The effectiveness of artificial intelligence in gamification assessment for meansuring learning outcomes of elementary school students at SDN Kemayoran 2 Bangkalan

Ika Dian Rahmawati[1*], *Rika* Wulandari[1], *Reskiya* Ramadani[1]

[1]Elementary School Teacher Education, Universitas Trunodjoyo Madura, Bangkalan, Jawa Timur, Indonesia

Abstract. This study aims to analyze the effectiveness of applying Artificial Intelligence (AI) in assessment gamification to improve student learning outcomes at SDN Kemayoran 2 Bangkalan. The research employed a quasi-experimental method with a nonequivalent control group design. The participants consisted of fourth-grade students divided into a control class and an experimental class. The instruments used included pretest and posttest assessments to measure students' understanding before and after the intervention, as well as data analysis using an independent samples t-test, normality test, and N-Gain test. The study results indicated an increase in the average learning outcome scores for both the control group (81.68) and the experimental group (86.63), with the latter showing a significantly greater improvement. The independent samples t-test yielded a significance value of 0.00 (< 0.05), confirming data were accepted. Furthermore, the normality test indicated that the data were normally distributed, as the significance value was ≥ 0.05. Additionally, the N-Gain test yielded an average score of 0.80 (high improvement category) and an effectiveness percentage of 80.41% (highly effective category). Therefore, the study concludes that the application of Artificial Intelligence in assessment gamification is proven effective and has a substantial impact on improving student learning outcomes at SDN Kemayoran 2 Bangkalan. This research implies that AI-based educational technologies can serve as innovative alternatives to support the learning process in elementary schools.

1 Introduction

The development of educational technology in the twenty-first century has brought significant changes to the field of education, affecting both the learning process and assessment systems. The rise of digital technologies has encouraged teachers to move beyond their traditional role as information transmitters and instead serve as facilitators and learning innovators who can integrate technology effectively. The use of technology in education can enhance the quality of learning by providing diverse, interactive, and easily accessible learning resources for students. With technological integration, learning can take place in a more flexible, personalized, and contextual manner aligned with students' needs [1].

One rapidly advancing innovation is the application of Artificial Intelligence (AI) in education. AI presents vast opportunities to develop more adaptive, personalized, and efficient learning and assessment systems. AI can analyze student performance data, provide instant feedback, and even recommend learning strategies tailored to individual abilities. The use of AI in the educational sector not only supports classroom learning processes but also transforms assessment systems by offering more intelligent, accurate, and responsive evaluation tools that align with students' progress [2].

Assessment, as a key component of the learning process, has also undergone major transformation. Assessment is no longer viewed solely as a tool for measuring final achievement but also as a means to improve the learning process through constructive feedback. Traditional assessments, such as written tests, are often perceived as monotonous and unengaging, which may diminish students' motivation to learn [3]. In fact, learning motivation plays a crucial role in students' success in achieving learning objectives. Motivation serves as the primary driver that determines the intensity, direction, and persistence of students' learning activities [4].

Therefore, there is a need for a more innovative and engaging assessment approach that aligns with the characteristics of today's learners, who are closely connected to digital technology. One emerging approach is assessment gamification. Gamification involves the incorporation of game mechanics, such as points, levels, badges, leaderboards, and rewards, into non-game environments like education. [5]. Through gamification, assessment process is no longer perceived as a stressful activity but instead becomes an enjoyable, competitive, and stimulating learning experience.

* Corresponding author: ika.rahmawati@trunojoyo.ac.id

Research by Nurfadilah and Rahman (2022) shows that gamification can increase student engagement, foster intrinsic motivation, and provide more meaningful learning experiences compared to conventional assessments [6].

Furthermore, the integration of gamification with Artificial Intelligence (AI) opens opportunities for more interactive, adaptive, and personalized forms of assessment. AI enables gamified systems to deliver questions that match students' ability levels, automatically analyze their learning progress, and provide immediate recommendations for improvement. As a result, students can learn through more personalized experiences while being continually encouraged to enhance their performance. This aligns with the perspective of Deterding et al. (2011), who assert that gamification supported by technology can increase students' cognitive and emotional engagement in the learning process [7].

Based on this background, the present study was conducted at SDN Kemayoran 2 Bangkalan to examine the effectiveness of AI-based gamified assessment in improving elementary students' learning outcomes. The effectiveness of the intervention was measured using the N-Gain test, which indicates the level of improvement by comparing students' pretest and posttest scores [8-9]. This study is expected to contribute to the development of innovative, technology-based assessment strategies that support the enhancement of primary education quality in Indonesia

2 Research methods

This research utilized a quantitative methodology with a Quasi-Experimental Design, specifically utilizing the Nonequivalent Control Group Design model.. The research was conducted in classes 4A and 4B at SDN Kemayoran 02 Bangkalan, located on Jl. Moh. Toha, Kemayoran, Bangkalan District, Bangkalan Regency, East Java 69116. Class 4A and 4B each consisted of 19 students. The study was carried out in early September 2025.

The research procedure followed three key stages: (1) preparation, (2) implementation, and (3) analysis (Arikunto, 2019). The data collection instruments included pretest and posttest results as well as student response questionnaires. The final phase consisted of analyzing the data through the Kolmogorov–Smirnov normality test, independent samples t-test, and N-Gain analysis to assess the intervention's effectiveness in enhancing students' comprehension and learning results. All analyses were conducted using SPSS version 20 [10].

3 Results and discussion

The implementation of this study consisted of three phases: planning, implementation, and data analysis.

3.1 Preparation stage

During the preparation stage, the researchers requested permission from the school principal by submitting an official letter of introduction from the university, which served as documentation for conducting the study at the school. Upon receiving approval and conducting preliminary observations, the researchers identified a key issue: teachers' limited motivation and knowledge regarding the use of Artificial Intelligence (AI) in gamified assessments. In response, the researchers determined an effective approach, testing the effectiveness of a gamified assessment platform for classroom evaluation. Subsequently, the researchers collaborated with the homeroom teachers to organize the research materials and schedule the implementation.

3.2 Implementation stage

In this stage, the researchers developed a pretest instrument to measure students' initial understanding of the material. Afterward, a posttest was designed to evaluate students' learning outcomes following the implementation of AI-based gamified assessment. Once the intervention and trial were completed, the researchers proceeded with data processing.

3.3 Data analysis stage

The data analysis of the learning outcomes obtained from the implementation of AI-based gamified assessment was conducted through the following procedures:

3.3.1 Pretest and posttest scores

Based on the SPSS results for mean, median, and mode, both the control and experimental groups experienced an increase in average learning outcomes. The average score for the control group was 81.68, whereas the experimental group scored 86.63. Therefore, it can be concluded that both groups exhibited improvement in learning outcomes, with the experimental group showing higher gains among Grade 4A and 4B students at SDN Kemayoran 02 Bangkalan.

3.3.2 Independent samples T-test

An independent samples t-test was performed to check for any significant differences in learning outcomes between the control and experimental groups. The interpretation criteria are as follows:
1. H0 is accepted and H1 is rejected if $p > 0.05$.
2. H0 is rejected and H1 is accepted if $p < 0.05$.
The following are the results of the **independent sample t-test**.

From the results of the independent samples t-test comparing the learning outcomes of both groups, the significance value was 0.00, which is less than 0.05. This suggests that the independent sample t-test regarding learning outcomes between the experimental group and control group is accepted.

Table 1. Statistics.

		Posttest (Control)	Prettest (Control)	Prettest (Experimental)	Posttest (Experimental)
N	Valid	19	19	19	19
	Missing	1	1	1	1
Mean		81.6842	55.1053	58.6316	86.6316
Median		87.0000	60.0000	60.0000	89.0000
Mode		93.00	60.00	53.00	80.00
Std. Deviation		15.54939	16.76933	17.42989	7.26644
Variance		241.784	281.211	303.801	52.801
Minimum		50.00	27.00	27.00	80.00
Maximum		100.00	87.00	87.00	100.00
Sum		1552.00	1047.00	1114.00	1646.00

Table 2. Independent sample test.

		Levene's Test for Equality of Variances		t-test for Equality of Means						
		F	Sig.	t	F	Sig.	Mean Difference	F	Sig. Lower	Upper
NILAI	Equal variances assumed	5.278	.028	-7.840	36	.000	-27.947	3.565	-35.177	-20.718
	Equal variances not assumed			-7.840	26.111	.000	-27.947	3.565	-35.273	-20.622

3.3.3 Normality test

The aim of the normality test is to assess whether the data follows a normal distribution. Data is considered normally distributed if the computed significance value is above the 5% alpha level (Significance > 0.05). Below are the results of the normality test:

Table 3. Tests of normality

Group	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Prettest (Control)	.194	19	.059	.941	19	.272
Posttest (Control)	.227	19	.011	.912	19	.082
Prettest (Experimental)	.215	19	.021	.904	19	.058
Posttest (Experimental)	.199	19	.046	.851	19	.007

a. Lilliefors Significance Correction

The normality test above was conducted using SPSS 20 with the Shapiro–Wilk method because the sample size was ≤ 30; if the sample size were ≥ 30, the Kolmogorov–Smirnov test would have been used instead. The results showed significance values for the control class, with a pretest significance value of 0.272 and a posttest significance value of 0.082. The experimental class obtained a pretest significance value of 0.058 and a posttest significance value of 0.007. Based on these results, it can be concluded that the data in this study are normally distributed because the significance values obtained are ≥ 0.05.

3.3.4 N-Gain test

he N-gain test evaluates the degree to which an instructional method can enhance student learning outcomes by analyzing the difference between pretest and posttest scores. The following criteria can be used to categorize improvement based on the N-gain score:

Table 4. Normalized gain criteria.

N-Gain Value	Interpretation
$0.70 \leq g \leq 100$	High
$0.30 \leq g \leq 0.70$	Medium
$0.00 < g < 0.30$	Low
$g = 0.00$	No improvement
$-1.00 \leq g \leq 0.00$	Decrease

Source: https://doi.org/10.1119/1.18809

In addition, to assess the level of effectiveness of the intervention, the criteria presented below maybe used as reference:

Table 5. Criteria for determining the level of effectiveness.

Percentage%	Interpretation
< 40	Not effective
40-55	Less effective
56-75	Moderately effective
>76	Effective

Source. https://doi.org/10.1119/1.18809

Table 6. Descriptive statistics.

	N	Minimum	Maximum	Mean	Std. Deviation
N-Gain_score	19	.50	1.00	.8042	.15199
N-Gain_persen	19	50.00	100.00	80.4197	15.19904
Valid N (listwise)	19				

Based on the results of the N-gain analysis, it was identified that there was an improvement in students' understanding through the application of artificial intelligence in gamified assessments to measure the learning outcomes of students at SDN Kemayoran 02 Bangkalan. The average N-gain score obtained from the pretest and posttest was 0.80, which falls within the range of $0.70 \leq g \leq 1.00$. This indicates a high level of improvement. Furthermore, this study evaluated the effectiveness of applying artificial intelligence in gamified assessments to measure the students' learning outcomes at SDN Kemayoran 2 Bangkalan using the criteria outlined in Table 5. The average percentage N-gain score was 80.41, which corresponds to the category of > 76. Therefore, it can be concluded that the application of artificial intelligence in gamified assessments is “highly effective” in improving the learning outcomes of students at SDN Kemayoran 02 Bangkalan. The average percentage N-gain score was 80.41, which corresponds to the category of > 76. Therefore, it can be concluded that the application of artificial intelligence in gamified assessments is “highly effective” in improving the learning outcomes of students at SDN Kemayoran 02 Bangkalan.

4 Conclusion

Based on the research conducted through three main phases: preparation, implementation, and data analysis, the following conclusions can be drawn:

1. Preparation Phase: The initial phase revealed that the main issue was the teachers’ low motivation and limited understanding of Artificial Intelligence (AI), particularly in the application of AI-based gamified assessments. To address this issue, the researchers implemented a strategy by testing the effectiveness of an AI-based gamified assessment platform in the learning process.
2. Implementation Phase: This phase involved the development of pretest and posttest instruments to measure differences in students’ understanding before and after the application of AI in gamified assessments. The collected data were then processed using statistical tests to determine the significance of the learning outcomes.
3. Data Analysis Phase: This phase yielded several important findings: • •. • •. Consequently, the implementation of Artificial Intelligence in gamified assessments was shown to be effective in enhancing the learning outcomes of fourth-grade students:
   - The average learning outcomes improved in both the control class (81.68) and the experimental class (86.63), showing a notable difference between the groups, with the experimental class achieving superior results.
   - The Normality Test indicated that the data was normally distributed, as the significance values were ≥ 0.05.
   - The Independent Sample t-test revealed a significance value of 0.00 (< 0.05), leading to the conclusion that the differences in learning outcomes between the experimental and control classes were acknowledged..
   - The N-Gain Test recorded an average score of 0.80 (indicating high improvement) and an effectiveness percentage of 80.41% (classified as highly effective).

Thus, the implementation of Artificial Intelligence in gamified assessments was shown to be effective in enhancing the learning outcomes of fourth-grade students at SDN Kemayoran 02 Bangkalan. This study demonstrates that technology-driven educational innovations can serve as a solution to low teacher and student motivation, while significantly enhancing both the learning process and learning outcomes.

The research team would like to express our deepest gratitude to all parties who supported and contributed to the completion of this research project, especially:
1. The Rector of Universitas Trunodjoyo Madura
2. The Head of LPPM, Universitas Trunodjoyo Madura
3. The Principal of SDN Kemayoran 02 Bangkalan
4. The Teachers of SDN Kemayoran 02 Bangkalan

May all assistance extended to the research team will be rewarded abundantly by Allah SWT. Finally, we hope that this report will be beneficial to all parties who make use of it.

References

1. S. Siahaan, Pemanfaatan teknologi informasi dalam pembelajaran abad 21. Jurnal Pendidikan dan Teknologi. **5**, 12–20 (2020)
2. B. Prasetyo, Artificial intelligence in education: Opportunities and challenges. Jurnal Teknologi Pendidikan. **22**, 101–112 (2020)
3. S. Arikunto, Prosedur penelitian: Suatu pendekatan praktik. Rineka Cipta. (2019)
4. A. M. Sardiman, Interaksi & motivasi belajar-mengajar. Rajawali Pers. (2018)
5. Z. Zainuddin, Gamification in education: A systematic review. Journal of Educational Multimedia. **4**, 88–102 (2021)
6. F. Nurfadilah, A. Rahman, Gamification in assessment: Increasing motivation and engagement in learning. Journal of Education and Instruction. **3**, 45–53 (2022)
7. S. Deterding, D. Dixon, R. Khaled, L. Nacke, From game design elements to gamefulness: Defining gamification. Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments. 9–15 (2011). https://doi.org/10.1145/2181037.2181040
8. R. R. Hake, Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses. American Journal of Physics. **66**, 64–74 (1998). https://doi.org/10.1119/1.18809
9. R. Sundayana, Statistika penelitian pendidikan. Alfabeta. (2016)
10. S. Santoso, Panduan lengkap SPSS versi 20. PT Elex Media Komputindo. (2020)
11. E. L. Deci, R. M. Ryan, Self-determination theory: Basic psychological needs in motivation, development, and wellness. Guilford Press. (2017)
12. I. Fathurrahman, Penerapan SPSS untuk analisis statistik penelitian pendidikan. Jurnal Statistik Edukasi. **4**, 19–30 (2020)

13. V. Guntoro, R. Putra, Gamification strategy to improve elementary students' motivation. International Journal of Learning Strategies. **5**, 201–214 (2022)
14. A. Hamzah, Penggunaan asesmen digital dalam pembelajaran daring. Jurnal Inovasi Teknologi Pendidikan. **8**, 25–34 (2021)
15. W. Holmes, M. Bialik, C. Fadel, Artificial intelligence in education: Promises and implications for teaching and learning. Center for Curriculum Redesign. (2019)
16. K. Kim, S. Park, Artificial intelligence-based assessment tools in primary education. Computers & Education. **170**, 104–226 (2021)
17. S. Li, C. Tsai, The effect of AI-supported assessment on students' learning performance. Journal of Educational Computing Research. **59**, 675–699 (2021)
18. R. E. Mayer, Evidence-based learning and teaching: A look at cognitive principles. Cambridge University Press. (2019)
19. T. Mulyani, Pembelajaran abad 21 dan tantangan teknologi. Jurnal Ilmiah Pendidikan Dasar. **7**, 98–110 (2020)
20. F. Rahman, Gamification and AI integration in classroom assessment. Journal of Digital Learning and Education. **3**, 77–89 (2020)
21. L. Rahmawati, Pengaruh gamifikasi terhadap motivasi belajar siswa sekolah dasar. Jurnal Pendidikan Dasar Nusantara. **7**, 157–166 (2021). https://doi.org/10.33487/mgr.v2i1.1731
22. I. Sainuddin, M. Rahmat, Gamified learning and its impact on students' achievement and motivation. International Journal of Instructional Technology. **4**, 55–66 (2021)
23. Sugiyono, Metode penelitian kuantitatif, kualitatif, dan R&D. Alfabeta. (2019)
24. S. Wulandari, Efektivitas model pembelajaran berbasis teknologi terhadap hasil belajar siswa. Jurnal Pendidikan Modern. **6**, 44–52 (2021)
25. A. Wijaya, Teknologi dan asesmen digital pada pendidikan dasar. Jurnal Teknologi Pendidikan Indonesia. **9**, 55–63 (2021)
26. A. Yudianto, S. Sumarni, AI-based learning analytics in elementary education. Journal of Smart Learning Environments. **9**, 1–12 (2022)