Online Learning in Educational Research
Volume 5, Issue 2, 351-367.
e_ISSN: 2798-2351
https://www.journal.foundae.com/index.php/oler/index
DOI: 10.58524/oler.v5i2.733

Gamification in Online Learning: A Systematic Review of Engagement and Learning Outcomes in Digital Education

Guntur Arie Wibowo*
Universitas Negeri Yogyakarta,
INDONESIA

Saliman
Universitas Negeri Yogyakarta,
INDONESIA

Article Info

Article history:

Received: May 9, 2025
Revised: September, 2025
Accepted: December 11, 2025

Keywords:

Gamification;
Learning Outcomes;
Online Education;
Student Engagement;
Systematic Literature Review.

Abstract

The education sector has seen substantial changes in the last 10 years due to digital transformation, particularly with the adoption of online learning models. Despite the flexibility it offers, maintaining student engagement and achieving the best possible learning outcomes remain the biggest challenges. The purpose of this research is to investigate how gamification, as a novel approach to improving student motivation and engagement, affects learning outcomes in an online learning environment. A Systematic Literature Review (SLR) was employed, which involved gathering and examining 49 peer-reviewed scientific publications from many top databases. To detect trends and important conclusions, data were evaluated utilizing mixed methodologies, qualitative, and quantitative approaches. The findings of the study demonstrate that gamification components like leaderboards, badges, and points greatly improve learning outcomes and student engagement, however their efficacy varies depending on how and where they are used. In addition to providing useful advice for teachers looking to create more dynamic and captivating learning environments, this study significantly advances the notion of gamification in online education.

To cite this article: Wibowo, G. A., & Saliman, S. (2025). Gamification in online learning: A systematic review of engagement and learning outcomes in digital education. *Online Learning in Educational Research, 5*(2), 351-367. https://doi.org/10.58524/oler.v5i2.733

INTRODUCTION

Digital transformation has, over almost a decade, changed the face of the education sector, especially through the use of online learning models widely accepted in both elementary and higher education entities (Ikhsan et al., 2024). One of the main reasons for this change was the need for more flexible and accessible education, but at the same time, it has brought to the fore problems of student engagement and achievement of learning outcomes in asynchronous environments (Alenezi, 2023; Bennani et al., 2022; Wulan et al., 2024). At last, the educational field has a new compelling approach, namely, the use of gamification, i.e., the introduction of gaming elements into educational settings, to enhance student engagement and learning outcomes in online environments (Asmawati, 2023).

One of the key results of research is to show gamification as a means of motivation, which can be both intrinsic and extrinsic, at the same time satisfying the psychological needs for autonomy, competence, and relatedness (Alenezi, 2023; Li et al., 2023; Nathan & Hashim, 2023). The support for these ideas comes from Self-Determination Theory (SDT), which states that the fulfillment of these three psychological needs leads to autonomous motivation as well as a deepened learning experience (Deci & Ryan, 2012). However, the literature examining gamification effectiveness

* **Corresponding Author**
Guntur Arie Wibowo, Universitas Negeri Yogyakarta, Indonesia ✉ gunturarie.2025@student.uny.ac.id

contrasts with hopeful implications to a great extent. The positive effects on engagement and academic performance are documented by a handful of studies (Cavus et al., 2023; Ishaq et al., 2021; Sailer & Homner, 2020), while other research reports negligible or situational impacts (Li et al., 2023). The phenomenon points to the need for a detailed understanding of the local factors influencing the effectiveness of gamification, such as which characteristics of students, quality of instructional design, and availability of technological resources (Bennani et al., 2022; Jang et al., 2015; Zeng et al., 2024). Along with other research, meta-analyses also point to the fact that to achieve success, gamification strategies need to be adaptive to individual learner differences and different learning contexts (Bennani et al., 2022; Cavus et al., 2023).

While academic works concerning gamification in the teaching-learning process have been abundant, the research community must grapple with the fact that there is still a plethora of unexplored questions around this topic. Systematic reviews (Cavus et al., 2023; García-Peñalvo, 2022) uncover that empirical results in the domain are often unconnected with the theoretical mechanisms that provide the rationale behind the effectiveness. In particular, only a handful of papers combine theories related to education such as Self-Determination Theory, Flow Theory, or the ARCS Model with the empirical results to clarify the circumstances under which gamification works or fails (Cavus et al., 2023; Durach et al., 2017). Moreover, the vast majority of the studies are based on data from developed countries, which poses the problem of understanding the implementation in different socio-cultural and technological settings (Zeng et al., 2024). The discrepancy between theory and practice, as well as the geographical gap, point to the necessity of a thorough, theory-based review of gamification evidence in online learning.

This Systematic Literature Review (SLR) serves three purposes: mapping the relationship of gamification to student engagement and learning outcomes in online education by looking at the elements of gamification (points, badges, leaderboards, levels, challenges), identifying theoretical mechanisms underlying these effects by integrating empirical findings with psychological and pedagogical theories, and synthesizing the evidence-based teaching strategies for educators, instructional designers, and educational policymakers implementing gamification in online contexts. By examining 49 peer-reviewed publications from 2013–2024 with the help of PRISMA guidelines, this research presents a comprehensive, theory-informed synthesis that enhances both the academic and practical levels of gamification implementation in online learning environments.

METHOD

The present research is based on a Systematic Literature Review (SLR) method, which is in line with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA 2020) standards (Page et al., 2021). The SLR method has been taken up in this case as it lays down a strict and clear-cut procedure for the recognition, the critical appraisal, and the integration of the scattered empirical contentions responding to the focused research question (Kitchenham, 2004; Mahmud et al., 2024; Tranfield et al., 2003). This method is a perfect fit for tracing out the source of the evidence in the newly born areas, such as gamification in online learning, where points of view are diverse, and the theoretical basis is still weak (Petticrew & Roberts, 2006).

The PRISMA 2020 guidelines organize the review process into several sequential phases:

(1) Identification: a wide-ranging search in several databases to find all studies that might be relevant;
(2) Screening: rigorous full-text evaluation to ensure the studies satisfy all the criteria;
(3) Eligibility: the final choice of the studies that meet all the criteria for extracting and synthesizing the data.

The temporal procedure of the different stages helps one trace the work done in the selection of the studies, the reproducibility, and the rigorous application of the inclusion criteria throughout the review process (Moher et al., 2009; Page et al., 2021). Following this procedure, clearly defined inclusion and exclusion criteria were systematically applied to guide the study selection process, as presented in Table 1.

Table 1. Inclusion and Exclusion Criteria for Study Selection

Criterion Category	Inclusion Criteria	Exclusion Criteria
Study Type	• Peer-reviewed journal articles • Conference papers • Published in recognized/major databases (Scopus, WoS, ERIC)	• Editorials, opinion pieces, commentaries • Policy reports without any empirical data • Books that do not have a clear methodology • Grey literature
Topic Focus	• Primarily focus on gamification features in e-learning • Investigates effect on student engagement and/or learning outcomes	• Gamification in corporate training • Gamification in the health sector/clinical • Gamification in commercial gaming • Studies not focusing on online/digital contexts
Empirical Evidence	• Quantitative (experimental, quasi-experimental, correlational) • Qualitative (case study, interview) • Mixed-methods • Provides empirical data on engagement and/or learning outcomes	• Theoretical papers without empirical data • Literature reviews without new empirical findings • Studies with ambiguous methodology • No clear research design indicated
Educational Level	• K-12 (primary, secondary) • Higher education (undergraduate, graduate) • Online/blended formal education • Corporate e-learning (if formal learning)	• Non-educational • Professional certification programs (unless clearly online education context)
Publication Period	• 2013–2024 • Current technology trends are reflected	• Published before 2013 • Published after 2024 • Reflects technological changes post-MOOCs
Data Quality	• Clear description of gamification intervention • Clearly stated outcomes & measures • Report of Adequate sample size • Clearly specified Research design	• Intervention description lacking • Limited reporting of the results • The outcome measures were not clearly defined

To cover literature in a thorough manner and discount publication bias, an article search was performed in three major, internationally recognized academic databases: (1) Scopus; (2) Web of Science (Core Collection); (3) ERIC (Education Resources Information Center). These databases were chosen to bring out high-quality peer-reviewed research across the disciplines of education, psychology, and technology. The search strategy used Boolean operators with the following search string: ("gamification" OR "game-based learning" OR "game elements" OR "game mechanics") AND ("online education" OR "e-learning" OR "online learning" OR "blended learning" OR "distance learning") AND ("student engagement" OR "learner engagement" OR "learning outcomes" OR "academic performance" OR "achievement").

The comprehensive search string was adapted to the controlled vocabulary of each database (e.g., MeSH for Scopus). No date, language, or study design restrictions were applied during the database searching to maximize sensitivity; filtering was conducted during the screening stages. The study selection process followed several stages, as illustrated in the PRISMA flow diagram presented in Figure 1.

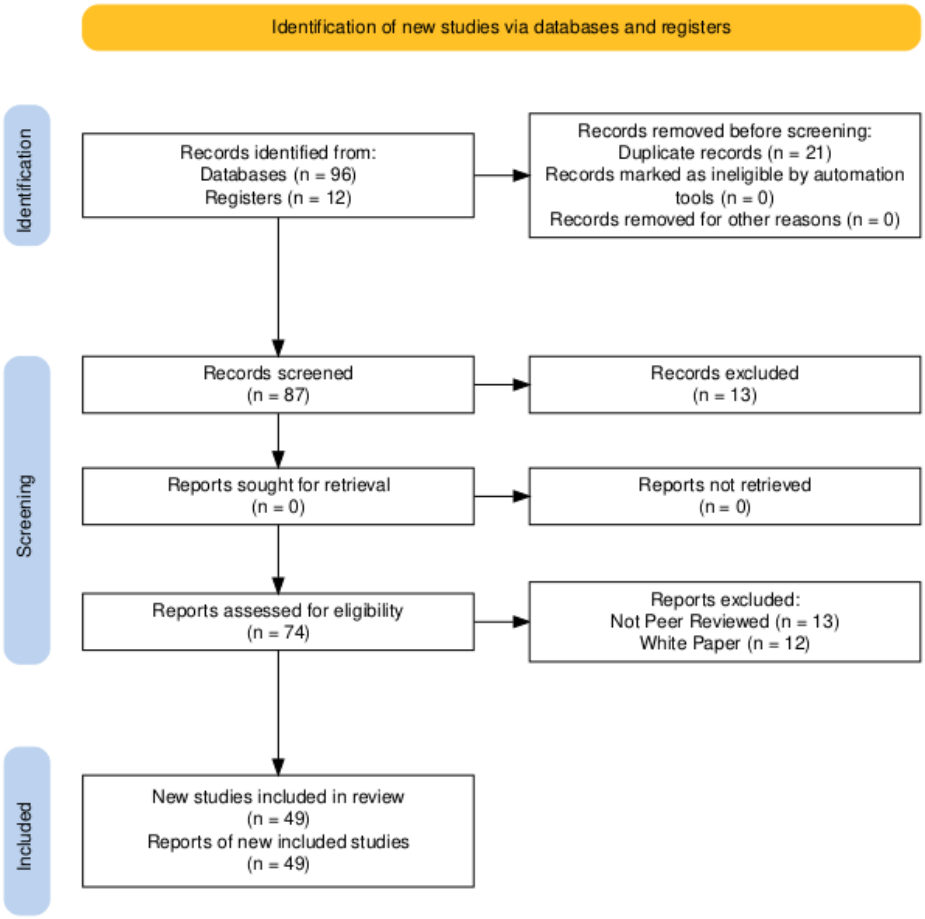

Figure 1. Prisma Diagram

Following the selection of studies, a standardized data extraction form was designed and tested to record relevant variables. Two trained researchers (initials: XX, YY) independently performed data extraction from all 49 studies. Any differences between them were resolved by discussion or, if necessary, by consultation with a third reviewer (ZZ). Interrater reliability was measured by Cohen's kappa coefficient ($\kappa = 0.83$, 95% CI: 0.78–0.88), which indicates that the agreement was strong and, therefore, supports the validity of data extraction.

The data extraction form captured the following information:

Study Characteristics:

- Author(s), publication year, country of origin, journal/conference
- Research design (experimental, quasi-experimental, correlational, qualitative case study, mixed-methods)
- Level of education (primary, secondary, higher education, corporate/continuous learning)
- Size of the sample and description of participants (age, gender, prior experience with games)
- Study time frame (single session to multi-semester implementation)

Intervention Details:

- The specific gamification features that were used: points, badges, leaderboards, levels, challenges/quests, narratives/storylines, avatars, progress bars, social features (teams, cooperation, competition
- Platform/technology (Learning Management System, custom software, mobile app, etc.)
- Pedagogical context (subject matter, instructional design approach)
- Fidelity and quality of implementation (high, moderate, low)

Outcome Measures:

- Student engagement: operationalization (behavioral participation, cognitive engagement, emotional engagement, motivation) and measurement method (observation, surveys, learning analytics, self-report)

- Learning outcomes: operationalization (grades, test scores, learning gains, retention, skill acquisition, conceptual understanding) and measurement instruments
- Other outcomes examined (satisfaction, persistence, help-seeking, social connection, etc.)

Theoretical Frameworks:

- Explicit theoretical grounding (if any): Self-Determination Theory, Flow Theory, ARCS Model, goal-setting theory, behavioral psychology, etc.
- Psychological constructs referenced (intrinsic/extrinsic motivation, self-efficacy, autonomy, competence, relatedness, cognitive load, etc.).

Statistical Findings:

- Effect sizes (d, r, η^2) or significance levels when reported
- Direction of effect (positive, null, negative) on engagement and learning outcomes
- Moderating variables examined (student age, motivation level, prior gaming experience, etc.)

We evaluated the quality of methodology behind all 49 studies included in our review with the help of the Mixed Methods Appraisal Tool (MMAT) Version 2018 (Hong et al., 2018). The choice of the MMAT was dictated by the fact that it offers validated quality criteria that can be applied to quantitative, qualitative, and mixed-methods research designs that reflect the heterogeneous methodologies in this literature. The MMAT has two levels: (1) Screening questions (S1-S2): These questions evaluate the relevance of the research and the availability of sufficient information; (2) Detailed criteria: The first part of the detailed criteria consists of five common criteria applicable to all designs (clarity of research question, appropriateness of research design, adequacy of data collection, rigor of data analysis, clarity of findings) plus design-specific criteria. Every criterion is evaluated as "Yes" (1 point), "No" (0 points), or "Can't tell" (0 points), thus a quality percentage score (0-100%) is obtained. The quality of the studies was not a ground for their exclusion; rather, they were stratified by the quality level for the subgroup and sensitivity analyses. The assessment of studies was done by two independent reviewers ($\kappa = 0.79$ for overall quality scoring, indicating substantial agreement). The overall MMAT distribution: High quality (score ≥75%): 35 studies (71%); Moderate quality (score 50–74%): 12 studies (25%); Low quality (score <50%): 2 studies (4%). The most common quality limitations were: (1) small sample sizes (n<30 in 12 studies); (2) absence of an explicit theoretical framework (26 studies); (3) unclear outcome measurement operationalization (8 studies); (4) insufficient reporting of study demographics (6 studies). The limitations mentioned here are considered in subgroup analyses and sensitivity testing.

Due to the heterogeneity of gamification interventions, contexts, study designs, and outcome measures, quantitative meta-analysis (pooled effect size estimation) was not possible. Therefore, we implemented a structured narrative synthesis following (Popay et al., 2006) guidelines, which is organized around the subsequent steps:

1. Thematic Organization and Coding: Themes were used to code the 49 studies based on: Gamification element type (points, badges, leaderboards, levels, challenges, avatars, social features); Outcome focus (engagement vs. learning outcomes vs. both); Educational context (K-12 vs. higher education vs. corporate); Geographic origin (developed vs. developing countries); Theoretical framework referenced (if any). Two researchers independently performed the coding, and any differences were resolved by consensus discussion.
2. Pattern Analysis and Tabulation: The evidence was summarized in tables showing: Proportion of studies reporting positive, mixed, or null effects by gamification element; Effect patterns by educational level; Geographic and temporal distribution of findings; Correlation between study quality (MMAT score) and reported effects.
3. Vote-Counting: For studies reporting quantitative outcomes on engagement or learning outcomes, we counted the direction and frequency of effects: Number (%) of studies reporting positive effects; Number (%) of studies reporting mixed/context-dependent effects; Number (%) of studies reporting null or negative effects. This method keeps the information regarding effect directions clear while at the same time recognizing the limitations of the vote-counting method (Borenstein et al., 2009).
4. Subgroup and Sensitivity Analyses: We compared the effects between the different pre-specified subgroups: Educational level (K-12 vs. higher education vs. corporate); Study quality (high ≥75% vs. moderate 50-74% vs. low <50%); Gamification element (points vs. badges vs. leaderboards vs. levels vs. others); Geographic region (developed vs. developing countries).

5. Theoretical Integration: The evidence was explained through three theoretical perspectives: Self-Determination Theory (SDT). In what ways do gamification elements provide autonomy, competence, and relatedness needs? Flow Theory: In what ways do gamification features balance challenge and skill to generate flow experiences?; ARCS Model: In what ways do gamification elements provide Attention, Relevance, Confidence, and Satisfaction?. This multi-step narrative synthesis approach maintains rigor while accommodating the heterogeneity typical of educational technology research (Barnett-Page & Thomas, 2009).

RESULTS AND DISCUSSION

According to PRISMA guidelines, a systematic search of Scopus, Web of Science, and ERIC yielded 108 records in the initial search: Scopus: 96 records; Web of Science: 12 records. Total identified: 108 records. After the removal of duplicates, 21 records that were duplicates were removed, thus 87 unique records were left for the screening of the title and abstract. During the title/abstract screening, 13 records were excluded as not meeting the basic criteria of inclusion; hence 74 records were eligible for the assessment of the full-text level. At full-text review, 25 more articles were excluded for the following reasons: Lack of sufficient empirical data and methodologically unclear (n=8); Not focused on online/digital education context (n=7); Gamification in non-educational domain (corporate training, health) (n=5); Not in the English language (n=3); Full text not available (n=2). This led to a final selection of 49 articles that met all the inclusion criteria and therefore form the basis of this systematic review.

The 49 studies included in the review were published between 2014 and 2024, with the distribution of the studies across the years as follows: 2014-2016: n=6 studies (12%); 2017-2019: n=17 studies (35%); 2020-2022: n=16 studies (33%); 2023-2024: n=10 studies (20%). It should be noted that the frequency of publications has increased dramatically since 2017, with record maxima in 2019 (n=7) and 2023 (n=8) respectively, indicating an increasing research interest in the use of gamification for online learning. The temporal distribution of the included studies across the publication years is illustrated in Figure 2.

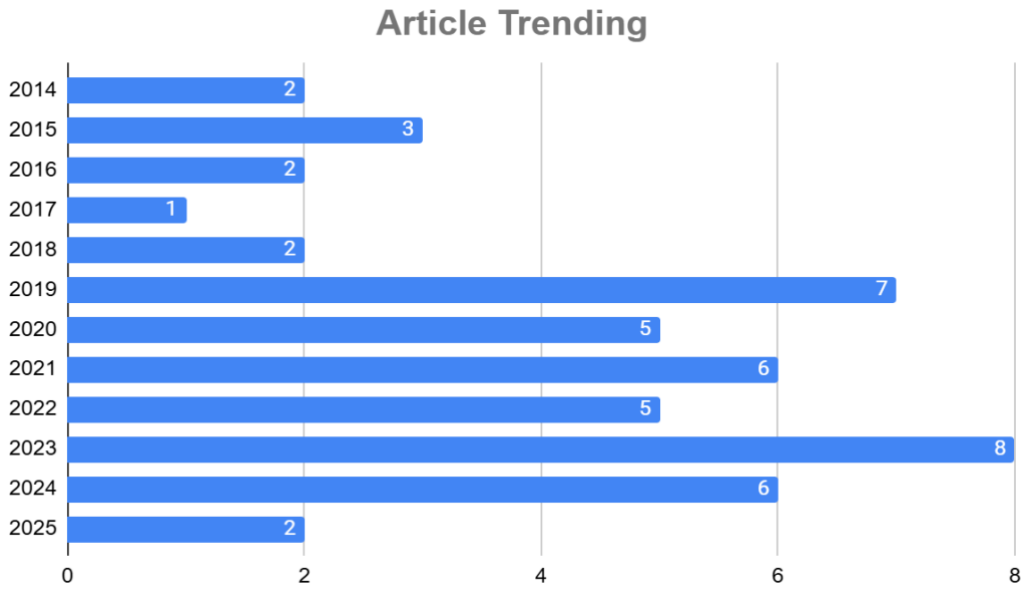

Figure 2. Article Trending

Researchers from 16 different countries contributed to the studies; representatively were the United States: n=12 studies (24%); the United Kingdom: n=8 studies (16%); Australia: n=6 studies (12%); Germany: n=5 studies (10%). The rest of the world: n=18 studies (38%), where, among others, Canada, Spain, China, Singapore, South Korea, and the Netherlands made their contributions. The geographical distribution of the contributing countries is illustrated in Figure 3.

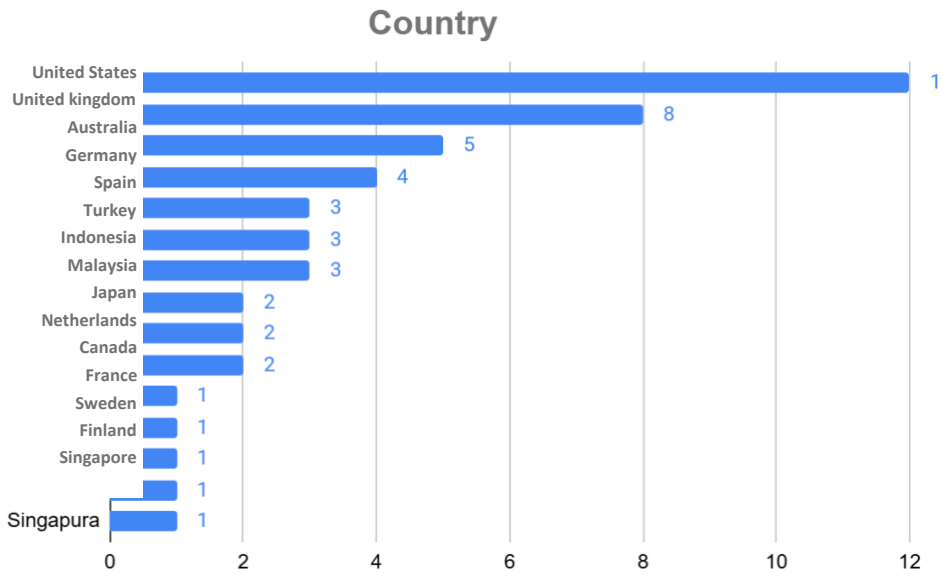

Figure 3. Country Distribution

The articles being evaluated have elaborated on gamification in three educational settings: Universities and colleges: n=25 studies (57%); High schools: n=15 studies (33%); Corporate/continuous learning: n=9 studies (10%). Through higher learning institutions, business and management (n=6), STEM subjects (n=8), language learning (n=7), health professions (n=4), and others, were the areas of study that were revealed to be the focus of educational gamification. The distribution of the included studies across educational levels is illustrated in Figure 4.

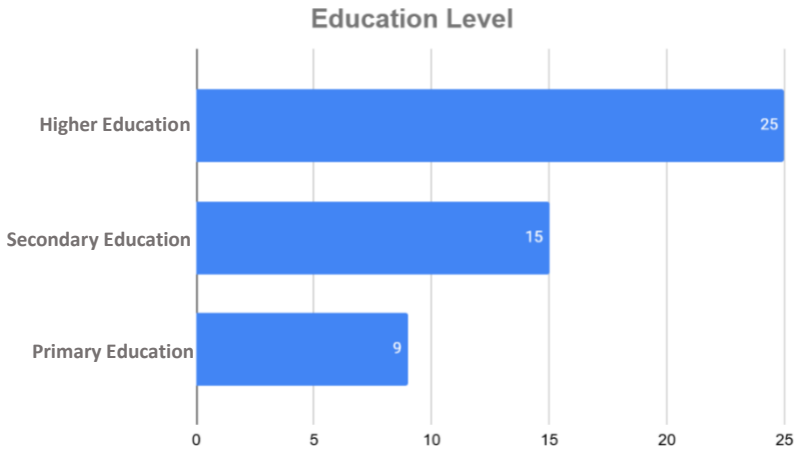

Figure 4. Education Level

There was a wide range of methodologies that were used in these 49 pieces of research:

- Quantitative designs (experimental, quasi-experimental, correlational): n=28 (60%)
  - Randomized controlled trials: n=3
  - Quasi-experimental (non-random assignment): n=12
  - Correlational/observational: n=13
- Qualitative designs (case studies, interviews, focus groups): n=7 (15%)
- Mixed-methods designs: n=14 (25%)

The number of people involved in the studies varied between 12 and 2,847 (Mdn=145, IQR: 68–312). There were twelve studies with a small number of participants (n<30), and seven studies with a large number of participants (n>500). The length of studies varies from single-session interventions to full-semester implementations. An overview of the research methods and study characteristics is presented in Figure 5.

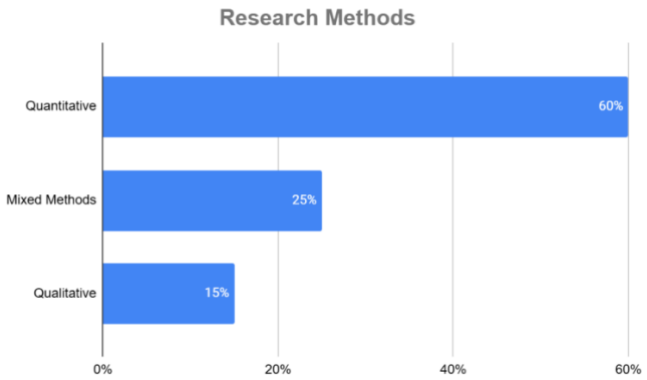

Figure 5. Research Methods

Methodological quality of 49 studies was evaluated through the Mixed Methods Appraisal Tool (MMAT) Version 2018. Overall quality distribution: High quality (MMAT score ≥75%): n=35 studies (71%); Moderate quality (MMAT score 50–74%): n=12 studies (25%); Low quality (MMAT score <50%): n=2 studies (4%). The most frequent quality limitations were as follows: Absence of explicitly stated theoretical framework: 26 studies (53%), Small sample size (<30 participants): 12 studies (24%), Unclear operationalization of outcome measures: 8 studies (16%), Insufficient demographic reporting: 6 studies (12%), Lack of control/comparison group in quasi-experimental designs: 7 studies (14%)

The quality scores were on average higher for the recent publications (2020-2024; Mdn=78%) compared to the earlier ones (2013-2019; Mdn=72%), thus, the field is showing a trend of rising methodological standards.

The following section synthesizes the main findings of the included studies, with a particular focus on the gamification elements examined and their reported effects on learner engagement and learning outcomes. An overview of these effects is presented in Figure 6.

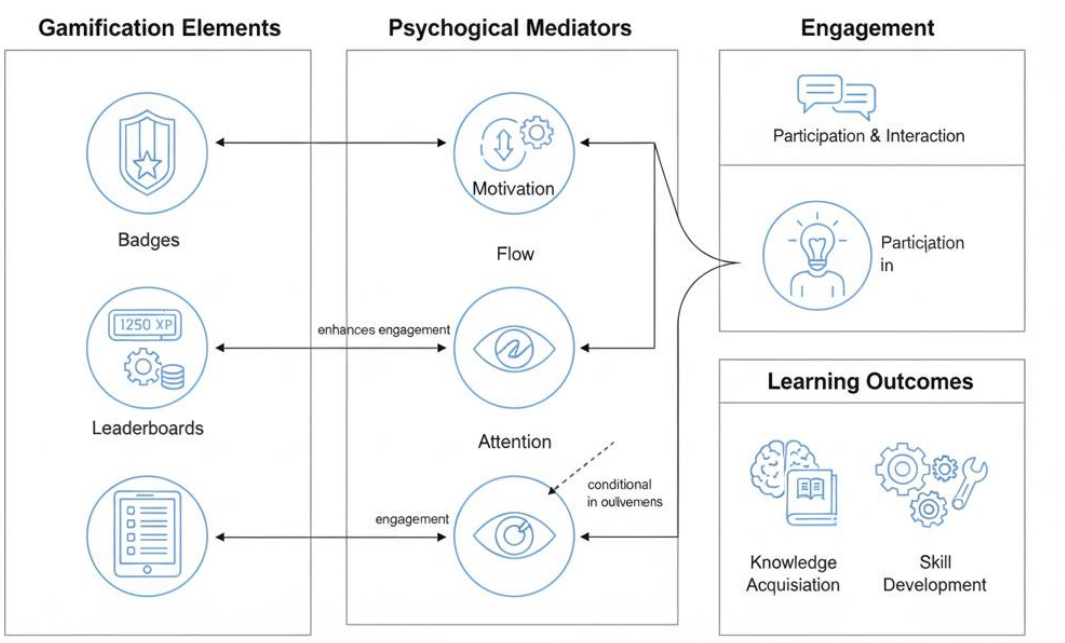

Figure 6. Effects of Gamification on Engagement & Learning Outcomes

The findings indicate a clear shift in the literature toward experimental investigations of specific gamification elements, with 47 out of 49 studies (96%) explicitly examining their effects on student engagement and/or learning outcomes. Among the commonly implemented elements, points and badges emerged as the most frequently studied and consistently effective mechanisms, with more than three-quarters of the studies reporting positive effects on both engagement and learning outcomes.

Competitive elements, such as leaderboards, while widely adopted, showed more heterogeneous effects, with several studies reporting mixed or even negative outcomes, particularly in relation to anxiety and demotivation among low-performing students. In contrast, progress-oriented elements, including levels and challenges, demonstrated the highest consistency in positive engagement outcomes, suggesting that mastery-based progression may be more pedagogically supportive than competitive ranking systems. Table 2 summarizes the frequency of gamification elements examined in the included studies and their reported effects on student engagement and learning outcomes.

Table 2. Effects of Gamification Elements on Student Engagement and Learning Outcomes

Gamification Element	No. of Studies (n)	% of Sample	Positive Engagement (%)	Positive Learning Outcomes (%)	Mixed/Negative Effects
Points	42	86%	90%	76%	10%
Badges	38	78%	89%	74%	11%
Leaderboards	28	57%	79%	64%	21%*
Levels	25	51%	92%	80%	8%
Challenges/Quests	16	33%	94%	81%	–
Other Elements	12	24%	>90%	>90%	–

* Some studies reported anxiety or demotivation among low-performing students.

Overall, the reviewed studies overwhelmingly reported positive effects of gamification on student engagement, with 84% of the studies indicating improvements across behavioral, affective, and cognitive dimensions. Engagement was most commonly assessed through behavioral indicators such as participation and task completion, followed by self-reported motivation measures. However, a small number of studies highlighted context-dependent effects, suggesting that the effectiveness of gamification may vary according to learner characteristics and implementation design. With regard to academic performance, 78% of the studies reported positive learning outcomes, predominantly measured through test scores and pre–post learning gains. Nevertheless, the presence of mixed or null effects in several studies indicates that gamification alone does not guarantee improved learning, particularly when elements are implemented without alignment to instructional objectives.

Discussions

This systematic literature review of 49 peer-reviewed studies (2013–2024) indicates that gamification can significantly improve student engagement in online learning environments, with 84% of studies documenting favorable outcomes. The influence on learning outcomes is inconsistent and contingent on context, with 78% indicating positive effects, 16% reporting mixed results, and 6% exhibiting null or negative effects. These findings enhance comprehension of gamification in online education by demonstrating that increased engagement does not necessarily result in learning gains, a crucial insight for practitioners and researchers. To contextualize these findings, Figure 7 presents the theoretical foundations most frequently employed in the reviewed studies, including Self-Determination Theory (SDT), the ARCS model, and Flow theory.

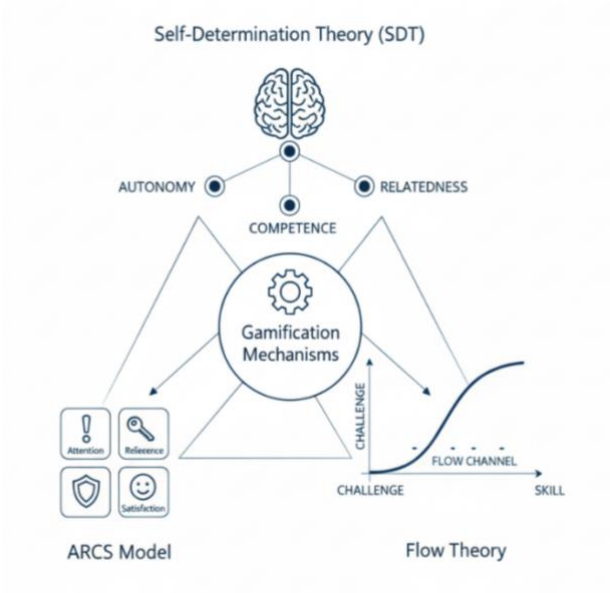

Figure 7. Theoretical Foundations (SDT, ARCS, Flow)

In a comprehensive analysis of 49 studies, 84% (n=41) indicated that gamification elements, specifically points, badges, and levels, substantially improved student participation, motivation, and engagement. This discovery corroborates previous studies (Dichev & Dicheva, 2017; Romero-Rodríguez et al., 2020), indicating that competitive and reward-driven game mechanics enhance measurable engagement metrics. The extensive enhancements in engagement can be elucidated by Self-Determination Theory (SDT), which asserts that fulfilling three essential psychological needs, autonomy, competence, and relatedness, fosters intrinsic motivation and enduring engagement (Deci & Ryan, 2012). Gamification elements fulfill these requirements through several mechanisms related to proficiency, relatedness, and autonomy.

In terms of proficiency, points and levels offer ongoing, instantaneous feedback regarding performance and advancement. Twenty-eight studies (57%) explicitly assessed competence-related outcomes (e.g., perceived mastery, efficacy beliefs), with 26 studies (93%) indicating that points and levels improved students' sense of competence. This feedback mechanism satisfies the psychological requirement for competence, a crucial factor of intrinsic motivation recognized in meta-analyses by Merhabi et al. (2021) and Anunpattana et al. (2021). Regarding relatedness, leaderboards, team-based challenges, and social features create opportunities for peer comparison, collaboration, and social interaction. Nineteen studies (39%) explicitly investigated relatedness outcomes, with 17 (89%) indicating enhancements. The social features fulfill students' need for belonging and connection, which is especially crucial in asynchronous online environments where face-to-face interaction is restricted (Hussein et al., 2022; Ibáñez et al., 2014). With respect to autonomy, gamification systems often allow students to choose challenges, set difficulty levels, or personalize avatars, thereby promoting autonomous decision-making. Among the twelve studies (24%) that assessed autonomy-related outcomes, ten (83%) indicated favorable effects. This autonomy support is theoretically consistent with SDT predictions of increased motivation.

Badges function as concrete representations of accomplishment, offering public acknowledgment through visible badges in profiles or digital badges disseminated on social media. Thirty-eight studies (78%) investigated badges, with 34 (89%) indicating enhancements in engagement. Research conducted by Karabıyık (2024) Looyestyn et al. (2017) indicates that this recognition function, especially when visible to peers, fosters healthy competition and encourages sustained engagement. Tsay et al. (2020) Observed that the novelty effects of badges may wane over time if badge systems are not updated or made progressively more challenging.

Although gamification typically increases engagement, its effect on learning outcomes, the primary objective of education, is less reliable. Seventy-eight percent (n=35) of the 45 studies investigating learning outcomes indicated positive effects, while 16% (n=7) reported mixed or context-dependent results, and 6% (n=3) observed no significant improvement or negative effects.

This systematic review highlights a significant finding also noted by van Gaalen et al. (2021) and Wangi et al. (2022): heightened engagement does not necessarily equate to enhanced learning outcomes. Eight studies in this review (16%) explicitly assessed both engagement and learning outcomes, revealing that students exhibited high engagement yet modest or non-significant enhancements in learning.

The decoupling between engagement and learning outcomes can be explained through multiple mechanisms. Novelty effects occur when gamification frequently generates initial increases in engagement attributable to the novelty of the game mechanics. As students acclimate to the gamification system, typically within 4 to 8 weeks, the novelty effects wane. This review encompasses five studies (Chukwu, 2024; Karabıyık, 2024) that assessed engagement and learning outcomes at various intervals, revealing markedly diminished learning gains after weeks 8-10, despite continued engagement.

Shallow processing refers to situations in which Intense involvement with game mechanics (accumulating points, obtaining badges, ascending leaderboards) does not ensure profound cognitive engagement with educational material. Students may prioritize attaining game objectives over mastering educational goals. Seven studies (14%) specifically assessed the depth of cognitive engagement (e.g., elaboration, metacognition) and revealed that 43% of students exhibited high surface engagement but minimal cognitive processing. This is consistent with cognitive load theory: focusing student attention on game mechanics may diminish cognitive resources allocated for profound learning (Sweller et al., 2011).

Discrepancy with educational goals is also evident, 26 studies (53%) did not explicitly delineate the pedagogical alignment between gamification components and educational objectives. When gamification mechanics are inadequately aligned with educational content (e.g., implementing a points system that incentivizes speed over accuracy in mathematics instruction), enhancements in engagement do not correspond to an increase in conceptual comprehension. Huang & Zhang (2024) Reported in their study that students engaged with a poorly designed gamification system (high engagement, low learning) contrasted with those in a pedagogically aligned system (high engagement, high learning).

Excellence in instructional design further explains this pattern, as the 35 studies indicating favorable learning outcomes (78%) generally incorporated explicit focus on pedagogical design: well-defined learning objectives, gamification elements tailored to learning progression, and synchronization of game feedback with content mastery. Conversely, the seven studies with mixed results frequently referred to gamification as an "add-on" without significant instructional redesign.

Well-designed gamified learning facilitates deep processing. The studies indicating favorable impacts on learning outcomes (n=35, 78%) exhibit common characteristics that imply mechanisms for success. Levels and incremental difficulty settings facilitate the maintenance of Flow, the ideal state in which challenge aligns with skill level (Csikszentmihalyi & Csikzentmihaly, 1990). Twenty-three studies (47%) utilized adaptive difficulty systems, while 22 (96%) indicated favorable learning outcomes. By mitigating both boredom (insufficient challenge) and anxiety (excessive challenge), gamification can facilitate prolonged cognitive engagement and learning. Points and progress indicators deliver prompt performance evaluation. Research in learning sciences indicates that prompt, precise feedback improves learning outcomes (Hattie & Timperley, 2007). Thirty-four studies (69%) employed points or progress visualizations to deliver real-time feedback on performance, enabling swift error correction and concept enhancement. Motivation augmented by gamification fosters task persistence and encourages help-seeking behavior. Eight studies (16%) assessed persistence (time-on-task, number of attempts on difficult problems) and determined that gamification enhanced help-seeking behaviors and retry attempts on challenging content, subsequently improving learning (mastery through practice).

Various gamification elements exhibit differing levels of effectiveness based on context. Ninety percent (n=38) of the studies analyzed indicated enhancements in engagement, while 76% (n=32) reported improvements in learning outcomes. Points serve as competence feedback, consistent with Self-Determination Theory and principles of behavioral reinforcement. The effectiveness is substantial across various age demographics and topics.

Emblems show high engagement effectiveness at 89%, while learning outcomes are comparatively lower at 74%. The recognition function of badges significantly enhances the

relatedness needs and motivation of socially-oriented learners. Nonetheless, effectiveness is contingent upon the significance and attainability of badges: inconsequential badges (easily obtained) or inadequately defined badges exhibit negligible educational impact. Levels are highly successful in engagement at 92% and learning outcomes (80%). Progress visualizations associated with levels enhance motivation and offer structured challenge advancement. The efficacy of levels may indicate that they integrate competence feedback (progress) with autonomy (students advance at their own pace). Leaderboards represent the most context-sensitive component. Although 79% of studies indicated enhancements in engagement, leaderboards exhibited the lowest rate of favorable learning outcomes at 64% and were the sole instances of adverse effects, including anxiety and diminished persistence among low-performing students. Six studies indicated that competitive leaderboards demoralize underperforming students by emphasizing performance disparities, thereby diminishing persistence and the propensity to seek assistance (Buckley & Doyle, 2016; Tsay et al., 2020). In 13 studies utilizing non-comparative leaderboards (displaying individual progress over time) or peer collaboration settings, leaderboards improved both engagement and learning outcomes. This finding underscores that the efficacy of leaderboards is contingent upon the implementation context and the attributes of the students.

The ARCS Framework offers a supplementary theoretical model for elucidating the mechanisms of gamification. The Attention, Relevance, Confidence, and Satisfaction model proposed by Keller (2009) explains how gamification intrinsically engages attention via novelty, interactive mechanisms, and prompt feedback. All 49 studies implicitly depended on attention capture; however, only four studies (8%) explicitly assessed attention through eye-tracking or task-focused metrics.

Gamification relevance by increasing the significance and intrinsic appeal of learning by incorporating game narratives and providing options in learning pathways. Only 12 studies (24%) addressed relevance by contextualizing gamification within authentic problem-solving; the majority of gamification was generic and not subject-specific. Research demonstrating contextual significance (e.g., business simulation games in management curricula, diagnostic games in medical training) indicated enhanced learning outcomes.

Confidence is enhanced through scaffolded difficulty, adaptive systems, and progress visualization. Twenty-three studies (47%) utilizing adaptive difficulty demonstrated significantly elevated confidence and persistence metrics. Satisfaction is Immediate rewards (points, badges) and progress visualization fulfill students' desire for recognition of their achievements. Nonetheless, intrinsic satisfaction derived from the learning process was infrequently assessed independently from extrinsic reward satisfaction. Research demonstrating a successful shift for students from extrinsic (game rewards) to intrinsic (mastery interest) satisfaction in learning exhibited the greatest long-term learning advancements.

Contextual Variables further Regulate the Efficacy of Gamification. Higher education studies (n=28) indicated an 80% positive impact on learning outcomes, whereas secondary education studies (n=16) reported a 73% positive impact. This disparity may indicate that higher education students exhibit superior self-regulation and are more adept at integrating game mechanics with profound learning, while higher education content may also be more conducive to gamification design, such as business cases and clinical simulations.

The Geospatial and Technological Framework also influences outcomes. In developed countries (n=38), 81% reported positive effects on learning outcomes. While in developing countries (n=11), 70% reported positive effects, this 11-percentage-point disparity corresponds with the findings of Cavus et al. (2023), which indicate that insufficient technological infrastructure limits the efficacy of gamification. In resource-constrained environments, technological malfunctions, inadequate internet connectivity, and device constraints diminish system reliability and user engagement.

Research Integrity further differentiates outcomes. Studies of superior quality (MMAT ≥75%): 83% beneficial impact on educational results. Studies of moderate quality: 67% exhibit positive effects. This gradient indicates that methodological rigor is associated with more robust evidence of effectiveness, thereby partially substantiating the assertion that well-structured gamification is effective.

The findings of this review are generally consistent with Previous Systematic Reviews and Meta-Analyses Zeng et al. (2024) conducted a meta-analysis of 87 studies, revealing a medium effect size (d=0.52) on learning outcomes, which aligns with this review's observation of predominantly positive yet variable effects. Cavus et al. (2023) reported that 71% of studies demonstrated positive effects, closely aligning with this review's rate of 78%. Puspitasari & Arifin (2023) conducted a meta-analysis on gamification in online learning, revealing positive effects on engagement and outcomes. However, effect sizes differed based on element type and context, aligning with the detailed element-specific analysis of this review.

This review extends prior research by explicitly delineating the decoupling of engagement and learning outcomes, elucidating theoretical frameworks such as Self-Determination Theory, Flow, and ARCS that clarify the conditions under which effects emerge, cataloging adverse effects and moderating variables including leaderboards and study quality, and examining geographic and infrastructural disparities.

LIMITATIONS

There are several limitations of this study that need to be considered in interpreting these findings. First, only research published in English was included in the review and as such, there is a chance that the evidence base from non-English-speaking regions (e.g., sub-Saharan Africa, central Asia, and some parts of South America) may not have been represented, and so geo-generalizability could be reduced. Second, the operationalization of "learning outcomes" differed significantly from one included study to another (ranging from grades and test scores to learning gains, retention, and skills assessments), which led to a considerable lack of measurement harmony and reduced comparability across results. Third, a considerable number of studies used self-report measures only on engagement and learning in response to gamification, which are vulnerable to social desirability bias and an overestimation of the effect. Fourth, potential publication bias should be taken into account; authors who report that their studies' findings are positive are more likely to have these studies published, possibly resulting in a skewed evidence base toward gamification. Fifth, the majority of studies are short duration (one academic semester or approximately 12 weeks), examining novelty effects, and no evidence on long-term outcomes exists. 6 More than half of the included studies did not explicitly address existing learning theories, which makes it difficult to explain why gamification influences the learning process.

The practical and policy implications of these results call for careful, pedagogically founded consideration in advocating gamification in education. To educators and course designers, it is important to understand that more engagement in learning doesn't necessarily mean more learning, so game elements should be purposefully aligned with learning objectives so as not to detract from the core content. The choice of a complementary gamification element depends on the learning objective, the student type, and who the learner is: points derive an effective formative feedback tool aiding in monitoring progress and levels, badges are suitable to signify milestones and build social cohesion, leaderboards should be implemented with great caution in cooperative and low-stakes contexts to avoid shaming their users. If adaptive difficulty is implemented, learning can be further improved as a result of maintained Flow, and designing reward structures that shift students gradually from an extrinsic to an intrinsic motivation has positive effects on long-term learning.

For instructional designers and educational technologists, pedagogical alignment should be the guiding principle, keeping gamified elements connected to the learning objectives, content pacing, and assessment, rather than superficially appended. The focus on technological infrastructure is also vital, especially in the context of resource-poor settings, as system dependability and offline availability can be the make-or-break factors for implementation success. Also, designers should expect novelty effects and incorporate updates to game mechanics to sustain motivation over time.

Also, a pedagogically aligned approach should be the lowest cost central priority, and what that means is gamified elements should be aligned with learning objectives, the content sequencing, and assessment, not layered on externally. Technological infrastructure also requires attention=20, particularly in resource-poor settings where system reliability and availability for use offline could

make or break implementation. Furthermore, game designers should expect to run into novelty effects and schedule periodic updates to game mechanics to sustain motivation.

From the policy and administration end, gamification should be considered a vehicle to increase engagement and not a panacea for all learning woes. Support for gamified platforms should therefore be coupled with professional development to enhance teachers' pedagogical capacity. Policy makers should also focus on funding for research in underexplored domains and disciplines, and for student populations, particularly at-risk learners or those with diverse linguistic backgrounds, to expand the evidence base. Finally, there is a need to set institutional guidelines for the use of gamification that promote pedagogical rigor, quality assurance, and ethical safeguards to protect against the implementation of toxic design approaches that unduly prioritize manipulation over transformative educational outcomes.

CONCLUSION

The analysis of 49 peer-reviewed studies (2013-2024) through a systematic review shows that gamification is a potent technique to attract students' attention in digital education. About 84% of the researched works have recorded the positive impact of this approach. Nevertheless, there exists a more complex interplay between engagement and learning outcomes. A significant portion (78%) of studies have shown that learning outcomes have been elevated; however, large differences have been observed, which depend on various factors such as the quality of the design, pedagogical alignment, choice of gamification elements, and situational factors. The key points indicate that the effectiveness of gamification depends on the interconnection of three main components, namely the theoretical foundation of learning psychology, including Self-Determination Theory, Flow Theory, and the ARCS Model; pedagogical considerations that emphasize the alignment between game mechanics and educational objectives; and contextual factors encompassing student characteristics, available technological resources, and the level of education. This investigation, however, is most important in showing that the increase in engagement does not necessarily lead to better learning, an outcome that has significant implications for practice. Educators should not be satisfied with merely adding a 'gamify-as-add-on' approach; instead, they should consciously incorporate gamification as a fundamental element of their pedagogical strategy, one that is in harmony with clear learning objectives and instructional design principles.

By way of mechanisms, future inquiries can illuminate the occurrence and reasons of gamification success or failure through long-term observational studies, individual differences analyses, and neurobiological research. Besides, investigations in less-represented geographical areas and different student groups will improve the generalizability and the practical aspect of the findings. In sum, gamification can be a powerful tool to make online learning more engaging if it is designed carefully and executed with pedagogical purposes. The discipline has evolved past the stage of naive excitement to an evidence-based understanding of the potential and limitations of gamification, which places it as one of the many valuable resources in the educational designer's toolkit.

AUTHOR CONTRIBUTIONS

GAW contributed to the literature search, study screening and selection process, data extraction, and synthesis of findings. He also led the writing of the original draft and the preparation of tables and figures. AAK was responsible for methodological design, data analysis strategy, and validation of the review process, and contributed substantially to the interpretation of results and manuscript revision. MR contributed to the conceptualization of the study, provided academic supervision throughout the research process, and critically reviewed the manuscript to ensure theoretical coherence and methodological rigor. S contributed to data curation, thematic organization of the reviewed studies, and supported the interpretation of findings from an educational and pedagogical perspective. He also participated in manuscript review and editing. NS contributed to methodological refinement, theoretical integration, and high-level academic review of the manuscript, particularly in strengthening the discussion and ensuring alignment with

educational research standards. All authors reviewed and approved the final manuscript and agreed to be accountable for all aspects of the work.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the academic support and institutional facilities provided by Universitas Pendidikan Indonesia (UPI) during the completion of this study. No external funding was received.

REFERENCES

Alenezi, A. (2023). Teacher perspectives on AI-driven gamification: Impact on student motivation, engagement, and learning outcomes. *Information Technologies and Learning Tools, 97*(5), 138–148. https://doi.org/10.33407/itlt.v97i5.5437

Anunpattana, P., Khalid, M. N. A., Iida, H., & Inchamnan, W. (2021). Capturing potential impact of challenge-based gamification on gamified quizzing in the classroom. *Heliyon, 7*(12), Article e08637. https://doi.org/10.1016/j.heliyon.2021.e08637

Asmawati, L. (2023). The development of puzzle games for early childhood based on the Banten local culture. *Jurnal Ilmiah Peuradeun, 11*(2), 531–550. https://doi.org/10.26811/peuradeun.v11i2.895

Barnett-Page, E., & Thomas, J. (2009). Methods for the synthesis of qualitative research: A critical review. *BMC Medical Research Methodology, 9*, 59. https://doi.org/10.1186/1471-2288-9-59

Bennani, S., Maalel, A., & Ben Ghezala, H. (2022). Adaptive gamification in e-learning: A literature review and future challenges. *Computer Applications in Engineering Education, 30*(2), 628–642. https://doi.org/10.1002/cae.22477

Borenstein, M., Cooper, H., Hedges, L., & Valentine, J. (2009). Effect sizes for continuous data. In H. Cooper, L. Hedges, & J. Valentine (Eds.), *The handbook of research synthesis and meta-analysis* (2nd ed., pp. 221–235). Russell Sage Foundation.

Buckley, P., & Doyle, E. (2016). Gamification and student motivation. *Interactive Learning Environments, 24*(6), 1162–1175. https://doi.org/10.1080/10494820.2014.964263

Cavus, N., Ibrahim, I., Okonkwo, M. O., Ayansina, N. B., & Modupeola, T. (2023). The effects of gamification in education: A systematic literature review. *BRAIN. Broad Research in Artificial Intelligence and Neuroscience, 14*(2), 211–241. https://doi.org/10.18662/brain/14.2/452

Chukwu, J. (2024). The effectiveness of gamification in online learning. *Journal of Online and Distance Learning, 3*, 53–65.

Csikszentmihalyi, M. (1990). *Flow: The psychology of optimal experience*. Harper & Row.

Deci, E. L., & Ryan, R. M. (2012). Motivation, personality, and development within embedded social contexts: An overview of Self-Determination Theory. In R. M. Ryan (Ed.), *The Oxford handbook of human motivation* (pp. 85–107). Oxford University Press.

Dichev, C., & Dicheva, D. (2017). Gamifying education: What is known, what is believed and what remains uncertain: A critical review. *International Journal of Educational Technology in Higher Education, 14*(1), 9. https://doi.org/10.1186/s41239-017-0042-5

Durach, C. F., Kembro, J., & Wieland, A. (2017). A new paradigm for systematic literature reviews in supply chain management. *Journal of Supply Chain Management, 53*(4), 67–85. https://doi.org/10.1111/jscm.12145

García-Peñalvo, F. J. (2022). Developing robust state-of-the-art reports: Systematic Literature Reviews. *Education in the Knowledge Society, 23*, Article E28600. https://doi.org/10.14201/eks.28600

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*(1), 81–112. https://doi.org/10.3102/003465430298487

Hong, Q. N., Gonzalez-Reyes, A., & Pluye, P. (2018). Improving the usefulness of a tool for appraising the quality of qualitative, quantitative and mixed methods studies: The Mixed Methods Appraisal Tool (MMAT). *Journal of Evaluation in Clinical Practice, 24*(3), 459–467. https://doi.org/10.1111/jep.12884

Huang, S., & Zhang, H. (2024). Gamification acceptance model towards online learning among college students: An empirical study based on mediation and moderation. *Asian Education*

and Development Studies, 13(2), 150–167. https://doi.org/10.1108/AEDS-11-2023-0152

Hussein, T., Cemal Nat, M., Hasan, H., & Mahdi, A. (2022). Development and evaluation of an online gamified assessment environment. 4th International Conference on Communication Engineering and Computer Science (CIC-COCOS'2022). https://doi.org/10.24086/cocos2022/paper.744

Ibáñez, M.-B., Di-Serio, Á., & Delgado-Kloos, C. (2014). Gamification for engaging computer science students in learning activities: A case study. *IEEE Transactions on Learning Technologies, 7*(3), 291–301. https://doi.org/10.1109/TLT.2014.2329293

Ikhsan, M., Rochaminah, S., & Mastura, A. (2024). Development of Geo-Math application by integrating GeoGebra applets to improve students' spatial ability. *Jurnal Ilmiah Peuradeun, 12*(3), 1129–1154. https://doi.org/10.26811/peuradeun.v12i3.1492

Ishaq, K., Mat Zin, N. A., Rosdi, F., Jehanghir, M., Ishaq, S., & Abid, A. (2021). Mobile-assisted and gamification-based language learning: A systematic literature review. *PeerJ Computer Science, 7*, e496. https://doi.org/10.7717/peerj-cs.496

Jang, J., Park, J. J. Y., & Yi, M. Y. (2015). Gamification of online learning. In C. Conati, N. Heffernan, A. Mitrovic & M. F. Verdejo (Eds.), *Artificial Intelligence in Education* (pp. 646–649). Springer.

Karabıyık, Ü. (2024). Investigation of the effect of gamified learning on motivation and success in math class. *Journal of Education and Practice, 8*, 32–52. https://doi.org/10.47941/jep.1645

Keller, J. M. (2009). *Motivational design for learning and performance: The ARCS model approach*. Springer.

Kitchenham, B. A. (2004). *Procedures for undertaking systematic reviews*. Technical Report EBSE-2004-01. Keele University & National ICT Australia.

Li, M., Ma, S., & Shi, Y. (2023). Examining the effectiveness of gamification as a tool promoting teaching and learning in educational settings: A meta-analysis. *Frontiers in Psychology, 14*, 1253549. https://doi.org/10.3389/fpsyg.2023.1253549

Looyestyn, J., Kernot, J., Boshoff, K., Ryan, J., Edney, S., & Maher, C. (2017). Does gamification increase engagement with online programs? A systematic review. *PLOS ONE, 12*(3), e0173403. https://doi.org/10.1371/journal.pone.0173403

Mahmud, S., Musfirah, A. M. R., Marlina, C. N., Fitria, S., Bujangga, H., Pranajaya, S. A., & Alfiyanto, A. (2024). Integrating Howard Gardner's multiple intelligences in Islamic education: A systematic review of Indonesian practices. *Jurnal Ilmiah Peuradeun, 12*(3), 1017–1050. https://doi.org/10.26811/peuradeun.v12i3.1215

Merhabi, M. A., Petridis, P., & Khusainova, R. (2021). Gamification for brand value co-creation: A systematic literature review. *Information, 12*(9). https://doi.org/10.3390/info12090345

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & The PRISMA Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Medicine, 6*(7), e1000097. https://doi.org/10.1371/journal.pmed.1000097

Nathan, L. S., & Hashim, H. (2023). A systematic review on gamified learning for improving language skills in ESL classrooms. *International Journal of Academic Research in Business and Social Sciences, 13*(4), (458-469). http://dx.doi.org/10.6007/IJARBSS/v13-i4/16873

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., … & Moher, D. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *Systematic Reviews, 10*(89). https://doi.org/10.1186/s13643-021-01626-4

Petticrew, M. & Roberts, H., 2006. *Systematic reviews in the social sciences: A practical guide*. Blackwell Publishing / Wiley.

Popay, J., Roberts, H., Sowden, A., Petticrew, M., Arai, L., Rodgers, M., Britten, N., Roen, K., & Duffy, S. (2006). Guidance on the conduct of narrative synthesis in systematic reviews. *ESRC Methods Programme*

Puspitasari, I., & Arifin, S. (2023). Implementation of gamification on learning motivation: A meta-analysis study. *International Journal of Progressive Sciences and Technologies, 40*(1), 356–360. https://doi.org/10.52155/ijpsat.v40.1.5596

Romero-Rodríguez, L. M., Ramírez-Montoya, M. S., & Aguaded, I. (2020). Determining factors in MOOCs completion rates: Application test in energy sustainability courses. *Sustainability, 12*(7), 7287. https://doi.org/10.3390/su12072893

Sailer, M., & Homner, L. (2020). The gamification of learning: A meta-analysis. *Educational*

Psychology Review, 32(1), 77–112. https://doi.org/10.1007/s10648-019-09498-w

Sweller, J., Ayres, P., & Kalyuga, S. (2011). *Measuring cognitive load in Cognitive Load Theory*. In J. Sweller, P. Ayres & S. Kalyuga (Eds.), *Cognitive Load Theory* (1st ed., pp. 71–85). Springer.

Tranfield, D., Denyer, D., & Smart, P. (2003). Towards a methodology for developing evidence-informed management knowledge by means of systematic review. *British Journal of Management, 14*(3), 207–222. https://doi.org/10.1111/1467-8551.00375

Tsay, C. H.-H., Kofinas, A. K., Trivedi, S. K., & Yang, Y. (2020). Overcoming the novelty effect in online gamified learning systems: An empirical evaluation of student engagement and performance. *Journal of Computer Assisted Learning, 36*(2), 128–146. https://doi.org/10.1111/jcal.12385

van Gaalen, A. E. J., Brouwer, J., Schönrock-Adema, J., Bouwkamp-Timmer, T., Jaarsma, A. D. C., & Georgiadis, J. R. (2021). Gamification of health professions education: A systematic review. *Advances in Health Sciences Education, 26*(2), 683–711. https://doi.org/10.1007/s10459-020-10000-3

Wangi, N. B. S., Chandra, N. E., & Ali, A. Z. (2022). Application of gamification-based online learning on student learning outcomes during the Covid-19 pandemic. *International Journal of Research and Review, 9*(11), 554–561. https://doi.org/10.52403/ijrr.20221172

Wulan, D. R., Nainggolan, D. M., Hidayat, Y., Rohman, T., & Fiyul, A. Y. (2024). Exploring the benefits and challenges of gamification in enhancing student learning outcomes. *Global International Journal of Innovative Research, 2*(7 SE-Articles), 1657–1674. https://doi.org/10.59613/global.v2i7.238

Zeng, J., Sun, D., Looi, C.-K., & Fan, A. C. W. (2024). Exploring the impact of gamification on students' academic performance: A comprehensive meta-analysis of studies from 2008 to 2023. *British Journal of Educational Technology, 55*(6), 2478–2502. https://doi.org/10.1111/bjet.13471