



How are we evaluating Gamification? A Systematic Mapping of Methods and Psychometric Instruments for Evaluation of Gamification in Educational Contexts

Vinicius Lopes, ICMC/USP, lopesvinicius@usp.br, 0000-0002-2534-4665

Seiji Isotani, ICMC/USP, sisotani@icmc.usp.br, 0000-0003-1574-0784

Diego Dermeval de Cunha Matos, NEES/UFAL, diego.matos@famed.ufal.br, 0000-0002-8415-6955

Luiz Rodrigues, NEES/UFAL, luiz.rodrigues@nees.ufal.br, 0000-0003-0343-3701

Abstract: Gamification has emerged as a widely adopted strategy in educational contexts, aiming to increase student motivation, engagement, and learning outcomes. However, evaluations of its effectiveness remain inconsistent, partly due to the lack of standardized methods and instruments. This study presents a systematic mapping of the literature to understand how gamification has been evaluated in educational settings. We identified common methodological approaches, evaluation criteria, and psychometric instruments used across recent studies. Findings reveal a strong reliance on custom-made instruments, which hinders comparability and replication. At the same time, we identified the adoption of validated instruments capable of assessing critical aspects like usability, motivation, and gameful experience. By organizing these instruments into conceptual domains, this study offers a clearer perspective on how gamification effectiveness can be assessed more rigorously. The results support the need for more structured evaluation practices that enhance the reliability and impact of gamification research in education.

Keywords: gamification; evaluation; education; psychometric instruments.

Como estamos avaliando a Gamificação? Um Mapeamento Sistemática de Métodos e Instrumentos Psicométricos para Avaliação da Gamificação em Ambientes Educacionais

Resumo: A gamificação tem se consolidado como uma estratégia amplamente adotada em contextos educacionais, com o objetivo de aumentar a motivação, o engajamento e os resultados de aprendizagem dos estudantes. No entanto, as avaliações sobre sua efetividade permanecem inconsistentes, em parte devido à ausência de métodos e instrumentos padronizados. Este estudo apresenta um mapeamento sistemático da literatura para compreender como a gamificação tem sido avaliada em ambientes educacionais. Foram identificadas as abordagens metodológicas, critérios de avaliação e instrumentos psicométricos mais comumente utilizados em estudos recentes. Os resultados revelam uma forte dependência de instrumentos desenvolvidos sob medida, o que dificulta a comparabilidade e a replicação entre estudos. Ao mesmo tempo, também foram identificados instrumentos psicométricos validados, capazes de avaliar aspectos críticos como usabilidade, motivação e experiência gamificada. Ao organizar esses instrumentos em domínios conceituais, o estudo oferece uma visão mais clara de como a efetividade da gamificação pode ser avaliada com maior rigor. Os resultados reforçam a necessidade de práticas avaliativas mais estruturadas, que aumentem a confiabilidade e o impacto das pesquisas em gamificação na educação.

Palavras-chave: gamificação; avaliação; educação; instrumentos psicométricos.

1. Introduction

Gamification, using game elements in non-game contexts, has been gaining popularity as a strategy to enhance engagement and motivation in various fields, such as marketing, healthcare, and work and educational environments [Deterding et al. 2011]. The core concept of gamification is to incorporate the elements of games into real-world contexts, to encourage specific behaviors or outcomes [Werbach, Hunter and Dixon 2012, Dicheva et al. 2015]. These elements range from levels, points, and badges to competition, cooperation systems, and even narrative and storytelling elements [Toda et al., 2019].



Several empirical studies have shown the positive effects of gamification on learners in the education and training contexts [Koivisto and Hamari 2019], while some studies also present negative effects, like indifference, loss of performance, undesired behavior, and declining effects over time [Toda, Vale and Isotani 2017]. However, no consensus exists on what factors provide the most effective gamification design and outcomes in educational contexts [Sailer and Homner, 2020]. Furthermore, many studies on gamification in education tend to evaluate this effectiveness through the development of custom-designed instruments, like surveys, interviews or by analyzing student data, making it difficult to compare results across different studies or to establish widely accepted standards for gamification in education [Klock et al. 2018, Reinheimer and Medina, 2020].

This paper presents a systematic literature mapping of studies that apply gamification strategies to an educational context and report their evaluation methodology. We systematically map and categorize the most prominent methods, criteria, and psychometric instruments used in these evaluations. With this research, we intend to obtain an overview of the state of the art on the evaluation of gamification in this context and contribute a standardized solution for these evaluations.

2. Contextualization

Studies show that traditional learning methods are less effective in engaging and motivating students with educational activities [Lee and Hammer 2011, Lo and Hew 2020]. This proves to be a challenge since motivation is perceived as a prominent factor in learning [Hanus and Fox 2015, Vansteenkiste et al. 2009]. Several alternative approaches have emerged to address this issue, such as the use of digital and serious games, as well as gamification, seeking a potential solution to make learning more engaging and enjoyable for students [Pessini et al. 2014, Dicheva et al. 2015].

By definition, gamification is the use of game-based elements in external contexts, such as workspaces, educational environments, health and lifestyle apps, or marketing campaigns [Deterding et al. 2011]. The primary concept behind gamification is the application of dynamics, mechanics, and elements found in games, like point-scoring systems, levels, challenges, and rewards, to non-game contexts. The main goal of this approach is to make these contexts more engaging and motivating by applying game elements that tap into people's intrinsic desires for competition, achievement, and recognition [Werbach, Hunter and Dixon 2012].

Gamification strategies have been widely used in educational contexts. The use of grades (scores and points) and achievements for good or average grades (rewards) can be recognized as gamification elements [Smith-Robbins 2011]. Kapp (2012) stated that gamification can impact student motivation and engagement, directly affecting their performance [Kapp 2012].

Several studies have been developing gamification strategies for education, while trying to understand their effects on student motivation, engagement, and performance [Dichev and Dicheva 2017]. However, while gamification is most prominent in the educational field, its effects and outcomes have the most mixed results [Borges et al. 2014, Darejeh and Salim 2016]. Positive and negative effects have been observed across several studies, with some attesting improvement in student motivation and engagement, while others stating the opposite [Seaborn and Fels 2015]. Among the adverse effects of gamification are indifference, loss of performance, undesired behavior and declining effects over time [Toda, Vale and Isotani 2017]. However, there is a lack of understanding of which factors moderate such effectiveness [Sailer and Homner, 2020], as well as the lack of standardization of an evaluation process of the gamification design and its outcomes.



In the literature, there are several empirically validated instruments, such as questionnaires, scales, and models that measure and evaluate psychometric factors, such as learnability, enjoyment, and immersion [Shaneyfelt et al. 2016]. Seeing as those instruments could prove useful for a potential standardization of these evaluation processes, we performed a systematic literature mapping (SLM) of the most recent studies that applied gamification to an educational context to understand how these evaluations are being performed, which criteria are being evaluated, and which instruments are used more frequently.

3. Related work

Systematic literature reviews on the subject exist in the literature, like Ogawa, Klock, and Gasparini's (2016) study. It analyzes how gamification was evaluated in studies ranging from 2010 to 2015. It focused on studies that evaluated gamification through comparative studies between gamification and non-gamification learning environments. A more recent study, from Klock et al. (2020), presents another systematic literature review in which the evaluation process, methods, criteria, and effects are analyzed and compared between gamification and no-gamification contexts. Reinheimer and Medina (2020) also performed a systematic literature mapping on the subject, ranging from 2016 to 2019. In this study, the authors categorize several evaluation points for gamification practices, like research methods, resources utilized, and criteria evaluated. Finally, Monteiro et al. (2021) reports another systematic mapping of gamification evaluation, presenting trends, comparisons, and gaps on the subject. Different from other studies, this one focuses solely on the evaluation of gamification performed in software engineering educational scenarios.

Since these studies range from 2010 to 2020, we decided it would be suitable to perform another analysis of this scope, ranging from the last 5 years (2020 to 2024). Also, since evaluation of gamification has been presenting mixed results in the literature [Sailer and Homner, 2020], it would be relevant to understand how different methods can affect different outcomes. We also differ from these other studies by focusing on psychometric instruments and their adaptability to various educational contexts, as we intend to contribute with a standardized solution for these evaluations in the future.

4. Research methodology

For this research, we decided to perform a Systematic Literature Mapping (SLM), designed by Kitchenham (2004), a structured method for collecting, classifying, and analyzing existing research related to a specific area of interest. It aims to provide an overview of the quantity and type of research available, highlighting gaps and trends to guide future investigations [Kitchenham et al. 2010]. We used this method to identify and categorize different strategies and approaches to evaluating gamification in educational environments, such as research methods, psychometric instruments, and evaluation criteria adopted. Therefore, we raised the following research questions:

- **RQ1:** What methodological approaches (e.g., quantitative, qualitative, mixed methods) have been adopted to evaluate gamification in educational environments?
- **RQ2:** What criteria (e.g., engagement, motivation, performance, usability) have been considered in evaluating the gamification of these environments?
- **RQ3:** What psychometric instruments (e.g., questionnaires, scales) have been adopted to evaluate gamification in educational environments?
 - **RQ3a:** Which psychometric instruments adopted for the evaluation in these studies were validated?



- **RQ3b:** Can these instruments be used for evaluation in other educational gamification environments?

To answer these research questions, we developed a search string composed of the terms ‘gamification’, ‘education’, and ‘evaluation’, and its synonyms. The final string used was the following: (gamification OR gameful) AND (evaluation OR experiment OR validation OR assessment OR psychometric) AND (education OR teaching OR learning). The research bases selected were ACM, IEEE, ScienceDirect, and Taylor & Francis, the most common bases for studies involving gamification and educational technologies. We sought studies that designed gamification strategies in educational contexts and reported their evaluation processes. Therefore, we used the following inclusion criteria (IC) and exclusion criteria (EC):

- **IC1:** *The study presents the use of gamification in an educational environment or context.*
- **IC2:** *The study reports the evaluation process of the adopted gamification approach.*
- **EC1:** *Does not address gamification concepts or misrepresents them as games.*
- **EC2:** *Does not apply gamification to an educational context.*
- **EC3:** *Does not describe any evaluation processes of the adopted gamification approach.*
- **EC4:** *The study is not available in the English language.*
- **EC5:** *Is not a full-text paper or presents research still in development.*
- **EC6:** *Is not a primary peer-reviewed study.*
- **EC7:** *Is not a full-access paper or not available through institutional access.*
- **EC8:** *Not published between 2020 and 2024.*
- **EC9:** *The study is duplicated in a different database.*

For this SLM, we used Parsifal, an online platform designed to assist researchers in conducting systematic literature reviews (SLRs). This platform allowed us to import metadata from all databases into a more user-friendly interface and more efficiently apply the inclusion and exclusion criteria. Table 1 reports all the stages that were conducted for this SLM.

Table 1 - Studies remaining at every stage of the systematic literature mapping

Research databases	Studies found	Duplicates removed	Metadata read	Papers available in full access	Full text read and data extraction
ACM	153	115	32	21	20
IEEE	561	449	66	66	63
Science Direct	222	157	26	26	26
Taylor & Francis	74	61	23	23	23
Total	1.010	782	147	136	132

By applying the search string to every database and using their filters for English-only (EC4) and full-text papers (EC5) from the years between 2020 and 2024 (EC8), we found 1.010 studies. The papers from these studies had their metadata downloaded and imported into Parsifal. Since this platform has the option to scan and remove duplicate papers between databases, we were able to arrive at 782 papers (EC9).



The next step of the paper assessment involved peer reviewing their metadata by reading the papers' title and abstract to determine if it aligns with the inclusion criteria (IC1 and IC2). In this stage, we excluded studies that did not apply gamification concepts to educational contexts (EC2) or misrepresented the concept of gamification, like studies that used educational games, serious games, escape rooms, virtual reality, or other similar concepts (EC1). We also excluded papers that presented secondary or tertiary studies, like systematic reviews and meta-analyses (EC6). By the end of this stage, we selected 147 papers for the extraction phase.

We could only access the content of 136 of these papers, either because they were full-access papers or through institutional login in the research databases (EC7). We performed a full-text read of these 136 papers to collect the data to answer our research questions. In the process, we identified and excluded 4 papers that did not meet EC3, by not reporting the evaluation process, or because they were only theoretical papers with no experiments involved. In the end, we extracted data from 132 papers, and the results were analyzed. The following section reports and discusses these findings.

5. Results and discussion

This section reports the data collected from the 132 papers that remained after applying the exclusion criteria and that met all the inclusion criteria for this study. A review of these papers revealed that 114 consisted of practical gamification experiments, while 18 studies not only conducted experiments but also validated a framework or model in the process, like [Polito and Temperini 2021] and [Bassanelli, Bucchiarone, and Gini 2024].

Regarding educational contexts, Figure 1 shows that STEM (Science, Technology, Engineering, and Mathematics education) is the context in which gamification experiments have been most prominent (n=75). Followed by language learning (n=10), health (n=9), business education (n=7), and biological and environmental studies (n=7). Other contexts include cooking, criminal justice, design, driving education, library services, memory, organizational communication, culinary, traffic knowledge, and information identification skills, each appearing only once (n=10).

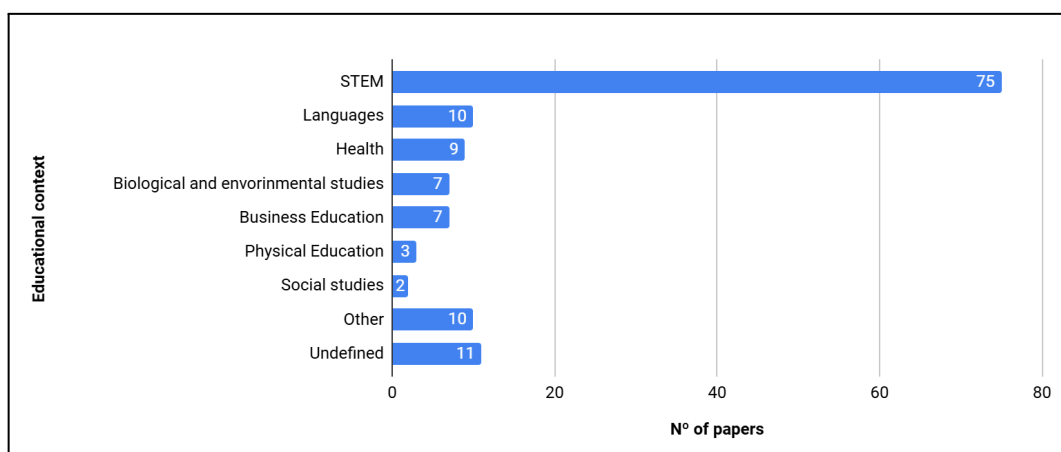


Figure 1 - Studies by educational context.

These results corroborate findings of Klock et al. (2020) that gamification approaches, while most predominant in STEM education, have been slowly expanding into other educational contexts, with other areas experimenting with their motivational potential to engage students with their experiments.

To answer **RQ1: What methodological approaches have been adopted to evaluate gamification in educational environments?**, we collected data referring to the type of research

that was conducted and which methodological approaches were used. In this data we perceived that 72 papers employed quantitative approaches, such as student log data, test scores, or survey quantitative data. Only 4 papers adopted a qualitative approach by using interviews, student observation, or asking open-ended questions in surveys/questionnaires. Finally, 56 papers employed a mixed-methods strategy, integrating both quantitative and qualitative methodologies.

Most studies incorporated multiple methodological approaches within their research (Figure 2). Several studies evaluated gamification through comparative studies ($n=45$), employing a control group and one or multiple experimental groups. Data collected from these groups were subsequently contrasted to derive findings, like in [Sprint and Fox 2020] and [Schatten and Schatten 2020]. A pre/post-test approach was also employed in some of the papers ($n=35$), in which a student evaluation was performed before and after applying a gamification strategy or technology. This method is most commonly used for evaluating student performance and learning outcomes, as seen in [Putz, Hofbauer and Treiblmaier 2020].

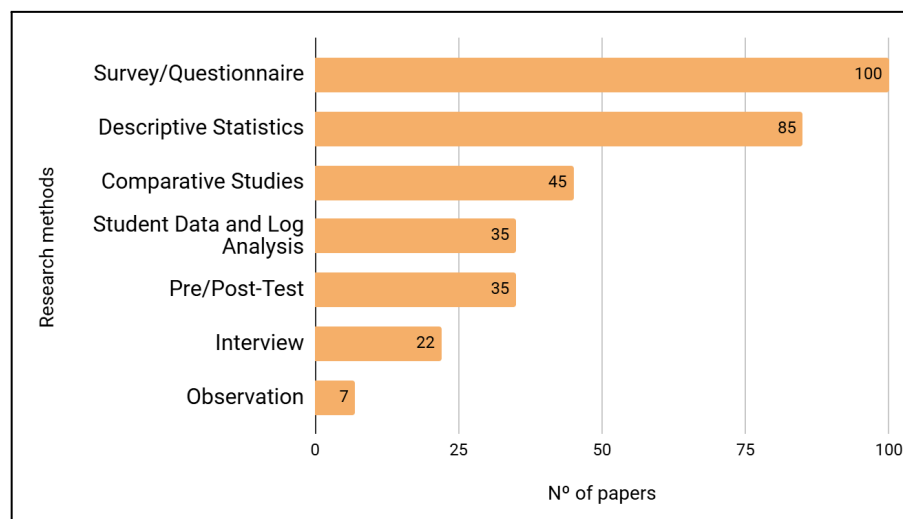


Figure 2 - Studies by research method.

Regarding data analysis, 85 studies reported using descriptive statistics, employing statistical methods like t-test, ANOVA, regression analysis, and utilizing tools such as IBM SPSS or comparable software to calculate results. For data collection, 35 studies used student data, like scores, attendance records, or interaction logs with learning management systems (LMS). The most common data collection method is using surveys or questionnaires ($n=100$) to assess an evaluation criteria, like motivation or engagement. Other data collection methods include student and teacher interviews ($n=22$) and class observation ($n=7$).

For **RQ2: What criteria have been considered in evaluating the gamification of these environments?**, we collected data referring to the evaluation criteria considered in these studies, as seen in Figure 3. This was done by analyzing the research method, objectives, or results, in which the evaluation criteria were clearly stated or reported. The most common criteria evaluated were *motivation* ($n=54$), *engagement* ($n=38$), and *performance* ($n=26$), which are the most associated terms with gamification and other educational strategies, as corroborated by Monteiro et al. (2021).

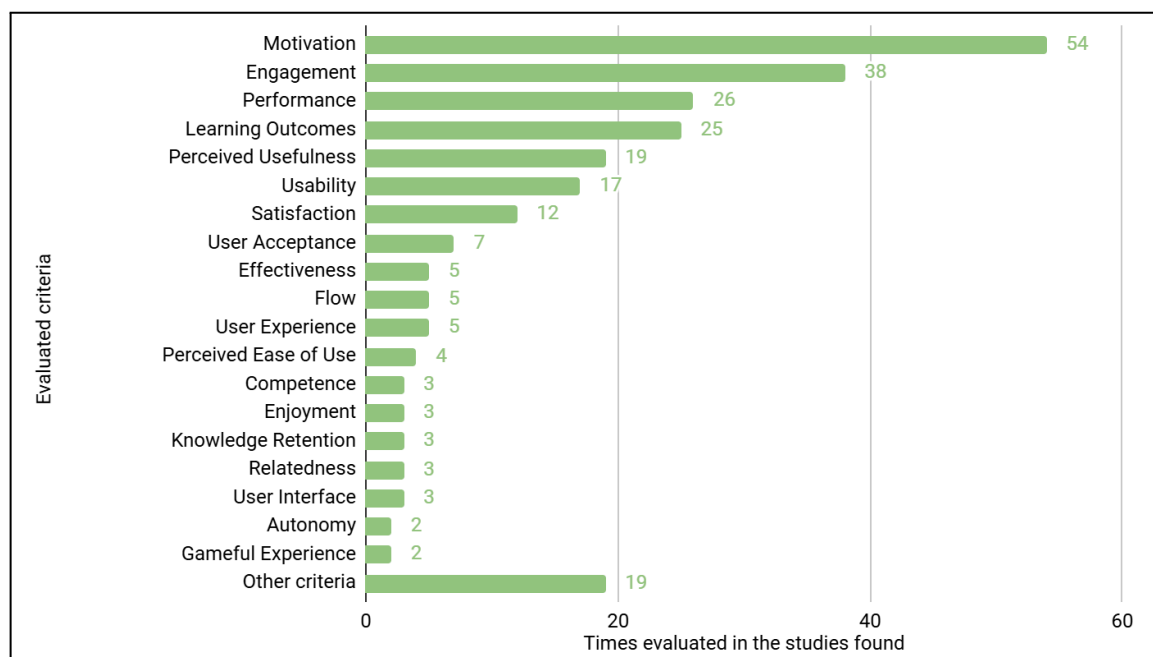


Figure 3 - Evaluated criteria for gamification analyzed in the studies found

Learning outcomes, the term used for assessing learning objectives achieved, is also a criteria highly considered during evaluations ($n=25$). Similar to *performance*, *knowledge retention* ($n=2$), and *knowledge acquisition* ($n=1$), it also considers learning effectiveness as an important criterion for measurement of the impact of gamification during evaluation. They are usually assessed through tests and knowledge assessments, as seen in [Sadeghi 2024] and [Kumar et al. 2023].

Criteria like *motivation*, *engagement*, *satisfaction* ($n=12$), and *relatedness* ($n=3$) tend to evaluate the learners' feelings towards the learning activities. In contrast, *gameful experience* ($n=2$), *flow* ($n=5$), *enjoyment* ($n=3$), and *fun* ($n=1$) are indicators of learners' perception towards the use of gamification. Both are usually evaluated through surveys/questionnaires that assess these criteria, like [Garaccione, Coppola, and Ardito 2024] and [Rodrigues et al. 2021].

Usability ($n=17$) and similar concepts, like *perceived usefulness* ($n=19$), *perceived ease of use* ($n=4$), *user acceptance* ($n=7$), and *user experience* ($n=5$), are commonly associated with the design of systems in which gamification was implemented in educational contexts. These systems also tend to be assessed through surveys and questionnaires that focus on user evaluation, like [Agustin et al. 2024].

Finally, for **RQ3: What psychometric instruments have been adopted to evaluate gamification in educational environments?**, we collected data regarding psychometric evaluation instruments used in these studies. Out of the 132 studies analyzed, 24 used methodologies that did not involve the use of psychometric instruments like comparative studies and student log and data analysis.

On the other hand, 56 studies reported creating and using their own instruments to assess gamification metrics, while 54 others reported using one or more validated instruments in the literature. Moving to the research subquestion **RQ3a: Which psychometric instruments adopted for the evaluation in these studies were validated?**, we collected all validated instruments found during data extraction in this research (Table 2). We categorized them across four domains, according to their evaluation goals and evaluation criteria analyzed across studies: system usability, learner motivation, gameful experience and educational impact.



Table 2 - Validated psychometric instruments for evaluation categorized across four different domains

System Usability	Learner Motivation
• Technology Acceptance Model (TAM) • System Usability Scale (SUS) • User Experience Questionnaire (UEQ) • Hedonic Motivation System Adoption Model (HMSAM) • Net Promoter Score • Perceived Usefulness and Ease of Use and User Acceptance of Information Technologies • Single Ease Question (SEQ) • Technology Acceptance Scale • User Acceptance Test • VADER sentiment analyzer	• Intrinsic Motivation Inventory (IMI) • Attention Relevance Confidence and Satisfaction Scale (ARCS) • Situational Motivation Scale (SIMS) • Unified Motive Scales (UMS) • Agentic Engagement Scale (AES) • Basic Psychological Needs in Exercise Scale (BPNES) • Basic Psychological Needs Satisfaction Scale • Big Five Inventory (BFI) • Creative Tendency Scale • Elliot's Revised Achievement Goals Questionnaire (AGQ-R) • Expectancy-Value-Cost Scale (EVC) • Motivation Toward Learning • Need for Relatedness at College Questionnaire (NRC-Q) • Perceived Locus of Causality in Physical Education Scale (PLOC) • Sport Motivation Scale (SMS-II-PE) • Stanford House Staff Wellness Survey
Educational Impact	Gameful Experience
• Cognitive Load Scale • NASA-TLX (NASA Task Load Index) • Metacognitive Awareness Inventory (MAI) • Reduced Instructional Materials Motivation Survey (RIMM) • Student Evaluation of Educational Quality (SEEQ) • Test Satisfaction Questionnaire	• Gameful Experience Scale (GAMEX) • Gameful Experience Questionnaire (GAMEFULQUEST) • EGameFlow • Flow Scale for Games (FSG) • Flow Short Scale (FKS) • Flow State Scale • FunQ • Game User Experience Satisfaction Scale (GUESS) • Gamification Perception Questionnaire • Short Flow Experience Scale (FSS) • User Engagement Scale (UEQ)

The most commonly used instruments are the *Technology Acceptance Model (TAM)*, which measures users' perceptions of usefulness and ease of use of a technology (n=9) [Masrom 2007], and the *System Usability Scale (SUS)*, that assesses the overall usability of a system through a 10-item questionnaire (n=8) [Brooke 1996]. The next most commonly used instrument was the *Intrinsic Motivation Inventory (IMI)*, which assesses participants' subjective experiences related to a target activity, particularly focusing on interest/enjoyment, perceived competence, and effort (n=4) [Markland and Hardy 1997].

This was followed by the *Gameful Experience Scale (GAMEX)*, developed by Eppmann, Bekk, and Klein (2018), which measures the psychological outcomes associated with gameful system use, such as playfulness, engagement, and perceived value (n=3). Another frequently used instrument was the *User Experience Questionnaire (UEQ)*, which captures users' perceptions of a system's usability and user experience across pragmatic and hedonic quality dimensions (n=3) [Laugwitz, Held and Schrepp 2008].



Regarding **RQ3b**: *Can these instruments be used for evaluation in other educational gamification environments?*, several instruments are designed for specific educational contexts, like the *Sport Motivation Scale (SMS-II-PE)* [Pelletier 2013], the *Basic Psychological Needs in Exercise Scale (BPNES)* [Vlachopoulos and Michailidou 2006], and the *Perceived Locus of Causality in Physical Education Scale (PLOC)* [Guedes and Yamaji 2024]. However, most instruments can be used in other educational contexts to evaluate different criteria with little to no adaptation. These findings reinforce the potential for standardizing gamification evaluation practices through the development of a model that supports consistency and comparability in gamification research.

Limitations for this study include the restricted number of databases consulted during the SLM, which may have limited the scope of relevant literature retrieved. Future research could address this by expanding the search to additional academic databases. Another limitation lies in the exclusive consideration of articles published in English, potentially overlooking valuable contributions in other languages, such as Spanish and Brazilian Portuguese. Furthermore, this study primarily focused on identifying evaluation instruments for gamification, without exploring other methodological approaches in depth, such as experimental testing procedures or statistical methodologies. These limitations represent opportunities for further investigation of gamification evaluation practices.

6. Conclusion

Gamification continues to grow as an educational strategy to enhance motivation, engagement, and learning outcomes. However, as this study reveals, the evaluation of such interventions still lacks uniformity and rigor. By systematically mapping 132 peer-reviewed studies from the most recent years, we identified not only the diversity of methods and criteria employed but also shed light on the disproportionate use of custom-made evaluation tools that limit replicability and comparative analysis across contexts.

A key finding is the substantial reliance on surveys and questionnaires as primary data collection methods. However, while 54 studies utilized validated psychometric instruments from the existing literature, other 56 studies reported developing their own instruments to assess evaluation metrics, indicating a significant gap in methodological consistency. Among these validated instruments, the System Usability Scale (SUS), Technology Acceptance Model (TAM), and Intrinsic Motivation Inventory (IMI) emerged as the most adopted tools, each aligned with widely recognized constructs like usability, acceptance, and motivation, respectively. The mapping also revealed that, although many of these instruments were initially designed for specific domains, their criteria are transferable and adaptable to broader educational contexts.

By organizing validated instruments into four domains: system usability, learner motivation, gameful experience, and educational impact, this work contributes as a valuable reference for researchers and practitioners seeking more standardized and psychometrically validated approaches. These instruments not only provide more reliable insights into gamification's effects but also allow for cross-study comparisons and meta-analyses that can deepen our understanding of when, how, and for whom gamification works.

Future work will employ these insights to develop a standardized evaluation model that aligns instruments with gamification design goals. This model will act as a toolkit for researchers and educators, reducing the reliance on ad-hoc instruments and improving the quality and comparability of gamification studies. Through this standardization effort, we hope to minimize the contradictory findings in the literature and establish clearer, evidence-based guidelines for gamification in education.



References

- AGUSTIN, F. E. M.; HIDAYAT, T.; AZHARI, M.; WAHYUNI, A.; HIDAYATI, N. *Gamification Analysis and Design on Arabic Learning for Indonesian Elementary School Student*. In: 2024 12th International Conference on Cyber and IT Service Management (CITSM). IEEE, 2024. p. 1–6.
- BASSANELLI, S.; BUCCHIARONE, A.; GINI, F. *Gamidoc: The importance of designing gamification in a proper way*. IEEE Transactions on Games, 2024.
- BORGES, S.; DURELLI, V. H.; REIS, H. M.; ISOTANI, S. *A systematic mapping on gamification applied to education*. In: Proceedings of the 29th Annual ACM Symposium on Applied Computing. ACM, 2014. p. 216–222.
- BROOKE, J. *SUS: A quick and dirty usability scale*. In: Usability Evaluation in Industry. 1996. p. 4–7.
- DAREJEH, A.; SALIM, S. S. *Gamification solutions to enhance software user engagement—A systematic review*. International Journal of Human-Computer Interaction, v. 32, n. 8, p. 613–642, 2016.
- DETERDING, S.; DIXON, D.; KHALED, R.; NACKE, L. *From game design elements to gamefulness: defining “gamification”*. In: Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments. 2011. p. 9–15.
- DICHEV, C.; DICHEVA, D. *Gamifying education: what is known, what is believed and what remains uncertain: a critical review*. International Journal of Educational Technology in Higher Education, v. 14, p. 1–36, 2017.
- DICHEVA, D.; DICHEV, C.; AGRE, G.; ANGELOVA, G. *Gamification in education: A systematic mapping study*. Journal of Educational Technology & Society, v. 18, n. 3, p. 75–88, 2015.
- GARACCIONE, G.; COPPOLA, R.; ARDITO, L. *Gamifying business process modeling education: a longitudinal study*. In: Proceedings of the 28th International Conference on Evaluation and Assessment in Software Engineering. 2024. p. 580–589.
- GUEDES, D. P.; YAMAJI, B. H. S. *Revised Perceived Locus of Causality in Physical Education Scale: tradução, adaptação transcultural, validade psicométrica e concordância dos formatos impresso e online para uso em escolares brasileiros*. Retos: Nuevas Tendencias en Educación Física, Deporte y Recreación, n. 56, p. 390–407, 2024.
- HANUS, M. D.; FOX, J. *Assessing the effects of gamification in the classroom: A longitudinal study on intrinsic motivation, social comparison, satisfaction, effort, and academic performance*. Computers & Education, v. 80, p. 152–161, 2015.
- KAPP, K. M. *The Gamification of Learning and Instruction: game-based methods and strategies for training and education*. Hoboken: John Wiley & Sons, 2012.
- KITCHENHAM, B.; PRETORIUS, R.; BUDGEN, D.; BRERETON, O. P.; TURNER, M.; NIAZI, M.; LINKMAN, S. *Systematic literature reviews in software engineering – a tertiary study*. Information and Software Technology, v. 52, n. 8, p. 792–805, 2010.
- KITCHENHAM, B. *Procedures for Performing Systematic Reviews*. Keele: Keele University, 2004.



- KLOCK, A. C. T.; GASPARINI, I.; PIMENTA, M. S.; HAMARI, J. *Tailored gamification: A review of literature*. International Journal of Human-Computer Studies, v. 144, p. 102495, 2020.
- KLOCK, A. C. T.; OGAWA, A. N.; GASPARINI, I.; PIMENTA, M. S. *Does gamification matter? A systematic mapping about the evaluation of gamification in educational environments*. In: Proceedings of the 33rd Annual ACM Symposium on Applied Computing. 2018. p. 2006–2012.
- KOIVISTO, J.; HAMARI, J. *The rise of motivational information systems: A review of gamification research*. International Journal of Information Management, v. 45, p. 191–210, 2019.
- KUMAR, K.; RAGHUWAIYA, K.; KHAN, G.; SINGH, B.; PRASAD, J.; BLANCO, A. *Kahoot gamified technology use to assess student performance*. In: 2023 IEEE Asia-Pacific Conference on Computer Science and Data Engineering (CSDE). IEEE, 2023. p. 1–6.
- LAUGWITZ, B.; HELD, T.; SCHREPP, M. *Construction and evaluation of a user experience questionnaire*. In: Symposium of the Austrian HCI and Usability Engineering Group. Berlin, Heidelberg: Springer Berlin Heidelberg, 2008. p. 63–76.
- LEE, J. J.; HAMMER, J. *Gamification in education: What, how, why bother?* Academic Exchange Quarterly, v. 15, n. 2, p. 146, 2011.
- LO, C. K.; HEW, K. F. *A comparison of flipped learning with gamification, traditional learning, and online independent study: the effects on students' mathematics achievement and cognitive engagement*. Interactive Learning Environments, v. 28, n. 4, p. 464–481, 2020.
- MARKLAND, D.; HARDY, L. *On the factorial and construct validity of the Intrinsic Motivation Inventory: Conceptual and operational concerns*. Research Quarterly for Exercise and Sport, v. 68, n. 1, p. 20–32, 1997.
- MASROM, M. *Technology acceptance model and e-learning*. Technology, v. 21, n. 24, p. 81, 2007.
- MONTEIRO, R. H. B.; DE ALMEIDA SOUZA, M. R.; OLIVEIRA, S. R. B.; DOS SANTOS PORTELA, C.; DE CRISTO LOBATO, C. E. *The diversity of gamification evaluation in the software engineering education and industry: trends, comparisons and gaps*. In: 2021 IEEE/ACM 43rd International Conference on Software Engineering: Software Engineering Education and Training (ICSE-SEET). IEEE, 2021. p. 154–164.
- OGAWA, A.; KLOCK, A. C. T.; GASPARINI, I. *Avaliação da gamificação na área educacional: um mapeamento sistemático*. In: Brazilian Symposium on Computers in Education (Simpósio Brasileiro de Informática na Educação – SBIE). v. 27, n. 1, p. 440, 2016.
- PELLETIER, L. G.; ROCCHI, M. A.; VALLERAND, R. J.; DECI, E. L.; RYAN, R. M. *Validation of the revised sport motivation scale (SMS-II)*. Psychology of Sport and Exercise, v. 14, n. 3, p. 329–341, 2013.
- PESSINI, A.; OLIVEIRA, H. C. D.; KEMCZINSKI, A.; HOUNSELL, M. D. S. *O uso de jogos sérios na educação em informática: um mapeamento sistemático*. Nuevas Ideas en Informática Educativa TISE, p. 537–541, 2014.



POLITO, G.; TEMPERINI, M. *A gamified web based system for computer programming learning*. Computers and Education: Artificial Intelligence, v. 2, p. 100029, 2021.

REINHEIMER, W. S.; MEDINA, R. D. *Gamification in the educational context: a systematic mapping of literature with a focus on the evaluation of gamification*. Revista Novas Tecnologias na Educação, v. 18, n. 2, p. 398–407, 2020.

SAILER, M.; HOMNER, L. *The gamification of learning: A meta-analysis*. Educational Psychology Review, v. 32, n. 1, p. 77–112, 2020.

SCHATTEN, M.; SCHATTEN, M. *A comparative study of gamification in programming education in a Croatian high school*. In: 2020 43rd International Convention on Information, Communication and Electronic Technology (MIPRO). IEEE, 2020. p. 700–704.

SEABORN, K.; FELLS, D. I. *Gamification in theory and action: A survey*. International Journal of Human-Computer Studies, v. 74, p. 14–31, 2015.

SHANEYFELT, T.; BAUM, K. D.; BELL, D.; FELDSTEIN, D.; HOUSTON, T. K.; KAATZ, S.; GREEN, M. *Instruments for evaluating education in evidence-based practice: a systematic review*. JAMA, v. 296, n. 9, p. 1116–1127, 2006.

SMITH-ROBBINS, S. *This game sucks: How to improve the gamification of education*. EDUCAUSE Review, v. 46, n. 1, p. 58–59, 2011.

SPRINT, G.; FOX, E. *Improving student study choices in CSI with gamification and flipped classrooms*. In: Proceedings of the 51st ACM Technical Symposium on Computer Science Education. 2020. p. 773–779.

TODA, A. M.; OLIVEIRA, W.; KLOCK, A. C.; PALOMINO, P. T.; PIMENTA, M.; GASPARINI, I.; CRISTEA, A. I. *A taxonomy of game elements for gamification in educational contexts: Proposal and evaluation*. In: 2019 IEEE 19th International Conference on Advanced Learning Technologies (ICALT). IEEE, 2019. p. 84–88.

TODA, A. M.; VALLE, P. H.; ISOTANI, S. *The dark side of gamification: An overview of negative effects of gamification in education*. In: Researcher Links Workshop: Higher Education for All. Cham: Springer International Publishing, 2017. p. 143–156.

VANSTEENKISTE, M.; SIERENS, E.; SOENENS, B.; LUYCKX, K.; LENS, W. *Motivational profiles from a self-determination perspective: the quality of motivation matters*. Journal of Educational Psychology, v. 101, n. 3, p. 671, 2009.

VLACHOPOULOS, S. P.; MICHAILIDOU, S. *Development and initial validation of a measure of autonomy, competence, and relatedness in exercise: The Basic Psychological Needs in Exercise Scale*. Measurement in Physical Education and Exercise Science, v. 10, n. 3, p. 179–201, 2006.

WERBACH, K.; HUNTER, D. *For the win, revised and updated edition: The power of gamification and game thinking in business, education, government, and social impact*. Philadelphia: University of Pennsylvania Press, 2020.