Evaluating Kahoot-Based Assessment in Arabic Language Final Exam at Madrasah Aliyah

Yulyana Dewi Nabilah[1*], M. Abdul Hamid[2], Nur Hasan[3]
[1, 2] State Islamic University of Maulana Malik Ibrahim Malang, Indonesia
[3] Islamic University of Malang, Inonesia

Abstract:

The use of digital tools such as Kahoot! in Arabic language examinations opens up new opportunities to improve assessment efficiency, increase student participation, and align assessment practices with contemporary educational demands. This study examines the validity, reliability, and implementation of a Kahoot-based assessment in the Arabic final exam at Madrasah Aliyah Bilingual Batu. The aim is to evaluate the quality of the test items and understand how students respond to the use of Kahoot in a formal exam setting. Using a descriptive quantitative method, the study analyzed item validity, reliability, and student perceptions through expert review, statistical tests, and a questionnaire. The findings show that the exam instrument achieved very high validity (94%) and strong reliability (α = 0.834). The student perception questionnaire also showed high validity (91%) and good internal consistency (α = 0.845). Students reported feeling more motivated, less anxious, and more comfortable during the test when using Kahoot, suggesting that digital game-based assessments can make Arabic exams more engaging and less stressful. However, several challenges were found, including unstable internet connections and limited time for answering questions. It indicates that Kahoot need more careful calibration of time limits, infrastructural readiness, and other supplementary assessment to achieve a balanced evaluation tools. Overall, the study shows that Kahoot has strong potential to help schools adopt digital assessment.

Keywords: *Kahoot, Digital Assessment, Arabic Learning, Student Motivation, Gamification*

Abstrak:

Penggunaan alat digital seperti Kahoot! dalam ujian bahasa Arab membuka peluang baru untuk meningkatkan efisiensi penilaian, meningkatkan partisipasi siswa, dan menyelaraskan praktik penilaian dengan tuntutan pendidikan kontemporer. Penelitian ini menelaah validitas, reliabilitas, dan penerapan asesmen berbasis Kahoot dalam ujian akhir Bahasa Arab di Madrasah Aliyah Bilingual Batu. Tujuan penelitian ini adalah menilai kualitas butir soal serta memahami bagaimana respons siswa terhadap penggunaan Kahoot dalam konteks ujian formal. Menggunakan metode deskriptif kuantitatif, penelitian ini menganalisis validitas butir, reliabilitas instrumen, dan persepsi siswa melalui telaah ahli, uji statistik, dan angket. Hasil penelitian menunjukkan bahwa instrumen ujian memiliki validitas sangat tinggi (94%) dan reliabilitas kuat (α = 0.834). Angket persepsi siswa juga menunjukkan validitas tinggi (91%) dan konsistensi internal yang baik (α = 0.845). Siswa merasa lebih termotivasi, kurang cemas, dan lebih nyaman ketika mengikuti ujian menggunakan Kahoot, yang mengindikasikan bahwa asesmen digital berbasis permainan dapat membuat ujian Bahasa Arab lebih menarik dan tidak terlalu menegangkan. Namun, penelitian ini juga menemukan beberapa kendala seperti ketidakstabilan internet dan keterbatasan waktu pengerjaan. Hal tersebut menunjukkan bahwa penggunaan Kahoot memerlukan penyesuaian waktu, kesiapan infrastruktur, serta dukungan asesmen lain untuk mencapai evaluasi yang lebih seimbang. Secara keseluruhan, penelitian ini menunjukkan bahwa Kahoot memiliki potensi kuat dalam mendukung penerapan asesmen digital.

Kata Kunci: *Kahoot, Penilaian Digital, Pembelajaran Bahasa Arab, Motivasi Siswa, Gamifikasi*

Corresponding author
Email: dnabila613@gmail.com
How to Cite:
Nabilah, Y. D., Hamid, M. A., & Hasan, N. (2025). Evaluating Kahoot-Based Assessment in Arabic Language

INTRODUCTION / المقدمة

In the digital era, the use of technology in education has transformed how teaching and learning processes are designed, implemented, and evaluated. Various media and applications can now serve as easily accessible learning tools for both educators and students, supporting the achievement of educational goals and overcoming limitations in traditional classroom settings (Riwanda et al., 2021). The wide availability of digital learning platforms enables teachers to create engaging and interactive learning environments that enhance students' knowledge, motivation, and critical thinking skills (Santoso et al., 2025). Among these platforms, applications such as Quizizz, Google Forms, and Kahoot! have gained popularity for their ability to integrate gamification elements into learning and assessment (Salamiyah et al., 2025; Wang & Tahir, 2020). Therefore, the presence of such digital tools represents a new era of educational transformation that emphasizes creativity, interactivity, and learner-centered engagement.

Kahoot! is a web-based interactive platform that allows educators to design quizzes and discussions in the form of game-based learning activities. Through its user-friendly interface, teachers can create, store, and manage various learning materials and questions on their Kahoot! account page (Abidah et al., 2023). This platform functions not only as an evaluation tool but also as a medium to attract and sustain learners' attention, encouraging participation and focus through visually engaging and competitive learning experiences (Bicen & Kocakoyun, 2022; Plump & LaRosa, 2020). This feature makes Kahoot! an effective bridge between learning and assessment, transforming traditional tests into enjoyable and motivating learning experiences.

In the context of Arabic language education at Madrasah Aliyah, the use of interactive and technology-based evaluation media is increasingly important. Arabic is often perceived as a difficult subject by students due to differences in linguistic structure, orthography, and phonological systems compared to their native language (Almelhes, 2024; Firdausiyah & Jannah, 2025; Yaccob et al., 2022). Consequently, teachers need to employ innovative strategies that can reduce learning anxiety, minimize boredom, and foster students' interest and engagement in Arabic learning (Hafana & Setyabudi, 2025). The application of Kahoot! in final semester examinations offers a promising alternative, making the evaluation process more enjoyable and interactive without diminishing the accuracy of assessment results (Ismail et al., 2019; Nabila & Rustam, 2024). Consequently, integrating Kahoot! into Arabic examinations has the potential to make language evaluation more dynamic and responsive to students' learning needs.

However, the implementation of Kahoot! in formal examination contexts raises several pedagogical and technical concerns. As an assessment tool, Kahoot! must ensure the validity and reliability of the data it generates, so that student scores accurately reflect their actual competencies. If these psychometric properties are not guaranteed, the resulting data may not serve as a valid basis for academic evaluation (Taber, 2022; Eltahir et al., 2021). Furthermore, contextual challenges such as infrastructure readiness, teachers' and students' digital literacy, and internet connectivity issues can hinder the effectiveness of technology integration in madrasah (Hasyim et al., 2024; Mohtar et al.,

2023). Hence, a comprehensive understanding of these challenges is essential to ensure that digital-based assessments are implemented effectively and equitably across educational contexts.

Given these considerations, it is crucial to conduct empirical research that examines not only the psychometric aspects but also the implementation process of Kahoot! in Arabic language examinations. Such research contributes to understanding how the platform is integrated into the actual classroom context, the challenges encountered, and the pedagogical outcomes achieved. Therefore, this study aims to describe the implementation of Kahoot! in final semester Arabic examinations at Madrasah Aliyah and to analyze the validity of its use as a digital evaluation medium. The findings of this article are expected to support the development of an effective, credible, and pedagogically sound model of technology-based assessment for Arabic language learning in the digital era.

Furthermore, this study argues that Kahoot!-based assessments can serve not only as an engaging evaluation tool but also as a catalyst for enhancing students' digital literacy, self-regulated learning, and motivation in Arabic language acquisition. Unlike previous studies that primarily focus on classroom engagement or general gamification effects, this research provides empirical evidence on the application of Kahoot! in high-stakes assessment contexts, specifically final semester examinations. The original contribution of this study lies in its integrated approach, combining analysis of implementation procedures, psychometric validity, and pedagogical outcomes, thereby offering a comprehensive model for technology-enhanced assessment in Islamic secondary education. The insights gained are expected to guide educators and policymakers in adopting digital assessment tools that maintain academic rigor while promoting interactive and student-centered learning experiences.

RESEARCH METHOD / المنهجية

This study employed a descriptive quantitative design, a method that seeks to describe existing data and phenomena systematically and objectively without manipulating the research variables (Creswell & Creswell, 2018; Waruwu, 2023). This approach was chosen because it enables the researcher to provide a factual and comprehensive overview of how Kahoot! was implemented in the final semester Arabic examination and to analyze the validity of its application as an evaluation medium at Madrasah Aliyah. Descriptive quantitative designs are widely applied in educational technology research to obtain clear and structured insights into implementation processes (Alieto et al., 2024; Slater & Hasson, 2025). So, this design choice ensures that the study captures the real conditions of Kahoot implementation in the examination context.

The population of this study consisted of 10th grade students at Madrasah Aliyah Bilingual Batu. The sample size is 21 students. A purposive sampling technique was used, selecting the participants based on their specific criteria aligned with the objectives of this study (Sugiyono, 2019), which is their involvement in Arabic final exam using the Kahoot! application. This technique is commonly used in educational research when the sample must represent individuals who meet specific requirements relevant to the research focus (Patton, 2002; Etikan et al., 2015). The research object was the implementation of Kahoot! in the examination process and the validity evaluation of the instrument used.

Data were collected using two instruments: student perception questionnaires and Arabic language final exam delivered via the Kahoot! platform. The instrument consisted of

15 items, covering three indicators: 1) Ease of use. 2) Student engagement and motivation. 3) Presentation and display quality of the Kahoot application (Rojabi et al., 2022; Rusliana et al., 2024). All items used a 4-point Likert scale (1 = Strongly Disagree, 4 = Strongly Agree). The questionnaire items were adapted and refined from previous research on gamification and digital assessment and were validated by experts in language teaching, educational technology, and assessment. In addition to the questionnaire, the Arabic language final exam contained forty multiple-choice questions on theme "التحيات والتعارف" (Greetings and Introduction) and "الأسرة والبيت" (Family and House). It's designed to evaluate students' reading, listening, and basic writing competencies in the final semester examination. The test was reviewed and validated by content specialists to confirm alignment with the Arabic curriculum and suitability for 10ᵗʰ grade students.

The data analysis procedure was conducted in two stages. The first stage focused on assessing the quality of the instruments through validity and reliability testing. The questionnaire and exam items were examined using expert judgment. This expert validation ensured that the instruments accurately reflected the constructs being measured and were suitable for use with eleventh-grade students. Following the expert review, instrument reliability was tested using Cronbach's Alpha, with a coefficient above 0.60 interpreted as indicating acceptable internal consistency. Reliability analysis was conducted using SPSS to maintain accuracy and strengthen the credibility of the measurement process. The second stage of data analysis involved descriptive statistical techniques to interpret the questionnaire responses and student exam scores. Through this process, the study presented a clear and systematic overview of how Kahoot was utilised in the assessment context and how students responded to its use as a digital evaluation tool.

FINDINGS AND DISCUSSION / نتائج البحث و المناقشة
Findings

This section presents the findings of the study on Kahoot-based assessment in the Arabic language final exam at Madrasah Aliyah. It covers the validity and reliability of both the exam items and student perception questionnaire, followed by an analysis of the implementation process. The results highlight students' performance, engagement, and experiences with the gamified assessment format, as well as technical and contextual challenges encountered during its use.

Validity and Reliability of Kahoot-Based Assessment in Arabic Language Final Exam at Madrasah Aliyah

The validity of the 40 items in the Arabic final exam was assessed through an expert judgment approach. Each exam item was reviewed using a five level rating scale, in which 1 indicates very poor, 2 (poor), 3 (adequate), 4 (good), and 5 indicates very good. The validity level for each item was calculated using the formula:

$$V = \frac{Tse}{Tsh} \times 100\%$$

V = validity persentage
Tse = empirical score assigned by experts
Tsh = maximum expected score

This calculation produces a percentage that is then classified into validity levels such as very good, good, or adequate, thereby showing how closely each item aligns with the intended validity score and level of assessment. To interpret the resulting percentages, the levels of item validity were categorized as follows:

Table.1. Validation Level and Interpretation

Percentage	Validity Level	Interpretation
81% - 100%	Very Valid	No revision required
61% - 80%	Valid	No revision required
41% - 60%	Quite Valid	No revision required
21% - 40%	Less Valid	Partial revision required
0% - 20%	Invalid	Full revision required

According to Table 1, any instrument with a validity percentage exceeding 60% is considered valid and does not require any revisions, meaning it is suitable for use in assessments. In contrast, instruments with a validity percentage below 60% are classified as invalid and must be revised to ensure alignment with learning objectives and assessment standards. This threshold serves as a guideline to guarantee that the items included in tests or questionnaires meet acceptable quality criteria, accurately measure the intended constructs, and provide reliable evaluation results.

The expert judgment showed the majority of the exam items demonstrated a very high level of validity. Most items obtained the maximum score from expert validators, resulting in validity percentages of 100% and falling into the "Very Good" scale. This suggests that the content, structure, and alignment of these items are highly consistent with the learning objectives of the Arabic subject. Even so, a smaller proportion of items showed slightly lower validity levels. Several items such as numbers 1, 5, 12, 17, 26, 31, 33, and 37 achieved a validity level of 80%, categorized as "Good." These items are still acceptable for use but may require refinements in wording clarity or alignment with indicators to reach optimal quality. Two items, specifically items number 23 and 25, obtained a validity percentage of 60%, falling under the category of "Adequate." The results of expert validation can be seen in the following table:

Table.2. Results of Expert Validation for Arabic Final Exam Items

No.	Tse	Tsh	Percentage	Validity Level
1	4	5	80%	Good
2	5	5	100%	Very Good
3	5	5	100%	Very Good
4	5	5	100%	Very Good
5	4	5	80%	Good
6	5	5	100%	Very Good
7	5	5	100%	Very Good
8	5	5	100%	Very Good
9	5	5	100%	Very Good
10	5	5	100%	Very Good
11	5	5	100%	Very Good

12	4	5	80%	Good
13	5	5	100%	Very Good
14	5	5	100%	Very Good
15	5	5	100%	Very Good
16	5	5	100%	Very Good
17	4	5	80%	Good
18	5	5	100%	Very Good
19	5	5	100%	Very Good
20	5	5	100%	Very Good
21	5	5	100%	Very Good
22	5	5	100%	Very Good
23	3	5	60%	Adequate
24	5	5	100%	Very Good
25	3	5	60%	Adequate
26	4	5	80%	Good
27	5	5	100%	Very Good
28	5	5	100%	Very Good
29	5	5	100%	Very Good
30	5	5	100%	Very Good
31	4	5	80%	Good
32	5	5	100%	Very Good
33	4	5	80%	Good
34	5	5	100%	Very Good
35	5	5	100%	Very Good
36	5	5	100%	Very Good
37	4	5	80%	Good
38	5	5	100%	Very Good
39	5	5	100%	Very Good
40	5	5	100%	Very Good
Total	**188**	**200**	**94%**	**Very Valid**

Based on Table 2, the amount score across all 40 items reached a total of 188 out of 200, resulting in an overall validity percentage of 94%. According to the criteria, this percentage places the exam items within the very valid category. This means that the Arabic final exam items does not require revision and valid for evaluating students' proficiency in Arabic at the end of the semester.

Following the validation of the 40 exam items, the student perception questionnaire was also evaluated using expert judgment. The questionnaire consisted of 15 items focusing on three core dimensions related to students' experience with Kahoot during the Arabic final examination: ease of use, student engagement and motivation, and the presentation quality of the Kahoot application. Each item was reviewed using the same five-level rating scale applied to the exam items, ranging from 1 (very poor) to 5 (very good). The validity percentage of each questionnaire item was calculated using the formula previously described, where the empirical score (Tse) was compared with the maximum possible score (Tsh) to calculate the validity percentage for each item. The validity classification also followed the same criteria used for the exam items. As previously

described at Table 1, percentage above 60% are valid and acceptable for use without revision. The table below is the result of expert validation for student questionnaire:

Table.3. Results of Expert Validation for Student Questionnaire

No.	Tse	Tsh	Percentage	Validity Level
1	5	5	100%	Very Good
2	4	5	80%	Good
3	5	5	100%	Very Good
4	5	5	100%	Very Good
5	4	5	100%	Very Good
6	5	5	100%	Very Good
7	5	5	100%	Very Good
8	4	5	80%	Good
9	4	5	80%	Good
10	5	5	100%	Very Good
11	5	5	100%	Very Good
12	4	5	100%	Very Good
13	3	5	60%	Adequate
14	5	5	100%	Very Good
15	5	5	100%	Very Good
Total	68	75	91%	Very Valid

Based on Table 1 and Table 3, the expert judgment analyzes that most questionnaire items fall within the very good (81–100%) validity level. From 15 items, 11 items interpretate in very good and 3 items in good levels. 1 item adequate. Only one item (Item 13) received a validity percentage of 60%, categorized as adequate. A total score of 68 out of 75 yields an overall validity percentage of 91%, placing the questionnaire in the very valid category. This means the student questionnaire is valid for use without revision required.

Following the validity analysis, a reliability test was carried out to determine the internal consistency of the 40-item Arabic final exam. Reliability was measured using Cronbach's Alpha, which is commonly used to assess the degree to which items within an instrument consistently measure the same construct. The reliability coefficient was computed using SPSS, producing a Cronbach's Alpha value that reflects the consistency of the exam items. A higher alpha value indicates stronger internal consistency, meaning that the items work together consistently to evaluate the intended skill or ability.

Before interpreting the reliability value, the classification criteria for Cronbach's Alpha must be established. The interpretation guide typically categorizes alpha values into levels ranging from unacceptable to excellent. The criteria used for this study are presented in the following table:

Table.4. Interpretation of Reliability Coefficients

Reliability Coefficient	Reliability Level
$r_{11} \leq 0,20$	Very Low

$0{,}20 < r_{11} 0{,}40$	Low
$0{,}40 < r_{11} 0{,}70$	Adequate
$0{,}70 < r_{11} 0{,}90$	High
$0{,}90 < r_{11} 1{,}00$	Very High

The criteria in Table 4 serve as the basis for determining whether the exam items demonstrate sufficient consistency. An alpha value above 0.40 indicates adequate reliability, while values above 0.70 reflects high reliability, and reliability coefficient of 0.90 means the exam items in the very high reliability level. Reliability coefficients falling within the high or very high category indicate that the exam items are stable and measure the intended construct consistently. Coefficients in the adequate or low categories suggest that some items may not function properly and require revision or replacement. This interpretative framework ensures that the reliability outcome aligns with recognized psychometric standards. The SPSS output for reliability of Arabic final exam item is shown in the table below:

Table.5.

Reliability Statistics

Cronbach's Alpha	N of Items
.834	40

Table 5 showed the result of reliability test produced by Cronbach's Alpha. The Alpha value is 0.834, which exceeds the commonly accepted threshold of 0.70 for high reliability. According to Table 4, values between 0.70 and 0.90 fall into the category of high reliability. Therefore, the obtained reliability coefficient of 0.834 places the exam in the highly reliable category. This suggests that the exam items measure the intended construct consistently and can be considered reliable for evaluating students' achievement in Arabic.

After testing the reliability of 40 item Arabic final exam, the reliability of the student perception questionnaire was also examined to determine whether its items consistently measured students' responses toward the use of Kahoot during the examination. Similar to the exam reliability test, the questionnaire reliability was analyzed using Cronbach's Alpha. The interpretation of reliability coefficients presented previously in Table 4, where values between 0.70 and 0.90 are categorized as high reliability level, and values exceeding 0.90 fall into the very high reliability. The SPSS output for the student questionnaire reliability result is presented in the following table:

Table.6.

Reliability Statistics

Cronbach's Alpha	N of Items
.845	15

Table 6 shows that the Cronbach's Alpha coefficient for the student questionnaire is 0.845. Based on the criteria presented in Table 3, any value exceeding the threshold of 0.70 indicates a high level of reliability. This result demonstrates that the items in the questionnaire consistently measure the intended constructs. Specifically, it confirms that

the questionnaire is reliable for assessing students' perceptions of the Kahoot application, including its ease of use, the degree of student engagement and motivation it generates, and the quality of its visual and interactive presentation. Therefore, the instrument can be confidently used in evaluating students' experiences with Kahoot.

The Implementation of Kahoot-Based Assessment in Arabic Language Final Exam at Madrasah Aliyah

The implementation of the Kahoot-based assessment in the Arabic language final examination based from a needs analysis conducted at Madrasah Aliyah Bilingual Batu. Observations revealed that teachers had limited access to effective digital assessment tools, and the existing school testing application could not process written Arabic, resulting in exclusive reliance on multiple-choice questions. This constraint produced imbalances in cognitive-level distribution across test items and contributed to student boredom, which was often accompanied by off-task behaviours such as looking at other students' answers. Manual scoring further increased the workload and delayed feedback. These contextual issues highlighted the necessity for a more efficient, engaging, and technologically supported assessment format. Against this background, the Arabic final examination was designed and delivered through Kahoot using forty multiple-choice questions across two thematic units: "التحيات والتعارف" (Greetings and Introduction) and "الأسرة والبيت" (Family and House). Both themes reflected the core competencies required at the end of the semester. These items assessed reading comprehension, listening comprehension, and recognition of basic writing forms.

A total of 21 students accessed the examination through kahoot.it using mobile phones. Throughout the implementation, student participation was active and consistent. Across the forty-item test, the number of correct answers per student ranged from 10 to 34, showing substantial variation in performance. The mean score of 24.71 indicates that students answered more than half of the items correctly. Questions 1, 3, 7, and 11 had high accuracy levels with a majority of students responding correctly. These items likely assessed foundational vocabulary or skills that had been reinforced more frequently during instruction. In contrast, items such as questions 19, 22, and especially question 24 answered correctly by only 6 students, indicated areas of difficulty.

The platform's interactive nature, visual appeal, and immediate feedback contributed to sustained attention. This pattern aligns with quantitative findings from the student questionnaire, in which the highest scores appeared on items 9, 13, and 14. These number received the most favourable responses, indicating that the visual design and structural variation afforded by Kahoot enhanced student motivation and reduced exam-related fatigue. Students also widely perceived the platform as user-friendly, readable, and aesthetically engaging. However, the implementation also encountered issues reflected in lower-scoring questionnaire items. Questions related to time allocation (item 2) and technical constraints (item 7) scored the lowest. Several students reported that the time limits felt too short for careful reading, and some experienced disruptions associated with internet connectivity or device performance.

The qualitative comments provided by students further enriched the questionnaire result. Thematic analysis revealed four prominent themes. First, many expressed enjoyment and increased motivation, noting that Kahoot created a more relaxed and pleasant exam atmosphere. For example, "Ujian jadi lebih santai dan menyenangkan." Second, recurring comments described technical obstacles, especially unstable

connections: "Jaringan sering putus jadi nggak sempat jawab." Third, students frequently addressed time management difficulties, stating that the countdown was too quick: "Waktunya terlalu cepat, kadang belum selesai baca soal." Fourth, students appreciated Kahoot's usability, particularly its readable design and straightforward navigation: "Tampilannya menarik dan mudah dipakai." Finally, some students expressed a desire for Kahoot to be used more frequently for both practice and formal tests, recognizing its motivational benefits: "Semoga bisa dipakai waktu belajar juga."

Generally, students indicated that Kahoot made the examination more interesting, enjoyable, and less monotonous, with many agreeing that the gamified format increased their motivation during the test. Statements related to the visual appeal of Kahoot, such as color layout, clarity of options, and user-friendly interface, received consistently high levels of agreement. These findings suggest that the digital game-based environment as Kahoot successfully supported students' emotional engagement and reduced test-related anxiety. However, some students expressed that the allotted answering time felt too short, making it difficult for them to fully process each question, especially when the Arabic text required careful reading. Some respondents also reported technical issues, such as unstable internet connection or device-related delays, which occasionally disrupted their performance.

Despite the challenges mentioned, the majority of students agreed that Kahoot is an effective assessment tool, with many stating their preference for its continued use in future examinations or classroom activities. Students also acknowledged that Kahoot encouraged them to remain attentive throughout the assessment and made the testing experience more enjoyable overall. In summary, the findings indicate that students' attitudes toward Kahoot-based assessment were predominantly positive, especially regarding engagement, enjoyment, and usability. At the same time, areas for improvement remain, mainly focused on connection stability, device readiness, and allocation of answering time.

Discussion

The findings of this study demonstrate that the implementation of Kahoot-based assessment in the Arabic final examination at Madrasah Aliyah Bilingual Batu provides substantial pedagogical benefits while simultaneously presenting several limitations that merit further consideration. The results of the validity and reliability analysis in this study indicate that the Kahoot-based Arabic final examination has supporting its use as a credible and systematic digital assessment tool (Altakhaineh, 2025; Mouaziz, 2025; Nurbayan, 2025). The very high validity percentage obtained by expert judgment for the exam items (94%) demonstrates that the content, linguistic structure, and cognitive alignment of the question items are consistent with the intended learning objectives. It is consistent with a theory that well-constructed multiple-choice items can achieve strong content validity when subjected to expert review (Haladyna et al., 2002). Similar results also showed that digital game-based assessment platforms such as Kahoot can maintain item-quality equivalence with traditional assessment methods when validation stages are appropriately executed (Ahmed et al., 2022), reinforcing the reliability of adopting digital game-based assessment platforms.

The reliability analysis further strengthens this conclusion. The exam's high reliability coefficient ($\alpha = 0.834$) indicates strong internal consistency, meaning that the items collectively measure the intended constructs with stability across student

responses. According to previous studies, an alpha value above 0.80 indicates that the set of items measures a shared construct consistently across participants (Alharbi et al., 2020; Alsalman et al., 2020; Alshammari, 2020). This high level of consistency aligns with research revealing that Kahoot-based assessments often demonstrate robust reliability due to their standardized scoring, controlled administration, and reduction of human scoring errors (Mohammad et al., 2022). Furthermore, the automated scoring mechanism in Kahoot minimizes the possibility of clerical inaccuracies, thereby enhancing precision and fairness in the evaluation process (Ismail et al., 2020). Taken together, these findings confirm that the exam instrument is both methodologically sound and pedagogically scalable.

In addition to exam items, the questionnaire obtained a very high validity percentage of 91%, indicating that the items were well aligned with the constructs of usability, engagement, and presentation quality measured in this study. A validity level above 90% is considered strong evidence that questionnaire items accurately capture the intended latent traits and function consistently across respondents, especially when supported by expert judgment (Alfahid et al., 2021; Hejaili, 2023; Yamani et al., 2024). These outcomes align with previous studies showing that well-developed perception questionnaires in technology-enhanced learning environments often exhibit high validity and reliability when they are derived from established constructs such as ease of use, motivation, and perceived usefulness (Arribathi et al., 2024; Parvez et al., 2023). Thus, the questionnaire's strong validity and reliability further strengthen the robustness of the conclusions drawn regarding students' positive perceptions of Kahoot and its pedagogical impact.

Furthermore, the high reliability of the student questionnaire ($\alpha = 0.845$) confirms that the instrument effectively captures coherent constructs related to students' perceptions of ease of use, engagement, and interface quality. Questionnaires measuring affective responses toward digital tools often face risks of inconsistency due to subjective interpretation or response bias (Pallant, 2020). Yet, the strong reliability indicator in this study implies that student perceptions toward Kahoot were stable across item forms and not influenced by random variability. These results support the argument that well-designed Likert-scale questionnaires can reliably capture learner affect and technology acceptance levels, even in contexts where digital literacy and technological readiness vary, such as school environments.

The high levels of validity and reliability obtained from both the exam items and student questionnaire indicate that Kahoot can function as an effective digital assessment tool within the context of Arabic language learning. These outcomes align with previous studies that reported Kahoot-based assessment generally yields high reliability and strong construct validity in language learning contexts (Özdemir, 2025; Rahim et al., 2025). Such consistency supports the Kahoot's appropriateness as part of a structured evaluation system, especially when combined with pedagogical planning and proper item construction procedures.

Kahoot's positive impact on students' motivation and engagement, as reflected in questionnaire responses and qualitative comments, is consistent with broader gamification research. Gamified assessment environments have been shown to enhance learners' enjoyment, persistence, and attentional focus by integrating competitive and interactive elements that stimulate emotional involvement (Almelhes, 2024; Dhika et al., 2023; Ismail et al., 2023). In the context of language learning, Kahoot has demonstrated

particular effectiveness in reducing anxiety and fostering a positive emotional climate that supports comprehension performance (Husin & Azmuddin, 2022). The student feedback in this study highlighting the platform's visual appeal, ease of use, and enjoyable atmosphere. It reinforces the notion that technology-supported game-based assessments provide motivational value often absent in conventional exam formats. This suggests that integrating Kahoot may help cultivate more favorable attitudes toward Arabic learning, which is essential given the linguistic complexity of the subject.

In the context of Arabic language education, these findings hold particular relevance. Students commonly perceive Arabic as a challenging subject due to its orthographic density, phonological features, and rich morphological system complexity (Alqahtani & Kamhi, 2023; Rahmayana & Halim, 2024). Because of these challenges, an emotionally supportive assessment environment becomes essential to maintain learner confidence and minimize affective barriers. Digital gamified tools such as Kahoot appear to offer such support by reducing exam-related tension, encouraging playful engagement, and promoting a less intimidating testing environment. These outcomes align with studies showing that reduced anxiety can facilitate better comprehension, memory retention, and overall performance among learners of Arabic as a foreign language (Almelhes, 2024). Therefore, the positive student responses observed in this study further validate the pedagogical value of integrating interactive digital platforms, particularly in settings where rote memorization and traditional paper-based assessments have historically dominated.

Despite these advantages, several practical challenges emerged during the implementation of the Kahoot-based examination. Students reported issues related to internet instability and insufficient time allocation for certain items. These difficulties mirror findings from previous studies, which identify technological infrastructure and time pressure as major barriers to the successful use of Kahoot in formal assessment contexts (Ismail et al., 2019). Time constraints may be especially problematic for learners processing Arabic text, as psycholinguistic research indicates that L2 readers of Arabic typically require longer processing time due to visual complexity, diacritics, and morphological structure (Abu-Rabia et al., 2022; Verhoeven et al., 2019). While Kahoot's fast-paced interface can heighten engagement, it may unintentionally introduce cognitive load that disadvantages learners who require more time to decode written Arabic, thereby affecting the accuracy of performance results.

Another notable limitation relates to the exclusive use of multiple-choice items, which stems from institutional constraints in processing written Arabic in digital format. This structure limits the assessment's ability to measure productive skills such as writing, structured responses, and communicative expression (Al-Razgan & Alotaibi, 2022; ElSayed et al., 2023). Research in language assessment emphasizes that a comprehensive evaluation of language proficiency requires a balanced approach that includes both receptive and productive competencies, especially within communicicative language teaching frameworks (Bravo Pineda & Baño Sánchez, 2024). Although Kahoot is highly effective for testing receptive skills, such as vocabulary recognition and reading comprehension, the platform alone cannot replace more complex forms of performance-based assessment. This limitation indicates that Kahoot is best positioned as a complementary tool rather than a substitute for traditional evaluation methods.

Overall, the results of this study support the growing consensus in the literature that Kahoot-based assessments can improve affective, motivational, and cognitive dimensions

of learning. At the same time, they highlight the need for careful calibration of time limits, infrastructural readiness, and supplementary assessment tools to achieve a balanced evaluation framework. Future research should explore hybrid assessment models that integrate Kahoot with performance-based tasks, examine long-term impacts on Arabic language proficiency, and evaluate the platform's effectiveness across diverse Islamic educational institutions.

CONCLUSION / الخلاصة

This study demonstrates that implementing Kahoot as an assessment tool for the Arabic language final exam at Madrasah Aliyah (Islamic Senior High School) provides significant pedagogical benefits. Instrument validity and reliability indicate that the exam questions and student perception questionnaire are of excellent quality, with validity percentages of 94% and 91%, respectively, and Cronbach's Alpha values of 0.834 and 0.845. Key findings indicate that Kahoot not only increases student engagement, motivation, and learning experience but also helps reduce test anxiety, making the testing environment more enjoyable and interactive. Furthermore, student scores demonstrate a wide variation in ability, indicating the platform's ability to differentiate levels of material mastery. A key lesson from this study is that the use of game-based digital media can be an effective and enjoyable alternative assessment tool, especially in complex subjects like Arabic, while maintaining the accuracy and reliability of the assessment.

The primary scientific contribution of this study lies in presenting an integrative model of Kahoot in a formal evaluation context, which combines analysis of validity, reliability, pedagogical implementation, and student perceptions. This study adds to the literature on gamification in education, demonstrating that Kahoot can enhance cognitive and affective engagement while supporting students' digital literacy development. However, the study has limitations, including technical constraints such as unstable internet connections, limited time to complete the questions, and the use of multiple-choice questions, which limited the measurement of students' productive abilities. Future research is recommended to explore hybrid assessment models that combine Kahoot with performance-based tasks, as well as examine their long-term impact on Arabic language acquisition and effectiveness across various madrasah and other Islamic educational contexts.

REFERENCES / المراجع

Abidah, Z., Rohman, M. F., & Rahmadian, Y. (2023). The Effect of Implementing Kahoot Application-Based E-Learning Media on Arabic Writing Skills. *Edu Journal Innovation in Learning and Education*, *1*(2), 162–179. https://doi.org/10.55352/edu.v1i2.773

Ahmed, A. A. A., Sayed, B. T., Wekke, I. S., Widodo, M., Rostikawati, D., Ali, M. H., Abdul Hussein, H. A., & Azizian, M. (2022). An Empirical Study on the Effects of Using Kahoot as a Game-Based Learning Tool on EFL Learners' Vocabulary Recall and Retention. *Education Research International*, *2022*(1), 9739147. https://doi.org/10.1155/2022/9739147

Alfahid, A., Bitar, H., Alrige, M., Abeeri, H., & Sulami, E. (2021). DoItRight: An Arabic Gamified Mobile Application to Raise Awareness about the Effect of Littering among Children. *International Journal of Advanced Computer Science and Applications*, *12*(12), 151–157. https://doi.org/10.14569/IJACSA.2021.0121220

Alharbi, A., Aloufi, S., Assar, R., & Meccawy, M. (2020). Virtual Reality Street-Crossing

Training for Children with Autism in Arabic Language. In *2020 International Conference on Innovation and Intelligence for Informatics, Computing and Technologies, 3ICT 2020*. https://doi.org/10.1109/3ICT51146.2020.9311981

Alieto, E., Abequibel-Encarnacion, B., Estigoy, E., Balasa, K., Eijansantos, A., & Torres-Toukoumidis, A. (2024). Teaching Inside a Digital Classroom: A Quantitative Analysis of Attitude, Technological Competence and Access Among Teachers Across Subject Disciplines. *Heliyon*, *10*(2), e24282. https://doi.org/10.1016/j.heliyon.2024.e24282

Almelhes, S. A. (2024). Gamification for Teaching the Arabic Language to Non-Native Speakers: A Systematic Literature Review. *Frontiers in Education*, *9*. https://doi.org/10.3389/feduc.2024.1371955

Alqahtani, R. B., & Kamhi, A. G. (2023). Do Reading Teachers Have Sufficient Phonological and Morphological Knowledge of Arabic? *Learning Disabilities*.

Al-Razgan, M., & Alotaibi, H. M. (2022). Integrating Mobile Games in Arabic Orthography Classrooms. *Arab World English Journal*, *2022-July*(Special Issue), 146–165. https://doi.org/10.24093/awej/call8.10

Alsalman, D., Ali, Z. M. B., Alnosaier, Z. F., Alotaibi, N. A., & Alanzi, T. M. (2020). Gamification for Diabetes Type 1 Management: A Review of the Features of Free Apps in Google Play and App Stores. In *Journal of Multidisciplinary Healthcare* (Vol. 13, pp. 425–432). https://doi.org/10.2147/JMDH.S249664

Alshammari, M. T. (2020). Evaluation of Gamification in E-Learning Systems for Elementary School Students. *TEM Journal*, *9*(2), 806–813. https://doi.org/10.18421/TEM92-51

Altakhaineh, A. R. M., & Sulaiman, N. A. (2025). Teaching English Vocabulary to Arabic-Speaking EFL Learners Through Total Physical Response and Kahoot. *International Journal of Technologies in Learning*, *32*(2), 169–182. https://doi.org/10.18848/2327-0144/CGP/v32i02/169-182

Arribathi, A. H., Awalia Rahmawati, F., Suhada, Supriyono, I. A., Anggraini Santoso, N., & Audiah, S. (2024). Evaluating Gamified Platform for Enhancing Arabic Language Proficiency Using TAM. In *2024 3rd International Conference on Creative Communication and Innovative Technology, ICCIT 2024*. https://doi.org/10.1109/ICCIT62134.2024.10701098

Bicen, H., & Kocakoyun, S. (2022). Perceptions of Students for Gamification Approach: Kahoot as a Case Study. *International Journal of Emerging Technologies in Learning (iJET)*, *13*(02), Article 02. https://doi.org/10.3991/ijet.v13i02.7467

Creswell, J. W., & Creswell, J. D. (2018). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches* (5th ed.). Sage.

Deterding, S., Dixon, D., Khaled, R., & Nacke, L. (2021). From Game Design Elements to Gamefulness: Defining "Gamification." *Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments*, 9–15. https://doi.org/10.1145/2181037.2181040

Dhika, M. A., Khairani, D., Masruroh, S. U., Fiade, A., Arifin, V., & Tsaqofi, W. A. (2023). Comparing GraphQL and ReST Architecture in Arabic Learning Games: A Quality of Service (QoS) Approach. In *2023 11th International Conference on Cyber and IT Service Management, CITSM 2023*. https://doi.org/10.1109/CITSM60085.2023.10455108

ElSayed, H., Aldegaither, S., AlArifi, N., Monawar, A., & Hamidalddin, A. (2023). LexiaQuest:

Exploring the Feasibility of a NLP-Based Screening Tool for Dyslexia Using Virtual Reality. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics): Vol. 14026 LNCS* (pp. 204–216). https://doi.org/10.1007/978-3-031-35927-9_15

Eltahir, M. E., Alsalhi, N. R., Al-Qatawneh, S., AlQudah, H. A., & Jaradat, M. (2021). The impact of Game-Based Learning (GBL) on Students' Motivation, Engagement and Academic Performance on an Arabic Language Grammar Course in Higher Education. *Education and Information Technologies*, *26*(3), 3251–3278. https://doi.org/10.1007/s10639-020-10396-w

Etikan, I., Musa, S. A., & Alkassim, R. S. (2022). Comparison of Convenience Sampling and Purposive Sampling. *American Journal of Theoretical and Applied Statistics*, *5*(1), 1–4. https://doi.org/10.11648/j.ajtas.20160501.11

Firdausiyah, A., & Jannah, I. M. (2025). Analysis of Students' Problems in Writing Arabic Sentences. *Al-Lahjah : Jurnal Pendidikan, Bahasa Arab, Dan Kajian Linguistik Arab*, *8*(1), 854–861. https://doi.org/10.32764/lahjah.v8i1.5159

Glasow, P. (2005). *Fundamentals of Survey Research Methodology*.

Hafana, N., & Setyabudi, M. A. (2025). Implementation of Kahoot! in Arabic Language Learning at Muhammadiyah 1 High School, Nganjuk. *YASIN*, *5*(4), 3664–3676. https://doi.org/10.58578/yasin.v5i4.6283

Haladyna, T. M., Downing, S. M., & Rodriguez, M. C. (2022). A Review of Multiple-Choice Item-Writing Guidelines for Classroom Assessment. *Applied Measurement in Education*, *15*(3), 309–334. https://doi.org/10.1207/S15324818AME1503_5

Hasyim, N., Arismunandar, Butarbutar, R., Ramli, A. M., & Nur, I. D. M. (2024). Mind Mapping of Teachers' Readiness for Online Teaching and Learning: A Reflective Study of Urban and Suburban Areas. *Cogent Education*, *11*(1). https://doi.org/10.1080/2331186X.2023.2292864

Hejaili, A. Al, & Newbury, P. (2023). LAA: Learn the Arabic Alphabet: Integrating Gamification Elements with Touchscreen Based Application to Enhance the Understanding of the Arabic Letters Forms. *Electronic Journal of E-Learning*, *21*(4), 353–365. https://doi.org/10.34190/ejel.21.4.3043

Husin, M. Z. M., & Azmuddin, R. A. (2022). Learner Engagement in Using Kahoot! Within a University English Proficiency Course. *Educational Process: International Journal*.

Ismail, M. A. A., Ahmad, A., Fakri, N. M. R. M., Nasir, M. K. M., & Rahim, A. A. A. (2022). Using Kahoot! As a Formative Assessment Tool in Medical Education: A Phenomenological Study. *BMC Medical Education*, *19*(1), 230. https://doi.org/10.1186/s12909-019-1658-z

Ismail, U. S., Makhtar, N. I., Chulan, M., & Ismail, N. (2023). A Model Framework for the Implementation of Gamification in Arabic Teaching in Malaysia. *Theory and Practice in Language Studies*, *13*(11), 2800–2805. https://doi.org/10.17507/tpls.1311.09

Mohammad, H., Tamimi, H., & Abuamara, F. (2022). An Educational Arabic Sign Language Mobile Application for Children with Hearing Impairment. *International Journal of Interactive Mobile Technologies*, *16*(20), 114–129.

https://doi.org/10.3991/ijim.v16i20.32427

Mohtar, S., Jomhari, N., Omar, N. A., Mustafa, M. B. P., & Yusoff, Z. M. (2023). The Usability Evaluation on Mobile Learning Apps with Gamification for Middle-Aged Women. *Education and Information Technologies*, *28*(1), 1189–1210. https://doi.org/10.1007/s10639-022-11232-z

Mouaziz, A. (2025). Investigating English and Arabic Teachers' and Students' Attitudes Towards the Use of Humanistic-Based Teaching Methods and Their Impact on Language Skills Development. In *Colloquium in Information Science and Technology Cist* (pp. 421–426). https://doi.org/10.1109/cist65886.2025.11224084

Nabila, A., & Rustam. (2024). Utilization of the Kahoot Application as an Assessment of Indonesian Language Learning in High Schools. *Didaktika: Jurnal Kependidikan*, *13*(1), 897–906. https://doi.org/10.58230/27454312.1363

Nurbayan, Y., & Sanusi, A. (2025). Integrating Character Values and EdTech Media: A Study of Arabic Teachers' Pedagogical Competence in Bandung Raya. *International Journal of Learning, Teaching and Educational Research*, *24*(8), 248–271. https://doi.org/10.26803/ijlter.24.8.11

Özdemir, O. (2025). Kahoot! Game-based Digital Learning Platform: A Comprehensive Meta-analysis. *Journal of Computer Assisted Learning*, *41*(1), e13084. https://doi.org/10.1111/jcal.13084

Pallant, J. (2020). *SPSS Survival Manual: A Step by Step Guide to Data Analysis using IBM SPSS* (7th ed.). Routledge. https://doi.org/10.4324/9781003117452

Parvez, M. T., Alsuhibani, A. M., & Alamri, A. H. (2023). Educational and Cybersecurity Applications of an Arabic CAPTCHA Gamification System. *Ingenierie Des Systemes d'Information*, *28*(5), 1275–1285. https://doi.org/10.18280/isi.280516

Patton, M. Q. (2002). *Qualitative Research & Evaluation Methods*. SAGE Publications. https://books.google.co.id/books?id=FjBw2oi8El4C

Plump, C. M., & LaRosa, J. (2022). Using Kahoot! in the Classroom to Create Engagement and Active Learning: A Game-Based Technology Solution for eLearning Novices. *Management Teaching Review*, *2*(2), 151–158. https://doi.org/10.1177/2379298116689783

Rahim, M., Mohammed, L. A., & Batool, S. (2025). The Effect of Kahoot as a Gamification-Based Assessment Tool on Primary Students' Academic Achievement in Mathematics. *International Journal of Technology in Education*, *8*(4), 977–997. https://doi.org/10.46328/ijte.1220

Rahmayana, L., & Halim, A. (2024). The Impact of Kahoot-Based Formative Assessment on Student Learning Outcomes at a Junior High School in Samarinda. *EnJourMe (English Journal of Merdeka) : Culture, Language, and Teaching of English*, *9*(1), 103–118. https://doi.org/10.26905/enjourme.v9i1.13150

Riwanda, A., Ridha, M., & Islamy, M. I. (2021). Increasing Arabic Vocabulary Mastery Through Gamification; is Kahoot! Effective? *LISANIA: Journal of Arabic Education and Literature*, *5*(1), 19–35. https://doi.org/10.18326/lisania.v5i1.19-35

Rojabi, A. R., Setiawan, S., Munir, A., Purwati, O., Safriyani, R., Hayuningtyas, N., Khodijah, S., & Amumpuni, R. S. (2022). Kahoot, is it Fun or Unfun? Gamifying Vocabulary Learning to Boost Exam Scores, Engagement, and Motivation. *Frontiers in Education*, *7*. https://doi.org/10.3389/feduc.2022.939884

Rusliana, N. A., Sufyadi, S., & Qomario, Q. (2024). Kahoot Utilization! To Support Game-

Based Learning. *Jurnal Indonesia Sosial Teknologi*, *5*(10), 4286–4297. https://doi.org/10.59141/jist.v5i10.7021

Salamiyah, A., Wangi, N. B. S., & Rizqi, M. R. (2025). The Effectiveness of Using Kahoott as a Media for Evaluating Arabic Language Learning Outcomes for First Grade Students of Madrasah Aliyah Matholi'ul Anwar Simo Sungelebak Karanggeneng Lamongan. *AJER: Advanced Journal of Education and Religion*, *2*(2), 136–147. https://doi.org/10.52166/ajer.v2i2.9570

Santoso, T. A., Armeridha, Y. A. S., & Pradana, D. A. (2025). The Influence of the Gemification Method on Vocabulary Development in Primary Schools. *JOURNAL OF TECHNOLOGY, EDUCATION & TEACHING (J-TECH)*, *2*(1), 158–167. https://doi.org/10.62734/jtech.v2i1.483

Slater, P., & Hasson, F. (2025). Quantitative Research Designs, Hierarchy of Evidence and Validity. *Journal of Psychiatric and Mental Health Nursing*, *32*(3), 656–660. https://doi.org/10.1111/jpm.13135

Sugiyono. (2020). *Educational Research Methods: Quantitative, Qualitative, Combination, R&D and Educational Research*. Alfa Beta.

Taber, K. S. (2018). The use of Cronbach's alpha when developing and reporting research instruments in science education. *Research in Science Education*, *48*(6), 1273–1296.

Wang, A. I., & Tahir, R. (2020). The Effect of Using Kahoot! For Learning – A Literature Review. *Computers & Education*, *149*, 103818. https://doi.org/10.1016/j.compedu.2020.103818

Waruwu, M. (2023). Educational Research Approaches: Qualitative Research Methods, Quantitative Research Methods and Mixed Methods.*Jurnal Pendidikan Tambusai*, *7*(1), 2896–2910. https://doi.org/10.31004/jptam.v7i1.6187

Yaccob, N. S., Abd. Rahman, S. F., Azlan Mohamad, S. N., Abdul Rahim, A. A., Khalilah Abdul Rashid, K., Mohammed Abdulwahab Aldaba, A., Md Yunus, M., & Hashim, H. (2022). Gamifying ESL Classrooms through Gamified Teaching and Learning. *Arab World English Journal*, *8*, 177–191. https://doi.org/10.24093/awej/call8.12

Yamani, A. Z., Alziyady, R., AlYami, R., Albelali, S. A., Abouhagar, L., Almulhim, J., Alsulami, A., Alfarraj, M., & Al-Zaidy, R. (2024). The KIND Dataset: A Social Collaboration Approach for Nuanced Dialect Data Collection. In *EACL 2024 - 18th Conference of the European Chapter of the Association for Computational Linguistics, Proceedings of the Student Research Workshop* (pp. 32–43). https://doi.org/10.18653/v1/2024.eacl-srw.3