

التحول الرقمي في تقويم اللغة العربية: تطوير وتقييم اختبار القراءة الإلكتروني المستند

على الإطار المرجعي الأوروبي المشترك للغات

سفيرة عين نجية^{١*}، إخوان محمودي^{٢-ب}، محمد إسماعيل^{٣-ج}

^١، ^٢، ^٣ جامعة دار السلام كونتور

¹safiraaina30@gmail.com، ²ihwanm@unida.gontor.ac.id، ³ismail@unida.gontor.ac.id

Received: Nov 25, 2026 Accepted: March 16, 2026 Published: April 28, 2026

المستخلص

أدى التطوير السريع للتكنولوجيا الرقمية إلى تحولات كبيرة في تعليم اللغات، بما في ذلك تقييم المهارات اللغوية. تهدف هذا البحث إلى تطوير وتقييم اختبار مهارات القراءة عبر الإنترنت باستخدام منصة كلاس ماركر Classmarker. المصممة على أساس الإطار المرجعي الأوروبي المشترك للغات (CEFR). منهج البحث المستخدمة هي البحث والتطوير بنموذج 4D. تضمنت عملية التطوير مراحل تحليل الاحتياجات، وتصميم بنود الاختبار بناء على مستويات الإطار المرجعي الأوروبي المشترك للغات (CEFR)، والتحقق من صحتها من قبل الخبراء، واختبار طلاب الصف الخامس في معهد دار السلام كونتور الحرم الأول للبنات. أظهرت نتائج الدراسة أن الاختبار المطور كان صحيحًا وموثوقًا وحظي بردود فعل إيجابية من المستخدمين من حيث التطبيق العملي وسهولة الاستخدام. وبالتالي، يمكن أن يكون تطوير الاختبارات القائمة على الإطار المرجعي الأوروبي المشترك للغات من خلال هذه المنصة الرقمية ابتكارًا في التحويل الرقمي لتقييم اللغة العربية، مع دعم تحسين جودة تقييم تعلم اللغة العربية في العصر الحديث.

الكلمات الرئيسية: اختبار مهارات القراءة، الإطار المرجعي الأوروبي المشترك للغات (CEFR)، كلاس ماركر (ClassMarker)، تحليل بنود الأسئلة.

Abstrak

The rapid evaluation of digital technology has driven significant transformations in language education, including language skills assessment. This study aims to develop and evaluate an online Arabic reading skills test using the ClassMarker platform, which is designed based on the Common European Framework of Reference for Language (CEFR). The research method used in Research and Development (R&D) with a 4D model. The development process included the stages of need analysis, test item design based on CEFR levels, expert validation, and testing of 5th grade students at KMI Darussalam Gontor for Girls 1st Campus. The result of study show that the developed test is valid and reliable, and received positive responses from users in terms of practicality and ease of use. Thus, the development of CEFR-based tests through this digital platform can be an innovation in the digital transformation of Arabic language assessment, while also supporting the improvement of the quality of Arabic language learning evaluation in the modern era.

Keywords: Reading Skills Test, Common European Framework of Reference for Languages (CEFR), ClassMarker, Item Analysis.

المقدمة

في عصر العولمة وتطور التكنولوجيا الرقمية، يتطلب تعلم اللغات الأجنبية، بما في ذلك اللغة العربية، التركيز ليس فقط على إتقان قواعد اللغة، بل أيضاً على تحقيق كفاءة لغوية قابلة للقياس وموحدة دولياً.¹ تعد مهارة القراءة إحدى المهارات الأساسية في تعلم اللغة العربية، حيث تلعب دوراً مهماً في بناء الفهم اللغوي ومهارات التفكير النقدي والوصول إلى الدراسات الإسلامية والنصوص الأكاديمية.² ومع ذلك، لا يزال تقييم مهارات القراءة باللغة العربية في العديد من المؤسسات التعليمية يهيمن عليه الاختبارات التقليدية التي لا تفي تماماً بمبادئ الصحة والموثوقية والتوحيد الدولي.³

أحد الأطر الدولية الأكثر استخداماً في تعلم اللغات وتقييمها هو الإطار المرجعي الأوروبي المشترك للغات.⁴ يوفر هذا الإطار وصفاً منهجياً ومتدرجاً لمستويات إتقان اللغة، مما يتيح قياساً موضوعياً ومقارناً للكفاءة اللغوية.⁵ على الرغم من أن الإطار المرجعي الأوروبي المشترك للغات قد تم اعتماده على نطاق واسع في تعلم اللغات الأجنبية، مثل اللغة الإنجليزية، إلا أن تطبيقه في تعلم اللغة العربية - لا سيما في تقييم مهارات القراءة - لا يزال محدوداً، خاصة في سياق تعلم اللغة العربية كلغة أجنبية في إندونيسيا.⁶ ويصبح هذا القيد أكثر وضوحاً عندما لا تكون أدوات التقييم المستخدمة متكاملة مع التكنولوجيا الرقمية التي تدعم الكفاءة والدقة في التقييم.

مع تزايد الحاجة إلى التقييم الرقمي، أصبحت الاختبارات عبر الإنترنت بديلاً مناسباً واستراتيجياً. منصة كلاس ماركر ClassMarker هي نظام تقييم عبر الإنترنت يوفر التقييم التلقائي وإدارة الأسئلة وتحليل نتائج الاختبار في الوقت الفعلي.⁷ إن دمج هذه المنصة مع إطار المرجعي الأوروبي المشترك للغات لديه القدرة على إنتاج أداة تقييم لمهارات القراءة ليست عملية وفعالة فحسب، بل موحدة وقائمة على الكفاءة أيضاً. لذلك، تهدف هذا البحث إلى تطوير اختبار مهارات القراءة باللغة العربية عبر الإنترنت استناداً إلى الإطار المرجعي الأوروبي المشترك للغات باستخدام منصة كلاس ماركر ClassMarker واختبار خصائصه ومدى ملاءمته كأداة تقييم لطالبات الصف الخامس في معهد دار السلام كونتور الحرم الأول للبنات.

¹ Syamfa Agny Anggara and Mustahar Ali Wardana, "Instrumen Penilaian Bahasa Arab di Era Digital," *JUPI (Jurnal Ilmiah Pendidikan Islam)* 3, no. 1 (2024): 101–8, <https://doi.org/10.58788/jipi.v3i1.4201>.

² Amila Sholiha et al., "Innovation in Reading Skills Assessment in Arabic Textbooks Based on HOTS Assessment," *Izdihar : Journal of Arabic Language Teaching, Linguistics, and Literature* 6, no. 2 (2023): 2, Arabic Textbook Grade XI by the Ministry of Religious 2020, <https://doi.org/10.22219/jiz.v6i2.27151>.

³ Salwa Azizah Rahman and Hikmah Maulani, "Pembelajaran Bahasa Arab Dalam Meningkatkan Keterampilan Kalam (Berbicara) Di SMP Ma'had Ihya As-Sunnah Tasikmalaya: Metode, Tantangan, Dan Solusi," *LANGUAGE: Jurnal Inovasi Pendidikan Bahasa Dan Sastra* 4, no. 2 (2024): 2, <https://doi.org/10.51878/language.v4i2.3933>.

⁴ Erfan Gazali and Hasan Saefuloh, "Development of an Arabic Receptive Proficiency Test Instrument Based on the Common European Framework of Reference for Languages," *Al-Ta'rib : Jurnal Ilmiah Program Studi Pendidikan Bahasa Arab IAIN Palangka Raya* 11, no. 2 (2023): 293–308, <https://doi.org/10.23971/altarib.v11i2.6721>.

⁵ Sudaryanto Sudaryanto and Pratomo Widodo, "Common European Framework of Reference for Languages (CEFR) Dan Implikasinya Bagi Buku Ajar BIPA," *Jurnal Idiomatik: Jurnal Pendidikan Bahasa Dan Sastra Indonesia* 3, no. 2 (2020): 2, <https://doi.org/10.46918/idiomatik.v3i2.777>.

⁶ Laily Fazlin Khairil and Siti Eshah Mokshein, "21st Century Assessment: Online Assessment," *International Journal of Academic Research in Business and Social Sciences* 8, no. 1 (2018): Pages 659-672, <https://doi.org/10.6007/IJARBS/v8-i1/3838>.

⁷ Temitayo Omolara Onasanya et al., "Impact of the ClassMaker App on the Performance of Undergraduates in Online Learning of ICT: The University of Ilorin Experience," *West African Journal of Open and Flexible Learning* 9, no. 2 (2021): 55–74, <https://wajofel.org/index.php/wajofel/article/view/72>.

نتائج البحث والمناقشة

مواصفات اختبار إلكتروني لمهارة القراءة العربية على الإطار المرجعي الأوروبي المشترك للغات (CEFR) باستخدام منصة كلاس ماركر (ClassMarker).

فإن الاختبار الإلكتروني لمهارة القراءة باللغة العربية الذي طوّرتها الباحثة عبر منصة كلاس ماركر ClassMarker يمتلك خصائص تقنية وتربوية متكاملة. وقد قامت الباحثة بتطوير هذا الاختبار بالاعتماد على نظرية الإطار الأوروبي المرجعي المشترك للغات (CEFR) التي تؤكد على وضوح مستويات كفاءة نمهارة القراءة وأصالة النصوص.⁸ كما تدعم ذلك نظرية الاختبار اللغوي بمساعدة الحاسوب (CALT) Douglas و Chapelle J، والتي تشدد على أن التقييم الإلكتروني يُعد جزءًا من CALT، أي دمج التكنولوجيا الحاسوبية في قياس اللغة بطريقة تهدف إلى تحسين الدقة والسرعة وموضوعية التقييم. وقد مكّن اعتماد الباحث على نظرية الإطار المرجعي الأوروبي المشترك للغات CEFR ودمجها مع منصة كلاس ماركر ClassMarker من تقديم أسئلة قراءة تغطي المستويات من (1A) إلى (2C)، مع الاستفادة من ميزات متعددة مثل: أمن الأسئلة، بنك الأسئلة، التصحيح الفوري، التقارير المتقدمة، وغيرها. إن هذا التكامل التكنولوجي جعل عملية التقييم أكثر دقة وعملياتية، وملائمة لاحتياجات تقييم تعلّم اللغة العربية في العصر الرقمي.

من الجانب التربوي، قامت الباحثة بصياغة الاختبار بالاستناد إلى أهداف مهارة القراءة، وهي فهم رسالة النص بدقة وفقًا لسياقه التواصلية. ويتوافق ذلك مع رؤية هنري غونتور تاريغان الذي يرى أن الهدف الرئيس من مهارة القراءة هو تنمية القدرة على التفكير، وتوسيع آفاق المعرفة، وإثراء المفردات.⁹ وفي هذا البحث، اختارت الباحثة نصوصًا تعكس سياقات اجتماعية وثقافية وأكاديمية متنوعة، بحيث لا يقتصر دور الاختبار على قياس قدرة المتعلم على فهم النص فحسب، بل يُمكن المتعلم كذلك من اكتساب معارف جديدة من خلال عملية القراءة. من الجانب اللغوي، قامت الباحثة بصياغة النصوص والأسئلة وفقًا للمفهوم الذي شرحه ريتشارد سي. أندرسون، بأن مهارة القراءة تشمل جانبين رئيسيين، وهما: فكّ الترميز (معالجة الشكل اللغوي) والفهم (استيعاب

⁸ Council of Europe Council for Cultural Co-operation Education Committee Modern Languages Division, *Common European Framework of Reference for Languages: Learning, Teaching, Assessment* (Cambridge University Press, 2001).

⁹ Henry Guntur Tarigan, *Membaca: Sebagai Suatu Keterampilan Berbahasa* (CV. Angkasa, 2021).

محتوى النص).¹⁰ ولذلك، فإن الاختبار المُعدّ لا يقيس القدرة على التعرف على المفردات والتراكيب فحسب، بل يقيس أيضاً القدرة على فهم محتوى النص.

بشكل عام، يُعدّ الاختبار المطوّر في هذا البحث ابتكارياً، لأنه يجمع بين المعايير الدولية للإطار المرجعي الأوروبي المشترك للغات (CEFR) وتقنية التقييم الرقمي عبر منصة كلاس ماركر ClassMarker. ففي هذا البحث، قامت الباحثة بدمج نظرية القراءة، ونظرية تقويم اللغة، ونظرية CEFR، إضافةً إلى تصميم التقييم القائم على التكنولوجيا الرقمية، مما أسهم في إنتاج أداة تقييم أكثر موضوعية وكفاءة وسهولة في الوصول. تُظهر نتائج البحث أن اختبار مهارة القراءة عبر الإنترنت الذي تم تطويره قد نجح في أن يكون نموذجاً مناسباً للتقويم الرقمي في تعليم اللغة العربية للناطقين بغيرها. فهذا الاختبار لا يقتصر على تحقيق أهداف قياس مهارة القراءة وفقاً للإطار المرجعي الأوروبي المشترك للغات، بل يأخذ أيضاً بعين الاعتبار الجوانب النفسية والمعرفية للمتعلمين. وبفضل هذه الخصائص، يمتلك هذا الاختبار إمكانات كبيرة للتوسّع والتطبيق في سياقات تعليمية أخرى، بوصفه جزءاً من التحول الرقمي في تقويم اللغة العربية.

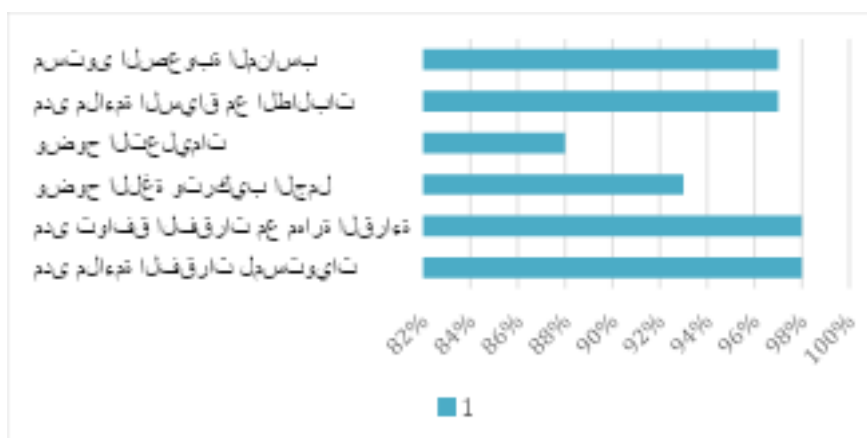
على الباحث أن يستخدم الحامشة كما ورد في المثال الآتي.¹¹ يجب ذكر الجداول على التوالي في النص. ويجب لكل الجدول عنوان وصفي، وإذا تم إعطاء قياسات رقمية فيجب تضمين الوحدات في عنوان العمود. لا ينبغي استخدام القواعد العمودية. وهذه الجداول لا بد أن توضح أو أن تكمل النص الخطوط وليس تكرار النص. صلاحية اختبار إلكتروني لمهارة القراءة العربية على الإطار المرجعي الأوروبي المشترك للغات (CEFR) باستخدام منصة كلاس ماركر (ClassMarker).

ومن المعلوم أن الاختبار الإلكتروني الذي تم تطويره قد مرّ بعدة مراحل من الاختبار. فقد أجري أولاً اختبار أولي من قِبل خبراء في مجالين، وهما: خبير في مادة تعليم اللغة العربية وخبير في تقويم التعلم. ثم تلت ذلك مرحلة التطبيق من خلال التجربة المحدودة والتجربة الواسعة.

¹⁰ Richard C. Anderson and And Others, *Becoming a Nation of Readers: The Report of the Commission on Reading* (University of Illinois, *Becoming a Nation of Readers*, P, 1985), <https://eric.ed.gov/?id=ED253865>.

استنادًا إلى نتائج تقييم الخبير في مجال مادة تعليم اللغة العربية، تبين أن فقرات الاختبار التي تم إعدادها قد حصلت على تصنيف "جيد جدًا" بنسبة بلغت ٩٥%. ويمكن استخدام هذه الفقرات مع إجراء بعض التعديلات والتحسينات الطفيفة، كما أنها تُعد مناسبة للتطبيق على الطالبات. ومن أجل توضيح هذه النتائج وتلخيصها، يعرض الباحثة النتائج في الشكل البياني الآتي:

الصور البياني ١. نتائج تقييم الخبراء في مجال اللغة العربية



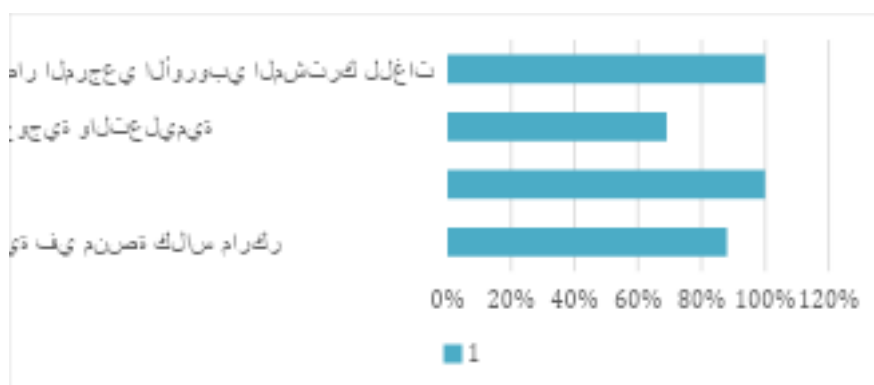
في الرسم البياني، بلغت نتيجة تقييم جانب مدى التوافق مع مستوى الإطار المرجعي الأوروبي المشترك للغات (CEFR) نسبة ٩٨% (ممتاز جدًا)، مما يدل على أن الفقرات الاختبارية الموضوعية قد جاءت متماشية مع أوصاف الكفاءة في كل مستوى من مستويات الإطار المرجعي الأوروبي المشترك للغات (CEFR). أما جانب مدى التوافق مع مهارة القراءة فقد حصل أيضًا على نسبة ٩٨% (ممتاز جدًا)، مما يشير إلى أن الأسئلة المطوّرة تقيس فعلاً عمليتي القراءة الأساسيتين، وهما: فك الترميز و فهم المحتوى.

كما حصل جانب وضوح اللغة وتركيب الجمل على نسبة ٩٣% (ممتاز جدًا)، مما يعني أن اللغة المستخدمة في النصوص واضحة عمومًا، رغم وجود بعض المواضيع البسيطة التي تحتاج إلى تحسين وتعديل. بينما حقق جانب وضوح التعليمات نسبة ٨٨% (جيد)، مما يدل على أن غالبية التعليمات مفهومة بشكل جيد، إلا أن بعضها يحتاج إلى مراجعة وتطوير.

وحصل جانب مدى ملاءمة السياق للطلاب على نسبة ٩٧% (ممتاز جدًا)، مما يدل على أن النصوص وموضوع الأسئلة مناسبة لخبرات الطلبة ومستواهم المعرفي واحتياجاتهم في الصف الخامس، وهو ما يدعم زيادة

مشاركتهم وفهمهم للنصوص. أما جانب مستوى الصعوبة المناسب فقد حقق أيضًا نسبة ٩٧% (ممتاز جدًا)، مما يشير إلى أن مستوى تعقيد الأسئلة وطول النصوص والمتطلبات المهارية في القراءة جاءت ملائمة ومناسبة. أما نتائج تقييم الخبير في مجال الاختبارات الإلكترونية فقد أظهرت أن المنصة المستخدمة قد حصلت على تصنيف "ممتاز جدًا"، ويمكن استخدامها دون الحاجة إلى أي مراجعة أو تعديل. ولتسهيل فهم النتائج وتوضيحها، قام الباحث بعرضها في شكل مخطط بياني على النحو الآتي:

الصور البياني ٢. نتائج تقييم الخبراء في مجال الاختبارات الإلكترونية



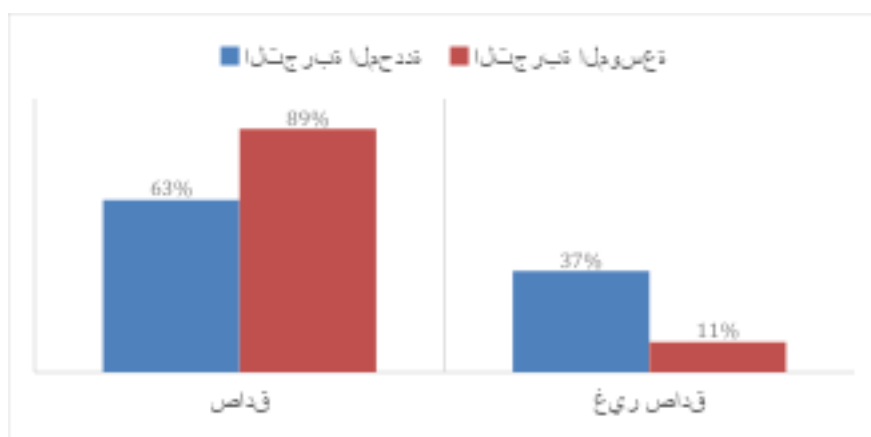
استنادًا إلى نتائج المخطط البياني أعلاه، تُظهر عملية التقييم أن جانب التقنيات في منصة كلاس ماركر ClassMarker قد حصل على نسبة ٨٨٪ (جيد)، مما يدل على أن المنصة توفر ميزات تقنية كافية لتنفيذ الاختبار الإلكتروني، وتتميز بواجهة استخدام سهلة، ونظام تنقل واضح، وكذلك إمكانية الوصول بسلاسة عبر أجهزة الحاسوب والهواتف المحمولة. بينما حصل جانب أمن الاختبار على نسبة ١٠٠٪ (ممتاز)، مما يعني أن منصة كلاس ماركر ClassMarker قادرة على توفير نظام آمن أمثل من خلال ميزات مثل عشوائية الأسئلة، وتحديد الوقت، ووضع المراقبة، وحماية بيانات المشاركين. أما الجانب التربوي/التعليمي فقد حصل على نسبة ٦٩٪ (متوسط)، وهذا يعني أنه على الرغم من أن معظم التعليمات المتعلقة بالاختبار كانت مناسبة، إلا أنه لا يزال هناك حاجة لتحسين وضوح التعليمات حتى تتم عملية أداء الاختبار بصورة أكثر فاعلية دون حدوث ارتباك أو سوء فهم. كما حصل جانب التوافق مع الإطار المرجعي الأوروبي المشترك للغات CEFR على نسبة ١٠٠٪ (ممتاز)، مما يدل على أن مستوى صعوبة النصوص متوافق تمامًا مع الوصف الوارد في الإطار المرجعي الأوروبي المشترك للغات CEFR.

بعد الانتهاء من عملية التقييم من قبل الخبراء، انتقل الباحث إلى المرحلة التالية، وهي إجراء التجربة المحدودة والتجربة الواسعة على طالبات الصف الخامس في قسم التعليم المتوسط بمعهد دار السلام كونتور

البحر الأول للبيانات. وفي كل من التجريبتين استخدمت الباحثة أربعة أنواع من التحليل، وهي: التحليل المتعلق بصدق الاختبار، والتحليل المتعلق بثبات الاختبار، وتحليل مستوى الصعوبة، وتحليل قوة التمييز. تشير صلاحية الاختبار إلى مدى قدرة فقرات الأسئلة على قياس الكفاءة المستهدفة. فالصلاحية ليست خاصية ملازمة للاختبار ذاته، بل ترتبط بالدرجات الناتجة عنه. فهذه الدرجات لا تُعد مجرد نتاج لعوامل أو ظروف تقنية معينة، بل تتأثر أيضًا بالطلبة الذين يجيبون عنها. ولذلك، يجب أن تكون هذه الدرجات ذات معنى ويمكن تفسيرها بشكل صحيح، وأن تمتلك دلالة كافية لاتخاذ القرارات بناءً على تلك النتائج.

في هذا التحليل، تم إجراء اختبار صلاحية الفقرات باستخدام الصلاحية التجريبية، والتي قيست من خلال العلاقة بين درجة كل فقرة والدرجة الكلية باستخدام برنامج SPSS. وقد حصلت الباحثة على قيم الارتباط بين فقرات الاختبار في التجربة المحدودة والتجربة الواسعة. ومن خلال المقارنة، تبين أن عدد الفقرات التي تمتلك صلاحية عالية جدًا قد ارتفع بشكل ملحوظ في التجربة الواسعة، من 63% إلى 89%. وفي المقابل، انخفضت الفقرات التي كانت صلاحيتها منخفضة بشكل كبير من 37% إلى 11%. ويُظهر ذلك وجود تحسن كبير في صلاحية الفقرات في التجربة الواسعة مقارنة بالتجربة المحدودة.

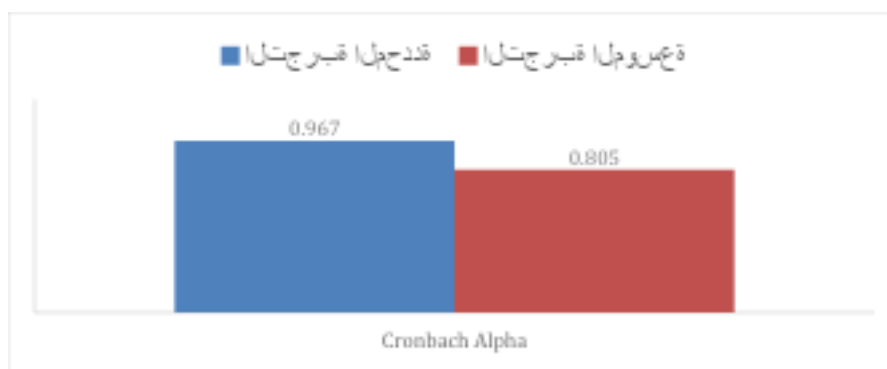
الصور البياني ٣. نتائج اختبار الصدق



بعد ذلك قامت الباحثة بحساب ثبات فقرات الاختبار لمعرفة مدى اتساق نتائج الاختبار إذا أُجري القياس مرة أخرى. وفي هذا التحليل، تم حساب الثبات باستخدام معامل ألفا كرونباخ من خلال برنامج SPSS، حيث يُستخدم معامل ألفا كرونباخ كمؤشر للثبات لقياس الاتساق الداخلي لأداة الاختبار.

من خلال مقارنة نتائج الاختبار في مرحلتى التجربة المحدودة والتجربة الواسعة، يتضح أن كلا المرحلتين أظهرتا درجةً عاليةً جداً من الثبات؛ إذ بلغ معامل الثبات (٠,٩٦٧) في التجربة المحدودة و(٠,٨٠٥) في التجربة الواسعة. وتدل هذه المعطيات على إمكانية تطبيق الاختبار بصفة متسقة على مجموعات مختلفة من المتعلمين.

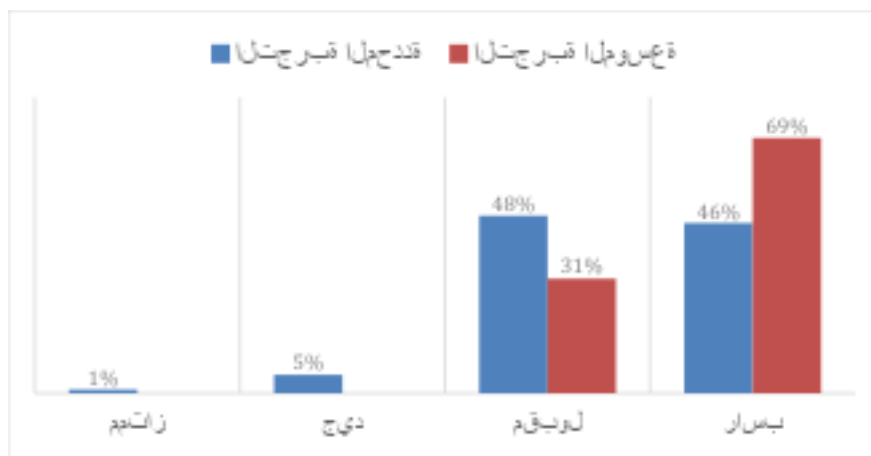
الصور البياني ٤. نتائج اختبار الثبات



بعد ذلك قامت الباحثة بإجراء تحليل التمييز. ويقصد بتمييز الفقرة قدرتها على التفريق بين الطالبات ذوات القدرة العالية والقدرة المنخفضة. وفي الاختبار المحدود أظهرت نتائج التحليل أن نسبة ١٪ من الفقرات كانت في فئة "جيد جداً"، و٥٪ في فئة "جيد"، و٤٨٪ في فئة "متوسط"، و٤٦٪ في فئة "ضعيف"، و١٠٪ في فئة "ضعيف جداً". وتشير هذه النتائج إلى أن غالبية الفقرات تقع في فئة "متوسط"، ومع ذلك ما يزال هناك عدد من الفقرات التي تمتلك قدرة تمييز منخفضة.

وفي الاختبار الواسع، تم الحصول على النتائج الآتية: ١٠٪ من الفقرات في فئة "جيد جداً"، و١٠٪ في فئة "جيد"، و٣١٪ في فئة "متوسط"، و٦٩٪ في فئة "ضعيف"، و١٠٪ في فئة "ضعيف جداً". وبشكل عام تشير نتائج تحليل التمييز إلى أن غالبية الفقرات ما تزال بحاجة إلى مراجعة إضافية لتحسين قدرتها على التمييز بين مستويات قدرات الطالبات.

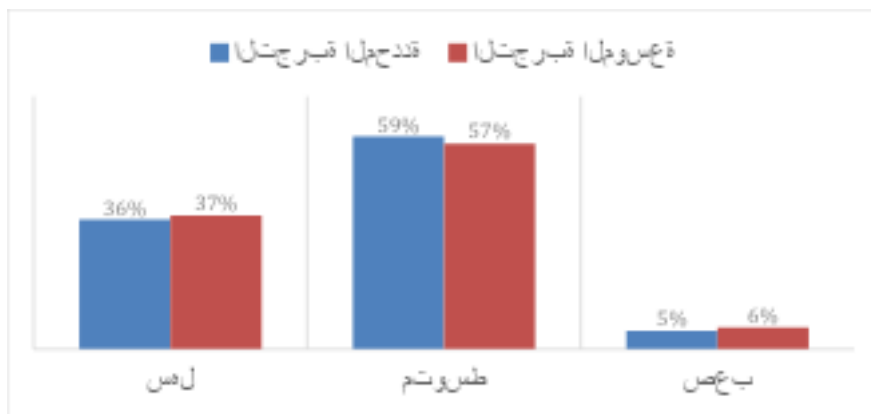
الصور البياني ٥. نتائج تحليل معاملة التمييز



المرحلة التالية التي قامت بها الباحثة هي تحليل مستوى صعوبة فقرات الاختبار. ففي هذا التحليل جرى تصنيف الفقرات بناءً على درجة صعوبتها، إذ إن مستوى الصعوبة يمثل النسبة المئوية التقريبية للطلاب الذين أجابوا عن الفقرة بشكل صحيح. وتُعد الفقرات السهلة جداً أو الصعبة جداً غير مناسبة لقياس قدرات الطلاب بفاعلية. وقد أظهرت نتائج التحليل أن غالبية الفقرات تقع في فئة الصعوبة المتوسطة، وهو ما يتوافق مع تفاوت قدرات الطالبات.

وفي مرحلة التجريب المحدود، بلغت نسبة الفقرات بحسب مستوى الصعوبة: ٣٦% في فئة السهلة، و ٥٩% في فئة المتوسطة، و ٥% في فئة الصعبة. أما في مرحلة التجريب الموسع، فقد وُجدت تغييرات طفيفة، إذ بلغت نسبة الفقرات السهلة ٣٧%، في حين بلغت نسبة الفقرات المتوسطة ٥٧%، وبلغت نسبة الفقرات الصعبة ٦%. وعند المقارنة بين التجريب المحدود والتجريب الموسع، يظهر أنّ هناك زيادة طفيفة في نسبة الفقرات السهلة (من ٣٦% إلى ٣٧%) وكذلك في الفقرات الصعبة (من ٥% إلى ٦%)، بينما انخفضت نسبة الفقرات المتوسطة من ٥٩% إلى ٥٧%. وعلى الرغم من وجود بعض التغيرات في أعداد الفقرات السهلة والمتوسطة، إلا أنّ توزيع مستويات الصعوبة بوجه عام ظلّ مستقرّاً بين التجريبيين.

الصور البياني ٦. نتائج تحليل مستوى صعوبة بنود الأسئلة



استنادًا إلى نتائج تحليل الصدق والثبات ومستوى الصعوبة والقوة التمييزية التي أجريت في مرحلتها التجريبية المحدودة والتجربة الواسعة، خلص الباحث إلى أن اختبار مهارة القراءة المبني على الإطار الأوروبي المرجعي المشترك (CEFR) يمتلك جودة كافية لقياس قدرة الطلاب في مهارة القراءة. وقد تبين أن غالبية فقرات الاختبار تقع ضمن الفئات المتوسطة إلى الجيدة، في حين تم تعديل الفقرات التي لم تستوف المعايير بشكل كافٍ وفقًا لمعايير تطوير الاختبارات. وبناءً على ذلك، فإن الاختبار المطور قد استوفى معايير الصلاحية ليكون أداة مناسبة لتقييم مهارة القراءة باللغة العربية.

الخاتمة

أنتجت هذه الدراسة اختبارًا عبر الإنترنت لمهارات القراءة باللغة العربية استنادًا إلى الإطار المرجعي الأوروبي المشترك للغات (CEFR) باستخدام منصة كلاس ماركر ClassMarker لطلاب الصف الخامس في معهد دار السلام كونتور الحرم الأول للبنات. أظهرت النتائج أن الأداة المطورة تتميز بخصائص تتوافق مع مواصفات مهارات القراءة في الإطار المرجعي الأوروبي المشترك للغات (CEFR) واعتُبرت قابلة للتطبيق بناءً على تقييم الخبراء والاختبارات التجريبية. أثبت تحليل العناصر أن الاختبار استوفى معايير الصلاحية والموثوقية العالية ومستوى الصعوبة النسبي والتمييز الجيد، مما مكّنه من قياس القدرة على القراءة بشكل موضوعي وموحد. يساهم دمج الإطار المرجعي الأوروبي المشترك للغات CEFR مع تقنية التقييم عبر الإنترنت من خلال ClassMarker بشكل كبير في تحسين جودة تقييم تعلم اللغة العربية، لا سيما في سياق المدارس الإسلامية الداخلية، من خلال توفير أداة فعالة ودقيقة وأقل

ذاتية. على الرغم من أن هذه الدراسة تقتصر على مهارات القراءة وسياق مؤسسة تعليمية واحدة، فإن هذه النتائج تفتح فرصًا لتطوير أدوات تقييم اللغة العربية القائمة على الإطار المرجعي الأوروبي المشترك للغات CEFR لمهارات لغوية أخرى وتطبيقها في سياق تعليمي أوسع.

المراجع

- Anderson, Richard C., and And Others. *Becoming a Nation of Readers: The Report of the Commission on Reading*. University of Illinois, Becoming a Nation of Readers, P, 1985. <https://eric.ed.gov/?id=ED253865>.
- Anggara, Syamfa Agny, and Mustahar Ali Wardana. "Instrumen Penilaian Bahasa Arab di Era Digital." *JIPi (Jurnal Ilmiah Pendidikan Islam)* 3, no. 1 (2024): 101–8. <https://doi.org/10.58788/jipi.v3i1.4201>.
- Division, Council of Europe Council for Cultural Co-operation Education Committee Modern Languages. *Common European Framework of Reference for Languages: Learning, Teaching, Assessment*. Cambridge University Press, 2001.
- Gazali, Erfan, and Hasan Saefuloh. "Development of an Arabic Receptive Proficiency Test Instrument Based on the Common European Framework of Reference for Languages." *Al-Ta'rib : Jurnal Ilmiah Program Studi Pendidikan Bahasa Arab IAIN Palangka Raya* 11, no. 2 (2023): 293–308. <https://doi.org/10.23971/altarib.v11i2.6721>.
- Khairil, Laily Fazlin, and Siti Eshah Mokshein. "21st Century Assessment: Online Assessment." *International Journal of Academic Research in Business and Social Sciences* 8, no. 1 (2018): Pages 659-672. <https://doi.org/10.6007/IJARBS/v8-i1/3838>.
- Onasanya, Temitayo Omolara, Justina Ojoma Attah, Bridget Idowu Otemuyiwa, and Samuel Adenubi Onasanya. "Impact of the ClassMaker App on the Performance of Undergraduates in Online Learning of ICT: The University of Ilorin Experience." *West African Journal of Open and Flexible Learning* 9, no. 2 (2021): 55–74. <https://wajofel.org/index.php/wajofel/article/view/72>.
- Rahman, Salwa Azizah, and Hikmah Maulani. "Pembelajaran Bahasa Arab Dalam Meningkatkan Keterampilan Kalam (Berbicara) Di SMP Ma'had Ihya As-Sunnah Tasikmalaya: Metode, Tantangan, Dan Solusi." *LANGUAGE : Jurnal Inovasi Pendidikan Bahasa Dan Sastra* 4, no. 2 (2024): 2. <https://doi.org/10.51878/language.v4i2.3933>.
- Sholiha, Amila, R. Umi Baroroh, and Razita binti Abdullah. "Innovation in Reading Skills Assessment in Arabic Textbooks Based on HOTS Assessment." *Izdihar : Journal of Arabic Language Teaching, Linguistics, and Literature* 6, no. 2 (2023): 2. Arabic Textbook Grade XI by the Ministry of Religious 2020. <https://doi.org/10.22219/jiz.v6i2.27151>.
- Sudaryanto, Sudaryanto, and Pratomo Widodo. "Common European Framework of Reference for Languages (CEFR) Dan Implikasinya Bagi Buku Ajar BIPA."

Jurnal Idiomatik: Jurnal Pendidikan Bahasa Dan Sastra Indonesia 3, no. 2 (2020): 2. <https://doi.org/10.46918/idiomatik.v3i2.777>.

Tarigan, Henry Guntur. *Membaca: Sebagai Suatu Keterampilan Berbahasa*. CV. Angkasa, 2021.