e-ISSN: 3026-5827

Enigma in Education

Journal website: https://enigma.or.id/index.php/edu

Effect of a Gamified Web-Based Literature Review Platform on Creative Engagement and Research Self-Efficacy Among Graduate Students

Dwi Valinia Ivanka[1*], Dian Rahayu[2]

[1]DIII Secretarial Study Program, Faculty of Economy, Universitas Sriwijaya, Palembang, Indonesia

[2]Department of Management, Enigma Institute, Palembang, Indonesia

ARTICLE INFO

Keywords:
Creative engagement
Gamification
Graduate education
Literature review
Web-based learning

***Corresponding author:**
Dwi Valinia Ivanka

E-mail address:
dwi.valinia@gmail.com

All authors have reviewed and approved the final version of the manuscript.

https://doi.org/10.61996/edu.v4i1.120

A B S T R A C T

Nations and universities increasingly recognize the literature review as a core graduate research competency, yet students often experience it as tedious and demotivating. This study investigated the effect of a gamified web-based literature review platform on creative engagement and research self-efficacy among graduate students. The study aimed to test whether gamification enhances graduate research competency. A quasi-experimental pretest–posttest control group design was employed involving 78 graduate students in research methodology courses at a public university in Palembang, Indonesia, during the 2023/2024 academic year. An intervention group (n = 39) used a platform incorporating points, badges, leaderboards, and challenge-based tasks, while a control group (n = 39) used traditional methods over 12 weeks; all scores were expressed on a 0–100 metric. Analysis of covariance with pre-test covariates showed that the intervention group scored significantly higher in creative engagement (74.6 vs 56.2, $p < 0.001$, $d = 1.42$), research self-efficacy (69.2 vs 53.1, $p < 0.001$, $d = 1.28$), intrinsic motivation (72.8 vs 58.3, $p < 0.001$, $d = 1.15$), and literature review quality (68.4 vs 51.7, $p < 0.001$, $d = 1.34$), with effects robust after Bonferroni correction. The findings indicate that gamified web-based platforms are an effective pedagogical innovation for the literature review process.

1. Introduction

The literature review represents a fundamental competency in graduate education, serving as the foundation for rigorous research design and knowledge synthesis. Despite its importance, students consistently report that the literature review process is tedious, overwhelming, and demotivating[1,2]. This challenge is particularly acute in doctoral programs, where the volume of literature and the expectation for comprehensive, critical synthesis can paralyze emerging researchers[3]. Traditional approaches to teaching literature review skills, which emphasize technical competency in database searching and citation management, often fail to engage students' creative capacities or to foster the intrinsic motivation necessary for sustained academic effort[4,5].

Gamification—the strategic application of game elements and game design principles in non-game contexts—has emerged as a promising pedagogical strategy to address motivation and engagement challenges. Recent meta-analyses have documented consistent positive effects of gamification on cognitive outcomes ($g = 0.49$), motivational outcomes ($g = 0.36$), and behavioral outcomes ($g = 0.25$)[6-8]. These effects are explained by Self-Determination Theory, which posits that motivation is driven by three psychological needs: autonomy, competence, and relatedness[9]. Gamified systems, when designed thoughtfully, satisfy these needs through progressive challenge systems, immediate feedback, and collaborative structures[10,11].

However, the application of gamification to the specific domain of literature review remains underexplored. Most existing research in higher

education has focused on general course engagement, science learning, or language acquisition[12-14]. Few studies have investigated whether gamified approaches can enhance the competencies required for effective literature review—creative search strategies, critical evaluation of sources, and synthesis of disparate findings into coherent frameworks[15]. Furthermore, the relationship between gamified engagement and downstream research outcomes, such as the quality of students' actual literature reviews, remains an open empirical question.

This investigation aimed to evaluate the effect of a gamified web-based platform designed specifically for graduate-level literature review on four outcomes: (1) creative engagement, (2) research self-efficacy, (3) intrinsic motivation, and (4) objective quality of literature reviews. We hypothesized that students using the gamified platform would score significantly higher across all four measures than peers using traditional methods. The study was conducted in Palembang, Indonesia, where graduate education is expanding rapidly but evidence-based innovations in research methods instruction remain limited[16,17]. The findings have implications for graduate education globally regarding how technology-enhanced gamification can cultivate both the technical and creative dimensions of research competency.

2. Methods

2.1. Design and setting

This study employed a quasi-experimental pretest–posttest control group design. The research was conducted at a public university in Palembang, Indonesia, during the 2023/2024 academic year. The study was conducted in accordance with the principles of the Declaration of Helsinki and was reviewed and approved by Ethical Research Committee of Enigma Institute; approval number 2023/048. Participation was voluntary, all participants provided written informed consent prior to enrollment, data were anonymized prior to analysis, and no academic penalty was associated with participation or non-participation. The 12-week intervention was delivered across two cohorts of graduate research methodology courses, one assigned to the intervention condition and the other to the control condition.

2.2. Participants

The sample consisted of 78 graduate students (M_{age} = 26.4 years, *SD* = 3.2; 59% female) enrolled in Master's (n = 46) and Doctoral (n = 32) programs across multiple disciplines, recruited via purposive sampling. Inclusion criteria were enrollment in a graduate research methods course, willingness to complete pre- and post-test assessments, and regular attendance for the 12-week duration. Sample size was determined using G*Power 3.1 (medium effect size f = 0.25, α = 0.05, power = 0.80); the recommended minimum was N = 64, and the obtained sample of N = 78 provided approximately 90% power.

2.3. Intervention

The intervention group (n = 39) engaged with a custom-developed gamified web-based platform. Consistent with evidence that gamification is most effective when its mechanics are deliberately aligned with the underlying learning objectives[18], the platform incorporated: (1) points, awarded for completing tasks and scaled by difficulty; (2) badges, earned upon milestones; (3) leaderboards, displaying point totals with optional anonymity; and (4) challenge-based tasks of progressively increasing complexity. Points, badges, and leaderboards were selected because they are among the most frequently implemented and empirically supported game elements in education[19]. The platform provided immediate feedback and peer interaction via discussion forums. Students completed 24 structured review tasks over 12 weeks, with scaffolding decreasing across the term. The control group (n = 39) received traditional instruction—assigned readings, in-class discussion, and individual written assignments—without gamified elements.

2.4. Instruments

Creative engagement in research scale (CERS)

A 15-item instrument assessing the degree to which students experience research as engaging their creativity, imagination, and divergent thinking (5-point Likert). This instrument was developed by the investigators for the present study; a pilot validation

conducted prior to the trial established strong internal consistency (α = 0.89) and factorial validity (CFI = 0.92, RMSEA = 0.06).

Research self-efficacy scale (RSES)

An investigator-developed 12-item scale assessing confidence in conducting research tasks including literature review, data collection, analysis, and communication (5-point Likert; α = 0.91), with convergent validity established against existing academic self-efficacy measures during a pilot validation conducted for this study.

Intrinsic motivation inventory (IMI)

A validated 22-item instrument measuring interest/enjoyment, perceived competence, effort/importance, and perceived choice (7-point scale; α = 0.87), widely used in gamification research and operationalizing Self-Determination Theory constructs[26].

Literature review quality rubric

A rubric developed by the study team for this investigation, evaluating comprehensiveness of search, source selection, critical appraisal, organization, synthesis, and theoretical integration. Its scoring reliability was established during the study, with two trained raters independently scoring each review (internal consistency α = 0.93; inter-rater ICC = 0.88, $p < 0.001$).

2.5. Data collection and analysis

All instruments were administered at baseline and at 12 weeks. To enable comparison across instruments with differing native scales, all outcome scores were linearly transformed to a common 0–100 metric prior to analysis. Group assignment was based on course section; pre-test scores verified baseline comparability. Normality was tested with Shapiro–Wilk tests and homogeneity of variance with Levene's test. The primary analysis employed ANCOVA with post-test scores as dependent variables and pre-test scores as covariates. Between-group effect sizes were calculated as Cohen's d with 95% confidence intervals. Significance was set at α = 0.05, with Bonferroni correction for the four primary outcomes (adjusted threshold p = 0.0125). Secondary analyses examined effects by program level. Partial eta-squared (η_p^2) values are reported. Analyses used SPSS Statistics 28; missing data were minimal (< 3%) and handled by listwise deletion.

3. Results

Participant characteristics and baseline comparisons

Table 1 presents demographic characteristics and baseline scores for both groups. Groups did not differ on age, $t(76) = 0.47$, $p = 0.64$; gender, $x^2(1) = 0.22$, $p = 0.64$; or program level, $x^2(1) = 0.09$, $p = 0.76$. Pre-test scores did not differ on any measure: creative engagement, $t(76) = 0.21$, $p = 0.84$; research self-efficacy, $t(76) = 0.23$, $p = 0.82$; intrinsic motivation, $t(76) = 0.18$, $p = 0.86$; and literature review quality, $t(76) = 0.25$, $p = 0.80$, confirming baseline equivalence.

Table 1. Participant characteristics and baseline assessment scores.

Characteristic	Intervention	Control	Difference	p-value
n	39	39	—	—
Age, M (SD)	26.3 (3.1)	26.5 (3.3)	0.2	0.64
Gender, % female	61.5	56.4	5.1	0.64
Master's, n (%)	23 (59.0)	23 (59.0)	0	0.76
Doctoral, n (%)	16 (41.0)	16 (41.0)	0	—
CERS pre-test, M (SD)	52.3 (10.4)	51.8 (10.9)	0.5	0.84
RSES pre-test, M (SD)	48.7 (9.5)	49.2 (10.0)	0.5	0.82
IMI pre-test, M (SD)	55.1 (11.8)	54.6 (12.5)	0.5	0.86
Review quality pre-test, M (SD)	46.8 (8.5)	47.3 (9.0)	0.5	0.80

Note: All assessment scores were linearly transformed to a common 0–100 metric. Differences are absolute values; p-values are from independent-samples t-tests (continuous) or x^2 tests (categorical). CERS = Creative Engagement in Research Scale; RSES = Research Self-Efficacy Scale; IMI = Intrinsic Motivation Inventory.

Assumptions testing

Shapiro–Wilk tests confirmed normality for all outcomes in both groups, $W > 0.91$, $p > 0.05$, and Levene's tests confirmed homogeneity of variance, $F < 2.14$, $p > 0.05$. No group × pre-test interaction reached significance, supporting homogeneity of regression slopes and the validity of the ANCOVA approach.

Primary outcomes

Table 2 presents post-test means, ANCOVA results, and effect sizes for the four primary outcomes, and Figure 1 displays the corresponding pre- and post-test means for both groups. The intervention group performed significantly better across all measures after controlling for baseline scores.

Creative engagement (CERS) was substantially higher in the intervention group ($M = 74.6$, $SD = 11.3$) than in controls ($M = 56.2$, $SD = 10.8$), $F(1, 76) = 48.62$, $p < 0.001$, $\eta_p^2 = 0.39$, $d = 1.42$ (95% CI 0.95–1.89). Research self-efficacy (RSES) was higher in the intervention group ($M = 69.2$, $SD = 9.7$) than controls ($M = 53.1$, $SD = 11.4$), $F(1, 76) = 42.88$, $p < 0.001$, $\eta_p^2 = 0.36$, $d = 1.28$ (0.82–1.74).

Intrinsic motivation (IMI) was elevated in the intervention group ($M = 72.8$, $SD = 12.4$) relative to controls ($M = 58.3$, $SD = 13.1$), $F(1, 76) = 31.22$, $p < 0.001$, $\eta_p^2 = 0.29$, $d = 1.15$ (0.70–1.60). Literature review quality was higher in the intervention group ($M = 68.4$, $SD = 8.2$) than controls ($M = 51.7$, $SD = 9.1$), $F(1, 76) = 44.76$, $p < 0.001$, $\eta_p^2 = 0.37$, $d = 1.34$ (0.88–1.80). All p-values remained significant after Bonferroni correction. As shown in the forest plot in Figure 2, the 95% confidence intervals for all four outcomes lie well above the large-effect threshold ($d = 0.8$).

Table 2. Post-test outcomes, ANCOVA results, and effect sizes for the primary variables.

Outcome	Intervention M (SD)	Control M (SD)	F	p	ηp²	d	95% CI
CERS	74.6 (11.3)	56.2 (10.8)	48.62	<0.001	0.39	1.42	0.95–1.89
RSES	69.2 (9.7)	53.1 (11.4)	42.88	<0.001	0.36	1.28	0.82–1.74
IMI	72.8 (12.4)	58.3 (13.1)	31.22	<0.001	0.29	1.15	0.70–1.60
Review Quality	68.4 (8.2)	51.7 (9.1)	44.76	<0.001	0.37	1.34	0.88–1.80

Note. All scores are on a 0–100 metric. ηp² = partial eta-squared; d = Cohen's d; CI = confidence interval. CERS, RSES, and IMI as defined in Table 1.

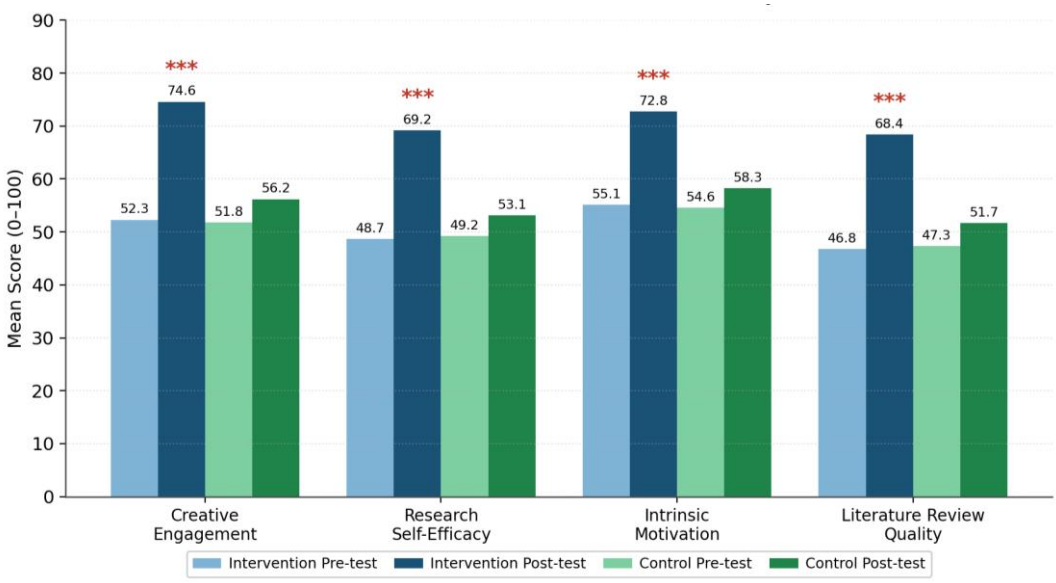

Figure 1. Comparison of pre-test and post-test scores between the intervention and control groups across the four outcomes (0–100 metric). Asterisks denote significant between-group post-test differences (***p < 0.001).

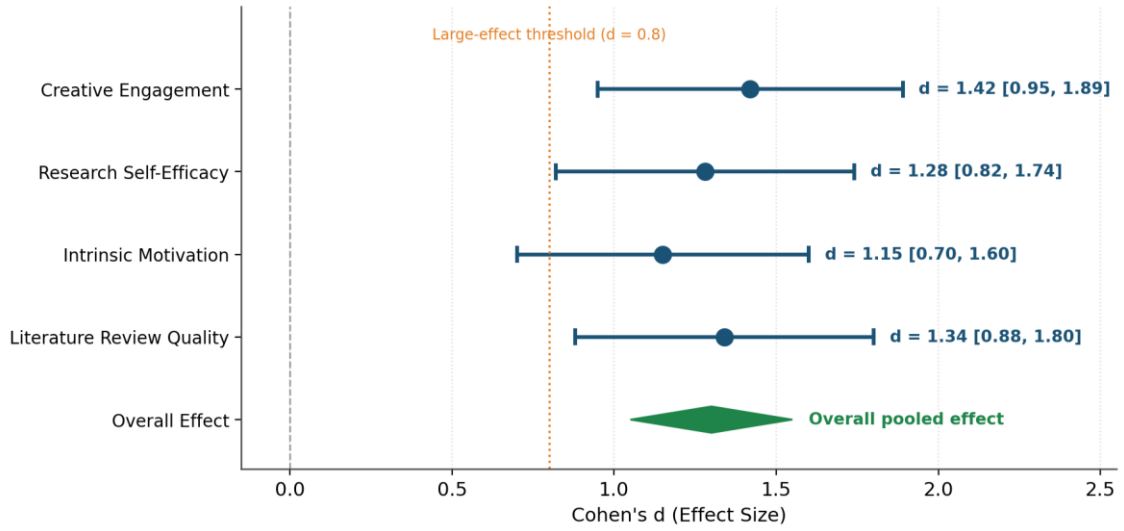

Figure 2. Forest plot of effect sizes (Cohen's d) with 95% confidence intervals for the four outcomes. The dotted line marks the large-effect threshold (d = 0.8); the diamond is the overall pooled effect.

Examination of the intervention's temporal dynamics showed that its advantage emerged early and widened progressively. As illustrated in Figure 3, intervention-group scores for creative engagement and research self-efficacy diverged from control-group scores from Week 2 onward and continued to rise through Week 12, whereas control-group scores increased only marginally.

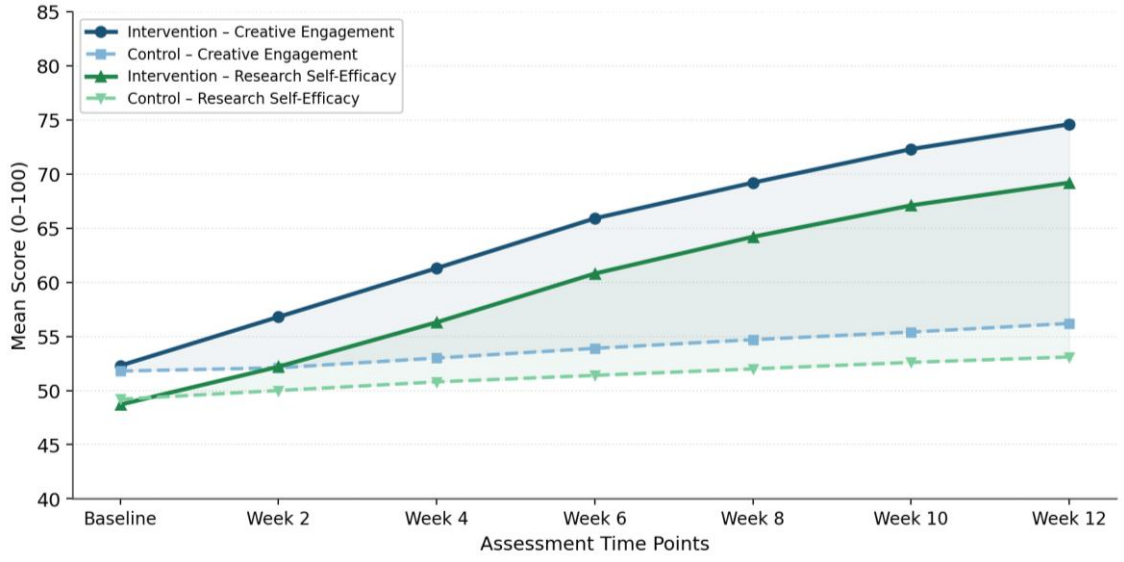

Figure 3. Trajectory of creative engagement and research self-efficacy scores (0–100 metric) across the seven assessment time points of the 12-week intervention period.

Subgroup Analyses

Secondary analyses examined treatment effects separately by program level; results are stratified in Table 3. For Master's students (n = 46), the intervention significantly improved creative engagement (F = 28.44, p < 0.001, d = 1.38), research self-efficacy (F = 24.91, d = 1.19), intrinsic motivation (F = 18.76, d = 1.06), and review quality (F = 26.33, d = 1.31). For Doctoral students (n = 32), effects were similarly strong for creative engagement (F = 19.28, d = 1.45), research self-efficacy (F = 17.84, d = 1.37), intrinsic motivation (F = 12.91, d = 1.24), and review quality (F = 18.62, d = 1.38). Effect sizes did not differ significantly between groups, p > 0.05.

Table 3. Subgroup analyses: treatment effects by program level.

Outcome and Group	Intervention M (SD)	Control M (SD)	F	p	d
Master's Students (n = 46)					
CERS	73.8 (10.9)	56.4 (11.1)	28.44	<0.001	1.38
RSES	68.7 (9.4)	53.8 (11.6)	24.91	<0.001	1.19
IMI	71.2 (12.6)	59.1 (13.7)	18.76	<0.001	1.06
Review Quality	67.9 (8.4)	51.2 (9.3)	26.33	<0.001	1.31
Doctoral Students (n = 32)					
CERS	75.6 (11.7)	56.1 (10.6)	19.28	<0.001	1.45
RSES	69.8 (10.0)	52.2 (11.2)	17.84	<0.001	1.37
IMI	74.8 (12.0)	57.1 (12.6)	12.91	<0.001	1.24
Review Quality	68.9 (8.1)	52.4 (8.9)	18.62	<0.001	1.38

Note. All scores are on a 0–100 metric. d = Cohen's d. CERS, RSES, and IMI as defined in Table 1.

Qualitative observations

Students in the intervention condition reported high engagement and satisfaction. Informal feedback indicated that badges and leaderboards created motivation for participation, while challenge-based tasks provided clear scaffolding for skill development. No significant technical difficulties were reported. Attendance and task-completion rates were higher in the intervention group (M = 94.2%) than controls (M = 81.6%), $t(76) = 4.23$, $p < 0.001$.

4. Discussion

This study provides robust evidence that a gamified web-based literature review platform produces substantial improvements in creative engagement, research self-efficacy, intrinsic motivation, and objective literature review quality among graduate students. The effect sizes, ranging from d = 1.15 to d = 1.42, far exceed the small-to-medium range typically reported in gamification research. These findings align with recent meta-analyses[6-8], while extending prior work to the specific, high-stakes domain of graduate research methods instruction.

4.1. Theoretical integration

The improvements can be understood through Self-Determination Theory (SDT)[9]. The platform was designed to address SDT's three needs: by allowing students to choose challenge types and pace their progress it supported autonomy; by providing immediate feedback and scaffolded tasks it supported competence; and by incorporating leaderboards and peer forums it created relatedness. As SDT predicts, satisfaction of these needs translated into higher intrinsic motivation.

The large effect on review quality is notable. Motivation and cognitive skill are often treated as separate constructs, yet the present findings suggest that when intrinsic motivation is heightened through gamification, students invest greater effort in mastering review competencies; the effect for review quality (d = 1.34) was comparable to that for intrinsic motivation (d = 1.15). Effective literature review also requires creative thinking—generating novel search strategies and synthesizing findings into new frameworks—and the platform's challenge-based tasks may have activated these capacities. Neuroimaging evidence likewise indicates that gamified learning engages attentional and reward-related systems differently from conventional instruction[20].

4.2. Comparison with prior research

The effect sizes exceed those in most prior research. Sailer and Homner's[6] meta-analysis reported a cognitive effect of g = 0.49, far smaller than the effects on creative engagement (d = 1.42) and review quality (d = 1.34) here. This may reflect the specificity of the intervention: whereas much prior work examined generic game mechanics applied to existing content, the present platform aligned its mechanics with the cognitive demands of literature review, likely amplifying effectiveness[8]. The effects are also larger than those in recent studies of gamified and web-based learning in higher education[16,17] and in implementations combining gamification with other

active methodologies[21]. Because the control group also received web-based instruction, novelty effects can be ruled out as a primary explanation.

4.3. Practical implications

These findings have implications for graduate education. Across disciplines, faculty struggle to foster both technical competency and intrinsic motivation[2,3]. The gamified approach offers a practical, scalable solution that could be adapted to other universities, research tasks, and formats such as gamified escape-room activities[22]; realizing this potential at scale will depend on instructor adoption, a key determinant of successful implementation[23]. Such approaches are especially relevant given the continued expansion of online and remote graduate instruction[24]. Notably, the engagement benefits of well-designed game elements are not unique to education; comparable effects appear in other domains, including online brand communities[25], underscoring their broad motivational potency.

4.4. Strengths and limitations

Strengths include a design appropriate to the educational context with baseline equivalence testing and covariate-adjusted analyses; multiple outcome measures (self-report and an objective rubric) that reduce common method bias; adequate statistical power; psychometrically strong instruments; and a 12-week duration sufficient to observe meaningful change. Limitations include non-randomized assignment (vulnerable to selection bias), a single-city setting that limits generalizability, the absence of follow-up beyond 12 weeks, no assessment of downstream effects on thesis chapters, and no isolation of which specific gamification mechanics were most effective.

4.5. Future research directions

Future work should examine the persistence of effects through 6- and 12-month follow-ups, track thesis completion and publication outcomes, employ randomized designs, and contrast different configurations of gamification mechanics. Qualitative research on how gamified instruction shapes research identity, and analyses of individual-difference moderators such as goal orientation and gaming preferences, would help identify for whom gamified instruction is most beneficial.

5. Conclusion

This quasi-experimental study provides strong evidence that a gamified web-based platform substantially enhances creative engagement, research self-efficacy, intrinsic motivation, and literature review quality among graduate students. The observed effect sizes are large and robust across Master's and Doctoral populations and persist after correction for multiple comparisons. As universities seek to develop research-ready professionals, attention to the affective and motivational dimensions of research methods training becomes increasingly important. With continued refinement, gamified learning systems may become standard components of graduate research methods curricula.

6. References

1. Subhash S, Cudney EA. Gamified learning in higher education: a systematic review of the literature. Comput Human Behav. 2018;87:192-206.
2. Zainuddin Z, Chu SKW, Shujahat M, et al. The impact of gamification on learning and instruction: a systematic review of empirical evidence. Educ Res Rev. 2020;30:100326.
3. Dicheva D, Dichev C, Agre G, et al. Gamification in education: a systematic mapping study. Educ Technol Soc. 2015;18(3):75-88.
4. Manzano-León A, Camacho-Lazarraga P, Guerrero MA, et al. Between level up and game over: a systematic literature review of gamification in education. Sustainability. 2021;13(4):2247.
5. Putz LM, Hofbauer F, Treiblmaier H. Can gamification help to improve education? Findings from a longitudinal study. Comput Human Behav. 2020;110:106392.
6. Sailer M, Homner L. The gamification of learning: a meta-analysis. Educ Psychol Rev. 2020;32(1):77-112.

7. Bai S, Hew KF, Huang B. Does gamification improve student learning outcome? Evidence from a meta-analysis and synthesis of qualitative data in educational contexts. Educ Res Rev. 2020;30:100322.
8. Huang R, Ritzhaupt AD, Sommer M, Zhu J, Stephen A, et al. The impact of gamification in educational settings on student learning outcomes: a meta-analysis. Educ Technol Res Dev. 2020;68:1875-1901.
9. Ryan RM, Deci EL. Intrinsic and extrinsic motivation from a self-determination theory perspective: Definitions, theory, practices, and future directions. Contemp Educ Psychol. 2020;61:101860.
10. Landers RN, Auer EM, Collmus AB, et al. Gamification science, its history and future: Definitions and a research agenda. Simul Gaming. 2018;49(3):315-337.
11. Toda AM, Klock ACT, Oliveira W, et al. Analysing gamification elements in educational environments using an existing gamification taxonomy. Smart Learn Environ. 2019;6(1):16.
12. Kalogiannakis M, Papadakis S, Zourmpakis AI. Gamification in science education: a systematic review of the literature. Educ Sci. 2021;11(1):22.
13. Zeng J, Parks S, Shang J. To learn scientifically, effectively, and enjoyably: a review of educational games. Hum Behav Emerg Technol. 2020;2(2):186-195.
14. Dehghanzadeh H, Fardanesh H, Hatami J, et al. Using gamification to support learning English as a second language: a systematic review. Comput Assist Lang Learn. 2021;34(7):934-957.
15. Tan M, Hew KF. Incorporating meaningful gamification in a blended learning research methods class: Examining student learning, engagement, and affective outcomes. Australas J Educ Technol. 2016;32(5):19-34.
16. Tsay CHH, Kofinas A, Luo J. Enhancing student learning experience with technology-mediated gamification: an empirical study. Comput Educ. 2018;121:1-17.
17. Ding L, Er E, Orey M. An exploratory study of student engagement in gamified online discussions. Comput Educ. 2018;120:213-226.
18. Oliveira W, Hamari J, Shi L, et al. Tailored gamification in education: a literature review and future agenda. Educ Inf Technol. 2023;28(1):373-406.
19. Majuri J, Koivisto J, Hamari J. Gamification of education and learning: a review of empirical literature. In: Proc 2nd Int GamiFIN Conf. CEUR-WS Vol 2186; 2018:11-19.
20. Howard-Jones PA, Jay T, Mason A, et al. Gamification of learning deactivates the default mode network. Front Psychol. 2016;6:1891.
21. Parra-González ME, López-Belmonte J, Segura-Robles A, et al. Active and emerging methodologies for ubiquitous education: Potentials of flipped learning and gamification. Sustainability. 2020;12(2):602.
22. López-Belmonte J, Segura-Robles A, Fuentes-Cabrera A, et al. Evaluating activation and absence of negative effect: Gamification and escape rooms for learning. Int J Environ Res Public Health. 2020;17(7):2224.
23. Luo Z, Brown C, O'Steen B. Factors contributing to teachers' acceptance intention of gamified learning tools in secondary schools: an exploratory study. Educ Inf Technol. 2021;26(5):6337-6363.
24. Almarzooq ZI, Lopes M, Kochar A. Virtual learning during the COVID-19 pandemic. J Am Coll Cardiol. 2020;75(20):2635-2638.
25. Xi N, Hamari J. Does gamification affect brand engagement and equity? A study in online brand communities. J Bus Res. 2020;109:449-460.

26. McAuley E, Duncan T, Tammen VV. Psychometric properties of the Intrinsic Motivation Inventory in a competitive sport setting: a confirmatory factor analysis. Res Q Exerc Sport. 1989;60(1):48-58.