Jurnal Inovasi dan Teknologi Pembelajaran
Vol 13, No 2, (2026), page 102-112
https://doi.org/10.17977/um031v13i22026p102
P-ISSN: 2406-8780 E-ISSN: 2654-7953
Open access: https://journal-fip.um.ac.id/index.php/jinotep/index

Establishing pre-intervention baseline equivalence for AI-supported gamified project-based learning in entrepreneurship education

Silvia Anita Prastica[a], Agus Wedi[b], Fikri Aulia[b], Made Duananda Kartika Degeng[b], Isma Addi Jumbri[c]

[a]*Master's Program in Educational Technology, Faculty of Education, Universitas Negeri Malang*
[b]*Department of Educational Technology, Faculty of Education, Universitas Negeri Malang*
[a,b]*Semarang Street No. 5, Sumbersari, Lowokwaru District, Malang City, East Java 65145, Indonesia*
[c]*Faculty of Technology Management and Technopreneurship, Universiti Teknikal Malaysia*
[c]*76100 Durian Tunggal, Melaka, Malaysia*
**Corresponding author, e-mail: silvia.anita.2401218@students.um.ac.id*

ARTICLE INFO

Article history:
Received: 21-05-2026
Revised 23-07-2026
Accepted: 13-08-2026

Keywords:
Baseline equivalence; Project-Based Learning; Gamification; Artificial Intelligence in education; Entrepreneurship education

This is an open access article under the Creative Commons Attribution-ShareAlike 4.0 International license.

Copyright © 2026 by Authors. Published by Universitas Negeri Malang.

ABSTRACT

Establishing pre-intervention baseline equivalence is essential for strengthening the internal validity of quasi-experimental studies. However, baseline comparability is often underreported in educational technology research, particularly in evaluations of emerging AI-supported instructional innovations. This study addresses this methodological gap by examining baseline comparability before implementing a newly developed AI-supported gamified Project-Based Learning (PjBL) model in entrepreneurship education. The novelty of the study lies in explicitly positioning baseline assessment as a methodological foundation for strengthening the interpretation of intervention effects in AI-enhanced learning research. A quasi-experimental design was employed involving an experimental group (n = 30) and a control group (n = 26). Pre-test data were analysed using descriptive statistics and the Mann-Whitney U test because the data were not normally distributed. The experimental and control groups recorded comparable mean pre-test scores of 85.33 and 85.19, respectively. No statistically significant baseline difference was found between the groups (U = 303.00, Z = −1.467, p = .142, r = .20). These findings support the baseline comparability of the two groups and provide a stronger methodological basis for interpreting subsequent differences in learning outcomes as being associated with the intervention rather than substantial pre-existing group differences.

INTRODUCTION

The rapid advancement of digital technology has transformed higher education by promoting student-centered, competency-based learning environments supported by emerging technologies, including artificial intelligence (AI) (Maphosa & Maphosa, 2023). This transformation is particularly vital in entrepreneurship education, where students must develop creativity, innovation, critical thinking, collaboration, and problem-solving skills alongside conceptual knowledge.

Project-Based Learning (PjBL) addresses these limitations by engaging students in projects that connect theoretical knowledge with real-world applications. Previous studies have demonstrated that PjBL enhances critical thinking, creativity, communication, collaboration, problem-solving skills, and learning achievement in higher education (Kokotsaki et al., 2016; Guo et al., 2020). Therefore, PjBL provides a relevant pedagogical foundation for integration with AI and gamification to support more engaging and meaningful entrepreneurship learning.

In entrepreneurship education, PjBL is particularly relevant because it allows students to develop business ideas, conduct market analyses, design products or services, and simulate real entrepreneurial processes (Pratama & Mulyana, 2021). This experiential learning approach bridges the gap between theory and practice, enabling students to apply conceptual knowledge in authentic and meaningful contexts. Nevertheless, the effectiveness of PjBL depends considerably on students' motivation, participation, collaboration, and readiness to manage complex project activities (Zen et al., 2022).

Previous studies have identified several challenges associated with PjBL, including inconsistent motivation and engagement throughout the learning process (Makkonen et al., 2021). Students may also experience difficulties in managing time, distributing tasks, and collaborating effectively in group projects. Unequal participation is another common concern, as some students may dominate project activities while others remain passive contributors (Hussein, 2021). These challenges can undermine the quality of collaboration and create differences in the learning experiences of individual group members (Lavoue et al., 2021).

Variations in students' prior knowledge and initial academic ability may further influence their capacity to participate effectively in project-based activities. Students with stronger prior knowledge may adapt more quickly to technology-supported learning, whereas those with lower initial abilities may struggle to understand tasks, contribute to group discussions, or complete assigned activities (Rafique et al., 2021). Consequently, differences observed after an intervention may reflect pre-existing disparities rather than the effects of the instructional model itself. This issue highlights the importance of assessing baseline ability before evaluating the effectiveness of PjBL-based interventions.

To address motivational and engagement challenges, educational technology research has increasingly explored gamification as a complementary instructional strategy. Gamification refers to the integration of game design elements such as points, badges, levels, leaderboards, and rewards into non-game educational contexts. The application of gamification in education has been shown to increase intrinsic motivation, strengthen student engagement, and create more enjoyable learning experiences (Hamari et al., 2014; Rivera et al., 2021; Cheng et al., 2021). In entrepreneurship education, gamification may support sustained participation, provide visible indicators of progress, and encourage students to complete project activities through structured challenges and feedback mechanisms (Huang & Wu, 2021). Alongside gamification, the emergence of artificial intelligence (AI), particularly generative AI, has created new possibilities for transforming higher education (Winkler et al., 2023). AI technologies are increasingly being integrated into learning environments to support personalised learning, adaptive instruction, and data-driven educational decision-making. Recent studies highlight that AI can enhance creativity, innovation, and decision-making processes in entrepreneurship education (Bell & Bell, 2023; Park et al., 2025). Moreover, AI-based systems can analyse student learning behavior, provide real-time feedback, and adapt learning content to individual learner needs (Zawacki-Richter et al., 2019; Crompton & Burke, 2023; Maier & Klotz, 2022). These capabilities make AI particularly relevant for supporting students with different levels of prior knowledge and learning readiness (Yang, 2022). The integration of AI and gamification into PjBL may provide complementary pedagogical benefits (Chen & Wang, 2022). PjBL offers authentic and collaborative learning experiences, gamification supports motivation and sustained participation, and AI enables personalised feedback and instructional adaptation. Together, these approaches may create a more responsive learning environment in which students are guided according to their performance while remaining actively involved in entrepreneurial projects.

Despite the growing interest in integrating PjBL, gamification, and AI in entrepreneurship education, previous intervention studies have primarily examined post-intervention outcomes, including learning achievement, engagement, motivation, and satisfaction. Baseline comparability is nevertheless decisive in quasi-experimental research because pre-existing group differences may provide an alternative explanation for observed intervention effects. Methodological guidance recommends assessing baseline equivalence when evaluating quasi-experimental studies, as pre-existing group differences may threaten the internal validity of intervention effects (Shadish et al., 2002; What Works Clearinghouse, 2022). Evidence from recent reviews on educational intervention indicates that baseline equivalence is not universally established or explicitly reported across quasi-experimental studies. For example, Bengochea and Sembiante (2024) found that baseline equivalence was demonstrated in most but not all of the 12 quasi-experimental studies included in their best-evidence synthesis, highlighting the importance of transparently assessing and reporting baseline comparability. Therefore, the present study examines the baseline pre-intervention comparability before evaluating an AI-supported gamified PjBL model in entrepreneurship education.

This methodological limitation is important because the effectiveness of an instructional intervention cannot be determined accurately when the experimental and control groups differ substantially at baseline. Without comparable initial conditions, observed post-intervention differences may be influenced by students' prior knowledge or academic ability rather than by the intervention. Experimental validity, therefore, depends partly on demonstrating that the groups are sufficiently equivalent before treatment implementation, or instructional intervention, and on controlling for pre-existing differences among students. As emphasised by Creswell (2014). However, baseline-equivalence analysis is frequently underreported in educational technology research, particularly in studies involving emerging AI-supported instructional models.

The limited attention given to pre-intervention conditions also restricts the development of adaptive learning strategies. Understanding students' initial abilities is necessary not only for establishing group equivalence but also for identifying differences in learning readiness that may require differentiated support. Baseline mapping can therefore provide both methodological evidence for experimental validity and pedagogical information for designing appropriate AI-supported learning activities.

To address these gaps, this study develops an AI-supported gamified PjBL model for entrepreneurship education by integrating constructivist pedagogy, project-based entrepreneurial activities, gamification mechanisms, and adaptive AI support within a unified instructional framework. Unlike previous studies that have generally examined PjBL, gamification, or AI separately, this study combines these components systematically to support personalised feedback, student engagement, project completion, and entrepreneurial competency development. Its novelty, therefore, lies not only in incorporating pre-intervention baseline equivalence into the evaluation process, but also in developing and evaluating an integrated pedagogical model specifically designed for AI-enhanced entrepreneurship education.

As an initial phase in evaluating the proposed framework, the recent study focuses specifically on students' pre-intervention abilities. It analyses and maps students' initial performance in entrepreneurship-learning and examines whether the experimental and control groups demonstrate comparable baseline conditions. Establishing this equivalence is necessary before subsequent analyses can attribute changes in learning outcomes to the AI-supported gamified PjBL intervention with greater confidence. Therefore, this study aims to establish pre-intervention baseline comparability before implementing the newly developed AI-supported gamified PjBL model in entrepreneurship education. Accordingly, it addresses the following research question: Is there a statistically significant difference in pre-intervention entrepreneurial learning abilities between the experimental and control groups?

METHOD

This study employed a quasi-experimental pre-intervention design with a baseline equivalence assessment before implementing an AI-supported gamified PjBL model in

entrepreneurship education (Cohen et al., 2018). The primary objective was to ensure statistical comparability between the groups before the treatment.

The research participants consisted of two existing undergraduate entrepreneurship classes at a higher education institution. Two existing entrepreneurship classes were selected as intact groups based on their enrollment in the same course during the study period. All students in the selected classes who met the study eligibility criteria were invited to participate. One class comprising 30 students was assigned to the experimental group, while the other class of 26 students was in the control group. Since the classes had been established before the study according to the university's academic schedule, individual random assignment was not feasible. Therefore, this study employed a quasi-experimental design using intact groups. Both groups followed the same entrepreneurship curriculum, were taught during the same academic semester, and had comparable academic characteristics, making them appropriate for baseline equivalence analysis before the intervention. This research was conducted in accordance with established ethical principles by ensuring participant confidentiality, protecting respondents' identities, and excluding personally identifiable information from the reporting of the findings.

The research instrument was a pre-test designed to measure students' basic capabilities in entrepreneurship learning. The pre-test instrument covered students' understanding of fundamental entrepreneurship concepts, basic analytical skills, and contextual problem-solving abilities relevant to the learning materials. The test was administered to all students before the instructional intervention, ensuring that the results reflected the students' baseline conditions and were not influenced by the instructional model. Before the data collection, the research instruments were reviewed by three experts with relevant expertise in entrepreneurship education, instructional design, and educational technology as part of the overall validation of the developed instructional products. The experts evaluated the relevance, clarity, appropriateness, and alignment of the instrument items with the learning objectives using a structured validation questionnaire. The validation score was calculated by dividing the total score obtained by the maximum possible score and multiplying the result by 100%, which was categorised as highly valid based on the validation criteria used in this study. In addition, the overall validation of the developed products, including the instructional model, learning materials, instruments, and LMS, achieved an average validity score of 99.11%, indicating that the instruments and developed instructional products were considered suitable for implementation in the study.

The present article reports only the pre-intervention phase of a larger quasi-experimental study designed to evaluate a newly developed AI-supported gamified Project-Based Learning (PjBL) model. During the baseline phase, both the experimental and control groups underwent identical research procedures. All participants completed the same entrepreneurship pre-test before any instructional intervention was introduced, ensuring that the scores reflected their initial abilities without being influenced by the instructional model. Following the baseline assessment, the experimental group received instruction using the AI-supported gamified PjBL model, which integrated project-based learning activities, gamification elements, and AI-assisted learning support. In contrast, the control group received entrepreneurship instruction through the institution's conventional teaching approach using the same course content and learning objectives but without the integration of AI, gamification, or the proposed PjBL model. The implementation and evaluation of these instructional interventions, however, are beyond the scope of the present article and will be reported in a subsequent study.

This study was conducted in sequential stages, including group assignment, pre-test administration, data collection, normality testing, and inferential statistical analysis. The first stage involved determining the experimental and control groups. The second stage involved administering the pre-test to all students to obtain baseline capability data. The third stage entailed collecting and processing the pre-test data. Subsequently, the data were analysed using descriptive and inferential statistics to examine the distribution of students' initial capabilities and to test for equivalence in baseline capabilities between the groups. For the overall research procedure, see Figure 1.

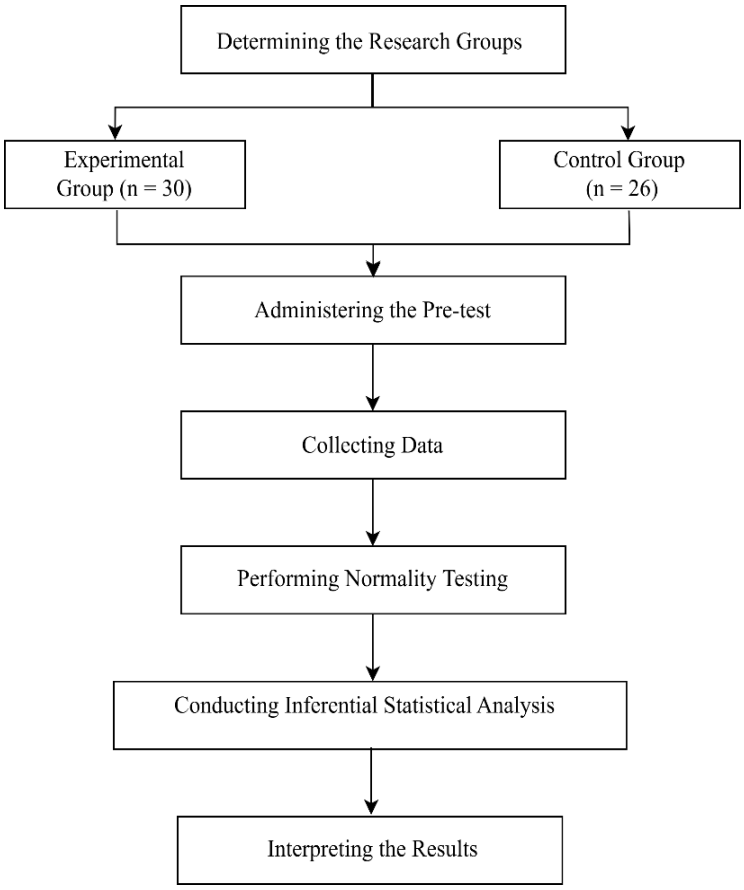


Figure 1. Research procedure

The data analysis was conducted quantitatively, utilising both descriptive and inferential statistics. Descriptive statistics were used to describe the distribution of students' initial capabilities within each group, including the mean, minimum, maximum, and variance. The mean value was calculated using the Formula 1.

$$\overline{X} = \frac{X}{N} \quad (1)$$

Information:
$\overline{X}$ = the Mean Score
X = the Total Scores
N = the Sample Size of Each Group

In addition to descriptive statistics, this study utilised inferential analysis to test the equivalence of initial capabilities between the experimental and control groups. Before hypothesis testing, the data were tested for normality using the Shapiro–Wilk test. Testing distributional assumptions is vital for selecting an appropriate statistical procedure when assessing baseline comparability between research groups (Wang et al., 2022). The results indicated that the data were not normally distributed, as evidenced by a p-value of less than 0.05. Therefore, the analysis proceeded using the non-parametric Mann–Whitney U test to compare the initial capabilities between the two groups.

The Mann–Whitney U test was selected because it is appropriate for comparing two independent groups with non-normal distributions. The decision criterion was determined by comparing the p-value with the 0.05 significance level. If the p-value was greater than 0.05, it was considered that there was no statistically significant difference in initial capabilities between the two groups. Conversely, if the p-value was less than 0.05, it was considered that a statistically significant difference in initial capabilities exists between the two groups. Through

this analytical procedure, this study sought to ensure that the experimental and control groups had equivalent baseline conditions before implementing the proposed model. Consequently, the research findings in the subsequent experimental phase can be interpreted more objectively, as differences in subsequent learning outcomes could be more confidently associated with the implemented learning model.

RESULT

This section presents the research findings based on students' pre-test data from both the experimental and control groups before the implementation of the proposed model. The findings focus on analysing the students' initial capabilities to ensure comparability between the experimental and control groups. Baseline abilities were measured using a pre-test instrument administered to both groups, and the results of the descriptive statistical analysis are presented in Table 1.

Table 1 presents the descriptive statistics of pre-test scores for both groups. The experimental group (M = 85.33) and the control group (M = 85.19) showed a negligible mean difference (Δ = 0.14), indicating similar initial abilities between the groups. In terms of score distribution, the experimental group showed greater variability (range: 35–100) compared with the control group (range: 65–90). This finding suggests greater heterogeneity in the students' prior abilities within the experimental class. Such variation reflects differences in students' initial learning readiness and may inform the design of adaptive instruction in Project-Based Learning environments.

Before the inferential analysis, a Shapiro–Wilk normality test was conducted to assess the data distribution. The results are presented in Table 2. Table 2 indicates that the significance values for both groups are below 0.05; thus, the data were considered non-normally distributed. Consequently, the analysis proceeded using the Mann–Whitney U test as a non-parametric alternative to the independent sample t-test. The results of the baseline equivalence testing are shown in Tables 3 and 4.

Table 1. Descriptive statistics of pre-test scores

Class	Sample (N)	Minimum Score	Maximum Score	Mean
Experimental Class Pre-Test	30	35	100	85.33
Control Class Pre-Test	26	65	90	85.19

Table 2. Shapiro–Wilk normality test results

Class	Sig. (p-value)	Description
Experimental Class Pre-Test	0.000	Non-Normal
Control Class Pre-Test	0.000	Non-Normal

Table 3. Mann–Whitney test ranks

Class	N	Mean Rank	Sum of Ranks
Experimental Class Pre-Test	30	31.40	942.00
Control Class Pre-Test	26	25.15	654.00
Total	**56**		

Table 4. Mann–Whitney test statistics

Statistic	Value
Mann–Whitney U	303.000
Wilcoxon W	654.000
Z	-1.467
Asymp. Sig. (2-tailed)	0.142
Effect Size (r)	**0.20**

Based on Table 4, the Mann–Whitney U test yielded a non-significant result (U = 303.00, Z = −1.467, p = .142), indicating no statistically significant difference between the pre-test scores of the experimental and control groups. The calculated effect size was small (r = .20), suggesting that the magnitude of the observed difference was limited (Cohen, 1988). Although the experimental group obtained a slightly higher mean score (M = 85.33) than the control group (M = 85.19), the difference was not statistically significant and was associated with only a small effect. These findings support the baseline comparability of the two groups before the implementation of the proposed AI-supported gamified PjBL model.

DISCUSSION

The results of this study indicate that there was no statistically significant difference in students' initial abilities between the experimental and control groups before the implementation of the instructional model. The Mann–Whitney U test result (p = 0.142 > 0.05) supports this conclusion, suggesting that both groups had comparable baseline conditions. Moreover, the average scores of the experimental group (M = 85.33) and the control group (M = 85.19) indicate only a slight difference (Δ = 0.14), which further reinforces the conclusion that both groups began with similar initial abilities.

This equivalence has methodological significance in quasi-experimental research because it increases confidence that subsequent differences in learning outcomes are associated with the instructional treatment rather than with substantial pre-existing differences between the groups. Creswell (2014) highlighted that baseline equivalence is crucial for preserving internal validity in experimental designs. Consequently, the validation of initial equivalence in this study provides an appropriate foundation for the subsequent evaluation of the AI-supported gamified PjBL framework. While the average scores were nearly identical, the descriptive findings indicate variation in the distribution of scores. The experimental group showed a broader range of scores (35–100) than the control group (65–90), suggesting greater variability in student abilities. This variation reflects differences in students' prior knowledge and learning readiness. In comparison, the control group demonstrated a more consistent distribution, indicating more uniform initial ability levels.

The descriptive results indicate variability in students' initial competencies, particularly within the experimental group, suggesting heterogeneous learning needs. This condition is relevant in Project-Based Learning environments, where collaborative tasks require balanced participation and coordination among learners. Such diversity highlights the importance of adaptive instructional support systems that can accommodate differences in students' readiness and ability levels.

In this context, Artificial Intelligence can function as a mechanism for adaptive learning through personalised feedback and dynamic task adjustment based on learner performance (Zawacki-Richter et al., 2019; Crompton & Burke, 2023). Similarly, gamification strategies may help sustain engagement across different ability levels by providing motivational support, such as rewards and progressive challenges (Hamari et al., 2014; Rivera et al., 2021). When integrated with PjBL, these approaches offer a complementary framework for addressing both cognitive and motivational diversity in entrepreneurship education.

Previous studies further indicate that variation in learner ability influences group dynamics in project-based settings. Students with higher prior knowledge tend to take more active roles in collaborative tasks, while those with lower competencies may experience difficulties in participation and interaction (Makkonen et al., 2021; Hussein, 2021). This reinforces the need for instructional designs that are responsive to diverse learner profiles and that provide equitable opportunities for participation and learning.

Statistical evaluation also supports the appropriateness of the analytical procedure. The Shapiro–Wilk test revealed that neither dataset followed a normal distribution ($p < 0.05$), supporting the choice of a non-parametric test. Thus, the Mann–Whitney U test was used because it is suitable for comparing two independent groups when the data are not normally distributed (Mishra et al., 2019). This ensured that the statistical analysis was appropriate for the data's distributional characteristics.

Additionally, employing intact groups without individual randomisation is a common feature of educational research conducted in classroom settings. The validation of statistical equivalence helps reduce concerns about selection bias and pre-existing group differences arising from the absence of random assignments. This enhances the study's methodological credibility and supports the application of quasi-experimental designs in authentic educational settings where randomisation is often impractical. From an educational perspective, these results emphasise the need to assess students' baseline abilities before introducing technology-supported instructional approaches. In entrepreneurship education, where results hinge on creativity, teamwork, and problem-solving, establishing comparable starting conditions is crucial for more accurately evaluating instructional effectiveness. Baseline mapping enables educators to gain insight into student diversity and to develop instructional strategies that respond to varying levels of readiness.

Additionally, the greater variability in the experimental group highlights the necessity for flexible instructional design. AI-supported systems can analyse students' learning behaviours and provide immediate, tailored feedback (Koedinger et al., 2022). This aligns with the tenets of adaptive learning, which aim to adjust learning experiences to learner characteristics and performance. Likewise, gamified educational settings promote ongoing engagement and motivation among students with different ability levels, thereby supporting more inclusive participation.

Overall, the results demonstrate that the experimental and control groups had statistically comparable baseline abilities, although the experimental group exhibited slightly greater variability in score distribution. These findings provide a robust methodological foundation for the subsequent experimental phase evaluating the effectiveness of the newly developed AI-supported gamified PjBL model. Establishing baseline comparability is particularly noteworthy in educational technology research because AI-supported and adaptive learning environments may personalise feedback, content, and task difficulty according to learners' initial abilities and performance. Unlike less adaptive educational interventions, differences in students' starting abilities may directly influence how they interact with and benefit from the technology. Without comparable baseline conditions, subsequent learning gains may reflect pre-existing differences rather than the effects of the instructional technology itself. Therefore, establishing baseline comparability strengthens the internal validity of the study and supports more credible causal interpretations of the intervention effects. Beyond its methodological contribution, this study highlights the importance of systematically assessing learners' initial abilities when evaluating AI-enhanced instructional innovations in higher education.

CONCLUSION

This study contributes methodologically by demonstrating the importance of establishing a pre-intervention baseline comparability as a prerequisite for evaluating a newly developed AI-supported gamified Project-Based Learning (PjBL) model in entrepreneurship education. Establishing baseline comparability between the experimental and control groups before the intervention strengthens the internal validity of the quasi-experimental design and provides a more robust basis for attributing subsequent differences in learning outcomes to the instructional intervention rather than to pre-existing group characteristics. These findings

underscore the importance of systematically assessing baseline comparability, particularly in studies involving AI-supported and adaptive instructional innovations, where learners' initial characteristics may influence their interaction with and response to technology. The principal contribution of this study, therefore, lies in providing a transparent methodological foundation for the subsequent evaluation of intervention effectiveness.

The equivalence of both groups provides a stronger foundation for evaluating the effectiveness of the proposed learning model in the next phase of the research. Because the groups started from comparable baseline conditions, the impact of the proposed model can be assessed more accurately, with reduced influence from pre-existing differences in students' initial abilities. Furthermore, this study highlights the importance of mapping students' initial capabilities when developing technology-supported learning environments. Such baseline data provide valuable evidence for designing adaptive learning strategies that are responsive to students' characteristics and levels of readiness throughout the instructional process.

Nevertheless, this study has several limitations. It focused exclusively on the pre-intervention phase and therefore did not evaluate post-intervention learning outcomes or the effectiveness of the proposed AI-supported gamified PjBL model. In addition, the study was conducted only in the Entrepreneurship course, which may limit the applicability of the findings to other courses and disciplinary contexts. Future research should evaluate the effectiveness of the AI-supported gamified PjBL model across diverse educational contexts by examining its impact on learning outcomes, student engagement, entrepreneurial competencies, and higher-order thinking skills. Future studies are also encouraged to investigate the model using larger and more diverse samples, different academic disciplines, and multi-institutional settings to examine the consistency and generalisability of its effectiveness while maintaining rigorous baseline comparability assessments.

Author contributions

The authors gratefully acknowledge the support and collaboration of Universitas Negeri Malang, Indonesia, and Universiti Teknikal Malaysia Melaka, Malaysia. The authors also express their sincere appreciation for the institutional assistance, academic support, and research facilities provided by both universities, which contributed to the successful completion of this study.

Author contributions

The authors made significant contributions to the study's conception and design. The authors were responsible for data analysis, interpretation, and discussion of results. The final manuscript was read and approved by the authors.

Funding

There was no specific grant for this research from any funding organisation in the public, private, or nonprofit sectors for this research.

Conflict of interest

The authors declare that there is no potential conflict of interest.

Data availability statement

All data are available from the authors.

REFERENCES

Bell, R., & Bell, H. (2023). Entrepreneurship education in the era of generative artificial intelligence. *Entrepreneurship Education, 6*(3), 229–244. https://doi.org/10.1007/s41959-023-00099-x

Bengochea, A., & Sembiante, S. F. (2024). Effective vocabulary interventions for young emergent bilinguals: A best-evidence synthesis. *Review of Education, 12*(1), 1-39. https://doi.org/10.1002/rev3.3458

Chen, L., & Wang, P. (2022). Gamification and artificial intelligence: A systematic mapping study. In *Proceedings of the International Conference on Learning and Collaboration Technologies* (pp. 102–115). Springer. https://doi.org/10.1007/978-3-031-05581-2_8

Cheng, V. W. S., Davenport, T., Johnson, D., Vella, K., & Hickie, I. B. (2021). Psychological interventions of virtual gamification within academic intrinsic motivation: A systematic review. *Journal of Affective Disorders*, 293, 444–465. https://doi.org/10.1016/j.jad.2021.06.070

Cohen, L., Manion, L., & Morrison, K. (2018). *Research methods in education* (8th ed.). Routledge. https://doi.org/10.4324/9781315456539

Creswell, J. W. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches* (4th ed.). Sage Publications. ISBN: 978-1-4522-2610-1

Crompton, H., & Burke, D. (2023). Artificial intelligence in higher education: The state of the field. *International Journal of Educational Technology in Higher Education, 20*(1), 1-22. https://doi.org/10.1186/s41239-023-00392-8

Guo, P., Saab, N., Post, L. S., & Admiraal, W. (2020). A review of project-based learning in higher education: Student outcomes and measures. *International Journal of Educational Research, 102*, 1-13. https://doi.org/10.1016/j.ijer.2020.101586

Hamari, J., Koivisto, J., & Sarsa, H. (2014). Does gamification work? A literature review of empirical studies. In *Proceedings of the 47th Hawaii International Conference on System Sciences* (pp. 3025–3034). IEEE. https://doi.org/10.1109/HICSS.2014.377

Huang, S. J., & Wu, C. H. (2021). Gamified project-based learning in higher education: Impacts on student engagement. *Computers & Education, 172*, 104256. https://doi.org/10.1016/j.compedu.2021.104256

Hussein, B. (2021). Addressing collaboration challenges in project-based learning: The student's perspective*. Education Sciences, 11*(8), 1-20. https://doi.org/10.3390/educsci11080434

Koedinger, K. R., McLaughlin, S., & Stamper, E. A. (2022). Automated feedback in AI-based learning systems. *International Journal of Artificial Intelligence in Education, 32*(3), 589–610. https://doi.org/10.1007/s40593-021-00270-5

Kokotsaki, M., Menzies, V., & Wiggins, A. (2016). Project-based learning: A review of the literature. *Improving Schools, 19*(3), 267–277. https://doi.org/10.1177/1365480216659733

Lavoue, E., Ju, Q., Hallifax, S., & Serna, A. (2021). Analyzing the relationships between learners' motivation and observable engaged behaviors in a gamified learning environment. *International Journal of Human-Computer Studies, 154*, 1-40. https://doi.org/10.1016/j.ijhcs.2021.102670

Maier, U., & Klotz, C. (2022). Personalized feedback in digital learning environments: Classification framework and literature review. *Computers and Education: Artificial Intelligence, 3*, 1-13. https://doi.org/10.1016/j.caeai.2022.100080

Maphosa, V., & Maphosa, M. (2023). Artificial intelligence in higher education: A bibliometric analysis and topic modeling approach. *Applied Artificial Intelligence, 37*(1), 1-23. https://doi.org/10.1080/08839514.2023.2261730

Makkonen, T., Tirri, K., & Lavonen, J. (2021). Engagement in learning physics through project-based learning: A case study of gifted Finnish upper-secondary-level students. *Journal for the Education of the Gifted, 44*(4), 247–274. https://doi.org/10.1177/1932202X211018644

Mishra, P., Pandey, C. M., Singh, U., Gupta, A., Sahu, C., & Keshri, A. (2019). Descriptive statistics and normality tests for statistical data. *Annals of Cardiac Anaesthesia, 22*(1), 67–72. https://doi.org/10.4103/aca.ACA_157_18

Park, J.-H., Kim, S.-J., & Lee, S.-T. (2025). AI and creativity in entrepreneurship education: A systematic review of LLM applications. *AI, 6*(5), 1-22. https://doi.org/10.3390/ai6050100

Pratama, R. P., & Mulyana, A. S. (2021). Implementasi model project-based learning pada pembelajaran kewirausahaan di perguruan tinggi. *Jurnal Pendidikan Bisnis dan Manajemen, 7*(1), 12–25. https://doi.org/10.17977/um003v7i12021p012

Rafique, G. M., Mahmood, K., Warraich, N. F., & Rehman, S. U. (2021). Readiness for online learning during COVID-19 pandemic: A survey of Pakistani LIS students. *The Journal of Academic Librarianship, 47*(3), 1-10. https://doi.org/10.1016/j.acalib.2021.102346

Rivera, E. S., & Garden, C. L. P. (2021). Gamification for student engagement: A framework. *Journal of Further and Higher Education, 45*(7), 999–1012. https://doi.org/10.1080/0309877X.2021.1875201

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). Experimental and quasi-experimental designs for generalized causal inference. Houghton Mifflin.

Wang, J., Li, X., & Liang, H. (2022). A new exact p-value approach for testing variance homogeneity. *Statistical Theory and Related Fields, 6*(1), 81–86. https://doi.org/10.1080/24754269.2021.1907519

What Works Clearinghouse. (2022). *What Works Clearinghouse procedures and standards handbook, version 5.0*. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.

Winkler, C., Hammoda, B., Noyes, E., & Van Gelderen, M. (2023). Entrepreneurship education at the dawn of generative artificial intelligence. *Entrepreneurship Education and Pedagogy, 6*(4), 1-11. https://doi.org/10.1177/25151274231198799

Yang, D. (2022). The impact of adaptive learning in entrepreneurial behavior for college students. *Frontiers in Psychology, 12*, 1-13. https://doi.org/10.3389/fpsyg.2021.797459

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education – where are the educators? *International Journal of Educational Technology in Higher Education, 16*(1), 1-27. https://doi.org/10.1186/s41239-019-0171-0

Zen, Z., Reflianto, Syamsuar, & Ariani, F. (2022). Academic achievement: The effect of project-based online learning method and student engagement. *Heliyon, 8*(11), 1-13. https://doi.org/10.1016/j.heliyon.2022.e11509