Contemporary Educational Technology, 2026, **18**(3), ep670
ISSN: 1309-517X (Online)

Review Article

Structural analysis of high-impact literature on gamification and game-based learning in higher education

Hugo Alejandro Muñoz-Bonilla [1*]
0000-0001-8757-3909

Juana María Ortega-Tudela [2]
0000-0003-3649-3524

[1] Corporación Universitaria Minuto de Dios – UNIMINUTO, Bogotá, COLOMBIA
[2] Universidad de Jaén, Jaén, SPAIN
[*] Corresponding author: hmuozbon@uniminuto.edu.co

Citation: Muñoz-Bonilla, H. A., & Ortega-Tudela, J. M. (2026). Structural analysis of high-impact literature on gamification and game-based learning in higher education. *Contemporary Educational Technology, 18*(3), Article ep670. https://doi.org/10.30935/cedtech/18870

ARTICLE INFO

Received: 3 May 2026

Accepted: 23 Jun 2026

ABSTRACT

Gamification research in higher education remains fragmented across conflicting theoretical frameworks, limiting evidence synthesis. This study examined conceptual structures and methodological patterns in 25 frequently cited articles (2011-2025) using Scopus, Web of Science, and PRISMA 2020 guidelines. Four conceptual categories emerged: tactical training (mean citations = 144.86), structural motivational gamification (172.14), user profile analysis (87.80), and adaptive training with artificial intelligence (123.83). Quasi-experimental designs predominated (68%, all non-probabilistic). Critical gaps included: demographic analysis (28%, n = 7), emerging technologies (4%, n = 1), and absent longitudinal follow-up. Domains concentrated on computer science (40%) and education (28%). High-impact research is organized around distinct epistemologies rather than unified framework. The dominance of scalable structural approaches over pedagogically integrated designs suggests a practice-evidence gap. Systematic integration of demographic analysis, emerging technologies, and longitudinal designs are necessary to advance the field.

Keywords: educational technology, gamification, higher education, instructional design

INTRODUCTION

According to Deci and Ryan (2000) there are three basic psychological needs that must be fulfilled in order for motivation to be enhanced: autonomy, competence, and relatedness (i.e., feeling accepted by others). Higher Education places demands on individuals whereby they must begin to function with a greater degree of self-direction to develop their professional identity; therefore these three needs will become increasingly pertinent.

Despite advanced cognitive abilities, university students often remain in emerging adulthood (Sussman & Arnett, 2014), characterized by ongoing maturation in planning, decision-making, and self-regulation. The fact that there is a gap between their individual intellectual maturity and their ongoing developing to emerging adulthood creates an obstacle for educators, traditional approaches may not adequately foster motivation and autonomy in complex learning environments.

Within this context, gamification—the use of game mechanics in non-game environments—has emerged as a strategy to enhance student engagement (Deterding et al., 2011), has been identified as a strategy to enhance motivation and participation of students in academically rigorous, structured learning environments. There are examples of evidence-based studies (e.g., the serious game (SG) TRIVIREC; Rodríguez & Mezquita, 2016), which demonstrate that gamification can enhance motivation and cooperative/collaborative learning by using structured designs. Additionally, there are studies that have indicated the value of play as a means to engage students with low-interest content (Toledo et al., 2017).

Research on gamification has proliferated, as evidenced by multiple systematic reviews. Marti-Parreno et al. (2016) identify effectiveness, engagement, and social interaction as central themes, with limited experimental rigor. Subhash and Cudney (2018) show that reward-based mechanisms can enhance motivation, although their effectiveness depends on context.

More recent studies indicate consolidation and diversification. Wang et al. (2022) and Behl et al. (2022) highlight thematic expansion toward personalization and engagement metrics, while Chugh and Turnbull (2023), together with Dallaqua et al. (2024), report sustained growth, particularly in higher education.

Despite field consolidation, the literature reveals three persistent gaps. First, demographic analysis remains limited, with insufficient attention to how gender and age variables moderate gamification effects. Second, emerging technologies—artificial intelligence (AI), Internet of things (IoT), augmented and virtual reality—are minimally explored despite their pedagogical potential. Third, the investigation of moderating variables across different educational contexts remains underdeveloped. These gaps prompt critical examination of the conceptual structures and methodological patterns that characterize high-impact gamification research in higher education (2011-2025). The study proposes three objectives:

- Identify the most impactful articles on gamification in higher education (2011-2025).
- Characterize bibliometric and methodological trends in Scopus and Web of Science (WoS).
- Identify the conceptual structures underlying high-impact research.

Theoretical Framework

Conceptual foundations

This study distinguishes three approaches to incorporating game elements into education. Gamification integrates game design elements into non-ludic environments while preserving the instructional structure. Game-based learning (GBL) uses complete game systems as the primary learning medium, and SGs are purpose-designed for educational outcomes (Vergara Rodríguez & Mezquita, 2016).

According to Schoebel et al. (2021), game-informed learning is viewed as part of an overarching ecosystem. The term "digital learning games" serves as a general term while still allowing for the establishment of analytical differences among other similar types of GBL. High-impact research is defined through an operational metric of publication with a high number of citations in Scopus and WoS, as per Dallaqua et al. (2024), due to the citation count serving as one measure of a given research's structural impact on a wide range of research agenda and methodology.

Progression of systematic reviews

Analyses of the evolution of the systematic reviews associated with GBL can be seen through a continual evolution of systematic review articles. Marti-Parreno et al. (2016) explained their examination of systematic reviews, addressing a set of core themes including engagement, effectiveness and poor methodological quality. In subsequent articles, Subhash and Cudney (2018) focused on the role of design elements and contextuality in producing differential outcomes.

Subsequent systematic reviews, including those conducted by Wang et al. (2022) and Behl et al. (2022), illustrate the growing diversification of the field towards personalization and analytics; while Chugh and Turnbull (2023) and Dallaqua et al. (2024) demonstrate continued growth and specialization among post-secondary institutions.

The work conducted within GBL has demonstrated an increasing body of literature but exhibits considerable gaps in variables regarding the integration of demographics, emerging technologies and moderating variables as well as a continued lack of understanding of the ways indexing systems impact research visibility. These deficiencies highlight the importance of a meta-investigative approach aimed at developing a framework for the conceptual and methodological structures of high-impact research.

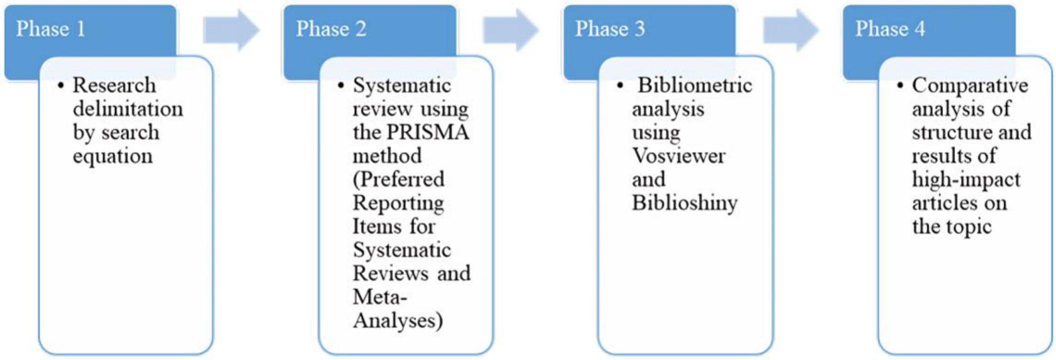

Figure 1. Research design (the authors' own elaboration)

METHOD

This is a meta-investigative exercise (research on research), so it does not directly evaluate the effectiveness of gamified interventions, but rather analyzes how the phenomenon has been studied, what methodological designs predominate, what gaps persist in the literature, and how indexing databases differ in visibility and methodological characteristics of publications.

From this perspective, the study does not aim to provide an exhaustive review of recent literature, but rather to examine the consolidated structures that have shaped the field. Accordingly, the selection criteria prioritize highly cited publications as an analytical strategy to identify the most influential contributions and the conceptual foundations underpinning the discipline. In line with Dallaqua et al. (2024), citation counts are considered a valid indicator of academic impact and for tracing the development of dominant research trajectories.

This study does not aim to perform a comparative analysis between databases. Instead, Scopus and WoS are used as complementary retrieval systems to identify high-impact publications. The unit of analysis is therefore the set of highly cited articles, rather than the databases themselves.

Within this framework, the search strategy adopts an educational technology ecosystem perspective, in which gamification and GBL are understood as related approaches within a broader field of practices that incorporate game mechanics into learning contexts. Following Schoebel et al. (2021), who propose the concept of digital learning games as an umbrella term, the use of the Boolean connector "OR" between these terms does not imply conceptual equivalence but rather aims to ensure comprehensive retrieval of high-impact literature. This decision reflects a deliberate analytical criterion: to capture the full spectrum of studies employing game mechanics and subsequently differentiate them through the four-category taxonomy developed in this study.

A documentary review study was designed with an analytical-descriptive approach in four phases: specialized search, PRISMA systematic review, bibliometric analysis (VOSviewer/Biblioshiny), and comparative structural analysis (**Figure 1**).

Data Collection

The search strategy employed a Boolean string composed of three conceptual parts combined using Boolean connectives, which can be broken down as follows: ("Motivation" OR "Student engagement" OR "Intrinsic motivation" OR "Extrinsic motivation") AND ("Gamification" OR "Game-based learning") AND ("Higher education" OR "University students").

The first part is intended to capture information about different types of motivation in educational settings. The second part of the chain identifies the concept of "gamification" or "game-based learning" to collect research dealing with different game related methods of learning. Using the "or" in each part does not indicate equal meaning or interchangeability; it serves to capture highly influential works that utilize a related concept, in this case both "gamification" and "game-based learning", so that both can be differentiated and analyzed using the taxonomy of four categories developed for the study. Finally, the third block delimits the search to studies focused on higher or university education.

The results obtained are subjected to a systematic review in the second phase of the study, following the PRISMA 2020 guidelines (Page et al., 2021) and Pardal-Refoyo et al. (2020) to identify relevant academic production on the subject. The sample universe is limited to the contents of Scopus and WoS, databases recognized for their impact and editorial rigor in the academic field (Dallaqua et al., 2024).

Following bibliometric precedent (Bornmann, 2014; Cheng et al., 2022), we selected articles with ≥ 20 citations to identify structural patterns and research trajectories. The cut-off of 20 citations as a minimum corresponded to the cut-off point for citations among the highest 20 ranked articles in each of the respective databases (i.e., when retrieved) (Bodily et al., 2019; Uthman et al., 2013).

A key limitation is temporal bias: articles from 2016-2018 had more time to accumulate citations than more recent publications, which could inflate their apparent impact. Furthermore, the potential exclusion of other innovative but less cited articles corroborates our limitations with the objective of this project (i.e., focusing on the impact of established citations rather than the impact of generating new citations, as in the case of studying popular topics in research). As a result of the aforementioned duplicates, as well as the exclusion based on the subsequent eligibility criteria, a final sample of 25 articles is obtained, which supports the selection methods described.

After applying an eligibility parameter control (article access, elimination of duplicates through identification of DOI and matching titles between Scopus and WoS, and exclusion of review articles), two additional studies were included (Ferriz-Valero et al., 2020; Murillo-Zamorano et al., 2021), selected for their relevance. This procedure guaranteed that the analysis would be based on studies with high influence in the scientific community, thus forming a final sample.

Content Analysis Protocol

The content analysis of the selected documents focuses on identifying the following variables:

- The associated database and its impact (number of citations).
- The methodological approach, type of research, and scope.
- Sample characteristics (size, type of population, and temporality).
- Theoretical frameworks employed (technology acceptance model [TAM], self-determination theory [SDT], game and playful architecture [GPA]) and implemented design elements (objectives, scores, badges, feedback, collaboration, challenges).
- The type of technology used in gamified interventions (hard, soft, mixed, emerging).
- The instruments and forms of measurement of motivation and student engagement reported in each study.
- The main findings and contributions to the design of gamified strategies and SGs.
- The disaggregated impact analysis by gender and age.
- The evidence of reporting statistical power components (effect sizes, confidence intervals, power)
- The identification of emerging conceptual categories that articulate gamified design elements and their associated theoretical frameworks (TAM, SDT, GPA).

Data extraction was performed through a coding matrix designed specifically for this study, in which the identified variables were systematically recorded.

Reliability Protocol

To guarantee coding reliability, a two-stage validation process was implemented following emerging recommendations on the appropriate use of large language models (LLMs) in systematic reviews (Dijk et al., 2023; Ge et al., 2024). First, a sample of 40% of the articles (10 of 25) was subjected to parallel coding through AI-assisted analysis, specifically Claude 3.5 Sonnet, Anthropic, as second coder. Structured prompts that replicated the coding matrix criteria were used. This approach is employed based on contemporary studies that validate the use of LLMs as assistants in data extraction for systematic reviews (Lieberum et al., 2024; Schmidt et al., 2025).

For objective variables (e.g., methodological design, sample size, and presence of demographic analysis), the level of agreement between AI coding and human coding was 94% while the level of agreement for interpretation variables (i.e., identification of theoretical framework and technological category) was 87%. These values are consistent with the benchmarks for this type of work (Jensen et al., 2025; Khraisha et al., 2024). The bulk of the discrepancies between coders were resolved through the use of primary source information; in the case of objective variables (n = 6), the human coder was able to find information in footnotes or participant sections where the AI could not and in the case of interpretive variables (n = 4), divergences arose from ambiguity in the source article itself, not due to a variable coder. All discrepancies were resolved through the consensus of the co-authors.

Three limitations of using AI-assisted coding are mentioned. The first limitation is that the AI's training data may have an over-representation of mainstream frameworks (TAM and SDT), making it possible that IT systems following less prevalent approaches are under-identified when being coded. The second limitation is that the AI has accurately identified only those frameworks for which there is an explicit name; therefore, the AI may not accurately identify those frameworks that are applied in an implicit manner. Lastly, the AI sometimes has confused constructs that are related to nature; for example, engagement vs. motivation. To help mitigate these limitations, all framework identifications by the AI required verification from a human; and coding prompts included operational definitions that require explicit identification of the theoretical frameworks being used. By combining algorithmic systematization with the expert judgement of the human coder (Fabiano et al., 2024), all final classifications were made by humans verified coder rather than the unaudited output of the AI.

In the cases where classification required interpretation (identification of theoretical frameworks or type of technology), specific search criteria were established: for the theoretical frameworks, their explicit mention was recorded in the foundation or discussion sections of each article, while technological classification was based on the description of the resources and tools reported in the methodological section.

Theoretical frameworks (TAM, SDT, GPA) were identified through keyword search in the full text of each article, recording their explicit mention in the theoretical framework or discussion. The type of technology was classified following the taxonomy proposed by Brynjolfsson and McAfee (2014) and Universidad VIU (2024): hard technologies (specialized hardware and software requiring advanced technological infrastructure), soft (methods, approaches, and tools that do not require specialized hardware), mixed (combination of gamification activities and tools), and emerging (AI, IoT, augmented reality, virtual reality, blockchain). Demographic variables (gender, age) were recorded when studies reported disaggregated analyses of these characteristics, regardless of the statistical significance of the findings.

Conceptual Taxonomy Development

After the structural analysis of the 25 most cited articles, a thematic categorization process was implemented to identify the underlying conceptual structures that organize research on gamification in higher education. This process involved three steps:

- Grouping by theoretical-methodological approach: Studies were classified according to their main emphasis on
  (a) GBL and simulation as a tactical training strategy,
  (b) structural motivational mechanics based on reward and feedback systems,
  (c) analysis of user profile and attitude as a moderating factor of effectiveness, or
  (d) adaptive systems with AI integration.
- Analysis of articulating theoretical frameworks: For each identified category, the predominant theoretical models (TAM – technology acceptance, SDT – self-determination, GPA – game design) and the specific gamified design elements employed (scores, badges, feedback, collaboration, challenges, competitiveness) were documented. The prevalence of each element and its association with reported impact factors (motivation, engagement, academic performance) were also recorded.

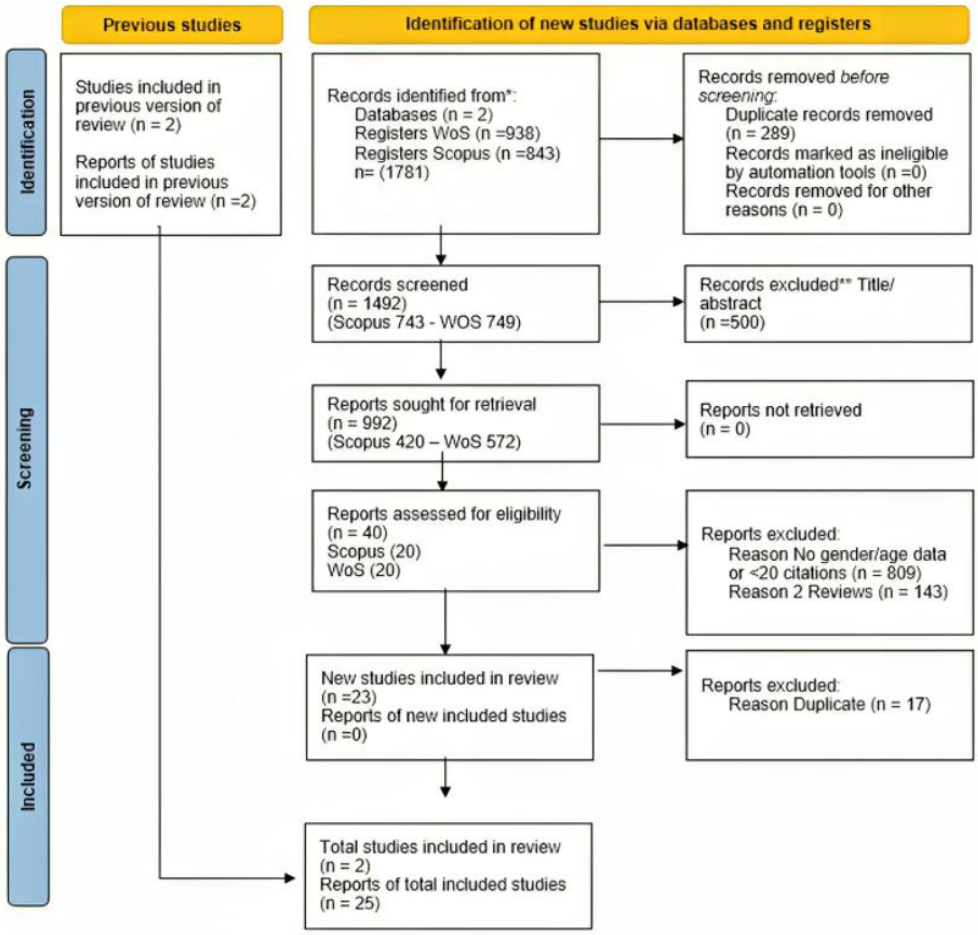

Figure 2. PRISMA methodology applied (adapted by the authors from Page et al., 2021)

- Differential bibliometric impact evaluation: The average citations and normalized citation rate were calculated for each conceptual category, with the objective of determining which approaches have generated greater academic influence in the field.

The resulting categorization would allow not only mapping the field, but identifying patterns of differential effectiveness, conceptual gaps, and strategic directions for future research.

Analysis Tools

Bibliometric analysis was performed using VOSviewer (version 1.6.18) for the construction of co-citation, co-authorship, and keyword co-occurrence networks, as well as Biblioshiny (bibliometrix package of R) for the analysis of production by country, temporal evolution, and thematic clustering. Descriptive data of the 25 articles were processed through descriptive statistics (frequencies, percentages, means) using Microsoft Excel and Jamovi 2.3. Correlations between methodological variables and citation level were evaluated through Spearman's Rho coefficient, appropriate for ordinal data and non-normal distributions.

This methodological approach allows a comprehensive understanding of trends and gaps in research on gamification in higher education, both from the bibliometric perspective and from the structural analysis of studies with greatest academic impact.

RESULTS

General Bibliometric Characterization

The search equation identified 1,781 articles. After applying the PRISMA process, two datasets were obtained for bibliometric analysis: 420 articles indexed in Scopus and 572 in WoS. The observation period spans from 2011 to 2025 (**Figure 2**).

Table 1. Comparative of production by indexed database

Source	Documents	
	Scopus	WoS
Sustainability (Switzerland)	17	24
Education and Information Technologies		21
Computers and Education	12	16
Education Sciences	13	15
Frontiers in Education	8	12
Education and Information Technologies	12	
International Journal of Game-Based Learning		11
International Journal of Educational Technology in Higher Education		10
Computer Applications in Engineering Education		9
Interactive Learning Environments		9
Ed-educational Technology Research and Development		8
International Journal of Emerging Technologies in Learning	8	
International Journal of Environmental Research and Public Health	8	
International Journal of Information and Education Technology	7	
Interactive Learning Environments	6	
Cogent Education	5	

The analysis with Biblioshiny (Aria & Cuccurullo, 2017) establishes that the mean of authors per article is 3.19 in Scopus, while in WoS it is slightly higher at 3.36. The mean citation per document is 21.18 for Scopus and 13.74 for WoS, evidencing a difference in visibility and impact of publications, according to the indexing database with which the publication journal is associated. It is identified that, during the observation period, a greater number of authors have published in journals indexed in WoS (486 more than in Scopus). The reviewed indexing sources indicate that Spain is the contribution, with an evident collaborative network with authors from Colombia and Mexico. The publication that leads dissemination is Sustainability (Switzerland) with 24 articles, followed by Education and Information Technologies with 21, and in third place Computers and Education with 16 articles (**Table 1**).

High-Impact Literature

From the bibliometric analysis, the 25 articles with greatest academic influence were selected (**Table 2**), with an average annual citation rate of 19.77 and normalized rate of 2.8. Publications from the 2018-2020 period concentrate the greatest impact, with a decreasing trend since 2021.

Methodological Patterns

In response to the first research objective on bibliometric trends and methodological characteristics, the analysis of the 25 most cited articles reveals consolidated methodological patterns in the field:

- **Approach and design:** 68% adopted a quantitative methodological approach, with deductive logic in the 72% of cases. The Quasi-experimental designs predominate (52%), all with non-probabilistic convenience samples. The 68% presented cross-sectional temporality, reflecting the preference for evaluations of immediate effects of gamified interventions.
- **Research scope:** There is evidence of a shift towards more robust approaches. Descriptive research shows a decrease compared to previous review reports. Currently, explanatory (48%) and associative (32%) approaches predominate, especially in the areas of psychology and computer science, indicating the methodological maturity of the field (**Figure 3**).

Analysis of Demographic Variables

Only 28% of studies (7 of 25) incorporated analysis by gender, constituting a critical gap identified. Of these, only two reported significant differences favorable to male participants (Ferriz-Valero et al., 2020; Kyewski & Krämer, 2018). Regarding the age, only two studies performed disaggregated analyses without finding significant differences. This scarce attention to the demographic variables limits the understanding of how the specific psycho-evolutionary characteristics of the university students moderate the effects of the gamification, representing an opportunity for future research. The Spearman rank correlations between citation count and methodological variables revealed no significant associations (all $p > .05$).

Table 2. Articles included in the structural analysis

Study	Article	Total citations	TCA*	Normalized TC**
1	Huang et al. (2019)	232	38.67	8.29
2	Yildirim (2017)	226	28.25	3.39
3	Kyewski and Krämer (2018)	218	31.14	2.65
4	Hew et al. (2016)	211	23.44	3.43
5	Troussas et al. (2020)	187	37.40	5.28
6	Tsay et al. (2018)	177	25.29	2.15
7	Topalli and Cagiltay (2018)	167	23.86	2.03
8	Huang and Hew (2018)	161	23.00	1.96
9	Fotaris et al. (2016)	158	17.56	2.57
10	Vanduhe et al. (2020)	152	30.40	4.29
11	Sanchez et al. (2020)	134	26.80	4.70
12	Zainuddin (2018)	133	19.00	2.98
13	Cózar-Gutiérrez and Sáez-López (2016)	132	14.67	2.15
14	Hwang et al. (2014)	129	11.73	2.57
15	Lin and Kaur (2018)	124	17.71	1.51
16	Sailer and Sailer (2021)	106	26.50	4.23
17	Chapman and Rich (2018)	103	14.71	1.25
18	Hou and Li (2014)	102	9.27	2.03
19	Eltahir et al. (2021)	72	18.00	2.88
20	Wiggins (2016)	92	10.22	1.90
21	Tsai et al. (2016)	90	10.00	1.86
22	Villagrasa et al. (2014)	76	6.91	1.73
23	Ferriz-Valero et al. (2020)	79	8.45	1.43
24	Nousiainen et al. (2018)	72	10.29	1.61
25	Murillo-Zamorano et al. (2021)	68	7.70	1.09

Note. *Citation rate per year & **Normalized citation rate

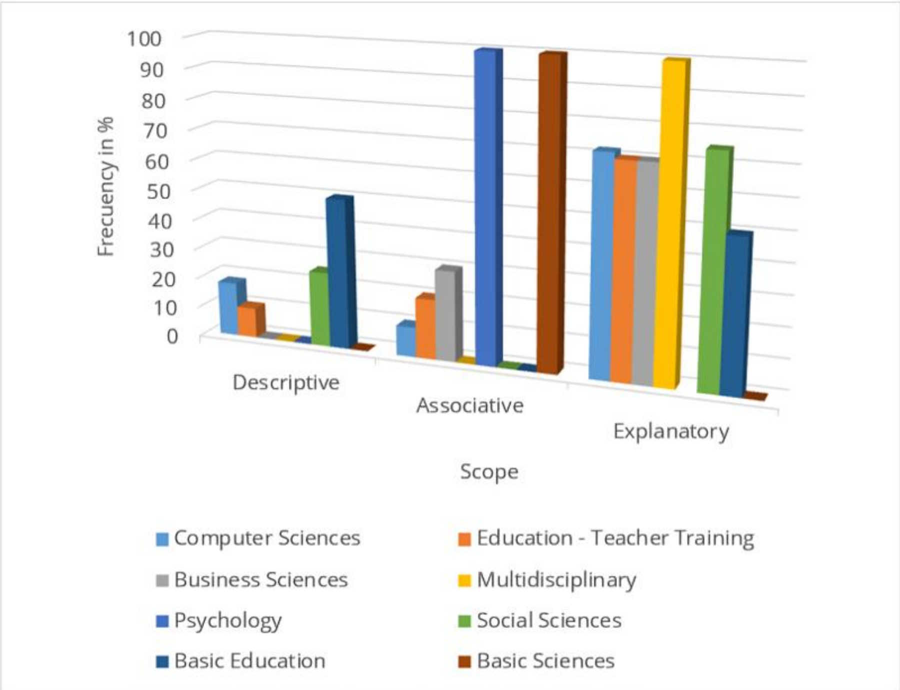

Figure 3. Behavior of research scope (the authors' own elaboration)

Effect sizes were negligible to weak: methodological approach (ρ = .265, p = .201), research design (ρ = -.044, p = .835), effect size reporting (ρ = .221, p = .288), sample size (ρ = .134, p = .534), temporality (ρ = −.185, p = .377), and research scope (ρ = .118, p = .574).

Weak correlations suggest that citation levels are largely independent of the methodological characteristics of the articles. Furthermore, the prevalence of quantitative studies among the most cited articles may reflect their dominance in the field rather than a differential advantage in impact. Analysis of

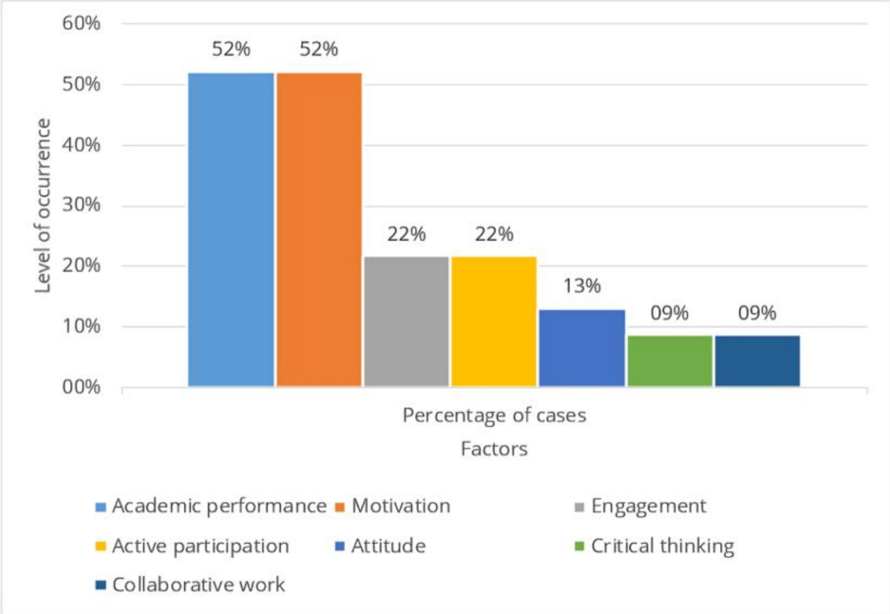

Figure 4. Prevalence of impact factors of gamification in gamified pedagogical processes (the authors' own elaboration)

technologies reveals a diversified distribution: hard technology predominates at 44% (specialized hardware and software requiring advanced infrastructure), followed by soft technology at 32% (methods and tools without specialized hardware), mixed technology at 20% (a combination of activities and tools), and only 4% employ emerging technologies (AI, IoT, augmented/virtual reality).

The Kruskal-Wallis test comparing citation counts across technology types (hard, $n = 10$, $M = 112.5$; soft, $n = 9$, $M = 127.4$; mixed, $n = 5$, $M = 116.4$; emerging, $n = 1$, $M = 132.0$) was not significant, $H(3) = 0.840$, $p = .840$, $\eta^2 < .001$ (negligible effect). However, with $N = 25$ and four groups, statistical power to detect medium effects was approximately 25-35%, meaning that a real but moderate association could remain undetected. Within this limitation, the results suggest that academic impact is largely independent of the technological sophistication of the intervention.

The low adoption of the emerging technologies (4%) represents a significant gap considering the current context (2011-2025) of proliferation of accessible gamification tools, constituting a research opportunity. Meanwhile, the analysis of the reported effects reveals prevalence of benefits in academic performance (52%) and motivation (52%), followed by improvements in engagement and active participation (21.7%) (**Figure 4**).

Specific Analysis of Effects on Motivation and Engagement

When observing motivation and student engagement as a missional element of a pedagogical process, a detailed analysis was performed of the 13 studies that reported effects on these variables:

- **Instruments employed:** The 62% used validated scales (IMI – intrinsic motivation inventory; MSLQ – motivated strategies for learning questionnaire), while 38% developed ad hoc questionnaires
- **Dimensions evaluated:** The 69% measured both intrinsic and extrinsic motivation; 23% focused exclusively on intrinsic motivation; 8% on academic engagement without differentiating motivational types.
- **Magnitude of effects:** The reported effect sizes vary considerably: 31% reported small effects ($d = 0.20$-0.49), 46% moderate effects ($d = 0.50$-0.79), and 23% large effects ($d \geq 0.80$). This variability suggests that effectiveness depends on contextual and specific design factors.
- **Identified moderators:** Three studies (Huang & Hew, 2018; Sailer & Sailer, 2021; Tsay et al., 2018) identified that the presence of immediate feedback and clear objectives positively moderated effects on intrinsic motivation. The collaboration showed mixed effects, benefiting students in technological areas more.

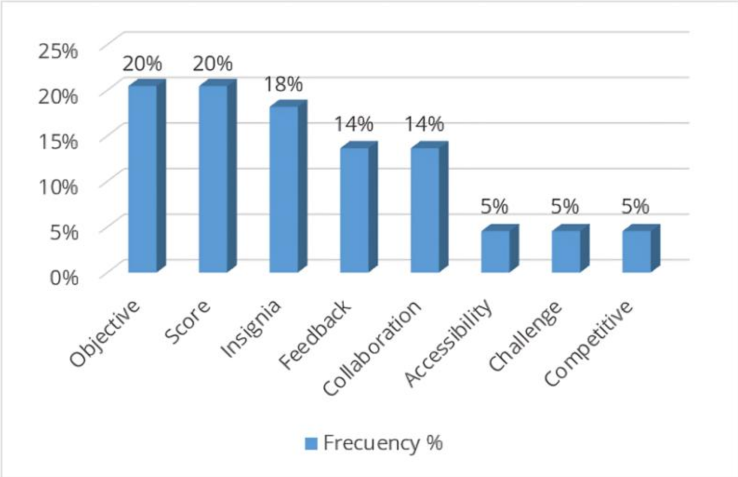

Figure 5. Prevalence of design elements associated with the impact of SGs and GBL (the authors' own elaboration)

Despite finding evidence supporting gamification as a positive motivator, it is clear that the degree to which it is effective, and the degree to which it supports improvement of educational outcomes, may be significantly influenced by the design of the gamification program and other contextual characteristics. As is evident from the variance report, there is a great degree of variability (heterogeneity) in the impacts of gamification found to occur. Many impacts vary in size – learning/collaboration from small to moderate; intrinsic motivation/academic performance - most impacts were small to moderate, but there were a few cases of substantial impact when several design elements were combined (scores + feedback + progressive challenges).

Association Between Design Elements and Reported Effects

The cross-analysis between design elements and effects on motivation reveals patterns:

- **Most effective combinations:** The studies that reported large effects ($d \geq 0.80$) on intrinsic motivation consistently employed the triad: clear objectives + immediate feedback + progressive challenges (Huang et al., 2019; Sailer & Sailer, 2021). This combination simultaneously satisfies needs of competence (SDT) and perceived usefulness (TAM).
- **Moderate effects:** Interventions based primarily on scores and badges without qualitative feedback showed moderate effects ($d = 0.50$-0.79), suggesting that extrinsic reward elements are less potent than those that support intrinsic psychological needs (Kyewski & Krämer, 2018; Tsay et al., 2018).
- **Moderating role of collaboration:** Collaboration showed heterogeneous effects: positive in computer science and engineering areas (where it is culturally valued), but neutral or negative in contexts where students prefer individual work (Chapman & Rich, 2018; Troussas et al., 2020).
- **Implications for design:** These findings suggest that the effectiveness of gamification does not depend on the quantity of elements implemented, but on their coherent articulation with specific psychological needs of the educational context. The most successful interventions integrated elements from at least two theoretical frameworks, especially combining TAM (to facilitate technological adoption) with SDT (to sustain intrinsic motivation in the long term).

Design Elements for GBL and SGs

The GBL and SGs design elements identified in the 25 most cited articles are articulated with three predominant theoretical frameworks: TAM (Davis, 1989), SDT (Deci & Ryan, 2000) and GPA (Schell, 2008). Simultaneously, the frequency with which specific design elements appear in the analyzed gamified interventions allows identifying that the prevalent elements are clear objectives (84%) and scores or leaderboards (76%), followed by immediate feedback (68%), badges or achievements (60%), collaboration (44%), progressive challenges (40%), personalized accessibility (28%) and competitiveness (24%) (**Figure 5**).

Table 3. Summary of groups according to associated theories

Element	TAM	SDT	GPA
Objective	√		
Score	√		√
Insignia		√	√
Feedback	√		
Collaboration		√	
Accessibility		√	
Challenge		√	
Competitive			√

Table 4. Thematic modeling

Conceptual category	Predominant approach	Characteristics and supporting metrics	Studies*
1. Tactical training (GBL and simulation)	GBL/SG	Focus on translating abstract concepts into reality through practice in safe environments. Perceptions of validity and ease predict 97.7% of variance in learning.	[8, 14, 6, 13, 21, 9, 7]
2. Structural motivational gamification	Gamification 100%	Use mechanics (points, leaderboards, badges) to foster persistence and reduce dropout. Focuses on self-efficacy and satisfaction (up to 74%) to regulate academic behavior.	[1, 2, 3, 4, 12, 16, 23]
3. User profile and attitude	Gamification + GBL	Analyze how personality, gender, and motivation influence acceptance. Attitudes and motivation impact perceived learning, even surpassing the "Flow" state.	[15, 17, 19, 24, 25]
4. Adaptive training and innovation (AI)	GBL + SG + AI integration	Systems that adjust content and feedback according to usage patterns. Includes human-AI collaboration, although warns that excessive dependence negatively affects learning if there is no self-reflection.	[5, 10, 11, 18, 20, 22]

Note. *Numbers refer to articles in **Table 2**

Table 3 synthesizes how the identified elements are linked with each theoretical framework. Elements associated with TAM (clear objectives, scores, feedback) are related to the perception of usefulness and ease of use. Those linked to SDT (challenges, badges, collaboration, accessibility) satisfy the psychological needs of competence, autonomy, and social relatedness. GPA elements (badges, competitiveness, scores) function as game mechanics that drive engagement. Some elements fulfill multiple theoretical functions, such as badges and scores that appear in more than one framework.

Conceptual Structures of High-Impact Research

The analysis of methodology and results allow generating a categorization of design trends of gamified tools within the set of top research (**Table 4**).

The systematization of empirical evidence grouped by conceptual category reveals a concentration of academic visibility in structural gamification. This category accumulates the greatest number of total citations in the corpus (1,205) and the highest average (172.14), suggesting that authors such as Huang et al. (2019) and Hew et al. (2016) occupy a prominent position in the citation network of the field, characterized by their use of badges, points, and leaderboards to regulate academic behavior

Tactical training concentrates 1,014 citations in the corpus. Authors such as Huang and Hew (2018) and Hwang (2014) are among the most cited within this cluster, whose studies examine how GBL approaches the translation of abstract concepts into practical realities

The adaptive training and AI category presents a lower citation average (123.83), consistent with its more recent publication dates. Troussas et al. (2020) and Vanduhe et al. (2020) are among the most cited studies in this cluster. The temporal distribution of citations in this category is consistent with an emerging research stream, integrating tools such as chatbots and AI assistants to offer 24/7 support and improve learning flow

The user profile and attitude category presents the lowest citation average in the corpus (87.80). This distribution is consistent with the more recent publication dates of the studies in this cluster and with the specialized focus of their research questions. Authors such as Nousiainen (2018) and Murillo-Zamorano (2021) are among the most cited within this group, contributing evidence on how personalization according to gender and user type moderates gamification effectiveness (**Table 5**).

Table 5. Impact by conceptual category

Conceptual category	Number of authors	Total citations	Average citations reached
Structural motivational gamification	7	1,205	172.14
Tactical training (GBL and simulation)	7	1,014	144.86
Adaptive training and innovation (AI)	6	743	123.83
User profile and attitude analysis	5	439	87.80
Total	25	3,401	136.04

DISCUSSION

According to the study's findings, the body of research on gamified learning in university education is organized around different epistemologies rather than a common framework. Each of these epistemologies has a unique theoretical basis, design rationale, and action mechanism. As discussed by Fernández-Velásquez et al.(2025), the differences among epistemologies lead to the existence of conflicting foundational constructs, all of which contribute to the conceptualization and evaluation of how 'effect' is measured.

There are three categories that display underlying conceptual conflicts:

(1) extrinsic vs. intrinsic motivation (category 1 and category 2);

(2) fixed vs. adaptive design (category 1 and category 2 vs. 3); and

(3) average effects vs. profile-dependent effects (categories 1, 2, and 3 vs. 4).

These conceptual conflicts are not purely theoretical; rather, they affect how interventions work, to whom they work, and under which circumstances. Thus, the field does not appear to be a collective aggregate of knowledge, but rather a site of coexistence and at times conflicting views.

Analyzing this corpus results in the identification of a possible contradiction between popularity and viability (effectiveness). Category 2 (structural motivational) concentrates the highest citation volume in the corpus. Notably, the studies in this category also tend to report more moderate effect sizes than those in category 1 (tactical training). This coexistence of high visibility and moderate reported effects is consistent with the interpretation that citation counts in this field may reflect factors such as scalability, ease of replication, and publication timing, rather than instructional effectiveness per se. A plausible reading is that structural gamification (points and badges) has gained citation momentum due to ease of application and dissemination, although this interpretation cannot be confirmed from citation data alone. The structural approach utilizes extrinsic motivation, which is more often associated with short-term engagement or use.

On the other hand, tactical training (category 1) creates a more effective use of a triad of clear objectives, immediate feedback, and progressively more challenging experiences. The triad incorporates intrinsic motivation and thus connects with greater learning-ability over time.

The conclusions drawn from the triad in tactical training suggest that a substantial amount of variance in learning can be accounted for, reinforcing a key point—the effectiveness of gamified elements is not a function of how many elements are used, but their coherent connection to psychological/principles within the gamification construct. Therefore, the use or development of elaborate scoring systems that do not include formative feedback represent limited opportunity and/or are the least efficient strategy.

The temporal analysis suggests that category 4 (adaptive/AI) is in an accelerated growth phase, particularly after 2020, while category 2 (structural motivational) shows a pattern consistent with what bibliometric evidence describes as a consolidation phase, with publications in gamified learning peaking in 2021 before leveling off (Sergeeva et al., 2024). This shift points toward a growing interest in personalization. However, the evidence indicates that algorithmic adaptation produces moderate effects and, in some cases, may compensate for weak instructional design rather than replace it. At the same time, some studies warn about the risks of excessive reliance on automation without fostering metacognitive reflection.

Category 3 (user profile and attitude) adds another layer of complexity by showing that effects are not uniform. What works in one context may not work in another: collaboration benefits certain disciplines, but generates resistance where individual work is the norm; competitiveness motivates some students, but discourages others. This reinforces the idea that "it works on average" can be misleading, as it often conceals important inequalities in outcomes.

These disparities also have been reflected in their methodological development. In terms of methodology, the field has been moving towards more explanatory and quasi-experimental designs but has become less consistent in its evolution over time. Category 1 (tactical training) exhibits the highest level of methodological maturity, particularly with regard to reporting effect size estimates (86%). Category 2 (structural motivational) follows closely (83%). In contrast, category 4 (adaptive/AI) is still consolidating its methodological standards (67%), and category 3 (user profile) remains largely exploratory due to its focus on heterogeneity (60%). Therefore, while the field is evolving, its evolution is not uniform.

A fundamental limitation across all categories is the heavy reliance on non-probabilistic sampling, thus limiting the external validity of research findings across all categories. Another source of concern is the lack of a significant relationship between methodological rigor and citation impact, which would suggest that visibility may rely more on the publication outlet or timing than on the quality of the research being published. This provides additional support for the earlier paradox between different aspects of research quality and different ways to assess such quality.

Further fragmentation of the field is evident in the limited focus on demographic variables in the literature. With the exception of category 3 (which has a more systematic approach), the other categories of evidence do not consider demographic variables. As a result, some studies examine average effects without considering diversity, while other studies examine diversity but have limited generalizability. In areas such as STEM education, this presents a challenge. Due to limited evidence regarding equity, the application of competitive mechanisms can unintentionally reinforce existing disparities.

The lack of integration between emerging technologies in category 4 is another significant gap. Despite advances in adaptive and AI systems in the last several years, they are still being developed relatively independently of the educational pedagogical and motivational approaches found in the other categories of evidence. This lack of integration constrains the development of more comprehensive proposals, based on consistent theoretical frameworks. Recent literature reviews of the field indicate this to be true as gamification and AI-supported learning continue to evolve along separate lines of inquiry (Akhmetova et al., 2025; Sergeeva et al., 2024).

Overall, the findings align with previous literature in recognizing that context matters but go a step further by showing that context is not a vague condition. It is structured by the interaction between conceptual approach, disciplinary norms, and user characteristics. From this perspective, the field of gamification in higher education appears not as a unified domain, but as a fragmented and evolving landscape, where different approaches coexist, interact, and sometimes compete.

Epistemological Limitations and Considerations

The classification system has interpretative attributes and could include 3, 5 or more categories to help understand theoretical tensions and how they affect theory and operationally what different predictive tests will develop from them (i.e., extrinsic vs. intrinsic motivation, adaptive vs. fixed design, and average/moderate effects).

Furthermore, this research has been limited on the technical side of gathering data that resulted in the collection of no more than 25 articles from each of Scopus and WoS. However, if additional databases (such as ERIC, IEEE, etc.) were included there may be other categories developed. The lack of grey literature (e.g., thesis, technical reports, etc.) may skew null effects and also create a temporal bias for stating that from 2016-2018 these studies had a longer period of time to accrue citations.

This is a synchronic taxonomy that was created from 2011-2015 and therefore based on the field being very dynamic some current categories may be combined/split/upgraded. Therefore, this taxonomy should be viewed as a current heuristic tool rather than as a permanent classification system.

CONCLUSIONS

The analysis shows that high-impact research on gamification in higher education is organized around distinct epistemological approaches with divergent priorities, reflecting a fragmented and non-linear development of the field.

The 25 articles were identified (1,781 initial → 420 Scopus, 572 WoS → 25 final) with an average annual citation rate of 19.77 (normalized 2.8), peak period 2018-2020. Spain leads production (collaboration with Colombia/Mexico), prevalent areas: computer science (40%) and pedagogy (28%).

Regarding methodological trends, quasi-experimental designs predominate (52%), explanatory scope (48%), reporting of effect sizes (76%), but with absolute dependence on non-probabilistic samples (100%). Scopus shows citation advantage over WoS (21.18 vs. 13.74), although this difference does NOT correlate with methodological quality ($p > 0.05$). Methodological maturity is unequal among categories: category 1 leads (86% reports effects), category 3 emerging methodology (50%). Four conceptual structures that organize high-impact research with differential bibliometric impact are identified:

- Category 1 – Tactical training (144.86 citations): More robust effects ($d = 0.60$-0.85), effective triad objectives + feedback + challenges.
- Category 2 – Structural motivational (172.14 citations): Highest citation average in corpus, but moderate effects ($d = 0.50$-0.79), emphasis on extrinsic motivation through points/badges.
- Category 3 – User profile (87.80 citations): Specialized, 80% includes demographic analysis vs. 28% overall, documents heterogeneity of effects.
- Category 4 – Adaptive/AI (123.83 citations): Emerging post-2020, algorithmic personalization, 50% uses emerging technologies vs. 4% overall.

Persistent Gaps

The analysis reveals three critical gaps that persist in the field. In first place, only the 28% analyzed gender/age, limiting severely the understanding of how the psycho-evolutionary characteristics of the university students moderate the effects. This gap is in contradiction with the principles of inclusive design and raises questions about the equity. In second place, only the 4% incorporated AI, virtual/augmented reality or blockchain, underutilizing the contemporary tools accessible in the period analyzed (2011-2025), what represents a significant gap considering the current context of proliferation of accessible gamification tools, constituting an opportunity of research. In third place, the 68% evaluated immediate effects without follow-up. No study of high impact examined the persistence beyond one semester, raising the critical question about the effect of novelty vs. the sustained impact.

Practical Implications

Only one study of high impact evaluated the persistence beyond one semester, what raises the critical question of the effect of novelty vs. the sustained impact. Collectively, the findings suggest that the designers should prioritize the theoretical integration TAM + SDT before the accumulation of mechanics, while the researchers need to incorporate demographic analyses, explore emerging technologies, and conduct longitudinal studies. Likewise, to consider aligning the choice of the journal with their objectives of scientific dissemination.

Limitations and Future Directions

The selection of Scopus/WoS excludes specialized databases (ERIC, IEEE Xplore). Bias toward cited articles may over-represent positive findings. Required: quantitative meta-analysis, grey literature analysis, examination of element-population-context interactions, and research on causal mechanisms.

Although the gamification shows documented benefits, the effectiveness is contextual and moderated by factors insufficiently understood. Future research must address:

(1) for whom effects operate,

(2) under what conditions,

(3) through what mechanisms, and

(4) for how long effects persist.

Additionally, citation counts throughout this analysis function as proxies for academic visibility within the indexed corpus, not as direct measures of instructional quality or pedagogical effectiveness. High citation frequency may reflect factors such as early publication, journal prestige, or methodological accessibility rather than superior empirical rigor (Bornmann, 2014; Dallaqua et al., 2024).

Author contributions: HAM-B: conceptualization, data curation, formal analysis, writing – original draft; **JMO-T:** supervision, writing – review & editing. Both authors approved the final version of the article.

Funding: The authors received no financial support for the research and/or authorship of this article.

Ethics declaration: This research was conducted in accordance with ethical principles established by the scientific community. As a documentary review study analyzing publicly available academic literature with no human subjects' involvement, ethical committee approval was not required.

AI statement: The authors used Claude's Sonnet 3.5 (Anthropic) as an independent observer's coding instrument for an inter-rater agreement on 40% of the data set to calculate the inter-rater agreement; all intercoder disagreements were resolved by human consensus after reviewing the primary sources to minimize epistemological constraints and possible algorithmic bias (e.g., over-identifying dominant frameworks [e.g., TAM and/or SDT]). Final conceptual categories were classified entirely by humans; the AI tool was used only as a confirmatory mechanism to validate reliability, not a tool to generate results. The authors comprised the writing and critical synthesis of this manuscript, supported by standardized checks of grammar and style, and therefore, all paraphrases were produced manually rather than through automated processes.

Declaration of interest: The authors declared no competing interest.

Data availability: Data generated or analyzed during this study are available from the authors on request.

REFERENCES

Akhmetova, A. I., Seitenova, S. S., Khodjaev, B. K., Jamoldinova, O. R., Yerkebaeva, S. Z., & Kazybayeva, K. U. (2025). Evolution of game-based learning research: A cross-database bibliometric analysis and visualization study (2015-2024). *Contemporary Educational Technology, 17*(3), Article ep585. https://doi.org/10.30935/cedtech/16451

Aria, M., & Cuccurullo, C. (2017). *bibliometrix*: An R-tool for comprehensive science mapping analysis. *Journal of Informetrics, 11*(4), 959-975. https://doi.org/10.1016/j.joi.2017.08.007

Behl, A., Jayawardena, N., Pereira, V., Islam, N., Del Giudice, M., & Choudrie, J. (2022). Gamification and e-learning for young learners: A systematic literature review, bibliometric analysis, and future research agenda. *Technological Forecasting and Social Change, 176*, Article 121445. https://doi.org/10.1016/j.techfore.2021.121445

Bodily, R., Leary, H., & West, R. E. (2019). Research trends in instructional design and technology journals. *British Journal of Educational Technology, 50*(1), 64-79. https://doi.org/10.1111/bjet.12712

Bornmann, L. (2014). How are excellent (highly cited) papers defined in bibliometrics? A quantitative analysis of the literature. *Research Evaluation, 23*(2), 166-173. https://doi.org/10.1093/reseval/rvu002

Brynjolfsson, E., & McAfee, A. (2014). *The second machine age: Work, progress, and prosperity in a time of brilliant technologies*. W. W. Norton & Company.

Chapman, J. R., & Rich, P. J. (2018). Does educational gamification improve students' motivation? If so, which game elements work best? *Journal of Education for Business, 93*(7), 315-322. https://doi.org/10.1080/08832323.2018.1490687

Cheng, S.-C., Hwang, G.-J., & Lai, C.-L. (2022). Critical research advancements of flipped learning: A review of the top 100 highly cited papers. *Interactive Learning Environments, 30*(9), 1751-1767. https://doi.org/10.1080/10494820.2020.1765395

Chugh, R., & Turnbull, D. (2023). Gamification in education: A citation network analysis using CitNetExplorer. *Contemporary Educational Technology, 15*(2), Article ep405. https://doi.org/10.30935/cedtech/12863

Cózar-Gutiérrez, R., & Sáez-López, J. M. (2016). Game-based learning and gamification in initial teacher training in the social sciences: An experiment with MinecraftEdu. *International Journal of Educational Technology in Higher Education, 13*(1), Article 2. https://doi.org/10.1186/s41239-016-0003-4

Dallaqua, M. F., Nunes, B., & Carvalho, M. M. (2024). Serious games research streams for social change: Critical review and framing. *British Journal of Educational Technology, 55*(2), 460-483. https://doi.org/10.1111/bjet.13404

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly, 13*(3), 319-340. https://doi.org/10.2307/249008

Deci, E. L., & Ryan, R. M. (2000). The "what" and "why" of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry, 11*(4), 227-268. https://doi.org/10.1207/S15327965PLI1104_01

Deterding, S., Dixon, D., Khaled, R., & Nacke, L. (2011). From game design elements to gamefulness: Defining "gamification". In *Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments*. https://doi.org/10.1145/2181037.2181040

Dijk, S. H. B. van, Brusse-Keizer, M. G. J., Bucsán, C. C., Palen, J. van der, Doggen, C. J. M., & Lenferink, A. (2023). Artificial intelligence in systematic reviews: Promising when appropriately used. *BMJ Open, 13*(7), Article e072254. https://doi.org/10.1136/bmjopen-2023-072254

Eltahir, Mohd. E., Alsalhi, N. R., Al-Qatawneh, S., AlQudah, H. A., & Jaradat, M. (2021). The impact of game-based learning (GBL) on students' motivation, engagement and academic performance on an Arabic language grammar course in higher education. *Education and Information Technologies, 26*(3), 3251-3278. https://doi.org/10.1007/s10639-020-10396-w

Fabiano, N., Gupta, A., Bhambra, N., Luu, B., Wong, S., Maaz, M., Fiedorowicz, J. G., Smith, A. L., & Solmi, M. (2024). How to optimize the systematic review process using AI tools. *JCPP Advances, 4*(2), Article e12234. https://doi.org/10.1002/jcv2.12234

Fernández-Velásquez, J. D. R., López-Regalado, O., & Fernández-Hurtado, G. A. (2025). Educational dualism in action: Systematic review of gamification and flipped classrooms' effects on young learners. *Contemporary Educational Technology, 17*(1), Article ep557. https://doi.org/10.30935/cedtech/15749

Ferriz-Valero, A., Østerlie, O., García Martínez, S., & García-Jaén, M. (2020). Gamification in physical education: Evaluation of impact on motivation and academic performance within higher education. *International Journal of Environmental Research and Public Health, 17*(12), Article 4465. https://doi.org/10.3390/ijerph17124465

Fotaris, P., Mastoras, T., Leinfellner, R., & Rosunally, Y. (2016). Climbing up the leaderboard: An empirical study of applying gamification techniques to a computer programming class. *The Electronic Journal of e-Learning, 14*(2), 95-110.

Ge, L., Agrawal, R., Singer, M., Kannapiran, P., De Castro Molina, J. A., Teow, K. L., Yap, C. W., & Abisheganaden, J. A. (2024). Leveraging artificial intelligence to enhance systematic reviews in health research: Advanced tools and challenges. *Systematic Reviews, 13*(1), Article 269. https://doi.org/10.1186/s13643-024-02682-2

Hew, K. F., Huang, B., Chu, K. W. S., & Chiu, D. K. W. (2016). Engaging Asian students through game mechanics: Findings from two experiment studies. *Computers & Education, 92-93*, 221-236. https://doi.org/10.1016/j.compedu.2015.10.010

Hou, H.-T., & Li, M.-C. (2014). Evaluating multiple aspects of a digital educational problem-solving-based adventure game. *Computers in Human Behavior, 30*, 29-38. https://doi.org/10.1016/j.chb.2013.07.052

Huang, B., & Hew, K. F. (2018). Implementing a theory-driven gamification model in higher education flipped courses: Effects on out-of-class activity completion and quality of artifacts. *Computers & Education, 125*, 254-272. https://doi.org/10.1016/j.compedu.2018.06.018

Huang, B., Hew, K. F., & Lo, C. K. (2019). Investigating the effects of gamification-enhanced flipped learning on undergraduate students' behavioral and cognitive engagement. *Interactive Learning Environments, 27*(8), 1106-1126. https://doi.org/10.1080/10494820.2018.1495653

Hwang, G.-J., Hung, C.-M., & Chen, N.-S. (2014). Improving learning achievements, motivations and problem-solving skills through a peer assessment-based game development approach. *Educational Technology Research and Development, 62*(2), 129-145. https://doi.org/10.1007/s11423-013-9320-7

Jensen, M. M., Danielsen, M. B., Riis, J., Kristjansen, K. A., Andersen, S., Okubo, Y., & Jørgensen, M. G. (2025). ChatGPT-4o can serve as the second rater for data extraction in systematic reviews. *PLoS ONE, 20*(1), Article e0313401. https://doi.org/10.1371/journal.pone.0313401

Khraisha, Q., Put, S., Kappenberg, J., Warraitch, A., & Hadfield, K. (2024). Can large language models replace humans in systematic reviews? Evaluating GPT-4's efficacy in screening and extracting data from peer-reviewed and grey literature in multiple languages. *Research Synthesis Methods, 15*(4), 616-626. https://doi.org/10.1002/jrsm.1715

Kyewski, E., & Krämer, N. C. (2018). To gamify or not to gamify? An experimental field study of the influence of badges on motivation, activity, and performance in an online learning course. *Computers & Education, 118*, 25-37. https://doi.org/10.1016/j.compedu.2017.11.006

Lieberum, J.-L., Töws, M., Metzendorf, M.-I., Heilmeyer, F., Siemens, W., Haverkamp, C., Böhringer, D., Meerpohl, J. J., & Eisele-Metzger, A. (2024). *Large language models for conducting systematic reviews: On the rise, but not yet ready for use – A scoping review*. medRxiv. https://doi.org/10.1101/2024.12.19.24319326

Lin, D. T. A., & Kaur, M. (2018). Kahoot! It: Gamification in higher education. *Pertanika Journal of Social Science and Humanities, 26*(1), 565-582.

Marti-Parreno, J., Mendez-Ibanez, E., & Alonso-Arroyo, A. (2016). The use of gamification in education: A bibliometric and text mining analysis. *Journal of Computer Assisted Learning, 32*(6), 663-676. https://doi.org/10.1111/jcal.12161

Murillo-Zamorano, L. R., López Sánchez, J. Á., Godoy-Caballero, A. L., & Bueno Muñoz, C. (2021). Gamification and active learning in higher education: Is it possible to match digital society, academia and students' interests? *International Journal of Educational Technology in Higher Education, 18*(1), Article 15. https://doi.org/10.1186/s41239-021-00249-y

Nousiainen, T., Kangas, M., Rikala, J., & Vesisenaho, M. (2018). Teacher competencies in game-based pedagogy. *Teaching and Teacher Education, 74*, 85-97. https://doi.org/10.1016/j.tate.2018.04.012

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., ... Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ, 372*, Article n71. https://doi.org/10.1136/bmj.n71

Pardal-Refoyo, J. L., Pardal-Peláez, B., Pardal-Refoyo, J. L., & Pardal-Peláez, B. (2020). Anotaciones para estructurar una revisión sistemática [Notes for structuring a systematic review]. *Revista ORL, 11*(2), 155-160. https://doi.org/10.14201/orl.22882

Rodríguez, D. V., & Mezquita, J. M. M. (2016). Diseño de juegos serios para reforzar conocimientos: Una experiencia educativa en secundaria [Designing serious games to reinforce knowledge: An educational experience in secondary education]. *Profesorado, Revista de Currículum y Formación del Profesorado, 20*(2), 238-254. https://doi.org/10.30827/profesorado.v20i2.10417

Sailer, M., & Sailer, M. (2021). Gamification of in-class activities in flipped classroom lectures. *British Journal of Educational Technology, 52*(1), 75-90. https://doi.org/10.1111/bjet.12948

Sanchez, D. R., Langer, M., & Kaur, R. (2020). Gamification in the classroom: Examining the impact of gamified quizzes on student learning. *Computers & Education, 144*, Article 103666. https://doi.org/10.1016/j.compedu.2019.103666

Schell, J. (2008). *The art of game design: A book of lenses*. A K Peters/CRC Press.

Schmidt, L., Hair, K., Graziosi, S., Campbell, F., Kapp, C., Khanteymoori, A., Craig, D., Engelbert, M., & Thomas, J. (2025). *Exploring the use of a large language model for data extraction in systematic reviews: A rapid feasibility study*. arXiv. https://doi.org/10.48550/arXiv.2405.14445

Schöbel, S., Saqr, M., & Janson, A. (2021). Two decades of game concepts in digital learning environments – A bibliometric study and research agenda. *Computers & Education, 173*, Article 104296. https://doi.org/10.1016/j.compedu.2021.104296

Sergeeva, O. V., Zvereva, E. V., Kosheleva, Y. P., Zheltukhina, M. R., Prokopyev, A. I., & Dobrokhotov, D. A. (2024). A comprehensive bibliometric review of gamified learning in higher education. *Contemporary Educational Technology, 16*(3), Article ep514. https://doi.org/10.30935/cedtech/14708

Subhash, S., & Cudney, E. A. (2018). Gamified learning in higher education: A systematic review of the literature. *Computers in Human Behavior, 87*, 192-206. https://doi.org/10.1016/j.chb.2018.05.028

Sussman, S., & Arnett, J. J. (2014). Emerging adulthood: Developmental period facilitative of the addictions. *Evaluation & the Health Professions, 37*(2), 147-155. https://doi.org/10.1177/0163278714521812

Toledo, P. A., González, C. S. G., Cruz, V. M., Rodríguez, M. H., Aciar, S. V., & Fardoum, H. (2017). Sistema de recomendación para juegos serios [Recommendation system for serious games]. In *Proceedings of the 5th International Congress on Video Games and Education*.

Topalli, D., & Cagiltay, N. E. (2018). Improving programming skills in engineering education through problem-based game projects with Scratch. *Computers & Education, 120*, 64-74. https://doi.org/10.1016/j.compedu.2018.01.011

Troussas, C., Krouska, A., & Sgouropoulou, C. (2020). Collaboration and fuzzy-modeled personalization for mobile game-based learning in higher education. *Computers & Education, 144*, Article 103698. https://doi.org/10.1016/j.compedu.2019.103698

Tsai, M.-J., Huang, L.-J., Hou, H.-T., Hsu, C.-Y., & Chiou, G.-L. (2016). Visual behavior, flow and achievement in game-based learning. *Computers & Education, 98*, 115-129. https://doi.org/10.1016/j.compedu.2016.03.011

Tsay, C. H.-H., Kofinas, A., & Luo, J. (2018). Enhancing student learning experience with technology-mediated gamification: An empirical study. *Computers & Education, 121*, 1-17. https://doi.org/10.1016/j.compedu.2018.01.009

Universidad VIU. (2024). Tipos de tecnologías existentes [Types of existing technologies]. *VIU España*. https://www.universidadviu.com/es/actualidad/nuestros-expertos/los-tipos-de-tecnologia-mas-representativos-que-debes-conocer

Uthman, O. A., Okwundu, C. I., Wiysonge, C. S., Young, T., & Clarke, A. (2013). Citation classics in systematic reviews and meta-analyses: Who wrote the top 100 most cited articles? *PLoS ONE, 8*(10), Article e78517. https://doi.org/10.1371/journal.pone.0078517

Vanduhe, V. Z., Nat, M., & Hasan, H. F. (2020). Continuance intentions to use gamification for training in higher education: Integrating the technology acceptance model (TAM), social motivation, and task technology fit (TTF). *IEEE Access, 8*, 21473-21484. https://doi.org/10.1109/ACCESS.2020.2966179

Vergara Rodríguez, D., & Mezquita, J. M. M. (2016). Diseño de juegos serios para reforzar conocimientos: Una experiencia educativa en secundaria [Designing serious games to reinforce knowledge: An educational experience in secondary school]. *Profesorado, Revista de Currículum y Formación del Profesorado, 20*(2), 238-254.

Villagrasa, S., Fonseca, D., Redondo, E., & Duran, J. (2014). Teaching case of gamification and visual technologies for education. *Journal of Cases on Information Technology, 16*(4), 38-57. https://doi.org/10.4018/jcit.2014100104

Wang, Y., Wang, Z., Liu, G., Wang, Z., Wang, Q., Yan, Y., Wang, J., Zhu, Y., Gao, W., Kan, X., Zhang, Z., Jia, L., & Pang, X. (2022). Application of serious games in health care: Scoping review and bibliometric analysis. *Frontiers in Public Health, 10*. https://doi.org/10.3389/fpubh.2022.896974

Wiggins, B. E. (2016). An overview and study on the use of games, simulations, and gamification in higher education. *International Journal of Game-Based Learning, 6*(1), 18-29. https://doi.org/10.4018/IJGBL.2016010102

Yildirim, I. (2017). The effects of gamification-based teaching practices on student achievement and students' attitudes toward lessons. *The Internet and Higher Education, 33*, 86-92. https://doi.org/10.1016/j.iheduc.2017.02.002

Zainuddin, Z. (2018). Students' learning performance and perceived motivation in gamified flipped-class instruction. *Computers & Education, 126*, 75-88. https://doi.org/10.1016/j.compedu.2018.07.003