BILINGUAL PUBLISHING GROUP

**Forum for Linguistic Studies**
https://journals.bilpubgroup.com/index.php/fls

**ARTICLE**

# The Effectiveness Gamification as an Assessment Method to Enhance Arabic Language Learning Competency in Indonesian Junior High Schools

***Mohammad Ahsanuddin*** [1*] **, *Moh. Ainin*** [1] **, *Moh. Fauzan*** [1] **, *Abd Rauf Hassan*** [2*] **, *Muhammad Ahsan Thoriq*** [1] **, *Farika Riskiyah*** [1] **, *Rico Fenda Pradana*** [1] **, *Arya Wahyu Pratama*** [1]

[1] *Arabic Language Education, Universitas Negeri Malang, Malang 65145, Indonesia*

[2] *Department of Foreign Languages, Universiti Putra Malaysia, Serdang 43400, Malaysia*

## ABSTRACT

This study aims to address the challenges of Arabic language learning in junior high schools by integrating Construct 2-based gamification as an assessment method. A quantitative research approach was employed using a true experimental design, specifically a post-test-only design with a non-equivalent group. The sample comprised 38 seventh-grade students from SMP Darul Faqih, divided into an experimental group (19 students) and a control group (19 students). Research instruments included tests, questionnaires, and observation sheets, all validated according to educational measurement theory. Data analysis involved normality tests, reliability assessments, and hypothesis testing using independent-sample t-tests or Mann-Whitney U Tests. The findings demonstrate that Construct 2-based gamification significantly enhanced students' Arabic language competence, with the experimental group achieving higher average scores compared to the control group (24.26 vs. 19.26). Interactive elements of gamification boosted student engagement, although no significant improvements were observed in listening skills and vocabulary acquisition. These results underscore the importance of a more integrated gamification design to optimize learning outcomes. This study contributes significantly to the development of technology-based assessment methods that align with the needs of the digital era, opening opportunities for integrating gamification into the educational curriculum in Indonesia.

***Keywords***: Gamification; Construct 2; Assessment Methods; Arabic Language Learning; Technology-based Education

*CORRESPONDING AUTHOR:
Mohammad Ahsanuddin, Arabic Language Education, Universitas Negeri Malang, Malang 65145, Indonesia;
Email: ohammad.ahsanuddin.fs@um.ac.id

ARTICLE INFO
Received: 9 June 2025 | Revised: 16 June 2025 | Accepted: 7 July 2025 | Published Online: 12 August 2025
DOI: https://doi.org/10.30564/fls.v7i8.10417

CITATION
Ahsanuddin, M., Ainin, M., Fauzan, M., et al., 2025. The Effectiveness Gamification as an Assessment Method to Enhance Arabic Language Learning Competency in Indonesian Junior High Schools. Forum for Linguistic Studies. 7(8): 764–775. DOI: https://doi.org/10.30564/fls.v7i8.10417

# 1. Introduction

The rapid development of digital technology has brought significant changes to various aspects of education, including Arabic language learning in junior high schools in Indonesia [1]–[3]. In today's digital age, the effectiveness of traditional teaching methods is being increasingly scrutinized, especially in terms of meeting the needs of students from the digital generation, who are more accustomed to digital interactions and prefer engaging, dynamic, and interactive learning environments [4],[5]. This issue is further compounded by a decrease in student motivation with traditional methods and the difficulty of sustaining active engagement during lessons [6],[7].

Arabic plays a crucial role in Indonesia's educational and socio-religious context. In many junior high schools, Arabic language learning serves as a means of deepening religious literacy and Islamic culture. However, the challenges faced in this area are complex, particularly low student motivation, which negatively impacts academic achievement [8]. Only 45% of junior high school students feel motivated to learn Arabic, largely due to the perception that the language is difficult and irrelevant [9]. Low motivation leads to decreased student engagement, resulting in diminished competence, especially in vocabulary and listening skills.

Based on preliminary diagnostics, the listening skills (maharah istima') of students at SMP Darul Faqih show a low level of proficiency. This is illustrated by the students' average pre-test score of 62.4, which is still below the set Minimum Mastery Criteria (KKM) of 75.

The majority of students struggle to comprehend spoken Arabic whether it is in the form of general directions, detailed information, or even the underlying meanings. This weak starting skill level stems from a variety of reasons, such as a lack of adequate vocabulary, insufficient exposure to authentic spoken Arabic, and minimal use of audio materials in previous lessons.

These areas of concern indicate that more effective and specific teaching methods need to be designed and implemented, especially those aimed at increasing the students' motivation and active participation in listening activities. To this end, the use of gamification has been adopted as an alternative strategy to improve students' listening skills gradually and continuously.

These challenges are further complicated by the implementation of the block scheduling system based on the latest curriculum policies. This system requires the completion of each subject within a short period, specifically one month and fifteen days [10]. In the context of language learning, the block scheduling system creates difficulties in maintaining student activity, as language requires continuous practice and comprehensive evaluation [11]. Therefore, innovative learning methods capable of addressing the challenges of the block scheduling system while enhancing learning effectiveness are urgently needed [12].

Gamification has emerged as an innovative solution to enhance student engagement by creating a more interactive learning environment [13]. Gamification, which integrates game elements into learning, has been proven to improve student motivation and learning effectiveness [12],[14]. Construct 2, a user-friendly game development platform, offers the opportunity to create engaging, effective, and digital generation-friendly assessment methods [15].

Based on the above, this study aims to address the challenges of Arabic language learning through Construct 2-based gamification. Specifically, the study measures the impact of gamification on students' vocabulary, listening skills, and overall competence in Arabic language learning at the junior high school level. This research also seeks to create an assessment method that is relevant to the needs of the digital era while maintaining student engagement within the block scheduling system.

Previous studies have demonstrated the effectiveness of gamification. Gamification increased learning motivation by 68% compared to conventional methods [16]. Gamification in assessments boosted student engagement by 72%, reduced evaluation anxiety by 45%, and improved vocabulary retention by 63% [17]. Other studies indicated that gamification in Arabic language learning increased student engagement and language proficiency [18].

This study offers a new contribution by linking Construct 2-based gamification to the local Indonesian context, particularly the block system in junior high school curricula. This approach provides ease of implementation for teachers, efficient assessment time, and more accurate evaluations [19]. Focusing on local relevance and 21st-century needs, this research not only contributes to the literature on gamification but also offers a technology-based learning model that can be replicated in various educational institutions across Indonesia.

# 2. Research Method

## 2.1. Research Approach and Design

This study employs a quantitative approach with an experimental design to examine the effect of gamification using the Construct 2 platform on Arabic language learning at SMP Darul Faqih, Indonesia. A quantitative approach is suitable for explaining causal relationships between variables [20]. The research design used is a true experimental design, specifically a post-test-only design with a non-equivalent group. This design allows for a comparison of learning outcomes between the experimental group using Construct 2-based gamification and the control group using

conventional learning methods [21].

This study measures the relationship between two types of variables: (1) independent variables consisting of Construct 2 and gamification, and (2) dependent variables consisting of vocabulary knowledge, listening skills, and student competence. The selection of these variables is based on the research objective to identify the impact of gamification on key aspects of Arabic language learning, supported by scholars who argue that gamification can enhance student engagement and motivation in learning [22].

## 2.2. Population and Sample

The research population consists of all VII (seventh-grade) students at SMP Darul Faqih, totaling 38 students. The sample is divided into two groups: (1) the experimental group, consisting of 19 students, and (2) the control group, consisting of 19 students. The focus of this study is to evaluate how effective a Construct 2-based approach to gamification is for learning the Arabic language. The class divisions were done with purposive sampling considering the equivalence of students' initial traits and the class level for which they were scheduled.

Both groups were assessed for baseline competencies through a pre-test evaluation of listening abilities and vocabulary knowledge, which was calculated descriptively. This experimental group was taught with interactive media designed on Construct 2 with gamification features like points, challenges, and feedback. The control group, on the other hand, was taught using conventional school teaching methods.

This method was implemented to make sure any variances in learning results were caused by the treatment given, as opposed to diversity in group characteristics at the onset. This approach strengthened the internal validity of the study [23].

## 2.3. Research Instruments

The instruments used in this study include (1) tests to measure vocabulary knowledge and listening skills in Arabic, (2) questionnaires to assess student competence in learning, and (3) observation sheets to measure student engagement during the learning process. These instruments were designed and validated based on the educational instrument development theory [24].

## 2.4. Research Procedure

The research process was carried out in several stages:

1. Preparation Stage: In this phase, the researcher created instructional materials using Construct 2 as well as research instruments, which included tests on vocabulary and listening comprehension. All instruments underwent preliminary validation with subject matter experts and media specialists to confirm their content relevance and appropriateness to the intended learning outcomes.
2. Implementation Stage: The research spanned over a period of four weeks with two groups, an experimental class and a control class. The experimental class received instruction in Arabic language through a gamification method implemented via Construct 2, aimed at enhancing student motivation and engagement. The control class, on the other hand, received instruction through traditional teaching methods that are prevalent in school settings.
3. Post-Test Stage: Following the interventions, a post-test was conducted with both groups to measure the outcomes of learning with an emphasis on vocabulary and listening skills. Accessible data from the post-test were assessed to determine the effectiveness of the interactive media created with Construct 2 based functions in comparison to the traditional teaching approaches used.

## 2.5. Data Analysis

Data analysis was performed in several stages: (1) *descriptive analysis* to calculate the mean, median, maximum and minimum values, as well as the standard deviation; (2) *assumptions testing*, including (a) normality test using the Kolmogorov-Smirnov test to examine the data distribution, and (b) reliability test using Cronbach's alpha coefficient to ensure instrument consistency; and (3) *hypothesis testing*, in which (a) if the data were normally distributed, an independent-sample t-test was used, and (b) if the data were not normally distributed, non-parametric tests, such as the Mann-Whitney U Test, were used. The results of the statistical tests will be interpreted to determine whether there are significant differences in vocabulary knowledge, listening skills, and student competence between the experimental and control groups.

# 3. Findings

## 3.1. Descriptive Analysis Results

This study evaluates the effect of Construct 2-based gamification as an assessment method for Arabic language learning among junior high school students by comparing the experimental group and the control group (each consisting of 19 students). The three variables studied include listening skills, vocabulary proficiency, and student competence (**Table 1**).

**Table 1.** Instrument validity results.

| Variable | Experimental Group | | | | Control Group | | | |
|---|---|---|---|---|---|---|---|---|
| | M (SD) | Min | Max | Skewness | M (SD) | Min | Max | Skewness |
| **Listening Skills** | 90.00 (16.997) | 40 | 100 | −1.897 | 89.95 (11.974) | 70 | 100 | −0.646 |
| **Vocabulary Mastery** | 66.84 (14.163) | 40 | 90 | 0.105 | 68.95 (12.865) | 40 | 90 | −0.308 |
| **Student Competition** | 24.26 (2.922) | 16 | 28 | −1.142 | 19.26 (2.663) | 12 | 22 | −1.223 |

Notes: m: mean; SD: Standard Deviation; min: minimum; max: maximum.

For the listening skills variable, the experimental group showed a skewness value of −1.897, while the control group had a skewness value of −0.646. Both values fall within the acceptable range (±2), but the significant difference indicates the need for further attention in the normality test. For the vocabulary mastery variable, the skewness values for both groups were relatively small (0.105 for the experimental group and −0.308 for the control group), indicating an approximately normal distribution. For the student competence variable, the skewness values were negative (−1.142 for the experimental group and −1.223 for the control group) but still within acceptable limits. However, the significant difference in mean scores for this variable (24.26 vs. 19.26) requires further analysis.

## 3.2. Normality Test

The normality test using the Shapiro-Wilk test indicated that the data were normally distributed for all research variables. The significance values for each variable in both groups were greater than 0.05, thus fulfilling the normality assumption (see **Table 2**). These data are suitable for analysis using parametric tests (independent-sample *t*-test).

**Table 2.** Assessment of Data Distribution.

| **Variable** | Shapiro-Wilk | | | |
|---|---|---|---|---|
| | Experimental Group | | Control Group | |
| | Statistic | Sig | Statistic | Sig |
| **Listening skills** | 0.899 | 0.058 | 0.911 | 0.095 |
| **Vocabulary mastery** | 0.926 | 0.148 | 0.942 | 0.286 |
| **Student competition** | 0.909 | 0.072 | 0.901 | 0.067 |

The results of the normality test using the Shapiro-Wilk test for the three research variables indicate that all data were normally distributed. For the listening skills variable, the experimental group obtained a significance value of 0.058, while the control group obtained a significance value of 0.095. For the vocabulary mastery variable, the normality test results for the experimental group showed a significance value of 0.148, while the control group obtained a significance value of 0.286. For the student competence variable, the experimental group obtained a significance value of 0.072, while the control group showed a significance value of 0.067. The significance values for this variable were relatively lower compared to the other variables.

Based on these analytical results, it can be concluded that all research variables in both groups meet the normality assumption, as they have significance values greater than 0.05. The fulfillment of this normality assumption is an important prerequisite for proceeding to parametric analysis using the independent samples t-test. Therefore, the research data are suitable for further analysis using the independent samples t-test to compare the differences between the experimental and control groups on these three variables.

## 3.3. Independent Samples *t*-test

The results of the *t*-test revealed the following findings in **Table 3**.

**Table 3.** Difference in the average distribution of data.

| Items | Variable | | | | | |
|---|---|---|---|---|---|---|
| | Experimental Group | | Control Group | | Statistical tests | |
| | N | M (SD) | N | M (SD) | *t*-test | *p*-value |
| | **Listening skills** | | | | | |
| Post-test | 19 | 90.00 (16.997) | 19 | 89.95 (11.974) | 0.011 | 0.991 |
| | **Vocabulary mastery** | | | | | |
| Post-test | 19 | 66.84 (14.163) | 19 | 68.95 (12.865) | −0.480 | 0.634 |
| | **Student competition** | | | | | |
| Post-test | 19 | 24.26 (2.922) | 19 | 19.26 (2.663) | 5.512 | 0.000* |

Notes: n: number; m: mean; SD: Standard Deviation. * *p*-value significant at the 0.05 level.

- Listening Skills: No significant difference was found between the experimental group (M = 90.00; SD = 16.997) and the control group (M = 89.95; SD = 11.974) with $t = 0.011$, $p = 0.991$ ($p > 0.05$).
- Vocabulary Mastery: No significant difference was found between the experimental group (M = 66.84; SD = 14.163) and the control group (M = 68.95; SD = 12.865) with $t = -0.480$, $p = 0.634$ ($p > 0.05$).
- Student Competence: A significant difference was found between the experimental group (M = 24.26; SD = 2.922) and the control group (M = 19.26; SD = 2.663) with $t = 5.512$, $p = 0.000$ ($p < 0.05$).

## 3.4. External Factors Analysis

This research described disparate starting abilities among students as one of the key external factors affecting the results of the study. That said, the students did implement a number of control strategies aimed at reducing bias in the interpretation of the results. To start with, the researcher used an experimental design with control and experimental groups so that traditional teaching methods could be contrasted with the with the Construct 2 gamification technique. In addition, all components of assessment such as daily grades, midterm and final exam scores, practical assignments, and even homework were analyzed thoroughly to ensure that intergroup differences were consistent. Last, the outcome of learning was classed using descriptive levels, 'excellent, good, fair, poor, and very poor', thus enabling a more neutral snapshot of mastery distribution (see **Table 4**).

The absence of random sample selection and baseline pre-testing are still areas of weakness that impact external variable control. In an ideal world, experimental studies would use strict randomization alongside equated baseline measurements for both groups, but in this study there were underlying class use restrictions as well as marked differences in starting ability that posed methodological concerns that must be accepted in addition to the findings (**Table 5**).

**Table 4.** Learning outcomes for experimental and control groups.

| Score Range | Category | Control Group | Experimental Group |
|---|---|---|---|
| | | **Class 7A** | **Class 7B** |
| 90 - 100 | Very Good | 15.8% (3 students) | 5.3% (1 student) |
| 80 – 89 | Good | 57.9% (11 students) | 31.6% (6 students) |
| 70 – 79 | Fair | 26.3% (5 students) | 47.4% (9 students) |
| 60 – 69 | Poor | 0% (0 student) | 15.7% (3 students) |
| < 60 | Very Poor | 0% (0 student) | 0% (0 student) |
| **Total** | | **100% (19 students)** | **100% (19 students)** |

**Table 5.** Improving Arabic language learning competencies.

| Groups | Assessment Criteria | | Statistical Tests | |
|---|---|---|---|---|
| | **n** | **M (SD)** | ***t*-test** | ***p*-value** |
| | **Daily Scores** | | | |
| Control | 19 | 78.5 (6.24) | 3.842 | 0.001 |
| Experimental | 19 | 72.3 (5.86) | | |
| | **Mid-term Test Scores** | | | |
| Control | 19 | 82.4 (5.92) | 4.156 | 0.000 |
| Experimental | 19 | 75.8 (5.43) | | |
| | **Final Test Scores** | | | |
| Control | 19 | 80.6 (6.13) | 3.967 | 0.000 |
| Experimental | 19 | 74.2 (5.75) | | |
| | **Practical Scores** | | | |
| Control | 19 | 79.8 (5.86) | 3.728 | 0.001 |
| Experimental | 19 | 73.5 (5.48) | | |
| | **Assignment Scores** | | | |
| Control | 19 | 81.2 (5.94) | 3.893 | 0.000 |
| Experimental | 19 | 76.4 (5.67) | | |

The differences in the students' starting skills remain particularly impactful concerning internal validity issues for listening and vocabulary skills. The data showed that the control group (Class 7A) showed a significant relative advantage, where 73.7% of students were classified as good to excellent, whereas only 36.9% of the students from the experimental group (Class 7B) were classified as good to excellent. This created a ceiling effect in the control group and a floor effect in the experimental group, which might mask the true effects of the Construct 2 gamification intervention on learning outcomes.

As far as the tests were concerned, comprehensive statistical assessments showed that there were consistent significant differences ($p < 0.05$) on all measures. The *t*-test

positive values pointed to the advantage of the control group. The gaps between groups appeared to be the most pronounced in midterm scoring ($t = 4.156$) and final examination ($t = 3.967$) suggesting that the ability gaps between the groups increased over time. This indicates that students with higher starting competencies are likely to retain or, even, strengthen their advantages irrespective of the teaching strategies employed.

This issue is particularly relevant to the assessment of listening skills and vocabulary because both are closely aligned with a student's overall academic proficiency. Students with weaker starting skills tend to need a longer acclimation period to improvement, complicating the ability to measure the full impact of gamification within a limited time frame.

With all the external factors at play, the fact that the experimental group demonstrated significant improvement in student competence with gamification in Construct 2 still provides an encouraging sign regarding the appropriateness of the intervention. Even with initial lower abilities, these outcomes demonstrate that there is potential value in more favorable contexts.

Future studies should incorporate matched or stratified random assignments based on prior competencies, including defined pre-tests, extend the duration of the study to allow for sufficient adaptation time, and adopt flexible approaches to account for the diversity of learners. Additionally, initial levels of competencies could be controlled using ANCOVA (Analysis of Covariance), allowing for the calculation of the genuine impact of the intervention.

Would you require assistance in adapting this into a journal-ready discussion section, or would a synthesis paragraph to close the presentation of findings suffice? I would appreciate the opportunity to assist with the entire paper to prepare it for publication.

The most prominent finding was observed in the student competence aspect, where the experimental group had an average score of 24.26 with a standard deviation of 2.922, significantly higher than the control group, which had an average score of 19.26 with a standard deviation of 2.663. The t-test results produced a $t$-value of 5.512 with a $p$-value of 0.000 ($p < 0.05$), indicating a highly significant difference in student competence between the experimental and control groups.

## 3.5. Implementation of Construct 2 Gamification

### 3.5.1 Before the Implementation of Gamification

Precisely to which degree of difficulty should the questions be set within the context of the gamification aimed at evaluating students' proficiency in the Arabic language within the Construct 2 framework? This is the crucial inquiry that the researcher addressed prior to implementing the gamification feature, especially considering the students' cognitive levels which were still significantly lower than the KKM. From the evaluated assessment instruments, a detailed item analysis was performed, with item analysis based on the revised Bloom's Taxonomy levels by Anderson and Krathwohl (2001) on remembering, understanding, and applying.

### 3.5.2 The Remembering Level

Within the measurement framework, recalling facts, recognizing information and retrieval processes were the most dominant subcategory. Terms and phrases identification accompanied by recognizing pictures "من هذه الصورة؟" and retrieval of relevant antonyms "ما الضد من هو؟" exemplify recall at the verbatim level. The retrieval of the verb "صباح الخير" and the nomino-verbal expression "كيف حالك" forms the core of the basic recall subset which items 1 to 5 sought to accomplish within the istima' components.

The emphasis on the remembering level resonates with Nation's (2001) principles of second language acquisition which state that mastery of a vocabulary list is fundamental to the development of more advanced skills. For learners under the KKM benchmark, focusing on improving their remembering skills solidifies foundational linguistic competencies that are critical to supporting advanced cognitive functioning.

### 3.5.3 Understanding Level

This component is assessed within 30–40% of the items. This aims to refine students' skills in interpretation, explanation, and inference. With regard to vocabulary, item 3 tested students' comprehension of the synonym of "الأستاذ" in contextual usage while item 4 tested the construction of sentences and their grammar. Sentence description: "He has many professional activities" necessitates some contextual understanding, reasoning, and construing to describe the job.

In the istima' component, tasks 6-10 required students to practice matching given audio stimuli to relevant diagrams. These tasks involved sophisticated conceptually creative comprehension far beyond rote memory retrieval and included semantic reasoning, synthesis of audio and visuals, and evaluative decision making grounded in a constructed hypothesis.

### 3.5.4 Applying Level

This cognitve level was used only in a limited scope, 10–20%, for learners who achieved the lower cognitive levels. Example "الأساتيذ ... في الملعب" assessed students' ability to apply rules of verb conjugation appropriate for plural subjects to their utterances. Such items demand the use of Arabic morphology and syntax knowledge in relevant communicative situations.

This limited proportion aligns with Vygotsky's (1978) zone of proximal development, which underlines the need for structure to be provided prior to the KKM threshold for learners to be able to use knowledge of language freely. This strategy ensures students are able to engage with pro-

ductive and effective learning experiences which, in turn, fosters the desire to learn and self-confidence when studying the language independently.

Construct 2 gamification successfully created a competitive environment that encouraged student participation. The average gamification score for students was 2.6 (on a 10-point scale), with the highest score being 6 and the fastest completion time recorded at 12.8 seconds (See **Figure 1**). Question Q5 had the highest success rate, while Q2 was the most difficult. The differences in students' strategic approaches to solving questions highlight the need to adjust the difficulty level of the questions to create a better balance.

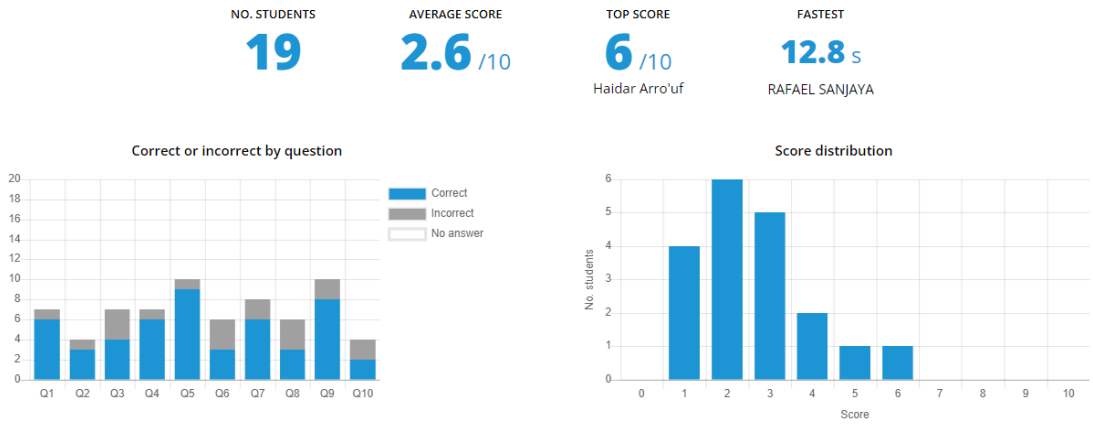


**Figure 1.** gamification assessment data.

Despite variations in the results, the application of gamification demonstrates potential in enhancing student engagement. The significant improvement in student competence in the experimental group reinforces the success of this method as an innovative approach to assessment in learning.

The implementation of gamification using Construct 2 in learning demonstrated an engaging, competitive dynamic among students. Out of a total of 19 participants, significant variation in achievement was found, with an average score of 2.6 on a 10-point scale. For example, a student with the initials H.A. achieved the highest performance with a score of 6/10, while another student with the initials R.S. excelled in the speed of completion, with the fastest time of 12.8 seconds. The score distribution revealed an interesting pattern, where the majority of students were concentrated in the 2–3 score range, with 6 students achieving a score of 2 and 5 students earning a score of 3. Item analysis showed varying levels of difficulty, with Question Q5 having the highest number of correct answers (9 correct answers), while Q2 was the most difficult, with only 3 correct answers. The pattern of student responses also demonstrated consistent participation, with at least 6 students answering each question, indicating a high level of engagement in the learning process through gamification. Interestingly, there was a considerable gap between the highest score (6) and the class average (2.6), which can be attributed to the varying difficulty levels of the questions. This highlights the need for improvement and adjustments in this gamified learning method. Nevertheless, overall, the data shows that the implementation of Construct 2 successfully created a competitive environment that encouraged active student participation. However, further refinement is needed to improve the average learning achievement.

The implementation of gamification using Construct 2 revealed an intriguing competition among the top three participants. From **Figure 2**, a student with the initials H.A. ranked first with an outstanding performance, achieving the highest score of 6 out of 10 and completing the test in 1 minute and 34 seconds. In second place, a student with the initials M.N.A. demonstrated a competitive performance with a score of 5 out of 10 and a nearly identical completion time of 1 minute and 44 seconds, just 10 seconds behind the first-place. Then, in third place, a student with the initials R.S. scored 4 out of 10 but with a longer time of 51.4 seconds. Interestingly, despite R.S. having a lower score, he exhibited better time efficiency compared to the two participants above him. This pattern suggests that each student applied a different strategy in approaching the questions, with H.A. and M.N.A. perhaps prioritizing accuracy with longer completion times, while R.S. focused on speed in decision-making. These three participants demonstrated a varying balance between accuracy and speed in completing the gamification challenge, reflecting the effectiveness of Construct 2 in accommodating diverse learning styles.

The analysis results show that the implementation of Construct 2 gamification as a method of assessment in Arabic language learning significantly improved student competencies, even though no significant differences were observed in listening skills and vocabulary. This suggests that the gamification method was more effective in enhancing overall student competency than conventional methods. This further strengthens the conclusion of the study, which indicates that, although the experimental group had lower learning outcomes, the implementation of Construct 2 gamification still successfully enhanced student competencies. The significant difference in student competencies, with a high t-value (5.512) and a very low p-

value (0.000), demonstrates that the application of Construct 2 gamification had a strong positive impact on improving student competencies in Arabic language learning at SMP Darul Faqih, Indonesia.

| No | Initials Students | Submission Time | Correct | Incorrect | Time (min:sec) |
|---|---|---|---|---|---|
| 1 | HA | 12:37 - 29 Aug 2024 | 6 | 3 | 1:34 |
| 2 | MNA | 12:36 - 29 Aug 2024 | 5 | 3 | 1:44 |
| 3 | RS | 12:36 - 29 Aug 2024 | 4 | 3 | 51.4 |
| 4 | MIR | 12:33 - 29 Aug 2024 | 4 | 3 | 1:12 |
| 5 | R | 12:33 - 29 Aug 2024 | 3 | 3 | 54.2 |
| 6 | MHAM | 12:33 - 29 Aug 2024 | 3 | 3 | 54.4 |
| 7 | HS | 12:35 - 29 Aug 2024 | 3 | 3 | 57.9 |
| 8 | VFRI | 12:37 - 29 Aug 2024 | 3 | 3 | 1:46 |
| 9 | RA | 12:33 - 29 Aug 2024 | 3 | 3 | 2:06 |
| 10 | I | 12:31 - 29 Aug 2024 | 2 | 3 | 35.1 |
| 11 | WA | 12:36 - 29 Aug 2024 | 2 | 3 | 35.6 |
| 12 | WK | 12:37 - 29 Aug 2024 | 2 | 3 | 47.0 |
| 13 | MA | 12:37 - 29 Aug 2024 | 2 | 3 | 51.6 |
| 14 | MKMA | 12:38 - 29 Aug 2024 | 2 | 3 | 56.4 |
| 15 | NR | 12:35 - 29 Aug 2024 | 2 | 3 | 1:04 |
| 16 | RHA | 12:32 - 29 Aug 2024 | 1 | 3 | 25.7 |
| 17 | NF | 12:36 - 29 Aug 2024 | 1 | 3 | 26.3 |
| 18 | HB | 12:32 - 29 Aug 2024 | 1 | 3 | 38.6 |
| 19 | MIK | 12:32 - 29 Aug 2024 | 1 | 3 | 1:04 |

**Figure 2.** Gamification assessment scores.

# 4. Discussion

## 4.1. Construct 2 Gamification as an Assessment Method

This study evaluates the impact of gamification using Construct 2 as an assessment method for Arabic language learning at SMP Darul Faqih, Indonesia, comparing the results between the experimental and control groups. The analysis shows that, although no significant differences were observed in listening skills and vocabulary, the experimental group demonstrated a significant improvement in student competency, with an average score of 24.26 compared to 19.26 in the control group ($p < 0.05$). The implementation of gamification also encouraged active participation and created an engaging, competitive environment, although there were variations in scores and outcome gaps, indicating the need for further adjustments. These results highlight the potential of Construct 2 gamification in enhancing student engagement and competency, although they may be influenced by external factors such as initial student ability.

Construct 2 gamification has shown potential as an innovative assessment method in Arabic language learning. This method offers a game-based approach that not only evaluates students' abilities but also increases their engagement during the assessment process. Gamification is the application of game elements in non-game contexts to increase motivation and engagement [25]. A recent study also supports this view, stating that gamification can enhance students' intrinsic motivation through rewards, challenges, and instant feedback [26].

The findings of this study indicate that the aspects of listening skills and vocabulary did not show significant differences between the experimental and control groups. However, this can be explained by external factors such as differences in students' initial abilities that may have affected the assessment results. Another study noted that the effectiveness of gamification can be influenced by students' initial readiness and the quality of learning design that supports their individual needs [13]. Thus, the impact of Construct 2 gamification on listening and vocabulary skills requires strengthening strategies based on differentiation.

The competitive and interactive elements within gamification help create an environment that encourages students to participate actively. The variation in gamification scores among students reflects individual abilities to respond to the challenges provided. Gamification effectively enhances engagement through direct reward mechanisms and precise feedback, allowing students to see their achievements instantly [22]. Consistent with these findings, elements such as leaderboards and badges can enhance healthy student competition, strengthening their commitment to learning tasks [27].

The success of Construct 2 gamification as an assessment method is also evident in the high participation rates of students in answering questions despite the gap between the highest and average scores. This indicates that the

method can capture students' interest, even when the difficulty of the questions varies. Studies state that technology-based gamification, such as Construct 2, can capture students' attention even in challenging contexts [28]. With further development, gamification can become a more adaptive assessment method tailored to students' abilities. The use of learning analytics could be the next step to personalize the gamification experience [29].

## 4.2. Effectiveness of Construct 2 Gamification in Arabic Language Learning

Construct 2 gamification in Arabic language learning has shown significant results, particularly in improving student competence within the experimental group. The t-test analysis revealed a substantial difference between the experimental and control groups ($t = 5.512$, $p < 0.05$), indicating that gamification significantly affected the students' learning outcomes. This aligns with the findings of previous researchers, who suggest that gamification can enhance intrinsic motivation by incorporating elements such as challenges, competition, and rewards [30]. They argue that competitive elements in gamification not only boost student motivation but also improve memory retention of the material being taught [31].

However, despite the improvements in student competence, the study found no significant differences in listening skills and vocabulary. This could be attributed to the short duration of the gamification implementation and a mismatch between the difficulty of the tasks and the students' initial skill levels. The success of game-based learning largely depends on task design appropriate for the learners' capabilities [32]. This observation is consistent with the findings, which highlighted that the success of gamification in language learning is influenced by the design of tasks that consider the learning context and students' background [13]. Therefore, developing more structured and level-based questions aligned with students' abilities is necessary to improve outcomes in these areas.

Gamification also has a positive impact on student engagement. Active participation is evident from the average scores per student, indicating consistency in answering each question. Well-designed gamification can foster engagement through interactivity and real-time feedback elements [26]. This is consistent with the findings of this study, which show high engagement levels despite varying learning outcomes. A study also indicates that the direct feedback elements in gamification help students understand their progress and enhance their learning motivation [13].

In addition, the competitive nature of the gamification process contributed to a more dynamic learning environment. Students with higher scores demonstrated increased motivation to complete tasks quickly, highlighting the relationship between challenge and performance. This corresponds to the flow theory, where students experience optimal learning when the difficulty level of tasks matches their skill levels [33]. Moreover, studies have shown that the competitive elements in gamification foster healthy competition and collaboration, which enhances the overall learning experience [34].

## 4.3. Gamification in Enhancing Student Competence

The use of Construct 2 gamification has proven to significantly contribute to the enhancement of student competence, particularly in Arabic language learning. The experimental group, which utilized gamification, demonstrated a much higher average competence (24.26) than the control group (19.26). These results support previous research, which found that gamification in education could improve learning outcomes through more engaging and interactive experiences [35]. A more recent study also shows that gamification enhances learning outcomes and strengthens student engagement with the material [36].

One key factor contributing to the improvement in student competence is the provision of quick and specific feedback within gamification. Students can instantly recognize their mistakes and achievements, encouraging them to learn from their mistakes. Rewards and feedback in a gamified environment can enhance intrinsic motivation, positively impacting student competence [9]. Additionally, research emphasizes that high-quality feedback, including that provided through technology, is a critical determinant of learning success [37].

Despite its benefits, the implementation of gamification still encounters challenges, particularly in matching the difficulty level of tasks with the students' abilities. The analysis of individual task scores reveals a notable gap between the highest and average class scores, suggesting the need for further refinement in task design. The success of gamification depends on striking the right balance between the challenge level and students' capabilities to prevent frustration or disengagement [9]. This view is supported by researchers, who emphasize the importance of personalizing gamification to better align with each student's needs [13].

In conclusion, Construct 2 gamification holds substantial promise in enhancing student competence, especially if accompanied by more adaptable task designs and continued professional development for teachers to maximize the effectiveness of this technology. It has been noted that proper teacher training in applying gamification technology is essential for its successful implementation in the classroom [31].

This study still offers ample room for further development, particularly in more specific areas of Arabic language proficiency. Therefore, future researchers are advised to:

1. **Deepen the Study of Vocabulary Mastery (Mufradat):** An in-depth investigation into the types, difficulty levels, and frequency of use of the taught vocabulary can provide a more comprehensive understanding of the instructional method's effectiveness. Additionally, analyzing vocabulary teaching strategies that

are contextual and based on real communicative needs can enrich the findings.

2. **Examine Listening Skills (Istima') More Specifically:** Listening is a crucial receptive skill that often receives limited attention. Future research could focus on the types of texts used (dialogues, monologues, news, etc.), audio duration, speech rate, and students' ability to capture both global and detailed information.
3. **Integrate Vocabulary with Listening:** Studies that combine vocabulary mastery with listening skills will offer a more holistic view of students' language competence development. For instance, exploring how vocabulary knowledge influences the ability to comprehend audio recordings could be a compelling focus.
4. **Develop More Varied Instruments:** It is recommended that future researchers employ a wider range of instruments, such as diagnostic vocabulary tests, listening journals, or student self-reflections, to obtain deeper and more accurate data.

Would you like help weaving this into your concluding section or aligning it with journal-specific formatting requirements? I can also help you craft a strong final paragraph that ties all insights together.

# 5. Conclusion

This study demonstrates that gamification using Construct 2 significantly enhances student competence in Arabic language learning, particularly within the experimental group, which achieved higher average scores than the control group (24.26 vs. 19.26). The interactive and competitive elements of gamification effectively promoted active student participation, fostering an engaging and innovative learning environment. However, the study also reveals that areas such as listening skills and vocabulary mastery did not show significant improvement, likely due to the short duration of the implementation and variations in students' initial abilities. These results underscore the need for further refinement of the gamification design to maximize its comprehensive impact.

The significance of this research lies in the integration of technology-based learning, particularly in the adoption of gamification as an innovative assessment method in schools. The findings suggest that gamification not only boosts student engagement but also facilitates a more enjoyable and effective learning experience. Additionally, this study lays the groundwork for developing more adaptive learning strategies, such as personalized task difficulty and dynamic assessments, to better address individual student needs.

# Author Contributions

Conceptualization, M.A. (Mohammad Ahsanuddin); methodology, M.A. (Moh. Ainin); software, M.A.T.; validation, M.F. and F.R.; formal analysis, A.R.H.; investigation, R.F.P.; resources, M.A. (Mohammad Ahsanuddin),M.A. (Moh. Ainin)., and M.A.T; data curation, F.R.,M.F., A.W.P, and R.F.P; writing—original draft preparation, M.Ah and M.A.T.; writing—review and editing, M.A. (Mohammad Ahsanuddin) and M.A.T; visualization, R.F.P. and A.W.P; supervision, A.R.H; project administration, F.R. and A.W.P; funding acquisition, M.A. (Mohammad Ahsanuddin). All authors have read and agreed to the published version of the manuscript.

# Funding

This work was supported by Lembaga Penelitian dan Pengabdian Kepada Masyarakat Universitas Negeri Malang Indonesia grant number 4.4.738/UN32.14.1/LT/2024.

# Institutional Review Board Statement

The study was conducted and approved by the Institutional Review Board (or Ethics Committee) of Universitas Negeri Malang (No. Reff : 1.8.5/UN32.14/PB/2025).

# Informed Consent Statement

Informed consent was obtained from all subjects involved in the study.

# Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

# Acknowledgements

The researcher would like to express deep appreciation to the LPPM (Lembaga Penelitian dan Pengabdian Masyarakat) Universitas Negeri Malang for the funding they provided, as well as the administrative support that was granted during the period of this study. Specifically, this aid helped in the meticulous planning, implementation, and finalization of this study. It is anticipated that the outcomes of this research will be significnt in driving forward scholarly discourse, especially within the context of educational pedagogy on the Arabic language.

# Conflicts of Interest

The authors declare no conflict of interest.

# References

[1] Haleem, A., Javaid, M., Qadri, M.A., et al., 2022. Understanding the role of digital technologies in education: A review. Sustainable Operations and Computers. 3, 275–285. Available from: https://linkinghub.elsevier.com/retrieve/pii/S2666412722000137

[2] Núñez-Merino, M., Maqueira-Marín, J.M., Moyano-Fuentes,

J., et al., 2020. Information and digital technologies of Industry 4.0 and Lean supply chain management: a systematic literature review. International Journal of Production Research. 58(16), 5034–5061. Available from: https://www.tandfonline.com/doi/full/10.1080/00207543.2020.1743896

[3] Verganti, R., Vendraminelli, L., Iansiti, M., 2020. Innovation and Design in the Age of Artificial Intelligence. Journal of Product Innovation Management. 37(3), 212–227. Available from: https://onlinelibrary.wiley.com/doi/10.1111/jpim.12523

[4] Guo, W., Li, W., 2022. A Workshop on Social Media Apps for Year-10 Students: An Exploratory Case Study on Digital Technology Education in Regional Australia. Online Journal of Communication and Media Technologies. 12(4), e202222–e202222. Available from: https://www.ojcmt.net/article/a-workshop-on-social-media-apps-for-year-10-students-an-exploratory-case-study-on-digital-technology-12237

[5] Iivari, N., 2020. Empowering children to make and shape our digital futures – from adults creating technologies to children transforming cultures. International Journal of Information and Learning Technology. 37(5), 279–293. Available from: https://www.emerald.com/insight/content/doi/10.1108/IJILT-03-2020-0023/full/html

[6] Martirosov, A.L., Alex, J., Doane, A., et al., 2023. Podcasts and videos and slides…oh my!: Traditional vs. nontraditional teaching methods in remote settings. Currents in Pharmacy Teaching and Learning. 15(6), 587–592. Available from: https://linkinghub.elsevier.com/retrieve/pii/S1877129723001442

[7] Carroll, F., Faruque, R., Hewage, C., et al., 2023. The Journey to Making 'Digital Technology' Education a Community Learning Venture. Education Sciences. 13(5), 428. Available from: https://www.mdpi.com/2227-7102/13/5/428

[8] Khuroidah, L., 2022. The Transformation of Arabic Learning Language Majors in High School. Studia Arabica. 13(2). Available from: https://jurnal.yudharta.ac.id/v2/index.php/studi-arab/article/view/3393

[9] Schüttpelz-Brauns, K., Hecht, M., Hardt, K., et al., 2020. Institutional strategies related to test-taking behavior in low stakes assessment. Advances in Health Sciences Education. 25(2), 321–335. Available from: http://link.springer.com/10.1007/s10459-019-09928-y

[10] Gahan, K., Talley, R., 1975. A Block Scheduling System. JONA: The Journal of Nursing Administration. 5(9), 39–41. Available from: http://journals.lww.com/00005110-197511000-00008

[11] Yang, C.D., 2023. Knowledge and Learning in Natural Language. Oxford University Press: Oxford, UK. Available from: https://academic.oup.com/book/48385

[12] Jiang, Z., Zhou, H., Tao, N., et al., 2024. Knowledge-Based Curved Block Construction Scheduling and Application in Shipbuilding. Journal of Shanghai Jiaotong University. 29(5), 759–765. Available from: https://link.springer.com/10.1007/s12204-022-2544-0

[13] Sailer, M., Homner, L., 2020. The Gamification of Learning: a Meta-analysis. Educational Psychology Review. 32(1), 77–112. Available from: http://link.springer.com/10.1007/s10648-019-09498-w

[14] Floryan, M., Chow, P.I., Schueller, S.M., et al., 2020. The Model of Gamification Principles for Digital Health Interventions: Evaluation of Validity and Potential Utility. Journal of Medical Internet Research. 22(6), e16506. Available from: https://www.jmir.org/2020/6/e16506

[15] Monteiro, R.H.B., Oliveira, S.R.B., De Almeida Souza, M.R., 2021. A standard framework for gamification evaluation in education and training of software engineering: an evaluation from a proof of concept. In Proceedings of the 2021 IEEE Frontiers in Education Conference (FIE), Lincoln, NE, USA, 13–16 October 2021; pp. 1–7. Available from: https://ieeexplore.ieee.org/document/9637232/

[16] Alsawaier, R.S., 2018. The effect of gamification on motivation and engagement. International Journal of Information and Learning Technology. 35(1), 56–79. Available from: https://www.emerald.com/insight/content/doi/10.1108/IJILT-02-2017-0009/full/html

[17] Helvich, J., Novak, L., Mikoska, P., et al., 2023. A systematic review of gamification and its assessment in EFL teaching. International Journal of Computer-Assisted Language Learning and Teaching. 13(1), 1–21. Available from: https://services.igi-global.com/resolvedoi/resolve.aspx?doi=10.4018/IJCALLT.322394

[18] Anjani, R.P., Sari, E., 2023. Perilaku sintaksis adjektiva CNN Indonesia Hastag Layar Demokrasi pada Agustus 2019. Sastranesia. 11(2), 3–7.

[19] Lampropoulos, G., Sidiropoulos, A., 2024. Impact of gamification on students' learning outcomes and academic performance: A longitudinal study comparing online, traditional, and gamified learning. Education Sciences. 14(4), 367. DOI: https://doi.org/10.3390/educsci14040367

[20] Creswell, J.W., 2014. Research Design: Qualitative, Quantitative, and Mixed Methods Approaches, 3rd ed. SAGE Publications: Thousand Oaks, CA, USA. pp. 1–270.

[21] Shadish, W.R., Cook, T.D., Campbell, D.T., 2002. Experimental and Quasi-Experimental Designs for Generalized Causal Inference. Houghton Mifflin and Company: Boston, MA, USA.

[22] Deterding, S., Dixon, D., Khaled, R., et al., 2011. From game design elements to gamefulness. In Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments, New York, NY, USA, 28 September 2011; pp. 9–15. Available from: https://dl.acm.org/doi/10.1145/2181037.2181040

[23] Gay, L.R., Mills, G.E., Airasian, P.W., 2011. Educational Research: Competencies for Analysis and Applications. Pearson Education: Boston, MA, USA.

[24] Fraenkel, J.R., Wallen, N.E., 2006. How to Design and Evaluate Research in Education. McGraw-Hill: New York, NY, USA.

[25] Kapp, K.M., 2012. The Gamification of Learning and Instruction: Game-Based Methods and Strategies for Training and Education. Pfeiffer: San Francisco, CA, USA.

[26] Hamari, J., Koivisto, J., Sarsa, H., 2014. Does gamification work? – A literature review of empirical studies on gamification. In Proceedings of the 2014 47th Hawaii International Conference on System Sciences, Waikoloa, HI, USA, 6–9 January 2014; pp. 3025 – 3034. Available from: http://ieeexplore.ieee.org/document/6758978/

[27] Philpott, A., Son, J.B., 2022. Leaderboards in an EFL course: Student performance and motivation. Computers & Education. 190, 104605. DOI: https://doi.org/10.1016/j.compedu.2022.104605

[28] Bouchrika, I., Harrati, N., Wanick, V., et al., 2021. Exploring the impact of gamification on student engagement and involvement with e-learning systems. Interactive Learning Environments. 29(8), 1244–1257. DOI: https://doi.org/10.1080/10494820.2019.1623267

[29] Liesa-Orús, M., Latorre-Cosculluela, C., Vázquez-Toledo, S., et al., 2020. The technological challenge facing higher education professors: Perceptions of ICT tools for developing 21st Century skills. Sustainability. 12(13), 5339.

[30] Zichermann, G., Cunningham, C., 2011. Gamification by Design: Implementing Game Mechanics in Web and Mobile Apps. O'Reilly Media, Inc.: Sebastopol, CA, USA.

[31] Bai, S., Hew, K.F., Huang, B., 2020. Does gamification improve student learning outcome? Evidence from a meta-analysis and synthesis of qualitative data in educational contexts. Educational Research Review. 30, 100322. DOI: https://doi.org/10.1016/j.edurev.2020.100322

[32] Gee, J.P., 2003. What video games have to teach us about learning and literacy. Computers in Entertainment. 1(1), 20. DOI: https://doi.org/10.1145/950566.950595

[33] Csikszentmihalyi, M., 1990. Flow: The Psychology of Optimal Experience. Harper & Row: New York, NY, USA.

[34] Nah, F.F.H., Zeng, Q., Telaprolu, V.R., et al., 2014. Gamification of education: A review of literature. In Proceedings of the International Conference on HCI in Business, Heraklion, Greece, 22–27 June 2014; pp. 401–409. Available from: http://link.springer.com/10.1007/978-3-319-07293-7_39

[35] Campo, L., Galindo-Domínguez, H., Bezanilla, M.J., et al., 2023. Methodologies for fostering critical thinking skills from university students' points of view. Education Sciences. 13(2), 132. DOI: https://doi.org/10.3390/educsci13020132

[36] Subhash, S., Cudney, E.A., 2018. Gamified learning in higher education: A systematic review of the literature. Computers in Human Behavior. 87, 192–206. DOI: https://doi.org/10.1016/j.chb.2018.05.028

[37] Wisniewski, B., Zierer, K., Hattie, J., 2020. The power of feedback revisited: A meta-analysis of educational feedback research. Frontiers in Psychology. 10, 3087. DOI: https://doi.org/10.3389/fpsyg.2019.03087