab.



الملخص

زَيْلَانِي نُوفِيَان، تطوير أداة التقويم التكويني الأصل لمهارة القراءة الاستدلالية باللغة العربية القائمة على تطبيق *Wayground* في معهد الشيخ جميل الرحمن الإسلامي ببنّتول .رسالة ماجستير: يوجياكرتا، برنامج الماجستير في تعليم اللغة العربية، كلية العلوم التربوية وتأهيل المعلمين، جامعة سونان كاليجاكا الإسلامية الحكومية، ٢٠٢٦ م.

انثقت هذه الدراسة من واقع تقويم مهارة القراءة باللغة العربية في معهد الشيخ جميل الرحمن الإسلامي ببنّتول، حيث ما يزال التقويم يميل إلى التركيز على الفهم الحرفي، ويُقَدَّرُ بالأساليب التقليدية، كما لم يُستفد منه على الوجه الأمثل بوصفه وسيلةً لتقديم التغذية الراجعة. وهدفت هذه الدراسة إلى: (١) إنتاج أداة للتقويم التكويني الأصل قائمة على تطبيق *Wayground* تتلاءم مع خصائص مهارة القراءة الاستدلالية باللغة العربية وسياق التعليم في المعهد الإسلامي؛ و(٢) وصف إجراءات تطوير الأداة؛ و(٣) التعرف على مدى صلاحية الأداة المطوّرة وقابليتها للتطبيق.

استخدمت الدراسة منهج البحث والتطوير وفق نموذج 4-D، الذي يشتمل على أربع مراحل، وهي: التعريف، والتصميم، والتطوير، والنشر. وتكوّنت عينة التجربة من تسعة عشر طالبًا من طلاب المستوى الأول. وجمعت البيانات من خلال الملاحظة، والمقابلة، والتوثيق، واستمارات تحكيم الخبراء، واستبانة استجابات المستخدمين، واستمارة متابعة التنفيذ، واختبار القراءة الاستدلالية. وقد خضع المنتج للتحكيم من قِبَل خبير في مادة اللغة العربية، ومعلّم/ممارس تربوي، وخبير في الوسائط

التعليمية. وحُلِّلت البيانات باستخدام التحليل الوصفي الكمي والكيفي، في حين حُلِّلت خصائص فقرات الاختبار باستخدام نموذج راش (*Rasch Model*) ، الذي شمل أحادية البعد، وملاءمة الفقرات للنموذج، والثبات، ومستوى صعوبة الفقرات.

وأظهرت نتائج الدراسة ما يأتي (١) أن أداة التقييم المطوّرة والقائمة على تطبيق *Wayground* تتسم بخصائص تكوينية، وأصيلة، واستدلالية، ورقمية، ومراعية لسياق المعهد الإسلامي. ويتكوّن المنتج من مؤشرات مهارة القراءة الاستدلالية، وجدول المواصفات، وفقرات الاختبار، ودليل التصحيح والتقدير، فضلاً عن تنفيذ عملية التقييم من خلال تطبيق *Wayground* (٢) أظهرت نتائج تقييم الخبراء أن الأداة صالحة للاستخدام، إذ حصلت على نسبة ٨٤٪ ضمن فئة «جيد جداً»، كما أظهرت نتائج التحكيم الخبروي باستخدام نسبة صدق المحتوى (CVR) ومؤشر صدق المحتوى (CVI) أن الأداة تتمتع بصدق المحتوى. وأظهر تحليل نموذج راش أن نسبة التباين المفسّر بالقياسات بلغت ٣٢٪، وهي نسبة تستوفي الحد الأدنى المطلوب لأحادية البعد. كما بلغ معامل ألفا كرونباخ ٠,٧٧، مما يدل على مستوى مقبول من الاتساق الداخلي (٣) أمكن تطبيق الأداة خلال أربع جلسات تجريبية، وبلغت نسبة تنفيذ الأنشطة التعليمية ١٠٠٪، كما أسهمت الأداة في تسهيل عرض الأسئلة، وجمع استجابات الطلاب، وإدارة نتائج التقييم. وبناءً على ذلك، يمكن القول إن أداة التقييم التكويني الأصيل القائمة على تطبيق *Wayground* صالحة بوصفها منتجاً أولياً للاستخدام المحدود في تقييم مهارة القراءة الاستدلالية باللغة العربية.



KEMENTERIAN AGAMA
UNIVERSITAS ISLAM NEGERI SUNAN KALIJAGA
FAKULTAS ILMU TARBIYAH DAN KEGURUAN
Jl. Marsda Adisucipto Telp. (0274) 513056 Fax. (0274) 586117 Yogyakarta 55281

PENGESAHAN TUGAS AKHIR

Nomor : B-2850/Un.02/DT/PP.00.9/08/2026

Tugas Akhir dengan judul : PENGEMBANGAN INSTRUMEN PENILAIAN FORMATIF AUTENTIK KETERAMPILAN MEMBACA INFERENSIAL BAHASA ARAB BERBASIS APLIKASI *FLYGROUND* DI PONDOK PESANTREN SYAIKH JAMILURRAHMAN BANTUL

yang dipersiapkan dan disusun oleh:

Nama : ZAILANI NOVIAN, S. Pd
Nomor Induk Mahasiswa : 24204021011
Telah diujikan pada : Selasa, 28 Juli 2026
Nilai ujian Tugas Akhir : A

dinyatakan telah diterima oleh Fakultas Ilmu Tarbiyah dan Keguruan UIN Sunan Kalijaga Yogyakarta

TIM UJIAN TUGAS AKHIR



Valid ID: 6a8e572987156

Ketua Sidang

Dr. Hj. R Umi Baroroh, S.Ag, M.Ag.
SIGNED



Valid ID: 6a8645d3c1473

Penguji I

Dr. Nurhadi, S.Ag, MA
SIGNED



Valid ID: 6a8d7ab23b4e2

Penguji II

Dr. Rohmatun Lukluk Isnaini, S.Pd.I., M.Pd.I.
SIGNED



Valid ID: 6a8f6d3515577

Yogyakarta, 28 Juli 2026
UIN Sunan Kalijaga
Dekan Fakultas Ilmu Tarbiyah dan Keguruan
Prof. Dr. Sigit Purnama, S.Pd.I., M.Pd.
SIGNED

NOTA DINAS PEMBIMBING

Kepada Yth.

Dekan Fakultas Ilmu Tarbiyah dan Keguruan UIN Sunan Kalijaga
Yogyakarta

Assalamu'alaikum wr. wb.

Setelah melakukan bimbingan, arahan, dan koreksi terhadap penulisan tesis yang berjudul:

**PENGEMBANGAN INSTRUMEN PENILAIAN FORMATIF AUTENTIK
KETERAMPILAN MEMBACA INFERENSIAL BAHASA ARAB BERBASIS
APLIKASI WAYGROUND DI PONDOK PESANTREN SYAIKH
JAMILURRAHMAN BANTUL**

yang ditulis oleh:

Nama : Zailani Novian

NIM : 24204021011

Jenjang : Magister (S2)

Program Studi : Pendidikan Bahasa Arab

Saya berpendapat bahwa tesis tersebut sudah dapat diajukan kepada Program Magister (S2) Fakultas Ilmu Tarbiyah dan Keguruan UIN Sunan Kalijaga untuk diujikan dalam rangka memperoleh gelar Magister Pendidikan (M.Pd.).

Wassalamu'alaikum wr. wb.

Yogyakarta, 01 Juli 2026

Pembimbing,

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

Dr. Hj. R Umi Baroroh, S.Ag M. Ag

KATA PENGANTAR

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

الْحَمْدُ لِلَّهِ رَبِّ الْعَالَمِينَ، وَالصَّلَاةُ وَالسَّلَامُ عَلَى أَشْرَفِ الْأَنْبِيَاءِ وَالْمُرْسَلِينَ، سَيِّدِنَا وَمَوْلَانَا مُحَمَّدٍ، وَعَلَى آلِهِ وَصَحْبِهِ أَجْمَعِينَ، أَمَّا بَعْدُ

Segala puji dan syukur penulis panjatkan ke hadirat Allah Swt. atas limpahan rahmat, taufik, hidayah, serta pertolongan-Nya sehingga penulis dapat menyelesaikan tesis yang berjudul “Pengembangan Instrumen Penilaian Formatif Autentik Berbasis Aplikasi Wayground untuk Keterampilan Membaca Inferensial Bahasa Arab di Pondok Pesantren Syaikh Jamilurrahman Bantul”. Sholawat dan salam semoga senantiasa tercurahkan kepada Nabi Muhammad saw., keluarga, para sahabat, serta seluruh umat beliau yang istiqamah mengikuti sunnahnya hingga akhir zaman.

Tesis ini disusun sebagai salah satu persyaratan untuk memperoleh gelar Magister Pendidikan pada Program Studi Magister Pendidikan Bahasa Arab, Fakultas Ilmu Tarbiyah dan Keguruan, Universitas Islam Negeri Sunan Kalijaga Yogyakarta. Penelitian ini dilatarbelakangi oleh pentingnya pengembangan instrumen penilaian yang mampu mengukur keterampilan membaca inferensial bahasa Arab secara lebih mendalam, sekaligus berfungsi sebagai sarana pemberian umpan balik selama proses pembelajaran. Melalui pengintegrasian prinsip penilaian formatif, autentik, dan teknologi digital berbasis aplikasi Wayground, instrumen yang dikembangkan diharapkan dapat menjadi alternatif penilaian yang layak, praktis, serta sesuai dengan karakteristik pembelajaran bahasa Arab di lingkungan pesantren.

Penulis menyadari bahwa penyelesaian tesis ini tidak terlepas dari bantuan, arahan, dukungan, doa, dan kerja sama dari berbagai pihak. Oleh karena itu, dengan penuh rasa hormat dan kerendahan hati, penulis menyampaikan terima kasih dan penghargaan yang sebesar-besarnya kepada:

1. Prof. Noorhaidi Hasan, M.A., M.Phil., Ph.D., selaku Rektor Universitas Islam Negeri Sunan Kalijaga Yogyakarta.
2. Prof. Dr. Sigit Purnama, S.Pd.I., M.Pd., selaku Dekan Fakultas Ilmu Tarbiyah dan Keguruan Universitas Islam Negeri Sunan Kalijaga Yogyakarta.
3. Dr. Dailatus Syamsiyah, S.Ag., M.Ag., Ketua Program Studi Magister Pendidikan Bahasa Arab Fakultas Ilmu Tarbiyah dan Keguruan Universitas Islam Negeri Sunan Kalijaga Yogyakarta yang telah memberikan pelayanan, arahan, selama penulis menempuh pendidikan.
4. Dr. Nasiruddin, M.S.I., M.Pd.I Sekretaris Program Studi Magister Pendidikan Bahasa Arab Fakultas Ilmu Tarbiyah dan Keguruan Universitas Islam Negeri Sunan Kalijaga Yogyakarta yang telah memberikan pelayanan, arahan, selama penulis menempuh pendidikan
5. Dr. Hj. R Umi Baroroh, S.Ag, M.Ag selaku dosen pembimbing tesis yang dengan penuh kesabaran telah meluangkan waktu, memberikan bimbingan, arahan, koreksi, motivasi, dan masukan akademik kepada penulis sejak proses penyusunan hingga penyelesaian tesis ini.
6. Dr. H. Zainal Arifin Ahmad, M.Ag., Selaku Dosen Penasihat Akademik yang telah memberikah nasehat dan arahan selama masa studi hingga penyusunan tesis ini.
7. Dr. Agung Setiawan, S.Pd.I., M.Pd.I Selaku Dosen ahli materi, , ahli media, dan praktisi yang telah bersedia memberikan penilaian, saran, serta masukan terhadap instrumen penilaian yang dikembangkan dalam penelitian ini.
8. Seluruh dosen Program Studi Magister Pendidikan Bahasa Arab UIN Sunan Kalijaga Yogyakarta yang telah membekali penulis dengan ilmu pengetahuan, pengalaman, wawasan, dan keteladanan selama proses perkuliahan.
9. Pimpinan, pengasuh, para ustadz, dan seluruh keluarga besar Pondok Pesantren Syaikh Jamilurrahman Bantul yang telah memberikan izin, bantuan, dan kesempatan kepada penulis untuk melaksanakan penelitian.

10. Guru bahasa Arab dan seluruh santri mustawa awwal Pondok Pesantren Syaikh Jamilurrahman Bantul yang telah berpartisipasi dan bekerja sama dengan baik selama proses observasi, wawancara, dan uji coba instrumen penelitian.
11. Ibu, istri, anak-anak dan keluarga tercinta, serta orang-orang terdekat penulis yang senantiasa memberikan doa, kasih sayang, nasihat, semangat, dan dukungan baik secara moral maupun material dalam setiap tahapan pendidikan penulis.
12. Teman-teman seperjuangan Program Studi Magister Pendidikan Bahasa Arab yang telah menjadi rekan belajar, berdiskusi, saling membantu, dan memberikan motivasi selama proses perkuliahan dan penyusunan tesis.
13. Semua pihak yang tidak dapat penulis sebutkan satu per satu yang telah memberikan bantuan dan dukungan dalam penyelesaian tesis ini.

Semoga segala bantuan, bimbingan, dukungan, dan kebaikan yang telah diberikan mendapatkan balasan pahala yang berlipat ganda dari Allah Swt. Penulis menyadari bahwa tesis ini masih memiliki keterbatasan dan belum sepenuhnya sempurna. Oleh karena itu, kritik dan saran yang konstruktif sangat diharapkan sebagai bahan perbaikan dan penyempurnaan pada masa yang akan datang.

Akhirnya, penulis berharap semoga tesis ini dapat memberikan manfaat bagi pengembangan keilmuan Pendidikan Bahasa Arab, khususnya dalam bidang penilaian keterampilan membaca inferensial, serta dapat menjadi referensi bagi guru, peneliti, dan lembaga pendidikan dalam mengembangkan penilaian formatif autentik berbasis teknologi digital.

Yogyakarta, 6 Juni 2026

Saya yang menyatakan,



Zailani Novian, S.Pd

NIM : 24204021011

DAFTAR ISI

PERNYATAAN KEASLIAN	i
PERNYATAAN BEBAS PLAGIASI.....	ii
MOTTO	iii
ABSTRAK	iv
PERSETUJUAN TIM PENGUJI UJIAN TESIS	viii
NOTA DINAS PEMBIMBING	viii
KATA PENGANTAR	ix
DAFTAR ISI	xiii
BAB I PENDAHULUAN	1
A. Latar Belakang Masalah	1
B. Rumusan Masalah.....	12
C. Tujuan Penelitian	13
D. Manfaat Penelitian	14
E. Kajian Pustaka	16
BAB V PENUTUP	182
A. Simpulan	182
B. Saran Pemanfaatan Produk	184
C. Saran Pengembangan Produk.....	186
DAFTAR PUSTAKA.....	187

BAB I PENDAHULUAN

A. Latar Belakang Masalah

Pembelajaran bahasa Arab memiliki posisi strategis dalam pendidikan Islam karena berfungsi sebagai alat utama dalam memahami sumber-sumber ajaran Islam serta khazanah keilmuan klasik dan kontemporer. Salah satu keterampilan bahasa yang memiliki peran sentral dalam pembelajaran bahasa Arab adalah keterampilan membaca (*maharah qirā'ah*), karena melalui kegiatan membaca peserta didik dapat mengakses, menafsirkan, dan mengkonstruksi makna teks secara mandiri.² Acep Hermawan menegaskan bahwa keterampilan membaca dalam pembelajaran bahasa Arab tidak sekadar melafalkan teks, tetapi mencakup proses pemahaman makna yang kompleks dan bertingkat. Oleh karena itu, penguasaan keterampilan membaca yang baik menjadi indikator penting keberhasilan pembelajaran bahasa Arab.³

Dalam kajian keterampilan membaca, para ahli membedakan empat tingkat pemahaman membaca, antara lain pemahaman literal, inferensial, kritis, dan evaluatif. Di antara tingkatan tersebut, membaca inferensial menuntut kemampuan kognitif yang lebih tinggi karena melibatkan penarikan kesimpulan

² Darminton Mondolalo and La Muda, "Keterampilan Membaca Pemahaman Level Menengah Paragraf Deskriptif Siswa Kelas VII SMP Karuna Dipa Palu The Intermediate Level Reading Comprehension Skills Descriptive Paragraph by Eighth Grade Karuna Dipa Palu Junior High School Students Jurnal Pembelajar," Jurnal Pembelajaran Bahasa Dan Sastra 1, no. 2 (2022): 263–270, <https://doi.org/https://doi.org/10.55905/jpbs.v1i2.34>.

³ Ade Arip Ardiansyah, Nanang Kosim, and Acep Hermawan, 'Al- Ta'rib The Use of Animated Videos in Increasing the Arabic Learning Interest', *Al-Ta'rib: Jurnal Ilmiah Program Studi Pendidikan Bahasa Arab*, 10.1 (2022), pp. 35–50, [doi:https://doi.org/10.23971/altarib.v10i1.3942](https://doi.org/10.23971/altarib.v10i1.3942) The.

implisit, penafsiran makna tersirat, serta pengaitan informasi teks dengan pengetahuan dan konteks pembaca. Snow menjelaskan bahwa pemahaman inferensial merupakan inti dari *reading comprehension*⁴, karena pembaca harus membangun makna yang tidak secara eksplisit tertulis dalam teks.⁵ Dalam konteks bahasa Arab, keterampilan membaca inferensial menjadi semakin penting mengingat struktur bahasa serta kekayaan makna leksikal dan gramatikalnya yang menuntut kemampuan penalaran dan interpretasi mendalam⁶.

Meskipun membaca inferensial memiliki peran penting, praktik pembelajaran bahasa Arab di berbagai lembaga pendidikan, termasuk pesantren, masih cenderung menitikberatkan pada pemahaman literal, seperti mengenali kosakata dan menjawab pertanyaan faktual. Hal ini sejalan dengan temuan penelitian Brown yang menyatakan bahwa asesmen membaca sering kali hanya mengukur kemampuan permukaan teks, bukan kemampuan penalaran dan interpretasi mendalam. Akibatnya, peserta didik kurang terlatih dalam mengembangkan kemampuan berpikir tingkat tinggi (*Higher Order Thinking Skills*) dalam membaca teks bahasa Arab.⁷

⁴ Catherine Snow, *Reading for Understanding: Toward an R&D Program in Reading Comprehension* (Rand Corporation, 2002).

⁵ Bunga Ayu Wulandari, Muhammad Sofwan, and R Septa Kusuma, 'Meningkatkan Kemampuan Membaca Pemahaman Literal Dan Inferensial Dengan Menggunakan Model Hands on Actifity', *JIIIP (Jurnal Ilmiah Ilmu Pendidikan)*, 6.5 (2023), pp. 2987–93.

⁶ Elinor Saiegh-Haddad and Rachel Schiff, 'Diglossic and Orthographic Features of Reading Comprehension in Standard Arabic: The Primacy of the Spoken Language', *Reading Research Quarterly*, 60.1 (2025), p. e598.

⁷ Maisa Septia AP, Chandra, and Inggria Kharisma, 'Kemampuan Berfikir Kritis Dalam Keterampilan Membaca Pemahaman Peserta Didik Kelas VI SD', *Pediaqu : Jurnal Pendidikan Sosial Dan Humaniora*, 4.3 (2025), pp. 5837–55.

Berdasarkan kedudukannya tersebut, membaca inferensial dipilih sebagai fokus penelitian karena menjadi penghubung antara pemahaman literal dan kemampuan membaca tingkat tinggi, khususnya analisis dan evaluasi. Kemampuan ini juga relevan dengan kebutuhan santri untuk memahami pesan yang tidak selalu dinyatakan secara langsung dalam teks keagamaan.

Salah satu faktor yang memengaruhi rendahnya pengembangan membaca inferensial adalah pendekatan penilaian yang digunakan guru. Penilaian pembelajaran bahasa Arab masih didominasi oleh penilaian sumatif yang berorientasi pada hasil akhir, seperti tes tertulis konvensional. Padahal, penilaian semestinya tidak hanya berfungsi sebagai alat ukur hasil belajar, tetapi juga sebagai sarana untuk memperbaiki proses pembelajaran.⁸ Black dan Wiliam (1998) menegaskan bahwa penilaian yang efektif adalah penilaian formatif yang mampu memberikan umpan balik berkelanjutan kepada peserta didik dan guru.⁹ Dengan demikian, penilaian formatif tidak hanya berperan sebagai alat evaluasi, tetapi juga sebagai strategi pedagogis utama yang secara langsung berkontribusi dalam pengembangan keterampilan membaca inferensial bahasa Arab.

Penilaian formatif menekankan proses pengumpulan informasi belajar peserta didik secara berkelanjutan untuk memberikan umpan balik yang konstruktif dan mendorong perbaikan pembelajaran. Menurut Popham, penilaian formatif berfungsi sebagai *instructional tool* yang membantu guru

⁸ Ghufuran Hasyim Achmad et al., "Penilaian Autentik Pada Kurikulum Merdeka Belajar Dalam Pembelajaran Pendidikan Agama Islam Di Sekolah Dasar," *EDUKATIF: Jurnal Ilmu Pendidikan* 4, no. 4 (2022): 5685–5699, <https://doi.org/10.31004/edukatif.v4i4.3280>.

⁹ Ayu Utami and Salito, "Implementasi Asesmen Berkelanjutan Untuk Meningkatkan Kualitas Pembelajaran Di Sekolah Dasar," *Jurnal Akuntansi, Manajemen, Dan Ilmu Pendidikan (JAMED)* 1, no. 3 (2025): 124–136.

menyesuaikan strategi pembelajaran sesuai dengan kebutuhan peserta didik.¹⁰ Dalam konteks membaca bahasa Arab, penilaian formatif memungkinkan guru untuk memantau perkembangan pemahaman inferensial peserta didik secara bertahap, bukan hanya menilai capaian akhir.

Penilaian formatif dipilih dalam penelitian ini karena permasalahan yang ditemukan tidak hanya berkaitan dengan rendahnya hasil membaca, tetapi juga dengan belum digunakannya hasil penilaian sebagai dasar pemberian umpan balik dan perbaikan pembelajaran. Oleh karena itu, fungsi formatif diperlukan agar penilaian dapat membantu guru mengetahui perkembangan dan kesulitan membaca santri selama proses pembelajaran berlangsung.

Penilaian pembelajaran bahasa Arab juga perlu terintegrasi dengan pendekatan autentik. Penilaian autentik menuntut peserta didik untuk menunjukkan kemampuan melalui tugas yang bermakna dan relevan dengan konteks penggunaan pengetahuan yang sesungguhnya. Wiggins menyatakan bahwa penilaian autentik memberikan gambaran yang lebih komprehensif terhadap kemampuan peserta didik karena menilai kinerja nyata¹¹, bukan sekadar jawaban tes.¹² Dalam konteks keterampilan membaca inferensial peserta didik mampu Menceritakan kembali isi bacaan dan menghasilkan ide dan gagasan¹³.

¹⁰ Moch Abduh, *Model Penilaian Formatif Pada Pembelajaran Abad Ke-21 Untuk Sekolah Dasar* (Perpustakaan Nasional Republik Indonesia, 2019).

¹¹ Grant Wiggins, 'The Case for Authentic Assessment', *Practical Assessment, Research, and Evaluation*, 2.1 (1990).

¹² Anggela Oktalia, 'Implementasi Assesmen Autentik Dalam Mata Pelajaran IPAS Kelas V Di SDN 32 Rejang Lebong', 2025.

¹³ Tsania Sausan Hasna and others, 'Innovation in Authentic Assessment of Maharah Qira' Ah in the Textbook of Al- ' Arabiyyah Baina Yadaik Volume II', II.1 (2025), pp. 63–77, doi:10.24042/albayan.v17i1.24681.

Penilaian autentik dipilih dalam penelitian ini karena kemampuan membaca inferensial tidak memadai apabila hanya diukur melalui pertanyaan faktual atau soal yang jawabannya tersedia secara langsung dalam teks. Kemampuan tersebut perlu dinilai melalui tugas menafsirkan makna tersirat, menarik kesimpulan implisit, menjelaskan hubungan antargagasan, serta mengaitkan isi bacaan dengan konteks kehidupan santri. Dengan demikian, pendekatan autentik memungkinkan instrumen mencerminkan kemampuan membaca santri secara lebih nyata dan bermakna.

Seiring dengan perkembangan teknologi, penilaian formatif dan autentik semakin berpeluang untuk diintegrasikan dengan media digital. Asesmen berbasis digital memungkinkan penyajian soal yang variatif, pemberian umpan balik secara cepat, serta pemantauan perkembangan belajar peserta didik secara sistematis. Menurut Redecker dan Johannessen, teknologi digital dapat meningkatkan efektivitas penilaian formatif karena mendukung interaksi, refleksi, dan pengambilan keputusan pedagogis berbasis data. Oleh karena itu, pemanfaatan teknologi digital dalam penilaian membaca inferensial menjadi kebutuhan yang relevan dengan tuntutan pembelajaran abad ke-21.¹⁴

Di antara berbagai media asesmen digital, aplikasi *Wayground* dipilih karena memiliki karakteristik yang sesuai dengan kebutuhan instrumen yang dikembangkan. Aplikasi ini mendukung beragam bentuk soal, termasuk pertanyaan terbuka yang dapat digunakan untuk menilai respons pemahaman,

¹⁴ Abduh, *Model Penilaian Formatif Pada Pembelajaran Abad Ke-21 Untuk Sekolah Dasar.*, hlm. 30-46.

pendapat, dan kemampuan berpikir tingkat tinggi.¹⁵ *Wayground* juga menyediakan laporan hasil pengerjaan yang memuat informasi mengenai akurasi, tingkat penyelesaian, jumlah peserta, dan respons terhadap soal, sehingga membantu guru memperoleh informasi hasil penilaian secara lebih cepat dan terorganisasi¹⁶. Karakteristik tersebut sejalan dengan prinsip penilaian formatif yang membutuhkan data hasil belajar untuk pemberian umpan balik dan tindak lanjut¹⁷. Adapun uraian lebih terperinci mengenai proses dan pertimbangan pemilihan *Wayground* disajikan pada tahap *design* dalam Bab IV.

Penelitian ini dilakukan di Pondok pesantren, pemilihan pondok pesantren sebagai konteks penelitian didasarkan pada pertimbangan akademik dan substantif. Pembelajaran bahasa Arab di pesantren memiliki keterkaitan yang kuat dengan penggunaan teks-teks keislaman sebagai sumber pembelajaran. Bahasa Arab tidak hanya dipelajari sebagai mata pelajaran atau objek akademik, tetapi juga digunakan sebagai sarana memahami ajaran Islam, kitab keagamaan, dan tradisi keilmuan Islam.

Kondisi tersebut menjadikan kebutuhan terhadap instrumen penilaian membaca inferensial yang autentik lebih nyata dan kontekstual. Selain itu, santri masih berada dalam proses penguatan kemampuan memahami bacaan sehingga memerlukan penilaian yang tidak hanya mencatat hasil akhir, tetapi juga memberikan informasi untuk membina perkembangan pemahaman mereka

¹⁵ Wayground, "Question Types Explained," *Wayground Help Center*, diakses 20 Januari 2026

¹⁶ Wayground, "Question Types Explained," *Wayground Help Center*, diakses 21 Januari 2026

¹⁷ Paul Black and Dylan Wiliam, 'Assessment and Classroom Learning', *Assessment in Education: Principles, Policy & Practice*, 5.1 (1998), pp. 7–74.

secara bertahap. Atas dasar itu, pondok pesantren dipilih karena memiliki kesesuaian yang lebih kuat dengan tujuan pengembangan instrumen dibandingkan lembaga yang pembelajaran bahasa Arabnya lebih berorientasi pada kajian akademik tingkat lanjut.

Pemilihan santri *mustawa awwal* (tingkat awal) sebagai subjek penelitian didasarkan pada pertimbangan pedagogis. Pada tingkat ini, santri sedang membangun dan memperkuat dasar pemahaman teks bahasa Arab sehingga masih membutuhkan bimbingan yang sistematis untuk bergerak dari pemahaman literal menuju pemahaman inferensial.¹⁸ Fase ini dipandang strategis karena pemberian penilaian formatif sejak tingkat awal memungkinkan guru mendeteksi kesulitan membaca, memberikan umpan balik, dan memperbaiki pola pemahaman sebelum santri memasuki materi bacaan yang lebih kompleks. Oleh karena itu, santri *mustawa awwal* dipilih bukan karena dianggap telah menguasai seluruh kemampuan membaca, tetapi karena berada pada tahap perkembangan yang relevan untuk penerapan dan pengujian instrumen.

Berdasarkan hasil observasi dan wawancara dengan guru di Pondok Pesantren Syaikh Jamilurrahman Bantul Pada Tanggal 10 Februari 2026 diketahui bahwa pembelajaran membaca bahasa Arab selama ini masih dilaksanakan secara sederhana, yaitu berfokus pada kegiatan membaca teks dan menerjemahkan secara langsung, tanpa perencanaan capaian pemahaman membaca yang terstruktur. Fokus pembelajaran pun lebih banyak ditekankan

¹⁸ Black and Wiliam, 'Assessment and Classroom Learning'.

pada pemahaman literal, sedangkan pengembangan membaca inferensial belum dilakukan secara sadar dan sistematis.¹⁹ Hasil observasi juga menunjukkan bahwa guru baru terlaksana sebagian dalam memberikan teks yang menuntut penafsiran makna tersirat, dan belum optimal dalam mendorong santri menarik kesimpulan implisit maupun mengaitkan isi teks dengan konteks dan pengetahuan sebelumnya. Kondisi ini menunjukkan adanya kesenjangan antara tuntutan keterampilan membaca tingkat tinggi dengan praktik pembelajaran yang berlangsung.

Temuan tersebut juga menunjukkan bahwa persoalan utama tidak hanya terletak pada metode pembelajaran, tetapi juga pada belum tersedianya perangkat penilaian yang memuat indikator membaca inferensial, butir soal yang terstruktur, pedoman penskoran, dan mekanisme umpan balik. Oleh karena itu, produk yang dikembangkan dalam penelitian ini secara khusus berbentuk instrumen penilaian.

Permasalahan yang sama juga tampak pada aspek penilaian membaca. Penilaian keterampilan membaca pada umumnya masih dilakukan melalui tes tertulis dan tanya jawab sederhana yang berfokus pada arti kata atau isi teks secara langsung, tanpa indikator pemahaman inferensial yang jelas. Penilaian lebih sering dilaksanakan pada akhir pembelajaran atau akhir materi, sehingga belum berfungsi sebagai penilaian formatif yang memantau proses belajar santri

¹⁹ Hasil observasi pembelajaran bahasa Arab di Pondok Pesantren Syaikh Jamilurrahman Bantul, 10 Februari 2026

secara berkelanjutan.²⁰ Hasil penilaian pun umumnya hanya digunakan untuk pemberian nilai, belum dimanfaatkan secara sistematis sebagai umpan balik untuk memperbaiki pemahaman membaca santri. Temuan observasi memperkuat kondisi tersebut, yakni penilaian selama proses pembelajaran baru terlaksana sebagian, pemberian umpan balik juga baru terlaksana sebagian, sedangkan penggunaan hasil penilaian untuk perbaikan pembelajaran masih belum terlaksana secara optimal.

Di samping itu, penilaian yang diterapkan belum sepenuhnya mencerminkan prinsip penilaian autentik. Berdasarkan hasil wawancara, tugas membaca selama ini lebih banyak berupa soal pemahaman langsung, sedangkan tugas yang menuntut interpretasi makna tersirat atau penarikan kesimpulan bacaan masih sangat jarang digunakan²¹. Hasil observasi juga menunjukkan bahwa penilaian berbasis tugas bermakna, seperti interpretasi teks dan penyimpulan isi bacaan, belum terlaksana secara optimal. Meskipun penilaian tidak hanya berbentuk pilihan ganda mulai tampak terlaksana sebagian, secara umum penilaian belum sepenuhnya mencerminkan kemampuan membaca santri yang sesungguhnya. Dengan demikian, penilaian yang berlangsung masih memerlukan penguatan agar lebih autentik dan lebih relevan dengan tuntutan pemahaman membaca inferensial.

Permasalahan berikutnya berkaitan dengan pemanfaatan teknologi digital dalam penilaian membaca. Guru menyatakan bahwa media digital belum

²⁰ Hasil observasi pembelajaran bahasa Arab di Pondok Pesantren Syaikh Jamilurrahman Bantul, 10 Februari 2026

²¹ Wawancara dengan Ustadz Fais Sofyan, Lc., Guru Bahasa Arab Pondok Pesantren Syaikh Jamilurrahman Bantul, Bantul, 10 Februari 2026.

dimanfaatkan secara optimal, dan penilaian masih didominasi oleh metode konvensional berbasis kertas²². Padahal, guru juga memandang bahwa media digital memiliki potensi besar untuk mendukung penilaian membaca, terutama dalam hal kemudahan penyajian soal, kecepatan umpan balik, dan efisiensi pengelolaan hasil penilaian.

Temuan observasi menunjukkan bahwa penggunaan media digital dalam penilaian membaca maupun pemanfaatannya untuk mendukung umpan balik masih berada pada kategori belum terlaksana optimal. Kondisi ini menunjukkan bahwa terdapat peluang besar untuk mengembangkan Instrumen penilaian yang lebih praktis, terstruktur, dan memanfaatkan dukungan teknologi digital secara lebih efektif.

Meskipun demikian, hasil observasi juga memperlihatkan adanya modal awal yang cukup potensial untuk pengembangan Instrumen penilaian di pesantren. Teks bacaan bernuansa keislaman telah terlaksana dengan baik dalam pembelajaran, dan proses pembelajaran pada dasarnya telah cukup sesuai dengan karakter santri. Selain itu, guru secara eksplisit menyatakan bahwa sangat diperlukan adanya Instrumen penilaian baru yang lebih terstruktur, praktis, dan mampu menilai pemahaman membaca secara lebih mendalam, khususnya membaca inferensial.²³

Guru juga berharap Instrumen yang dikembangkan mudah digunakan, sesuai dengan karakter santri pesantren, mampu memberikan umpan balik

²² Wawancara dengan Ustadz Fais Sofyan, Lc., Guru Bahasa Arab Pondok Pesantren Syaikh Jamilurrahman Bantul, Bantul, 10 Februari 2026.

²³ Wawancara dengan Ustadz Fais Sofyan, Lc., Guru Bahasa Arab Pondok Pesantren Syaikh Jamilurrahman Bantul, Bantul, 10 Februari 2026.

formatif, serta didukung oleh media digital yang memudahkan proses penilaian²⁴ Hal ini menunjukkan bahwa kebutuhan pengembangan Instrumen penilaian bukan hanya datang dari analisis teoritis, tetapi juga merupakan kebutuhan nyata di lapangan.

Kondisi tersebut menuntut penelitian yang tidak berhenti pada kegiatan mendeskripsikan masalah, tetapi menghasilkan produk yang dapat digunakan untuk menjawab kebutuhan lapangan.²⁵ Oleh karena itu, pendekatan penelitian dan pengembangan dipilih karena memungkinkan peneliti menyusun instrumen berdasarkan kebutuhan pengguna, memvalidasinya melalui penilaian ahli, mengujicobakannya dalam pembelajaran, serta menyempurnakannya berdasarkan hasil pengujian empiris. Dengan demikian, penggunaan istilah “pengembangan” dalam penelitian ini menunjukkan proses sistematis untuk menghasilkan instrumen yang memiliki dasar teoretis, kelayakan isi, dan keterterapan praktis.²⁶

Berdasarkan kajian teoretis dan temuan empiris tersebut, pembelajaran membaca bahasa Arab di Pondok Pesantren Syaikh Jamilurrahman Bantul masih menghadapi beberapa permasalahan, yaitu dominannya pemahaman literal, belum optimalnya pengembangan membaca inferensial, belum tersedianya instrumen yang terstruktur, belum berfungsinya penilaian sebagai sarana umpan balik, masih terbatasnya tugas membaca yang autentik, serta belum optimalnya

²⁴ Wawancara dengan Ustadz Fais Sofyan, Lc., Guru Bahasa Arab Pondok Pesantren Syaikh Jamilurrahman Bantul, Bantul, 10 Februari 2026.

²⁵ Sugiyono, *Metode Penelitian Kualitatif Kuantitatif Dan R&D* (Alfabeta, 2019).

²⁶ Sivasailam Thiagarajan, Dorothy S Semmel, and Melvyn I Semmel, ‘Instructional Development for Training Teachers of Exceptional Children’ (Minneapolis, Minnesota: leadership training institute/special education ..., 1974).

pemanfaatan teknologi digital. Dominannya pertanyaan literal menuntut pengembangan instrumen yang secara khusus mengukur keterampilan membaca inferensial; belum optimalnya umpan balik menuntut penerapan fungsi formatif; terbatasnya tugas bermakna menuntut pendekatan autentik; sedangkan penilaian yang masih konvensional menuntut dukungan media digital yang praktis dan terorganisasi.

Wayground dipilih untuk mendukung penyajian instrumen, pengumpulan respons, dan pengelolaan hasil penilaian, sementara konteks Pondok Pesantren Syaikh Jamilurrahman Bantul dipilih karena memiliki kebutuhan lapangan dan karakter pembelajaran teks keislaman yang sesuai dengan produk yang dikembangkan. Atas dasar itu, penelitian berjudul “Pengembangan Instrumen Penilaian Formatif Autentik Keterampilan Membaca Inferensial Bahasa Berbasis Aplikasi *Wayground* Arab di Pondok Pesantren Syaikh Jamilurrahman Bantul” dilakukan sebagai upaya menghasilkan instrumen yang valid, praktis, kontekstual, dan dapat mendukung perbaikan pembelajaran membaca bahasa Arab.²⁷

B. Rumusan Masalah

Berdasarkan latar belakang penelitian yang telah diuraikan, maka rumusan masalah dalam penelitian ini adalah sebagai berikut:

²⁷ Machmudah Umi and Abdul Wahab Rosyidi, ‘Active Learning Dalam Pembelajaran Bahasa Arab’ (Malang: UIN Malang Press, 2008).

1. Bagaimana karakteristik Pengembangan Instrumen penilaian formatif autentik keterampilan membaca inferensial bahasa Arab berbasis Aplikasi *Wayground* di Pondok Pesantren Syaikh Jamilurrahman Bantul?
2. Bagaimana prosedur pengembangan Instrumen penilaian formatif autentik keterampilan membaca inferensial bahasa Arab berbasis Aplikasi *Wayground* di Pondok Pesantren Jamilurrahman Bantul?
3. Bagaimana kelayakan Instrumen penilaian formatif autentik berbasis Aplikasi *Wayground* dalam pembelajaran keterampilan membaca inferensial bahasa Arab di Pondok Pesantren Syaikh Jamilurrahman Bantul?

C. Tujuan Penelitian

Penelitian ini bertujuan untuk:

1. Mengembangkan Instrumen penilaian formatif autentik keterampilan membaca inferensial bahasa Arab berbasis Aplikasi *Wayground* di Pondok Pesantren Syaikh Jamilurrahman Bantul.
2. Mendeskripsikan prosedur pengembangan Instrumen penilaian formatif autentik keterampilan membaca inferensial bahasa Arab berbasis Aplikasi *Wayground* di Pondok Pesantren Syaikh Jamilurrahman Bantul.
3. Mengetahui kelayakan Instrumen penilaian formatif autentik keterampilan membaca inferensial bahasa Arab berbasis Aplikasi *Wayground* di Pondok Pesantren Syaikh Jamilurrahman Bantul.

D. Manfaat Penelitian

1. Manfaat Teoretis

Secara teoretis, hasil penelitian ini dapat memberikan kontribusi dalam kajian penilaian pembelajaran bahasa Arab, khususnya dalam bidang penilaian formatif dan autentik. Penelitian ini memperkaya khazanah teori asesmen bahasa dengan menghadirkan Instrumen penilaian formatif autentik berbasis digital yang secara spesifik dirancang untuk menilai keterampilan membaca inferensial bahasa Arab, sebuah aspek yang masih relatif terbatas pembahasannya dalam penelitian-penelitian sebelumnya. Dengan demikian, penelitian ini dapat memperluas pemahaman konseptual mengenai integrasi penilaian formatif, penilaian autentik, dan teknologi digital dalam pembelajaran keterampilan membaca bahasa Arab.

Selain itu, penelitian ini memberikan landasan teoretis mengenai peran penilaian formatif sebagai bagian integral dari proses pembelajaran, bukan sekadar alat ukur hasil belajar. Instrumen yang dikembangkan menegaskan fungsi penilaian sebagai sarana pemberian umpan balik (*feedback*) dan penguatan proses berpikir tingkat tinggi dalam membaca inferensial. Dengan demikian, penelitian ini dapat menjadi rujukan teoretis bagi pengembangan Instrumen penilaian sejenis pada keterampilan bahasa lainnya maupun pada konteks pembelajaran bahasa asing yang berbeda.

2. Manfaat Praktis

Secara praktis, penelitian ini diharapkan memberikan manfaat bagi berbagai pihak yang terlibat dalam pembelajaran bahasa Arab, khususnya di lingkungan pesantren.

- a. Bagi Guru Bahasa Arab, hasil penelitian ini dapat menjadi pedoman praktis dalam merancang dan menerapkan penilaian formatif autentik berbasis digital untuk menilai keterampilan membaca inferensial peserta didik. Instrumen yang dikembangkan dapat membantu guru memberikan umpan balik yang lebih terarah, sistematis, dan berkelanjutan, sehingga proses pembelajaran membaca bahasa Arab menjadi lebih efektif dan bermakna.
- b. Bagi Peserta Didik (Santri), penerapan Instrumen penilaian formatif autentik berbasis digital diharapkan dapat membantu mereka memahami proses berpikir dalam membaca teks bahasa Arab, meningkatkan kemampuan menarik kesimpulan implisit, serta mendorong keterlibatan aktif dalam pembelajaran. Umpan balik yang diperoleh melalui penilaian formatif dapat menjadi sarana refleksi untuk memperbaiki pemahaman membaca secara berkelanjutan.
- c. Bagi Lembaga Pendidikan Pesantren, penelitian ini dapat menjadi referensi dalam pengembangan sistem penilaian pembelajaran bahasa Arab yang lebih adaptif terhadap perkembangan teknologi dan tuntutan pembelajaran abad ke-21, tanpa mengabaikan karakteristik dan nilai-nilai khas pesantren.

d. Bagi Peneliti Selanjutnya, hasil penelitian ini dapat dijadikan rujukan atau pijakan awal untuk penelitian lanjutan yang berkaitan dengan pengembangan penilaian formatif autentik, keterampilan membaca tingkat lanjut, atau integrasi teknologi digital dalam pembelajaran bahasa Arab maupun bahasa asing lainnya.

E. Kajian Pustaka

Dalam kajian pustaka, penelusuran terhadap penelitian terdahulu dilakukan untuk mengidentifikasi sejauh mana topik yang diteliti telah dibahas oleh peneliti sebelumnya. Telaah ini bertujuan untuk menemukan titik persamaan dan perbedaan antara penelitian terdahulu dengan penelitian yang akan dilakukan, sehingga peneliti dapat menentukan aspek-aspek yang belum dikaji secara optimal. Dengan demikian, hasil penelitian yang dilakukan diharapkan dapat melengkapi temuan sebelumnya atau menghadirkan kebaruan ilmiah yang relevan. Berdasarkan penelusuran tersebut, ditemukan beberapa penelitian yang memiliki keterkaitan dengan topik penelitian ini.

Salah satu penelitian yang relevan adalah tesis yang ditulis oleh Esthiningtyas Sheilla Purwandari dengan judul *Pengembangan Instrumen Penilaian Berbasis HOTS pada Kitab Durusul Lughah al-'Arabiyyah li Ghairi Nāthiqīnā Bihā Jilid 1*. Penelitian ini bertujuan untuk mengembangkan instrumen penilaian yang berorientasi pada keterampilan berpikir tingkat tinggi (*Higher Order Thinking Skills*) berdasarkan kitab *Durusul Lughah al-'Arabiyyah li Ghairi Nāthiqīnā Bihā Jilid 1*. Jenis penelitian yang digunakan adalah penelitian pengembangan (*Research and Development*) dengan Model 4-

D yang dikemukakan oleh Thiagarajan, Semmel, dan Semmel, yang meliputi tahapan *define, design, develop*, dan *disseminate*.

Teknik pengumpulan data dalam penelitian tersebut meliputi observasi, wawancara, angket, dan dokumentasi. Analisis data dilakukan dengan menggunakan pendekatan deskriptif kualitatif dan deskriptif kuantitatif. Sumber data penelitian melibatkan pengajar bahasa Arab, santri kelas I wustha di pondok pesantren, serta para ahli yang terdiri atas ahli materi dan ahli instrumen. Selain itu, kitab *Durusul Lughah al-'Arabiyyah li Ghairi Nāthiqīnā Bihā* Jilid 1 juga dijadikan sebagai sumber utama dalam pengembangan instrumen.

Hasil penelitian menunjukkan bahwa instrumen penilaian berbasis HOTS yang dikembangkan memperoleh tingkat kelayakan yang sangat baik berdasarkan hasil validasi para ahli dan praktisi. Skor validasi yang diperoleh dari para validator dan praktisi menunjukkan bahwa produk yang dikembangkan layak digunakan. Selain itu, hasil analisis menunjukkan bahwa instrumen penilaian tersebut memenuhi aspek validitas, memiliki tingkat reliabilitas yang tinggi dengan indeks reliabilitas sebesar 0,784, serta memiliki tingkat kesukaran butir instrumen yang bervariasi. Temuan ini menunjukkan bahwa instrumen penilaian berbasis HOTS yang dikembangkan telah memenuhi kriteria kelayakan secara teoretis maupun empiris.²⁸

Persamaan Penelitian Terdahulu dengan Penelitian ini, Secara konseptual dan metodologis, penelitian yang dilakukan oleh Esthiningtyas Sheilla

²⁸ Esthiningtyas Sheilla Purwandari, 'Pengembangan Instrumen Penilaian Berbasis Hots Pada Kitab Durusul Lughoh Al-Arabiyyah Li Ghairi Nathiqina Biha Jilid 1' (UIN Sunan Kalijaga Yogyakarta, 2023).

Purwandari memiliki beberapa persamaan dengan penelitian yang akan dilakukan penulis. Pertama, kesamaan dari sisi pendekatan penelitian, yaitu sama-sama menggunakan pendekatan penelitian dan pengembangan (Research and Development) dengan Model 4-D. Hal ini menunjukkan bahwa kedua penelitian berorientasi pada pengembangan produk pendidikan yang melalui tahapan analisis kebutuhan, perancangan, pengembangan, dan pengujian kelayakan. Kedua, kesamaan dari sisi fokus penilaian, yakni sama-sama menempatkan penilaian sebagai aspek sentral dalam pembelajaran bahasa Arab. Baik penelitian terdahulu maupun penelitian ini bertujuan untuk meningkatkan kualitas asesmen pembelajaran bahasa Arab melalui pengembangan perangkat penilaian yang lebih bermakna. Ketiga, kesamaan dari sisi ranah kognitif, yaitu sama-sama mengarah pada pengembangan kemampuan berpikir tingkat tinggi. Penelitian terdahulu menekankan HOTS secara umum, sedangkan penelitian ini menempatkan membaca inferensial sebagai bagian dari kemampuan berpikir tingkat tinggi dalam keterampilan membaca bahasa Arab.

Perbedaan Penelitian Terdahulu dengan Penelitian Pertama, perbedaan pada fokus keterampilan yang dinilai. Penelitian terdahulu berfokus pada pengembangan instrumen penilaian berbasis HOTS secara umum dalam kitab *Durusul Lughah al-'Arabiyyah*, tanpa menitikberatkan pada satu keterampilan bahasa secara spesifik. Sementara itu, penelitian ini secara khusus memfokuskan diri pada keterampilan membaca inferensial bahasa Arab, yaitu kemampuan memahami makna tersirat dan menarik kesimpulan implisit dari teks.

Kedua, perbedaan pada fungsi penilaian. Penelitian terdahulu lebih menekankan pada kelayakan instrumen penilaian dari aspek validitas, reliabilitas, dan tingkat kesukaran, yang cenderung bersifat evaluatif. Sebaliknya, penelitian ini menempatkan penilaian sebagai penilaian formatif, yang berfungsi memberikan umpan balik berkelanjutan dan mendukung proses pembelajaran membaca inferensial secara bertahap.

Ketiga, perbedaan pada bentuk produk yang dikembangkan. Penelitian terdahulu menghasilkan instrumen penilaian, sedangkan penelitian ini mengembangkan Instrumen penilaian formatif autentik berbasis digital yang mencakup kerangka konseptual, prosedur pelaksanaan, jenis tugas, rubrik penilaian, serta mekanisme umpan balik.

Keempat, perbedaan pada integrasi teknologi. Penelitian terdahulu belum secara eksplisit menempatkan teknologi digital sebagai bagian dari sistem penilaian formatif, sedangkan penelitian ini mengintegrasikan teknologi digital sebagai sarana penguatan fungsi penilaian autentik dan formatif dalam pembelajaran membaca inferensial.

Dengan demikian, penelitian yang akan dilakukan penulis tidak mengulang penelitian terdahulu, melainkan melengkapi dan memperluas kajian dengan menghadirkan pendekatan penilaian yang lebih komprehensif, kontekstual, dan berorientasi pada proses pembelajaran. Penelitian ini diharapkan mampu mengisi celah penelitian yang belum terjamah, khususnya dalam pengembangan Instrumen penilaian formatif autentik berbasis digital untuk keterampilan membaca inferensial bahasa Arab di lingkungan pesantren.

Penelitian relevan berikutnya adalah artikel jurnal yang ditulis oleh Faizmailatus Sofaa dkk. dengan judul *Innovation of Language Receptive Skills Based on Authentic Assessment*²⁹. Penelitian ini bertujuan untuk menganalisis permasalahan penilaian yang terdapat dalam buku Bahasa Arab kelas XI yang diterbitkan oleh Kementerian Agama Republik Indonesia tahun 2020, khususnya yang berkaitan dengan keterampilan reseptif bahasa, yaitu keterampilan menyimak dan membaca.

Penelitian tersebut menggunakan pendekatan kualitatif dengan metode studi pustaka. Sumber data primer berasal dari buku Bahasa Arab kelas XI terbitan Kementerian Agama Republik Indonesia tahun 2020, sedangkan data sekunder diperoleh dari berbagai literatur yang relevan dengan topik penilaian dan keterampilan reseptif bahasa. Teknik analisis dilakukan melalui telaah isi terhadap bentuk, tujuan, dan posisi penilaian keterampilan reseptif dalam buku ajar tersebut.

Hasil penelitian menunjukkan bahwa penilaian keterampilan reseptif dalam buku Bahasa Arab kelas XI masih diposisikan dalam ranah pengetahuan, bukan sebagai keterampilan berbahasa. Padahal, keterampilan menyimak dan membaca secara konseptual merupakan bagian dari keterampilan berbahasa yang seharusnya dinilai melalui pendekatan kinerja. Berdasarkan temuan tersebut, peneliti merekomendasikan penerapan penilaian autentik yang memposisikan keterampilan reseptif sebagai keterampilan, sekaligus tetap

²⁹ Faizmailatus Sofa and others, 'Innovation of Language Receptive Skills Based on Authentic Assessment', *Izdihar : Journal of Arabic Language Teaching, Linguistics, and Literature*, 6.1 (2023), pp. 31–44, doi:10.22219/jiz.v6i1.24029.

menilai unsur kebahasaan pada ranah pengetahuan serta aspek sikap melalui berbagai bentuk penilaian, seperti portofolio, tugas, kinerja, dan wawancara yang disesuaikan dengan kebutuhan pembelajaran.

Persamaan Penelitian Terdahulu dengan Penelitian yang Dilakukan Penulis. Penelitian yang dilakukan oleh Faizmailatus Sofaa dkk. memiliki beberapa persamaan substantif dengan penelitian yang akan dilakukan penulis. Pertama, kesamaan pada pendekatan penilaian, yaitu sama-sama menekankan pentingnya penilaian autentik dalam pembelajaran bahasa Arab. Kedua penelitian berpandangan bahwa penilaian keterampilan berbahasa tidak dapat direduksi menjadi penilaian pengetahuan semata, melainkan harus merepresentasikan kinerja nyata peserta didik dalam menggunakan bahasa.

Kedua, kesamaan pada fokus keterampilan reseptif, khususnya keterampilan membaca. Penelitian terdahulu menempatkan membaca sebagai bagian dari keterampilan reseptif yang perlu dinilai secara tepat, sedangkan penelitian ini secara lebih spesifik memfokuskan pada keterampilan membaca inferensial bahasa Arab sebagai bagian dari keterampilan reseptif tingkat lanjut.

Ketiga, kesamaan pada orientasi perbaikan sistem penilaian, di mana kedua penelitian sama-sama bertujuan untuk mengoreksi praktik penilaian bahasa Arab yang belum selaras dengan hakikat keterampilan berbahasa. Dengan demikian, baik penelitian terdahulu maupun penelitian ini berangkat dari kepedulian yang sama terhadap peningkatan kualitas asesmen pembelajaran bahasa Arab.

Perbedaan Penelitian Terdahulu dengan Penelitian yang dilakukan Penulis serta Posisi Gap Penelitian. Meskipun memiliki kesamaan, terdapat perbedaan

mendasar yang menunjukkan posisi kebaruan dan celah penelitian (research gap) dalam penelitian yang dilakukan penulis.

Pertama, perbedaan dari sisi tujuan penelitian. Penelitian oleh Faizmailiatus Sofaa dkk. bertujuan untuk menganalisis dan mengkritisi sistem penilaian dalam buku teks Bahasa Arab, sehingga bersifat evaluatif dan konseptual. Sementara itu, penelitian ini bertujuan untuk mengembangkan Instrumen penilaian formatif autentik berbasis digital, sehingga bersifat konstruktif dan aplikatif.

Kedua, perbedaan dari sisi metode penelitian. Penelitian terdahulu menggunakan pendekatan kualitatif dengan studi pustaka, tanpa melakukan pengembangan produk penilaian. Sebaliknya, penelitian ini menggunakan pendekatan penelitian dan pengembangan (R&D) untuk menghasilkan sebuah Instrumen penilaian yang dapat diuji kelayakan dan keterterapannya dalam pembelajaran nyata.

Ketiga, perbedaan dari sisi fungsi penilaian. Penelitian terdahulu menekankan penilaian autentik sebagai koreksi terhadap kesalahan konseptual penilaian keterampilan reseptif, namun belum secara eksplisit mengkaji fungsi formatif penilaian. Penelitian ini secara khusus menempatkan penilaian sebagai penilaian formatif yang berfungsi memberikan umpan balik berkelanjutan dan mendukung proses pengembangan keterampilan membaca inferensial.

Keempat, perbedaan dari sisi ruang lingkup dan konteks penelitian. Penelitian terdahulu berfokus pada analisis buku Bahasa Arab kelas XI secara nasional, sedangkan penelitian ini dikembangkan dan diterapkan dalam konteks

pembelajaran bahasa Arab di pondok pesantren, yang memiliki karakteristik peserta didik, tujuan pembelajaran, dan lingkungan akademik yang berbeda.

Dengan demikian, penelitian yang dilakukan penulis menempati posisi sebagai kelanjutan dan penguatan praktis dari temuan penelitian terdahulu. Jika penelitian Faizmailatus Sofaa dkk. menegaskan perlunya penilaian autentik untuk keterampilan reseptif, maka penelitian ini menjawab kebutuhan tersebut secara operasional melalui pengembangan Instrumen penilaian formatif autentik berbasis digital yang secara khusus dirancang untuk menilai keterampilan membaca inferensial bahasa Arab di lingkungan pesantren.

Penelitian relevan selanjutnya adalah tesis yang disusun oleh Mutmainna B. dengan judul *Pengembangan Instrumen Penilaian Autentik untuk Mengukur Keterampilan Menyimak dan Keterampilan Komunikasi Peserta Didik Kelas IV Sekolah Dasar*. Penelitian ini bertujuan untuk mengembangkan sebuah Instrumen penilaian autentik yang digunakan untuk mengukur keterampilan menyimak dan keterampilan komunikasi peserta didik dalam pembelajaran Bahasa Indonesia di kelas IV sekolah dasar.³⁰

Penelitian tersebut menggunakan pendekatan penelitian pengembangan dengan mengadopsi Instrumen ADDIE, yang meliputi lima tahapan utama, yaitu *analysis, design, development, implementation, dan evaluation*. Penelitian dilaksanakan di SD Muhammadiyah Bayen, Kabupaten Sleman, Yogyakarta, dengan subjek penelitian peserta didik kelas IV yang berjumlah 30 orang.

³⁰ Mutmainna.B, 'Pengembangan Model Penilaian Autentik Untuk Mengukur Keterampilan Menyimak Dan Keterampilan Komunikasi Peserta Didik Kelas Iv Sekolah Dasar' (UIN Sunan Kalijaga Yogyakarta., 2023).

Teknik pengumpulan data yang digunakan meliputi observasi, wawancara, angket, dan dokumentasi. Data penelitian diperoleh dari hasil penilaian para ahli desain dan ahli bahasa, respon guru terhadap Instrumen penilaian yang dikembangkan, serta hasil penilaian keterampilan menyimak dan keterampilan komunikasi peserta didik. Analisis data dilakukan dengan menggunakan pendekatan deskriptif kuantitatif, yang didasarkan pada kategori hasil belajar peserta didik.

Hasil penelitian menunjukkan bahwa Instrumen penilaian autentik yang dikembangkan memiliki tingkat kelayakan yang sangat baik. Hal ini ditunjukkan oleh hasil penilaian para ahli desain, ahli bahasa, serta respon positif dari guru. Dengan demikian, Instrumen penilaian autentik keterampilan menyimak dan keterampilan komunikasi yang dikembangkan dinyatakan layak untuk digunakan dalam pembelajaran Bahasa Indonesia di sekolah dasar.

Persamaan Penelitian Terdahulu dengan Penelitian yang Dilakukan Penulis Penelitian yang dilakukan oleh Mutmainana B. memiliki sejumlah persamaan konseptual dan metodologis dengan penelitian yang akan dilakukan penulis. Pertama, kesamaan dari sisi jenis dan tujuan penelitian, yaitu sama-sama merupakan penelitian pengembangan yang bertujuan menghasilkan Instrumen penilaian autentik. Kedua penelitian tidak hanya menganalisis fenomena pembelajaran, tetapi juga berorientasi pada pengembangan produk penilaian yang dapat diterapkan dalam praktik pembelajaran.

Kedua, kesamaan pada pendekatan penilaian, yaitu sama-sama mengadopsi konsep penilaian autentik sebagai pendekatan utama dalam mengukur

keterampilan berbahasa. Penelitian terdahulu menilai keterampilan menyimak dan komunikasi, sedangkan penelitian ini menilai keterampilan membaca inferensial bahasa Arab, yang sama-sama menuntut performa nyata peserta didik.

Ketiga, kesamaan pada prosedur pengembangan produk, di mana kedua penelitian melibatkan proses validasi oleh para ahli dan respon praktisi (guru) untuk menilai kelayakan Instrumen penilaian yang dikembangkan. Hal ini menunjukkan bahwa kedua penelitian menempatkan aspek kelayakan teoretis dan praktis sebagai bagian penting dalam pengembangan Instrumen asesmen.

Perbedaan Penelitian Terdahulu dengan Penelitian yang dilakukan Penulis serta Posisi Gap Penelitian. Meskipun memiliki sejumlah persamaan, terdapat perbedaan mendasar yang menunjukkan posisi kebaruan dan celah penelitian (research gap) dalam penelitian yang dilakukan penulis. Pertama, perbedaan pada bidang studi dan objek pembelajaran. Penelitian Mutmainana B. dikembangkan dalam konteks mata pelajaran Bahasa Indonesia di tingkat sekolah dasar, sedangkan penelitian ini difokuskan pada pembelajaran Bahasa Arab di lingkungan pondok pesantren. Perbedaan konteks bahasa dan institusi ini berimplikasi pada karakteristik keterampilan berbahasa, jenis teks, serta tujuan pembelajaran yang berbeda secara substansial.

Kedua, perbedaan pada keterampilan bahasa yang dikaji. Penelitian terdahulu berfokus pada keterampilan menyimak dan keterampilan komunikasi secara umum, sementara penelitian ini secara spesifik mengkaji keterampilan

membaca inferensial bahasa Arab, yang menuntut kemampuan berpikir tingkat tinggi dalam menafsirkan makna tersirat teks.

Ketiga, perbedaan pada fungsi penilaian. Penelitian Mutmainana B. mengembangkan Instrumen penilaian autentik yang berorientasi pada pengukuran keterampilan, namun belum secara eksplisit menekankan fungsi formatif penilaian. Sebaliknya, penelitian ini secara khusus menempatkan penilaian sebagai penilaian formatif, yang berfungsi memberikan umpan balik berkelanjutan untuk mendukung proses pembelajaran membaca inferensial.

Keempat, perbedaan pada integrasi teknologi. Penelitian terdahulu belum mengintegrasikan teknologi digital sebagai bagian dari sistem penilaian, sedangkan penelitian ini mengembangkan Instrumen penilaian autentik berbasis digital, di mana teknologi dimanfaatkan untuk memperkuat fungsi formatif melalui penyajian tugas, pemberian umpan balik, dan pemantauan perkembangan belajar peserta didik.

Kelima, perbedaan pada Instrumen pengembangan. Penelitian terdahulu menggunakan Instrumen ADDIE, sementara penelitian ini dirancang menggunakan Instrumen pengembangan yang disesuaikan dengan kebutuhan pengembangan penilaian formatif autentik dalam konteks pesantren. Perbedaan ini menunjukkan bahwa penelitian ini tidak mereplikasi penelitian sebelumnya, melainkan mengembangkan pendekatan metodologis yang relevan dengan tujuan dan konteks penelitian.

Dengan demikian, penelitian yang dilakukan penulis menempati posisi sebagai pengembangan lanjutan dan spesifik dari penelitian terdahulu. Jika

penelitian Mutmainana B. menegaskan kelayakan Instrumen penilaian autentik untuk keterampilan berbahasa di sekolah dasar, maka penelitian ini mengisi celah penelitian dengan mengembangkan Instrumen penilaian formatif autentik berbasis digital yang secara khusus dirancang untuk menilai keterampilan membaca inferensial bahasa Arab di lingkungan pesantren.

Berdasarkan telaah terhadap berbagai kajian teoretis dan penelitian terdahulu yang relevan, dapat disimpulkan bahwa pengembangan penilaian pembelajaran bahasa, khususnya dalam konteks pembelajaran bahasa Arab, telah memperoleh perhatian yang cukup signifikan. Penelitian-penelitian terdahulu telah mengkaji penilaian autentik, penilaian berbasis *Higher Order Thinking Skills* (HOTS), serta pemanfaatan media digital dalam asesmen pembelajaran bahasa. Namun demikian, kajian tersebut masih menunjukkan keterbatasan tertentu yang membuka ruang bagi pengembangan penelitian lebih lanjut.

Penelitian pertama menunjukkan kontribusi penting dalam pengembangan instrumen penilaian berbasis HOTS pada pembelajaran bahasa Arab, khususnya melalui pendekatan penelitian dan pengembangan. Meskipun demikian, penelitian tersebut masih berfokus pada pengembangan instrumen penilaian sebagai produk tunggal dan belum secara eksplisit menempatkan penilaian sebagai proses formatif yang terintegrasi dalam pembelajaran. Selain itu, fokus penilaian belum diarahkan secara spesifik pada keterampilan membaca inferensial, melainkan pada HOTS secara umum dalam konteks buku ajar tertentu.

Penelitian kedua memberikan landasan konseptual yang kuat mengenai pentingnya penilaian autentik untuk keterampilan reseptif, yaitu menyimak dan membaca, serta mengkritisi praktik penilaian yang masih memposisikan keterampilan reseptif dalam ranah pengetahuan. Namun, penelitian tersebut bersifat analitis dan konseptual, tanpa dilanjutkan pada pengembangan produk atau Instrumen penilaian yang operasional dan dapat diimplementasikan dalam praktik pembelajaran. Selain itu, fungsi formatif penilaian dan pemanfaatan teknologi digital belum menjadi fokus utama kajian tersebut.

Penelitian ketiga menunjukkan bahwa pengembangan Instrumen penilaian autentik dapat dilakukan secara sistematis dan dinyatakan layak digunakan dalam pembelajaran keterampilan berbahasa. Akan tetapi, penelitian tersebut dikembangkan dalam konteks mata pelajaran Bahasa Indonesia di sekolah dasar, dengan fokus pada keterampilan menyimak dan komunikasi secara umum. Penelitian tersebut belum mengintegrasikan penilaian formatif sebagai mekanisme umpan balik berkelanjutan, serta belum mengaitkannya dengan keterampilan membaca inferensial dan konteks pembelajaran bahasa Arab.

Sintesis terhadap ketiga penelitian terdahulu tersebut menunjukkan bahwa hingga saat ini belum ditemukan penelitian yang secara komprehensif mengembangkan Instrumen penilaian formatif autentik berbasis digital yang secara khusus dirancang untuk menilai keterampilan membaca inferensial bahasa Arab, terutama dalam konteks pendidikan pesantren. Penelitian-penelitian sebelumnya cenderung memisahkan antara fokus keterampilan, fungsi penilaian, pendekatan autentik, dan pemanfaatan teknologi, sehingga belum

menghasilkan Instrumen penilaian yang terintegrasi dan berorientasi pada proses pembelajaran.

Dengan demikian, terdapat celah penelitian (*research gap*) yang jelas pada aspek integrasi konseptual dan operasional antara: (1) keterampilan membaca inferensial sebagai target kemampuan kognitif tingkat tinggi, (2) penilaian formatif sebagai mekanisme pedagogis yang berfungsi memberikan umpan balik berkelanjutan, (3) penilaian autentik sebagai pendekatan evaluatif yang menilai kinerja nyata peserta didik, (4) pemanfaatan teknologi digital sebagai penguat sistem penilaian, dan (5) konteks pesantren sebagai lingkungan pembelajaran yang memiliki karakteristik khas. Celah inilah yang menjadi dasar dilakukannya penelitian ini.

Oleh karena itu, penelitian ini diarahkan untuk mengembangkan Instrumen Penilaian Formatif Autentik Berbasis Digital untuk Keterampilan Membaca Inferensial Bahasa Arab di Pondok Pesantren Jamilurrahman Bantul sebagai upaya mengisi kekosongan kajian tersebut. Instrumen yang dikembangkan diharapkan tidak hanya berkontribusi secara teoretis dalam pengembangan ilmu penilaian pembelajaran bahasa Arab, tetapi juga memberikan kontribusi praktis dalam meningkatkan kualitas pembelajaran membaca inferensial di lingkungan pesantren secara berkelanjutan dan bermakna.

BAB V

PENUTUP

A. Simpulan

Berdasarkan hasil penelitian dan pembahasan mengenai pengembangan Instrumen penilaian formatif autentik berbasis aplikasi *Wayground* untuk keterampilan membaca inferensial bahasa Arab di Pondok Pesantren Jamilurrahman Bantul, dapat dirumuskan beberapa simpulan sebagai berikut.

Pertama, karakteristik Instrumen penilaian formatif autentik berbasis aplikasi *Wayground* yang dikembangkan dalam penelitian ini ditandai oleh keterpaduan beberapa unsur utama, yaitu formatif, autentik, inferensial, digital, dan kontekstual pesantren. Instrumen ini bersifat formatif karena dirancang tidak hanya untuk menghasilkan skor akhir, tetapi juga untuk menyediakan informasi yang dapat digunakan guru dalam memberikan umpan balik dan tindak lanjut pembelajaran. Instrumen ini bersifat autentik karena menggunakan teks dan tugas yang bermakna serta dekat dengan pengalaman belajar santri di pesantren. Instrumen ini berorientasi pada keterampilan membaca inferensial karena fokus pengukurannya diarahkan pada kemampuan memahami makna tersirat, menarik kesimpulan, menafsirkan maksud penulis, mengevaluasi isi bacaan, dan merefleksikan nilai teks. Selain itu, Instrumen ini berbasis aplikasi *Wayground* sebagai media digital yang mendukung pelaksanaan penilaian secara lebih praktis dan terorganisasi. Di samping itu, Instrumen ini memiliki kekhasan konteks pesantren karena teks, muatan nilai, dan orientasi penilaiannya

disesuaikan dengan karakteristik santri dan lingkungan pembelajaran keagamaan. Dengan demikian, Instrumen yang dikembangkan bukan sekadar instrumen soal digital, melainkan suatu sistem penilaian yang dirancang untuk mendukung pembelajaran membaca inferensial bahasa Arab secara lebih bermakna.

Kedua, pengembangan instrumen dilakukan menggunakan model 4-D yang meliputi tahap *define*, *design*, *develop*, dan *disseminate*. Tahap *define* menghasilkan identifikasi kebutuhan terhadap instrumen yang lebih terstruktur, formatif, autentik, dan mampu mengukur keterampilan membaca inferensial. Tahap *design* menghasilkan rancangan indikator, kisi-kisi, butir soal, rubrik penilaian, dan desain implementasi melalui *Wayground*. Tahap *develop* menghasilkan produk yang telah melalui validasi ahli, revisi, uji coba terbatas, dan analisis empiris menggunakan model Rasch. Adapun tahap *disseminate* dilakukan secara terbatas melalui finalisasi dan penyampaian produk pada lingkungan penelitian. Prosedur tersebut menghasilkan produk awal yang dikembangkan secara sistematis berdasarkan kebutuhan lapangan, kajian teoretis, validasi ahli, dan hasil uji coba.

Ketiga, hasil penilaian ahli dan pengujian empiris menunjukkan bahwa instrumen memiliki kelayakan dengan skor rata-rata 84% dengan kategori sangat Baik, kemudian melalui *expert judgment* CVR/CVI hasilnya Instrumen Valid. Analisis Rasch menunjukkan varians yang dijelaskan oleh ukuran sebesar 32,0%, sehingga memenuhi batas minimal unidimensionalitas, sedangkan nilai Alpha Cronbach sebesar 0,77 menunjukkan konsistensi internal yang memadai.

Mayoritas butir menunjukkan kesesuaian dengan model *Rasch*, meskipun masih terdapat beberapa butir yang perlu direvisi karena belum sesuai atau berada pada tingkat kesukaran ekstrem. Reliabilitas person dan item yang masih rendah juga menunjukkan perlunya pengujian lebih lanjut dengan jumlah dan karakteristik responden yang lebih beragam. Instrumen dapat diterapkan dalam empat pertemuan uji coba pada 19 santri dengan dukungan aplikasi *Wayground*. Oleh karena itu, instrumen dapat dinyatakan layak sebagai produk awal untuk digunakan secara terbatas dalam penilaian keterampilan membaca inferensial bahasa Arab di Pondok Pesantren Syaikh Jamilurrahman Bantul, dengan tetap memerlukan penyempurnaan butir dan pengujian pada sampel yang lebih luas.

Secara keseluruhan, penelitian ini menegaskan bahwa pengembangan Instrumen penilaian formatif autentik berbasis aplikasi *Wayground* memberikan kontribusi nyata terhadap pembelajaran membaca bahasa Arab di pesantren, khususnya dalam menghadirkan penilaian yang tidak hanya berfungsi sebagai alat evaluasi, tetapi juga sebagai sarana pemantauan, umpan balik, dan penguatan kemampuan membaca inferensial santri. Dengan demikian, penelitian ini menghasilkan sebuah Instrumen penilaian yang memiliki dasar teoritis, dasar empiris, dan dasar praktis untuk digunakan serta dikembangkan lebih lanjut dalam pembelajaran bahasa Arab.

B. Saran Pemanfaatan Produk

Berdasarkan hasil penelitian yang telah dilakukan, terdapat beberapa saran pemanfaatan produk yang dapat dikemukakan sebagai berikut.

Pertama, Instrumen penilaian formatif autentik berbasis aplikasi

Wayground ini dapat dimanfaatkan oleh guru bahasa Arab di lingkungan pesantren sebagai alternatif penilaian membaca yang lebih variatif, khususnya untuk menilai keterampilan membaca inferensial. Produk ini dapat digunakan tidak hanya untuk memperoleh nilai, tetapi juga untuk membantu guru memantau kemampuan santri, memberikan umpan balik, dan merancang tindak lanjut pembelajaran secara lebih terarah.

Kedua, produk ini dapat dimanfaatkan sebagai perangkat penilaian pendukung dalam pembelajaran teks-teks bahasa Arab yang bernuansa keislaman. Karena Instrumen ini disusun dengan mempertimbangkan konteks pesantren, maka penggunaannya akan lebih relevan apabila diterapkan pada materi bacaan yang memiliki kedekatan dengan pengalaman belajar santri, baik dalam pembelajaran qirā'ah maupun dalam kajian teks keagamaan.

Ketiga, pemanfaatan produk ini sebaiknya disertai dengan penyesuaian terhadap kondisi teknis di lapangan, terutama terkait ketersediaan perangkat dan kestabilan jaringan internet. Dengan demikian, penggunaan aplikasi *Wayground* dapat berlangsung secara lebih lancar dan mendukung proses penilaian secara optimal.

Keempat, produk ini dapat dijadikan dasar awal untuk pengembangan instrumen penilaian serupa pada keterampilan berbahasa Arab yang lain atau pada jenjang pendidikan yang berbeda. Dengan kata lain, Instrumen yang dihasilkan dalam penelitian ini tidak hanya bermanfaat pada konteks penelitian saat ini, tetapi juga berpotensi dikembangkan lebih lanjut sebagai rujukan dalam inovasi penilaian bahasa Arab berbasis digital.

C. Saran Pengembangan Produk

Berdasarkan hasil penelitian ini, terdapat beberapa saran yang dapat diajukan untuk pengembangan produk maupun penelitian lanjutan. Pertama, produk yang dikembangkan dalam penelitian ini masih dapat disempurnakan, terutama pada aspek kualitas butir soal, pemerataan tingkat kesukaran, dan penguatan daya pembeda instrumen. Oleh karena itu, pengembangan berikutnya perlu diarahkan pada revisi lebih lanjut terhadap butir-butir yang masih menunjukkan kelemahan berdasarkan hasil analisis empiris.

Kedua, penelitian lanjutan disarankan melibatkan jumlah subjek yang lebih besar dan konteks lembaga yang lebih beragam agar kualitas Instrumen dapat diuji secara lebih luas. Dengan melibatkan sampel yang lebih banyak, hasil analisis instrumen akan menjadi lebih stabil dan kekuatan generalisasi Instrumen akan semakin meningkat.

Ketiga, penelitian berikutnya dapat mengembangkan Instrumen penilaian ini ke dalam bentuk yang lebih komprehensif, misalnya dengan menambahkan fitur umpan balik otomatis yang lebih terstruktur, pengayaan bank soal, atau integrasi dengan pelaporan hasil belajar yang lebih rinci. Dengan demikian, fungsi formatif Instrumen dapat diperkuat tidak hanya pada tahap penilaian, tetapi juga pada tahap tindak lanjut pembelajaran.

Keempat, penelitian lanjutan juga dapat mengadaptasi Instrumen ini pada keterampilan bahasa Arab lainnya, seperti menyimak, berbicara, atau menulis, sehingga pengembangan penilaian autentik berbasis digital dalam pembelajaran bahasa Arab menjadi lebih luas dan berkelanjutan.

DAFTAR PUSTAKA

- Abduh, Moch, *Model Penilaian Formatif Pada Pembelajaran Abad Ke-21 Untuk Sekolah Dasar* (Perpustakaan Nasional Republik Indonesia, 2019)
- Achmad, Ghufuran Hasyim, and others, 'Penilaian Autentik Pada Kurikulum Merdeka Belajar Dalam Pembelajaran Pendidikan Agama Islam Di Sekolah Dasar', *EDUKATIF: Jurnal Ilmu Pendidikan*, 4.4 (2022), pp. 5685–99, doi:https://doi.org/10.31004/edukatif.v4i4.3280
- Adriantoni, and others, 'Menilai Bukan Sekedar Menghitung: Peran Asesmen Formatif Dalam Meningkatkan Kualitas Pembelajaran', 11 (2025), pp. 221–29
- Anderson, Lorin W, and David R Krathwohl, *A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives: Complete Edition* (Addison Wesley Longman, Inc., 2001)
- Andrian, Dedek, and others, 'Implementasi Formatif Dan Sumatif Assesmen Dalam Meningkatkan', *Jurnal Pengabdian Masyarakat Bangsa*, 2.3 (2024), pp. 479–85
- Arifin, Zainal, *Evaluasi Pembelajaran* (Bandung: Remaja Rosdakarya, 2009), VIII
- Arip Ardiansyah, Ade, Nanang Kosim, and Acep Hermawan, 'Al- Ta'rib The Use of Animated Videos in Increasing the Arabic Learning Interest', *Al-Ta'rib: Jurnal Ilmiah Program Studi Pendidikan Bahasa Arab*, 10.1 (2022), pp. 35–50, doi:https://doi.org/10.23971/altarib.v10i1.3942 The
- Azhar, Muhammad, and others, 'E-ISSN : 2797-1910 Evaluasi Pembelajaran Bahasa Arab Berbasis Teknologi Digital', 6.1 (2025), pp. 78–99
- Azwar, Saifuddin, 'Metode Penelitian Psikologi Edisi II' 2018, 2018
- Badriyah, Nurul Lailatul, 'Analisis Instrumen Penilaian Hasil Belajar Mata Pelajaran Gambar Teknik Siswa Kelas X Sekolah Menengah Kejuruan Program Keahlian Bangunan' 2018, 2018
- Barrett, Thomas C, 'Taxonomy of Reading Comprehension', *Teaching Reading in the Middle Class* 1976, 1976, pp. 51–58
- Black, Paul, and Dylan Wiliam, 'Assessment and Classroom Learning', *Assessment in Education: Principles, Policy & Practice*, 5.1 (1998), pp. 7–74
- Borich, Gary D, *Observation Skills for Effective Teaching: Research-Based Practice* (Routledge, 2016)
- Branch, Robert Maribe, and İlhan Varank, *Instructional Design: The ADDIE Approach* (Springer, 2009), DCCXXII
- Citra, Cahyani Amildah, and Brilliant Rosy, 'Keefektifan Penggunaan Media Pembelajaran Berbasis Game Edukasi Quizizz Terhadap Hasil Belajar Teknologi Perkantoran Siswa Kelas X SMK Ketintang Surabaya', *Jurnal Pendidikan Administrasi Perkantoran (JPAP)*, 8.2 (2020), pp. 261–72

- Creswell, John W, and J David Creswell, *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches* (Sage publications, 2017)
- Damaianti, Vismaia, and others, 'Pengembangan Instrumen Literasi Membaca Autentik Berbantuan Aplikasi Quizizz', *Journal Visipena*, 15.2 (2024), pp. 219–35
- Daryanto, Mulyo Rahardjo, 'Model Pembelajaran Inovatif', *Yogyakarta: Gava Media*, 2012
- Dick, Walter, Lou Carey, and James O Carey, 'The Systematic Design of Instruction' 1985, 1985
- Djaali, D, and P Muljono, 'Pengukuran Dalam Pendidikan' (Jakarta: Program Pascasarjana, 2000)
- Van Doren, Charles, and Mortimer J Adler, *How to Read a Book: The Classic Guide to Intelligent Reading* (Blackstone Audio, 2012)
- Esthiningtyas Sheilla Purwandari, 'PENGEMBANGAN INSTRUMEN PENILAIAN BERBASIS HOTS PADA KITAB DURUSUL LUGHOH AL-ARABIYYAH LI GHAIRI NATHIQINA BIHA JILID 1' (UIN Sunan Kalijaga Yogyakarta, 2023)
- Faizmailatus Sofa, and others, 'Innovation of Language Receptive Skills Based on Authentic Assessment', *Izdiyar : Journal of Arabic Language Teaching, Linguistics, and Literature*, 6.1 (2023), pp. 31–44, doi:10.22219/jiz.v6i1.24029
- Fikri, Ahmad Futunul, and R Umi Baroroh, 'HOTS Based Maharah Qira'ah Summative Assessment Instrument Innovation on the Ability of Class VIII Students of Diponegoro Middle School Depok Sleman Yogyakarta', *INTAJIYA: Journal of Ta'lim Al-Lughah Al-'arabiyah*, 2.1 (2025), pp. 19–38
- Fuad, Ahmad, 'Metodologi Pengajaran Bahasa Arab' 2016, 2016
- Gall, Meredith Damien, Walter R Borg, and Joyce P Gall, *Educational Research: An Introduction* (Longman publishing, 1996)
- Habibi, Mukhlas, Fina Aulika Lestari, and Yusmicha Ulya Afif, 'Implementasi Penilaian Autentik Kurikulum 2013 Pada Mata Pelajaran PAI Dan Budi Pekerti Di SDN 1 Bangunrejo Ponorogo', *Qalamuna - Jurnal Pendidikan, Sosial, Dan Agama*, 13.2 (2021), pp. 833–52, doi:10.37680/qalamuna.v13i2.1114
- Hadjar, Ibnu, 'Dasar-Dasar Metodologi Penelitian Kuantitatif Dalam Pendidikan', *Jakarta: Raja Grafindo Persada*, 1996
- Hasanah, Mamluatul, and others, 'Arabic Performance Curriculum Development: Reconstruction Based On ACTFL And Douglas Brown Perspective', *Ijaz Arabi Journal of Arabic Learning*, 4.3 (2021), pp. 779–801, doi:10.18860/ijazarabi.v4i3.11900
- Hasmawati, Ahmad Mukhtar, 'Asesmen Dalam Kurikulum Merdeka Perspektif Pendidikan Agama Islam Institut Agama Islam As ' Adiyah Sengkang', *Indonesian Journal of Innovation Multidisipliner Research*, 1.3 (2023), pp. 197–211
- Hasna, Tsania Sausan, and others, 'Innovation in Authentic Assessment of Maharah Qira ' Ah in the Textbook of Al- ' Arabiyyah Baina Yadaik Volume II', II.1 (2025), pp. 63–77, doi:10.24042/albayan.v17i1.24681
- Hermawan, Acep, 'Metodologi Pembelajaran Bahasa Arab, Cet Ke-1, Bandung: PT', *Remaja Rosdakarya*, 135 (2011)

- , *PENILAIAN PEMBELAJARAN BAHASA ARAB PRINSIP DAN OPERASIONALNYA*, ed. by Rika Indrawati, 1st edn (PT Remaja Rosdakarya)
- Hidayat, Rahmat, 'Implementasi Pembelajaran Berbasis HOTS (Higher Order Thinking Skills) Pada Mata Pelajaran Pendidikan Agama Islam Dan Budi Pekerti Di SMA Negeri 4 Palu', 2022
- Huda, Asrul, and others, 'Pengembangan Modul Pelatihan Desain Grafis Dan Digital Printing Berbasis Flip Book', *JAVIT: Jurnal Vokasi Informatika*, 2022, pp. 114–20, doi:10.24036/javit.v2i3.127
- Izzan, H Ahmad, *Metodologi Pembelajaran Bahasa Arab* (Humaniora Utama Press, 2011)
- Jailani, M Syahrani, 'Teknik Pengumpulan Data Dan Instrumen Penelitian Ilmiah Pendidikan Pada Pendekatan Kualitatif Dan Kuantitatif', *IHSAN: Jurnal Pendidikan Islam*, 1.2 (2023), pp. 1–9
- Kalyana, Dyani, and others, 'Implementasi Asesmen Autentik Berbasis Kompetensi Dalam Pembelajaran', *Jurnal Pengabdian Masyarakat Dan Riset Pendidikan*, 4.2 (2025), pp. 10521–31
- Kintsch, Walter, 'The Role of Knowledge in Discourse Comprehension : A Construction-Integration Model', 95.2 (1988), pp. 163–82
- Kunandar, Kunandar, 'Penilaian Autentik (Penilaian Hasil Belajar Peserta Didik Berdasarkan Kurikulum 2013)', *Jakarta: Rajawali Pers*, 2013, pp. 16–18
- Kurnia, Asep Dudi, 'Implementasi Penilaian Formatif Dalam Pembelajaran Bahasa Inggris Tatap Muka Terbatas Untuk Meningkatkan Motivasi Belajar Siswa', *STRATEGY : Jurnal Inovasi Strategi Dan Model Pembelajaran*, 2.1 (2021), pp. 67–77
- Lawshe, Charles H, 'A Quantitative Approach to Content Validity.', *Personnel Psychology*, 28.4 (1975)
- Lestari, Tika Wahyu, 'Kahoot! And Quizizz: A Comparative Study on the Implementation of e-Learning Application toward Studentsâ?? Motivation', *Karya Ilmiah Dosen*, 2.2 (2019)
- Linda, Rosa, and Muhammad Syafriansyah, 'Pengembangan Media Pembelajaran Bahasa Arab Melalui Aplikasi Canva', *JSR: Jurnal Sains Riset*, 13.1 (2023), pp. 30–40, doi:10.47647/jsr.v13i1.856
- Mahmudah, and Agus Susilo, 'Penilaian Ideal Dan Evaluasi Efektif Dalam Pembelajaran Bahasa Arab', in *International Conference on Islamic Civilization and Humanities*, 2024, pp. 787–99
- Marlina, Sinthia, Amperawati Turnip, and Wiputra Cendana, 'Implementasi Penilaian Formatif Autentik Era Pembelajaran Daring Berbasis Permainan Digital Sederhana Kelas II Sekolah Dasar', *Trapsila: Jurnal Pendidikan Dasar*, 3.1 (2021), pp. 1–11
- Matlan, Siti Joanna, and Siti Mistima Maat, 'Penggunaan Aplikasi Quizizz Sebagai Alternatif Penilaian Formatif Dalam Pengajaran Dan Pembelajaran Matematik (The Use of Quizizz Application as an Alternative to Formative Assessment in Teaching and Learning Mathematics)', *Jurnal Dunia Pendidikan*, 3.4 (2021), pp. 217–27
- Merta, I Made Endra Danu, I Made Suarjana, and Luh Putu Putrini Mahadewi, 'Analisis Penilaian Autentik Menurut Pembelajaran Kurikulum 2013 Pada Kelas IV SD No. 4

- Banyuasri', *Mimbar PGSD Undiksha*, 3.1 (2015)
- Mondolalo, Darminton, and La Muda, 'Keterampilan Membaca Pemahaman Level Menengah Paragraf Deskriptif Siswa Kelas VII SMP Karuna Dipa Palu The Intermediate Level Reading Comprehension Skills Descriptive Paragraph by Eighth Grade Karuna Dipa Palu Junior High School Students Jurnal Pembelajaran', *Jurnal Pembelajaran Bahasa Dan Sastra*, 1.2 (2022), pp. 263–70, doi:<https://doi.org/10.55905/jpbs.v1i2.34>
- Mueller, J F, 'Authentic Assessment Toolbox. Retrieved February 1, 2015', 2014
- Mufti, Ali, and others, 'Socrative , Quizizz , and Google Form as Online-Based Evaluation Tools for Maharah Al- Qirā ' Ah Learning', 2022, pp. 120–36, doi:10.18196/mht.v4i2.14229
- Munir, M A, and M Ag, 'Perencanaan Sistem Pengajaran Bahasa Arab', *Teori Dan Praktik*, 2019
- Muthi'ah Erwandi, Asilah, Nidya Chandra Uji Utami, and Imaningtyas, 'Pengaruh Strategi Pembelajaran Directed ReadingThinking Activity Terhadap Kemampuan Membaca Inferensial Siswa Kelas IV SD', *Pendas : Jurnal Ilmiah Pendidikan Dasar*, 10.2 (2025), pp. 221–33
- Mutmainna.B, 'PENGEMBANGAN MODEL PENILAIAN AUTENTIK UNTUK MENGUKUR KETERAMPILAN MENYIMAK DAN KETERAMPILAN KOMUNIKASI PESERTA DIDIK KELAS IV SEKOLAH DASAR' (UIN SUNAN KALIJAGA YOGYAKARTA., 2023)
- Oktalia, Anggela, 'Implementasi Assesmen Autentik Dalam Mata Pelajaran IPAS Kelas V Di SDN 32 Rejang Lebong', 2025
- Pamungkas, Meylinda Ambar, and Tri Joko Raharjo, 'Pengaruh Penggunaan Quizizz Paper Mode Terhadap Penilaian Formatif Dan Motivasi Belajar Peserta Didik Kelas VI Dalam Kurikulum Merdeka Sekolah Dasar', *INNOVATIVE: Journal Of Social Science Research*, 4.1 (2024), pp. 12449–60
- Plomp, Tjeerd, 'Educational Design Research: An Introduction', *Educational Design Research*, 1 (2013), pp. 11–50
- Pusparani, Herlina, 'Media Quizizz Sebagai Aplikasi Evaluasi Pembelajaran Kelas vi Di Sdn Guntur Kota Cirebon', *Tunas Nusantara*, 2.2 (2020), pp. 269–79
- Putri, Arliani Gunawan, Arrin Nur Octavialis, and Irma Savitri Sadikin, 'IMPROVING STUDENTS'READING SKILL THROUGH COLLABORATIVE LEARNING', *PROJECT (Professional Journal of English Education)*, 2.6 (2019), pp. 861–68
- Rejeki, Pinte, 'Efektifitas Gabungan Tes Subjektif Dan Tes Objektif Dalam Mengevaluasi Hasil Belajar Fisika Siswa Smp Negeri 11 Banda Aceh', *Jurnal Ilmiah Mahasiswa Pendidikan Fisika*, 1.3 (2016)
- Richey, Rita C, and James D Klein, 'Design and Development Research', *Mahwah, NJ* 2007, 2007
- Riyanto, Bambang, 'Pendesainan Asesmen Inovatif Melalui Pembelajaran Pemodelan Menggunakan Konteks Biaya Rumah Tangga', *Prosiding SENPIKA Volume*, 1.2020 (2023), pp. 1–15

- Saekoko, Neti, and others, 'Peran Evaluasi Formatif Dalam Meningkatkan Kualitas Pembelajaran Di Era Digital', *Jurnal Ilmiah Literasi Indonesia*, 1.2 (2025), pp. 336–50, doi:doi.org/10.63822/8t7k4h35
- Saiegh-Haddad, Elinor, and Rachel Schiff, 'Diglossic and Orthographic Features of Reading Comprehension in Standard Arabic: The Primacy of the Spoken Language', *Reading Research Quarterly*, 60.1 (2025), p. e598
- Septia AP, Maisa, Chandra, and Inggria Kharisma, 'Kemampuan Berfikir Kritis Dalam Keterampilan Membaca Pemahaman Peserta Didik Kelas VI SD', *Pediaqu : Jurnal Pendidikan Sosial Dan Humaniora*, 4.3 (2025), pp. 5837–55
- Snow, Catherine, *Reading for Understanding: Toward an R&D Program in Reading Comprehension* (Rand Corporation, 2002)
- Subali, Bambang, and Pujiyati Suyata, 'Panduan Analisis Data Pengukuran Pendidikan Untuk Memperoleh Bukti Empirik Kesahihan Menggunakan Program Quest', *Lembaga Penelitian Dan Pengabdian Pada Masyarakat UNY*, 2011
- Sudijono, Anas, 'Pengantar Evaluasi Pendidikan', 2013
- Sugiyono, *Metode Penelitian Kualitatif Kuantitatif Dan R&D* (Alfabeta, 2019)
- Suharman, Suharman, 'Tes Sebagai Alat Ukur Prestasi Akademik', *AT-TA'DIB: Jurnal Ilmiah Prodi Pendidikan Agama Islam*, 2018, pp. 93–115
- Suharsimi, Arikunto, 'Prosedur Penelitian Suatu Pendekatan Praktik', *Jakarta: Rineka Cipta*, 134 (2006), p. 252
- Sumintono, Bambang, 'Aplikasi Pemodelan Rasch Pada Assessment Pendidikan', 2015
- Syamsul, Rizal, 'Reading Skill Teori Dan Praktik Pengukurannya', *Yogyakarta: Samudra Biru*, 2018
- Thiagarajan, Sivasailam, Dorothy S Semmel, and Melvyn I Semmel, 'Instructional Development for Training Teachers of Exceptional Children' (Minneapolis, Minnesota: leadership training institute/special education ..., 1974)
- Umamah, Ade, 'Analisis Kemampuan Membaca Teks Bahasa Arab Siswa Kelas XI Madrasah Aliyah Pondok Pesantren DDI As-Salman Allakuang Kabupaten Sidrap (Studi Perspektif Psikolinguistik)', 2024
- Umi, Machmudah, and Abdul Wahab Rosyidi, 'Active Learning Dalam Pembelajaran Bahasa Arab' (Malang: UIN Malang Press, 2008)
- Utami, Ayu, and Salito, 'Implementasi Asesmen Berkelanjutan Untuk Meningkatkan Kualitas Pembelajaran Di Sekolah Dasar', *Jurnal Akuntansi, Manajemen, Dan Ilmu Pendidikan (JAMED)*, 1.3 (2025), pp. 124–36
- Wibawa, Ramadhan Prasetya, Rohana Intan Astuti, and Bayu Aji Pangestu, 'Smartphone-Based Application "Quizizz" as a Learning Media', *Dinamika Pendidikan*, 14.2 (2019), pp. 244–53
- Widiyoko, Eko Putro, 'Evaluasi Program Pembelajaran: Panduan Praktis Bagi Pendidik Dan Calon Pendidik', *Yogyakarta: Pustaka Pelajar*, 2013
- Wiggins, Grant, 'The Case for Authentic Assessment', *Practical Assessment, Research, and*

Evaluation, 2.1 (1990)

- Wildan, Wildan, 'Pelaksanaan Penilaian Autentik Aspek Pengetahuan, Sikap Dan Keterampilan Di Sekolah Atau Madrasah', *Jurnal Tatsqif*, 15.2 (2017), pp. 131–53
- Wulandari, Bunga Ayu, Muhammad Sofwan, and R Septa Kusuma, 'Meningkatkan Kemampuan Membaca Pemahaman Literral Dan Inferensial Dengan Menggunakan Model Hands on Actifity', *JIIP (Jurnal Ilmiah Ilmu Pendidikan)*, 6.5 (2023), pp. 2987–93
- Yamin, M., 'Pelaksanaan Penilaian Autentik Kurikulum 2013 Pada Pembelajaran Tematik SD Gugus Markisa Kota Banda Aceh', *Elementary Education Research*, 3.1 (2018)
- Yamin, M Yamin M, 'Pelaksanaan Penilaian Autentik Kurikulum 2013 Pada Pembelajaran Tematik SD Gugus Markisa Kota Banda Aceh', *Elementary Education Research*, 3.1 (2018)
- Zakiyah, Nida, and Yunus Abu Bakar, 'Sistem Penilaian Pembelajaran Transformatif Berparadigma Holistik Dalam Kurikulum Merdeka', *At-Tahsin : Jurnal Manajemen Pendidikan Islam*, 5.1 (2025), pp. 120–35

