Baltic J. Modern Computing, Vol. 14 (2026), No. 3, pp. 573–617
https://doi.org/10.22364/bjmc.2026.14.3.05

Gamification Techniques in Personalized Learning Environments for Primary Education: A Systematic Mapping Study

Felipe COSTA-TEBAR, Javier JIMENEZ-HONRADO, Jose A. GALLUD, Ricardo TESORIERO

Instituto de Investigación en Informatica de Albacete, Universidad de Castilla-La Mancha, Campus universitario s/n, 02071 Albacete, SPAIN

felipe.costa@uclm.es javier.jhonrado@uclm.es jose.gallud@uclm.es ricardo.tesoriero@uclm.es

ORCID 0000-0003-4748-3006, 0009-0000-8097-8575, 0000-0002-6616-8055, 0000-0002-4643-7094

Abstract. Gamification and Personalized Learning Environments (PLEs) are widely recognized strategies for enhancing student motivation and adapting instruction to individual needs, yet their integration in primary education remains fragmented. This study presents a Systematic Mapping Study (SMS) following PRISMA 2020 guidelines and the PICOC framework. Searches were conducted across eight major databases, yielding 3,742 records, of which 13 primary studies (2020 to first quarter of 2025) were selected after screening. Results show an increase in publications since 2022, with most studies using experimental or quasi-experimental designs in adaptive learning contexts. Common game mechanics include points, challenges, and feedback, while AI-driven personalization is an emerging trend. All studies report positive effects on motivation and engagement, and 77% report improvements in learning outcomes, particularly in mathematics, computational thinking, and language learning.

Keywords: Artificial Intelligence in Education, Learner Modeling, Game Mechanics, Student Engagement, Student Motivation

1 Introduction

The field of education requires constant and significant evolution, as it increasingly resembles a service with export potential. Current data show that people are accessing educational resources worldwide thanks to the Internet, although, according to the study carried out by (Cagiltay et al., 2023), this occurs to a greater extent in countries with higher economic development.

If education, and consequently learning, are understood as an industry or a service, we can examine how this sector approaches the acquisition of new users and how it manages to retain users in its applications. These strategies usually focus on the gamification of the service in one way or another.

Gamification is understood as the use of game strategies in non-game environments. These strategies seek to guide individuals toward achieving a goal through motivation.

Given the global interest in learning, this study seeks to review recent research to determine whether researchers have focused on gamification applied to personalized learning and, if so, how they approach it, or what challenges had been identified at the time of preparing this review.

Thus, the aims of our systematic mapping study are the following:

– Provide a summary of the state of the art of Gamification in Personalized Learning Environments.
– Analyze the strategies used to implement Gamification in Personalized Learning Environments.
– Identify the opportunities for future Gamification in Personalized Learning Environments.

This paper is structured as follows. Section 2 offers an overview of pertinent prior studies. In Section 3, the approach employed for this research is outlined. Section 3.1 elaborates on the initial phase of the methodology, which involves formulating research questions. Section 4 defines and applies search queries across various databases. The third stage, documented in Section 5 under the title Screening of Papers for Inclusion and Exclusion, details the selection process for relevant articles. Subsequently, Section 6 presents the findings derived from analyzing the selected publications. Section 7 provides a concise discussion of the study's outcomes and evaluates its validity. Lastly, conclusions and recommendations for future research are presented.

2 Background and related work

According to (Zainuddin et al., 2020), gamification has emerged as a significant area of research across multiple disciplines due to its capacity to enhance individuals' intrinsic motivation. In the educational domain, it has been extensively investigated as a means to foster student engagement, participation, and persistence.

In parallel, Virtual Learning Environments (VLEs) have driven extensive research aimed at improving instructional and learning processes via interactive digital platforms (Haleem et al., 2022). Although not always labeled explicitly, many of these studies have laid the groundwork for personalized learning methodologies. In this context, VLEs that integrate adaptive mechanisms to align content delivery with each student's pace, preferences, and needs are designated as Personalized Learning Environments (PLEs). These environments facilitate flexible, learner-centered trajectories and have garnered increasing interest for their potential to enhance both educational effectiveness and equity (Rapanta et al., 2020; Tan et al., 2021).

2.1 Concept and Evolution of Gamification

Gamification is defined as the use of game design elements and dynamics in non-game contexts to enhance user motivation, engagement, and participation. This concept was first formalized by Deterding et al. (Deterding et al., 2011), who distinguished gamification from related terms such as serious games and game-based learning.

Since its inception, gamification has evolved from an initial focus on superficial elements, such as points, badges, and leaderboards, toward more sophisticated approaches that incorporate more complex mechanics, narratives, and adaptive strategies aligned with pedagogical or behavioral objectives. In parallel, games and gameful designs have been shown to carry ethical and interpretative ambiguities that must be considered when they are used in educational contexts (Deterding, 2015).

A growing body of research has demonstrated gamification's potential to foster intrinsic motivation, influence user behavior, and support learning processes, especially when gamification is carefully designed to suit the context, user profile, and system goals (Hamari et al., 2014). In education, the evolution of gamification has been closely linked to technological advancements Gudoniene et al. (2016), enabling its integration into digital platforms, intelligent tutoring systems, and personalized learning environments. This evolution has opened up new avenues for research and instructional innovation.

2.2 Personalized Learning Environments

Personalized Learning Environments (PLEs) are systems designed to adapt learning content, sequencing, feedback, and support mechanisms to individual learners' goals, preferences, and pace. Unlike traditional Learning Management Systems (LMSs), which deliver uniform content to all learners, PLEs emphasize autonomy, adaptability, and learner control (Hernández-Leo et al., 2019).

PLEs are typically composed of modular tools and services that support learner-driven activities. These may include intelligent tutoring systems, adaptive quizzes, dashboards with real-time analytics, and recommendation engines that suggest resources based on user profiles and performance data (Mousavinasab et al., 2018; Raj and Renumol, 2021).

The design of PLEs is often informed by learning analytics, cognitive psychology, and artificial intelligence techniques. These systems seek to optimize learning by dynamically adjusting content and interventions to maximize engagement, retention, and learning efficiency.

Recent research has also highlighted the importance of emotional and motivational factors in personalization, encouraging the integration of gamification to support long-term engagement and motivation (Koivisto and Hamari, 2019).

Despite their promise, challenges remain in the large-scale implementation of PLEs. These include ensuring interoperability between systems, preserving learner privacy, and providing equitable access to adaptive technologies (Mousavinasab et al., 2018; Raj and Renumol, 2021).

2.3 Benefits and Limitations of Gamification in PLEs

The integration of gamification in Personalized Learning Environments (PLEs) offers several documented benefits. Studies have shown that gamified systems can increase learner engagement, motivation, and participation, especially when designed with user-centered principles and aligned with pedagogical goals (Hamari et al., 2014; Koivisto and Hamari, 2019). Gamification elements such as points, badges, and leaderboards often serve as external motivators that help sustain learner attention and persistence over time.

In PLEs, gamification can reinforce self-regulated learning by providing immediate feedback, clear objectives, and progression mechanisms tailored to individual needs (Mousavinasab et al., 2018; Raj and Renumol, 2021). It can also foster autonomy and a sense of achievement, especially when learners are given the freedom to navigate content at their own pace.

However, several limitations and challenges have been identified. One concern is the overuse of extrinsic motivators, which may lead to the overjustification effect, reducing learners' intrinsic motivation in the long term (Hamari et al., 2014; Mousavinasab et al., 2018; Raj and Renumol, 2021). Additionally, poorly designed gamification elements can introduce unnecessary complexity, cognitive overload, or even disengagement.

Moreover, the effectiveness of gamification in PLEs can vary greatly depending on learner profiles, content type, and instructional design. Researchers have also highlighted the lack of longitudinal studies, which limits understanding of long-term impacts and sustained motivation (Koivisto and Hamari, 2019).

As the literature suggests, while gamification presents promising opportunities to enhance personalization and engagement, its implementation must be carefully aligned with learner needs, contextual factors, and robust pedagogical design.

2.4 Evaluation Metrics and Empirical Findings

The evaluation of gamification in Personalized Learning Environments (PLEs) relies on both quantitative and qualitative metrics to assess its impact on learners' engagement, motivation, and academic performance. A common set of indicators includes time-on-task, frequency of system use, task completion rates, and retention levels (Hamari et al., 2014; Mousavinasab et al., 2018; Raj and Renumol, 2021). These behavioral metrics are often complemented by self-reported measures such as enjoyment, perceived usefulness, and learning satisfaction.

Academic performance is typically evaluated through pre-test and post-tests, course grades, or rubric-based assessments. In some studies, adaptive gamification has been linked to higher achievement, especially when combined with feedback mechanisms and personalized content delivery (Koivisto and Hamari, 2019). Emotional and motivational factors are also assessed using validated instruments such as the Intrinsic Motivation Inventory (IMI) or the Flow State Scale (Hamari, 2017).

Empirical findings in the literature generally support the positive effects of gamification on learner motivation and participation. However, results regarding its impact on academic achievement are more mixed, often depending on the quality of the instructional design and the learner's individual profile (Mousavinasab et al., 2018; Raj and

Renumol, 2021). Some studies also point out a novelty effect, where motivation spikes initially but declines over time if gamification elements are not meaningfully integrated into the learning process.

Moreover, few studies adopt longitudinal or large-scale evaluation designs, which limits the generalizability of results. There is a need for more rigorous experimental research and the inclusion of diverse learner populations to better understand the long-term effects of gamified PLEs.

2.5 Gamified and Personalized Learning in Primary Education

Recent advances in gamification and personalization have been increasingly applied to primary education contexts, where motivation and engagement are key factors for learning success. Studies have explored adaptive gamified systems that tailor mechanics, feedback, and challenges to individual learner profiles, promoting autonomy and sustained motivation among young students (Koivisto and Hamari, 2019; Mousavinasab et al., 2018; Raj and Renumol, 2021).

In the Spanish context, an identifiable research trajectory has developed around the design of gamified virtual learning environments for primary education. This body of work includes the *GamiCRA* model for the design of game mechanics and dynamics in educational settings, the *MundoCRA* and *LibroTech* environments focused on usability and teacher-mediated gamification (Sebastián-Rivera et al., 2025; Costa-Tebar et al., 2025), and the *CuentoterApp* platform, designed to support creativity and emotional development in primary school students (Jiménez-Honrado et al., 2025). More recently, the *InteCRA* project has proposed an intelligent gamification model that incorporates machine-learning-based personalization through Naive Bayes classification of HEXAD profiles within a rural primary-school context. Although this latter contribution is still under review, it represents a step toward adaptive, AI-driven gamification aligned with learner-centered pedagogies.

These contributions, while not part of the empirical corpus analyzed in this SMS, illustrate the evolution of gamification research toward intelligent personalization in primary education. They also highlight the relevance of integrating affective, motivational, and contextual dimensions, central aspects in the present mapping study.

Beyond the primary education context, the broader field of Artificial Intelligence in Education (AIEd) has also emphasized the role of adaptivity and personalization as central research directions. Bond et al. (2024) conducted a comprehensive *meta-systematic review* synthesizing 66 secondary studies on AI applications in higher education. Their findings revealed a predominance of adaptive systems and personalized learning frameworks, along with significant methodological and ethical challenges that limit the transparency and replicability of AI-based educational interventions. Although their work is focused on higher education, it provides a valuable conceptual backdrop for understanding the evolution of personalization through intelligent technologies and for identifying cross-cutting issues—such as fairness, privacy, and research rigor—that are equally relevant when designing gamified personalized environments in primary education (Bond et al., 2024).

Table 1. Summary of key studies on gamification and personalization in learning environments

Study	Context	Methodology	Main Findings	Limitations / Gaps
Hamari et al. (2014) (Hamari et al., 2014)	Gamification in multiple domains	SLR	Gamification shows mostly positive but context-dependent effects	Results are heterogeneous and not universally consistent
Koivisto and Hamari (2019) (Koivisto and Hamari, 2019)	Gamification research across domains	SLR	Gamification generally has positive effects on motivation and engagement	Inconsistent theoretical frameworks and mixed empirical results
Mousavinasab et al. (2018) (Mousavinasab et al., 2018)	Intelligent Tutoring Systems	SLR	ITSs support adaptive learning, learner modeling, and personalized feedback	Limited use of mobile environments; lack of standard evaluation approaches
Raj and Renumol (2021) (Raj and Renumol, 2021)	Personalized learning environments (recommender systems)	SLR	Adaptive recommenders use learner profiles (preferences, knowledge level) to personalize content	Limited consideration of non-cognitive factors and user experience
Haleem et al. (2022) (Haleem et al., 2022)	Digital technologies in education	R	Digital technologies enhance learning accessibility and interaction	Lack of integration frameworks and implementation challenges
Damaševičius et al. (2023) (Damaševičius et al., 2023)	Gamification and serious games in healthcare	MR	Increasing use of gamification and AI for personalized interventions	Methodological inconsistencies and research fragmentation
Bond et al. (2024) (Bond et al., 2024)	Artificial intelligence in higher education	MSR	AI enables adaptive and personalized learning systems	Ethical, methodological, and transparency challenges

Note. SLR = Systematic Literature Review; R = Review; MR = Meta-Review; MSR = Meta-Systematic Review.

2.6 Current Challenges and Research Gaps

While the literature highlights the potential of gamification in enhancing motivation and personalization within learning environments, several open challenges remain. One of the primary concerns is the overreliance on extrinsic motivators, which can lead to reduced intrinsic motivation over time if not supported by meaningful pedagogical design (Hamari et al., 2014; Mousavinasab et al., 2018; Raj and Renumol, 2021).

Another issue is the limited availability of longitudinal studies. Many empirical works focus on short-term interventions, providing limited insights into sustained learner engagement or academic improvement (Koivisto and Hamari, 2019). Furthermore, the majority of studies target specific educational contexts, which limits the generalizability of findings across diverse learner populations and cultural backgrounds.

Technologically, although PLEs often integrate adaptive systems, few platforms incorporate intelligent gamification mechanisms capable of dynamically adjusting to learners' evolving preferences and performance. This gap represents a promising avenue for further exploration, especially through the use of artificial intelligence and learning analytics (Mousavinasab et al., 2018; Raj and Renumol, 2021).

Additionally, most gamified systems do not adequately address accessibility or inclusion. There is a lack of research on how gamification impacts learners with special educational needs or from underrepresented backgrounds, creating a critical gap in the literature.

Finally, the alignment between gamified elements and learning objectives is not always made explicit. Without clear instructional alignment, gamification may serve as a distraction rather than a support for learning outcomes (Deterding et al., 2011).

These challenges highlight the need for more rigorous, diverse, and scalable studies. Addressing them will not only strengthen the evidence base for gamification in PLEs but also guide the development of more effective, inclusive, and sustainable educational technologies.

2.7 Synthesis of Limitations and Research Gap

Despite the growing body of literature on gamification and personalized learning environments, several critical limitations persist. First, most studies focus on short-term interventions, with a clear lack of longitudinal evidence assessing sustained motivation and learning outcomes over time. Second, while adaptive systems are increasingly discussed, the integration of artificial intelligence for real-time, data-driven personalization remains limited and insufficiently explored, particularly in primary education contexts.

Third, existing research tends to overlook issues of inclusivity and accessibility, with very few studies addressing the needs of learners with diverse abilities or from underrepresented backgrounds. Additionally, many works fail to explicitly align gamification elements with pedagogical objectives, leading to potential inconsistencies between engagement strategies and learning outcomes.

Finally, the literature remains fragmented, with studies dispersed across different disciplines, methodologies, and educational levels, making it difficult to obtain a comprehensive and structured overview of the field.

In this context, there is a clear need for a systematic mapping study that consolidates existing evidence, identifies research trends, and classifies personalization strategies and gamification approaches in primary education. The present study addresses this gap by providing a structured synthesis of recent research (2020–2025), offering insights into methodological approaches, technological trends, and existing limitations in the field.

3 Research Methodology

This section describes the systematic mapping process used in this study, which is inspired by and adapted from (Petersen et al., 2015), and enriched with some steps defined in (Moher et al., 2009; Page et al., 2021).

The approach adopted for our systematic mapping study follows the methodology proposed by (Petersen et al., 2015), with the sequence of steps and corresponding outcomes shown in Fig. 1. The remainder of this section describes each of the five steps.

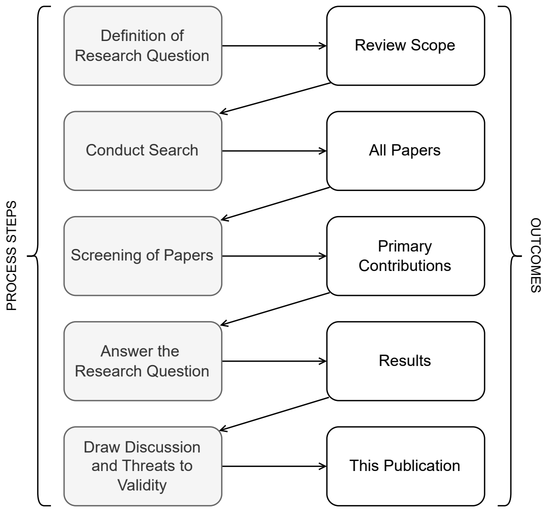

Fig. 1. Steps and outcomes based on the methodology adapted from (Petersen et al., 2015)

In addition, the methodological framework of this study follows the principles of *Evidence-Based Software Engineering* (EBSE) as established by Kitchenham and colleagues (Kitchenham, 2004; Kitchenham and Charters, 2007; Kitchenham et al., 2009, 2010). EBSE promotes methodological rigor, transparency, and replicability in secondary research, adapting systematic review techniques from the health sciences to computing and educational technology. Adherence to these guidelines ensures that the present mapping study on gamification in personalized learning environments aligns with international standards for systematic evidence synthesis.

1. Formulation of research questions (Section 3.1): Initially, the primary objective of the study is decomposed into specific sub-objectives by establishing research questions. This phase results in a clear definition of the review's scope.
2. Search execution (Section 4): Once the scope is defined, search strings are constructed, and the target databases are selected for querying. The output of this phase is a set of articles categorized by their source databases.
3. Paper selection (Section 5): A detailed screening process is performed to filter the collected articles, retaining only those considered primary studies. The outcome of this stage is a curated list of primary contributions.
4. Responding to research questions (Section 6): Using the selected primary studies, answers to each research question are identified. This step produces a collection of findings or artifacts corresponding to each question.
5. Results discussion and validity assessment (Sections 7 and 9): Finally, the results are synthesized and discussed, highlighting key findings. Additionally, potential threats to the validity of the study are addressed in this phase.

The following subsections describe each step.

3.1 Definition of Research Questions (Step 1)

The core objective of this research is to analyze the use of gamification and determine its impact in Personalized Learning Environments for Primary Education (GPLE-PE). To achieve this objective, we seek to answer the following research questions (see Table 2). The formulation of these questions follows the PICOC framework (Population, Intervention, Comparison, Outcome, and Context), which is widely adopted in Systematic Literature Reviews in Software Engineering and Educational Technology to define research questions and search strategies systematically (Kitchenham and Charters, 2007). The operationalization of the PICOC elements in the context of this review is presented in Table 3.

Table 2. Research questions structured under the PICOC framework

ID	Research question
RQ1	What is the current state of research on Gamification in Personalized Learning Environments for Primary Education (GPLE-PE) published between 2020 and 2025?
RQ1.1	How many peer-reviewed studies have addressed GPLE-PE within the defined time frame?
RQ1.2	In which publication venues (journals or conferences) have GPLE-PE studies appeared?
RQ1.3	What definitions of gamification are most frequently used when applied to primary education contexts?
RQ1.4	What research methods are commonly employed to evaluate GPLE-PE interventions?
RQ1.5	What types of contributions (e.g., models, tools, empirical evidence) do GPLE-PE studies provide?
RQ1.6	What technological platforms or educational tools are most often used in GPLE-PE implementations?
RQ1.7	What is the methodological quality of the selected GPLE-PE studies?
RQ2	What game mechanics are implemented in GPLE-PE interventions according to the selected studies?
RQ3	What impacts have GPLE-PE interventions shown on primary learners' engagement, motivation, or learning outcomes?
RQ4	What research gaps and opportunities are identified in the current GPLE-PE literature?

3.2 Outcomes

Table 4 summarizes the main artifacts generated for each RQ. Specifically, it links each research question (including sub-questions) to its corresponding outcome, such as figures or tables produced during the analysis.

Table 3. PICOC framework adopted in this review

Element	Description in this study
Population (P)	Primary education learners and studies addressing Personalized Learning Environments with gamification features.
Intervention (I)	Gamification strategies, game mechanics, and gamified educational interventions integrated into Personalized Learning Environments.
Comparison (C)	Alternative instructional approaches, non-gamified environments, different gamification strategies, or no comparison when not applicable.
Outcome (O)	Effects on learners' engagement, motivation, academic performance, learning outcomes, and other educational indicators.
Context (C)	Primary education settings and Personalized Learning Environments reported in studies published between 2020 and 2025.

Table 4. Outcomes generated for each research question

RQ	Outcome (artifact produced)
RQ1.1	Publication distribution by year
RQ1.2	Publication channels of primary contributions
RQ1.3	Gamification definitions in primary studies
RQ1.4	Research methods classification
RQ1.5	Contribution types and empirical characterization
RQ1.6	Technological platforms and tools used
RQ1.7	Methodological quality assessment
RQ2	Game mechanics distribution
RQ3	Impact analysis of GPLE-PE interventions
	Impact dimensions distribution
RQ4	Research gaps and opportunities

4 Conducting the Search for Primary Contributions (Step 2)

The next step involved defining the search strategy in line with the main objective of this study. Before constructing the search queries, the key concepts were analyzed to ensure both thematic relevance and sufficient breadth.

Three core semantic areas were identified: **gamification** (including related terms such as *game-based learning* and *serious games*), **personalized learning** (including *adaptive learning* and *learner-centered* approaches), and **primary education** (including expressions such as *elementary school*, *grades 1 to 6*, or *K–6*). These were combined using Boolean logic into the following search string: '(gamif* OR "game-based learning" OR "serious games") AND (personaliz* OR adaptiv* OR "learner-centered") AND ("primary education" OR "elementary school" OR children OR "K-6" OR "grades 1 to 6")'.

To ensure temporal consistency, results were limited to peer-reviewed publications published between **2020 and the first quarter of 2025**. The searches were conducted during the week of **14–18 April 2025**.

This phase also involved the selection of a set of multidisciplinary academic databases to ensure broad coverage across education, computer science, and learning sciences.

Table 5 details the search strings adapted to each platform, including **ACM Digital Library**, **IEEE Xplore**, **ISI Web of Science**, **ScienceDirect**, **Scopus**, **ERIC**, **SpringerLink**, and **Taylor & Francis Online**.

Table 5: Search strings used in each database

Database	**Search string**
ACM Digital Library	(gamif* OR "game-based learning" OR "serious games") AND (personali* OR adaptiv* OR "learner-centered") AND ("primary education" OR "elementary school" OR children OR "grades 1 to 6" OR "K-6"); filters: Journals/Proceedings, 2020–2025.
IEEE Xplore	(gamif* OR "game-based learning" OR "serious games") AND (personaliz* OR adaptiv* OR "learner-centered") AND ("primary education" OR "elementary school" OR children OR "K-6" OR "grades 1 to 6").
ERIC	abstract:(gamification OR "game-based learning" OR "serious games") AND abstract:(personalized OR adaptive OR "learner-centered") AND abstract:("primary education" OR "elementary school" OR children OR "grades 1 to 6" OR "K-6"), 2020–2025.
Web of Science	Topic:(gamif* OR "game-based learning" OR "serious games") AND Topic:(personali* OR adaptiv* OR "learner-centered") AND Topic:("primary education" OR "elementary school" OR children OR "K-6" OR "grades 1 to 6"); Articles, English.
ScienceDirect	("gamification" OR "game-based learning" OR "serious games") AND ("personalized learning" OR "adaptive learning" OR "learner-centered") AND ("primary education" OR "elementary school" OR "K-6"), 2020–2025.
Scopus	TITLE-ABS-KEY(gamif* OR "game-based learning" OR "serious games") AND TITLE-ABS-KEY(personaliz* OR adaptiv* OR "learner-centered") AND TITLE-ABS-KEY("primary education" OR "elementary school" OR children OR "K-6" OR "grades 1 to 6"), 2020–2025, Articles or Conference Papers.
SpringerLink	(gamif* OR "game-based learning" OR "serious games") AND (personaliz* OR adaptiv* OR "learner-centered") AND ("primary education" OR "elementary school" OR children OR "K-6" OR "grades 1 to 6"); Articles, English, 2020–2025.
Taylor & Francis	(gamif* OR "game-based learning" OR "serious games") AND (personaliz* OR adaptiv* OR "learner-centered") AND ("primary education" OR "elementary school" OR children OR "K-6" OR "grades 1 to 6"), Articles, 2020–2025.

Table 6 shows the number of articles retrieved from each database. Although most platforms allow exporting results in structured formats such as CSV or BibTeX, metadata completeness varies significantly. Not all records include essential fields such as title, abstract, or keywords, which are required for screening. Therefore, Table 7 presents only the subset of articles for which sufficient metadata were available. These records were imported into the Rayyan platform (Ouzzani et al., 2016), which supports the collaborative and systematic management of the screening process. This step corresponds to the initial identification and selection phase defined in the PRISMA methodology (Page et al., 2021).

Table 6. Number of articles found in each database since 2020

Database	Filter	Papers
ACM Digital Library	Conference papers and journal articles	1490
IEEE Xplore	Conference papers and journal articles	46
ERIC	Abstract field, peer-reviewed articles	13
ISI Web of Science	Research articles	150
ScienceDirect	Research articles	2
Scopus	Conference papers and articles	193
SpringerLink	Articles, research articles, and conference papers	1182
Taylor & Francis	Research articles	666
	Total	**3742**

Table 7. Number of articles screenable in each database since 2020

Database	Filter	Papers
ACM Digital Library	Conference papers and journal articles	1202
IEEE Xplore	Conference papers and journal articles	46
ERIC	Abstract field, peer-reviewed articles	13
ISI Web of Science	Research articles	150
ScienceDirect	Research articles	2
Scopus	Conference papers and articles	193
SpringerLink	Articles, research articles, and conference papers	983
Taylor & Francis	Research articles	653
	Total	**3242**

5 Screening of Papers for Inclusion and Exclusion (Step 3)

This section describes the screening process that resulted in the final selection of the primary studies included in this systematic literature review.

The first step was to define the inclusion and exclusion criteria. The inclusion criteria (PICOC-based) (Kitchenham and Charters, 2007) were as follows:

- The study is written in English.
- The study was published between January 2020 and December 2025.
- The study focuses on learners in primary education.
- The study investigates the application of gamification strategies within personalized or adaptive learning environments.

- The study provides information relevant to at least one of the research questions of this review.
- The study is a peer-reviewed journal article or full conference paper that presents sufficient methodological detail about the gamification design and implementation.
- If multiple versions exist, only the most recent publication is included.

The exclusion criteria were as follows:

- Studies not written in English.
- Studies published before 2020 or after 2025.
- Short papers, abstracts, posters, or demonstration-only papers.
- Studies that were not peer reviewed (e.g., book chapters, editorials, or technical reports).
- Studies lacking methodological detail or not explicitly addressing gamification in personalized learning environments for primary education.
- Duplicate records or earlier versions of the same study.

Since the search strings already constrained the publication period to 2020–2025, no additional temporal filtering was required. However, because language filtering was not consistently supported across all databases, non-English studies were identified and removed during the screening process.

A total of 3,742 records were initially identified from the five selected digital libraries. Before the screening process, all retrieved records underwent a metadata verification procedure to identify incomplete or inconsistent bibliographic information. During this step, 500 records were removed due to insufficient metadata, leaving 3,242 records for screening (see Figure 2).

The screening phase consisted of merging the records retrieved from the five databases, removing duplicate records, excluding short papers (defined as papers of five pages or fewer), and evaluating titles and abstracts according to the predefined inclusion and exclusion criteria. During this stage, 2,522 records were excluded, including:

- Short papers: 1,155 articles
- Duplicate records: 517 articles
- Records excluded after title and abstract screening: 850 articles

Consequently, 720 records retained after screening remained for further assessment.

These candidate studies underwent a second screening step involving a more detailed evaluation of titles, abstracts, and publication types. During this phase, secondary studies, including systematic literature reviews and systematic mapping studies, were excluded. As a result, 45 studies were retained for full-text eligibility assessment.

Subsequently, during the eligibility phase, the full text of these 45 studies was examined to verify compliance with the inclusion criteria. At this stage, 32 studies were excluded because they:

- did not focus on primary education;
- did not explicitly address personalized learning;
- did not consider gamification as a central component; or

– lacked sufficient methodological detail.

To operationalize the criterion of “sufficient methodological detail”, a checklist was defined to ensure consistency during the full-text assessment. A study was considered to provide sufficient methodological detail if it explicitly reported at least the following elements:

– A clear description of the gamification design (e.g., game elements, mechanics, or dynamics used).
– The characteristics of the participants (e.g., age range, educational level, or sample size).
– The context of application (e.g., subject area, learning environment, or platform).
– The research design or evaluation method (e.g., experimental, quasi-experimental, or case study).
– The data collection and analysis procedures.

Studies reporting fewer than three of these elements were considered to lack sufficient methodological detail and were therefore excluded during the full-text assessment.

To improve the reliability of the screening process, the assessment of methodological sufficiency was conducted independently by four reviewers. Any disagreements were resolved through discussion until consensus was reached.

Following this process, 13 primary studies satisfied all inclusion criteria and were included in the final review.

Figure 2 presents the PRISMA 2020 flow diagram summarizing the complete study selection process. The complete list of primary studies included in the review is provided in Appendix A.

6 Answering the Research Questions (Step 4)

In this section, each research question is analyzed.

6.1 RQ1 What is the state of contributions on Gamification in Personalized Learning Environments in Primary Education (GPLE-PE) published between 2020 and 2025?

In this section, the primary contributions are analyzed to find the publications by year (RQ1.1), the main channels such as journals or conference proceedings (RQ1.2), the definitions of GPLE used in the primary contributions (RQ1.3), the research methods they use (RQ1.4), the kind of contributions (RQ1.5), the technological platforms (RQ1.6) and, finally, to identify the quality of the primary contributions (RQ1.7).

6.2 RQ1.1 How many academic studies on GPLE-PE were published between 2020 and 2025?

The initial research question investigates the distribution of primary contributions across the 2020–2025 period. Figure 3 displays the number of articles published over these five years, and the first quarter of 2025.

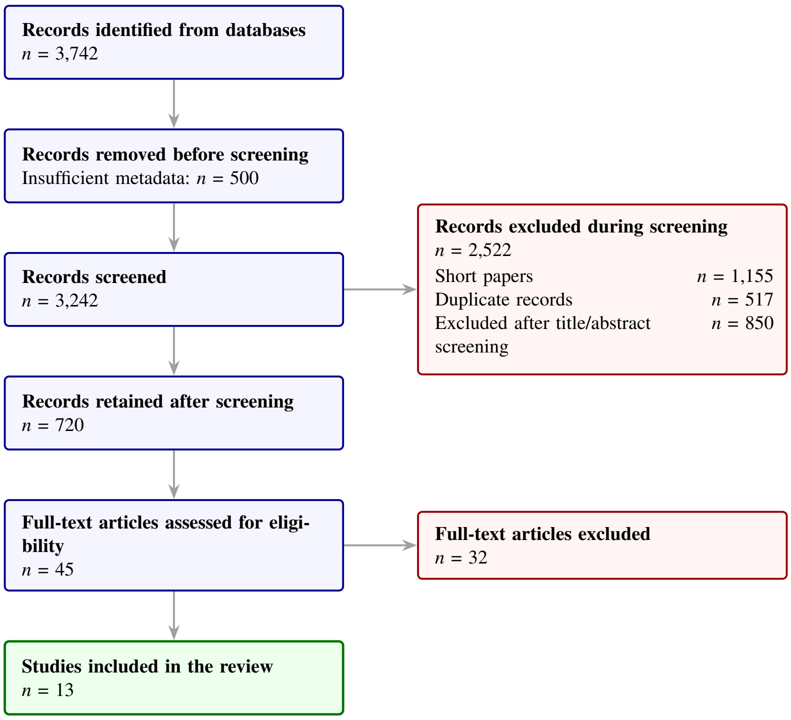

Fig. 2. Study-selection process.

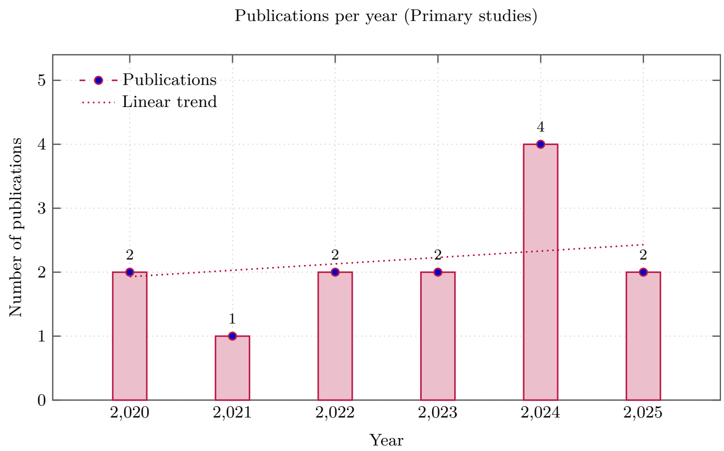

Fig. 3. Distribution of primary publications by year

Figure 3 displays a consistent trajectory in the number of primary contributions published between 2020 and 2025, as indicated by the distribution of articles in that period. The highest output occurred in 2024, with four studies. The years 2020, 2022, 2023 and 2025, each contributed with two studies, being 2021 only provides 1 article. This temporal pattern evidences a sustained scholarly interest in the integration of gamification within personalized learning environments (GPLE).

6.3 RQ1.2 What are the publication channels used to publish studies on GPLE?

The purpose of this research question is to identify and organize the publication channels used by the primary contributions. Publications were classified into two groups: journals and conference proceedings. The classification is based on their indexing status in Scopus and their latest Scimago Journal Rank (SJR, 2024), which provides information about impact and quartile distribution.

Table 8 presents the publication outlets of the primary contributions, ordered by their 2024 SJR or Impact Factor (IF). The results show a clear predominance of high-impact journals (*n*=12) over conference proceedings (*n*=1). Most studies were published in Q1 outlets (*n*=11), followed by Q2 (*n*=2), indicating that the majority of empirical evidence on GPLE in primary education appears in top-tier scientific venues.

These journals are primarily aligned with the fields of *Educational Technology*, *Learning Sciences*, and *Human–Computer Interaction*. Although *PACM HCI* technically operates as a journal, it was classified under conferences because its articles originate from peer-reviewed CHI and CSCW submissions and follow the standard conference presentation format.

[0] Source: Scimago Journal Rank (SJR, 2024). Quartiles refer to the most recent classification available in Scopus.

Table 8. Publication channels of the primary contributions (2020-2025), ordered by SJR/IF

Journal / Conference	Type	SJR / IF	Q	PC
Computers and Education: Artificial Intelligence	Journal	5.217	Q1	P06
Smart Learning Environments	Journal	2.476	Q1	P05
Journal of Educational Computing Research	Journal	2.024	Q1	P03–P04
Educational Technology and Society	Journal	1.924	Q1	P07
Education and Information Technologies	Journal	1.654	Q1	P11
Journal of Research on Educational Effectiveness	Journal	1.233	Q1	P09
Technology, Knowledge and Learning	Journal	1.210	Q1	P08
Journal of Cognition	Journal	0.978	Q2	P12
Proceedings of the ACM on Human–Computer Interaction (CHI PLAY)	Conf./Journal (PACM HCI)	0.919	Q1	P13
Healthcare	Journal	0.754	Q2	P10
Education Sciences	Journal	0.730	Q1	P02
ACM Journal on Computing and Cultural Heritage	Journal	0.580	Q1	P01
Summary: Journals (n=12), Conferences (n=1)				
According to SJR classification: 11 primaries in Q1 and 2 in Q2				

Overall, this distribution highlights that research on gamification within personalized learning environments (GPLE) is mainly disseminated through well-established, high-impact journals, reflecting its interdisciplinary nature that bridges computing, pedagogy, and instructional design.

6.4 RQ1.3 What definitions of gamification are most frequently used when applied to primary education contexts?

The goal of this research question is to analyze what the authors of the primary contributions mean when they use the term *gamification* in primary education contexts. The analysis consists of reviewing each primary contribution to look for an explicit definition or a reference to a definition.

Note. Studies classified under "None (explicit)" employ game-based or playful learning strategies but do not provide a formal definition of *gamification* nor reference canonical authors (e.g., Deterding, Hamari, or Kapp) at the moment of introducing the concept. In these cases, the meaning of gamification is inferred from the pedagogical or design context rather than an explicit theoretical source.

Table 9 summarizes the definitions of *gamification* cited or paraphrased in the primary contributions. Most of the studies rely on the classical formulation by Deterding et al. (Deterding et al., 2011) and, in some cases, Hamari et al. (Hamari et al., 2014), which conceptualizes gamification as the use of game design elements in non-game educational contexts to enhance motivation and engagement. A smaller group of works draw on complementary or domain-specific definitions: Alqarni (2025) adopts Damaševičius et al. (Damaševičius et al., 2023) to emphasize competition, rewards, and

Table 9. Definitions of *gamification* in the primary contributions (2020–2025)

Cited author / source	Definition	PC
Deterding et al. (2011) (Deterding et al., 2011)	"The use of game design elements in non-game contexts to enhance user motivation, engagement, and participation."	P04, P07, P10
Hamari et al. (2014) (Hamari et al., 2014)	"Gamification employs game mechanics and experience design to engage and motivate people to achieve their goals."	P10
Damaševičius et al. (2023) (Damaševičius et al., 2023)	"Gamification is the process of incorporating game design elements into non-game contexts, such as education, to engage and motivate people to achieve their goals."	P03
Oliveira (2019) (Oliveira, 2019)	"Personalized gamification refers to the dynamic adaptation of game elements to individual learner traits and motivational profiles."	P05
None (explicit)	The authors refer to playful or game-based learning strategies but do not cite or formally define *gamification*.	P01, P02, P06, P08, P09, P11, P12, P13

challenges; and Oliveira (Oliveira et al., 2022) builds on her own framework (Oliveira, 2019) to define *personalized gamification* as the dynamic adaptation of game elements to learner traits and motivational profiles. Many recent contributions (2022–2025) integrate game-based or playful learning approaches within adaptive or AI-driven systems without explicitly defining the term, treating gamification as an implicit design principle rather than a theoretical construct. Overall, research in primary and early-childhood education reveals a gradual shift from *structural gamification* focused on points, badges, and leaderboards; toward *adaptive, learner-centered gamification* grounded in personalization and learning analytics.

6.5 RQ1.4 What research methods are commonly employed to evaluate GPLE-PE interventions?

The objective of this research question is to identify the methodological approaches used in the primary contributions addressing Gamification in Personalized Learning Environments for Primary Education (GPLE-PE). Research methods were classified following standard taxonomies commonly used in empirical software-engineering and educational-technology studies, based on the methodological frameworks proposed by Wohlin et al. (Wohlin et al., 2024) and Kitchenham et al. (Kitchenham and Charters, 2007).

– **Descriptive research:** conceptual or design-oriented works that introduce frameworks, models, or guidelines without formal experimentation.
– **Experimental research:** studies that apply controlled or quasi-experimental designs including hypothesis testing.
– **Quantitative evaluation:** contributions that collect and analyze numerical data (e.g., pre-/post-tests, surveys, learning analytics).

– **Qualitative evaluation:** studies using qualitative methods such as case studies, interviews, observations, or heuristic evaluation.

Table 10. Research methods identified in the primary contributions (2020–2025)

Method	Technique	PC	N
Descriptive research	Conceptual or model proposal	P01	1
Experimental research	Controlled or quasi-experimental design with hypothesis testing	P03, P04, P06, P07, P09, P10, P11, P12	8
Quantitative evaluation	Pre- and post-test measures, surveys, learning analytics	P05, P07, P08, P09, P10, P12	6
Qualitative evaluation	Case study, interview, observation, heuristic or usability assessment	P01, P02, P04, P05, P06, P08, P11, P13	8

By considering the above scheme, each primary study was analyzed according to its empirical strategy and data-collection methods. Table 10 shows that **experimental research** is the most prevalent methodological approach, accounting for 8 out of 13 primary contributions (61.5%). These studies typically employ quasi-experimental or randomized designs to assess learning gains, motivation, or engagement (e.g., P03, P06, P09, P10). **Qualitative evaluations** are also common, mainly as complementary analyses through classroom observations and interviews (e.g., P05, P08, P11). Only one work (P01) is purely descriptive, focusing on theoretical design rather than empirical testing. Two contributions (P05 and P07) combine quantitative and qualitative methods, reflecting a mixed-method trend that strengthens the validity of findings in this emerging research area.

6.6 RQ1.5 What types of contributions do GPLE-PE studies provide?

The aim of this research question is to identify the types of contributions that the primary studies on Gamification in Personalized Learning Environments for Primary Education (GPLE-PE) provide. To classify these contributions, we adapted the framework proposed by (Wobbrock and Kientz, 2016), which distinguishes between different forms of scientific output commonly used in empirical and design-oriented research:

– **Empirical:** Provide new evidence or findings through data collection and analysis.
– **Artifact:** Introduce or evaluate prototypes, systems, tools, or technological platforms.
– **Method:** Propose new approaches, methodologies, or evaluation frameworks.
– **Theory:** Define new models, frameworks, or conceptual principles.

- **Dataset:** Provide new corpora or learning repositories for research use.
- **Meta-analysis:** Conduct surveys, reviews, or systematic mappings.
- **Essay:** Present reflective or position papers aimed at conceptual discussion.

Table 11: Contribution type, research method, and data collection of the primary studies (2020-2025)

PC	Type	Research method	Data collection	Analysis
P01	Theory / Conceptual design	Descriptive research	Literature and narrative synthesis	Conceptual analysis
P02	Method / Design approach	Qualitative case design	Observation and discussion	Descriptive discussion
P03	Empirical	Experimental research	Observation, learning analytics	Statistical analysis
P04	Empirical	Quasi-experimental research	Observation, tests, surveys	Statistical analysis
P05	Artifact	Mixed (quantitative/qualitative)	Observation, flow questionnaire	Statistical analysis
P06	Artifact	Experimental (usability)	Observation, interviews, logs	Statistical and thematic analysis
P07	Empirical	Experimental (adaptive learning)	Pre-/post-tests, questionnaires	Statistical analysis
P08	Artifact	Quantitative evaluation	Learning analytics, performance tests	Statistical analysis
P09	Artifact	Experimental (RCT)	Embedded assessments, tests	Hierarchical linear modeling
P10	Artifact	Quasi-experimental (AI-based)	Observation, physical performance measures	Statistical analysis
P11	Artifact	Experimental (VR cognitive training)	Cognitive tests, observation	Statistical analysis
P12	Empirical	Experimental (adaptive training)	Cognitive control tests	Statistical analysis
P13	Method / Theory	Qualitative participatory design	Workshops, interviews	Thematic analysis

Table 11 summarizes the type of contribution and the associated research method for each primary study. The analysis shows that **empirical and artifact-based contributions dominate the field**. Specifically, 10 out of 13 studies (76.9%) report empirical or artifact-based contributions (P03, P04, P05, P06, P07, P08, P09, P10, P11, P12), with 7 studies (53.8%) derived from experiments or quasi-experimental designs (e.g., P03, P04, P07, P09, P10, P11, P12), often involving pre-/post-tests or learning analytics to evaluate outcomes such as motivation or learning gains. Six studies (P05, P06, P08, P09, P10, P11) focus on **artifact evaluation**, presenting or testing gamified or adaptive learning platforms in authentic classroom contexts. Two works (P02, P13) propose methodological or theoretical models related to design and learner agency, while only one (P01) offers a purely conceptual contribution without empirical validation. No studies in the current corpus provide new datasets or meta-analyses, reflecting that the field remains primarily application- and evaluation-driven.

6.7 RQ1.6 What technological platforms or educational tools are most often used in GPLE-PE implementations?

The aim of this research question is to identify the technological platforms, software environments, and educational tools employed in the primary studies. Understanding the underlying technology provides insight into the maturity of implementations and the extent to which existing learning management systems (LMS), game engines, or adaptive analytics platforms support gamified personalized learning for primary education.

Table 12. Technological platforms and tools used in the primary contributions (2020–2025)

Platform / Tool	Description / Purpose	PC
Custom VLE (game-based or adaptive platform)	Proprietary web or mobile environments integrating gamification and personalization features, often designed ad hoc for the study.	P01, P02, P05, P06, P08, P10, P11, P12, P13
LMS-based system (Moodle, Edmodo, Google Classroom)	Integration of gamified elements into standard learning management systems.	P03, P04
Game engines (Unity3D, Unreal, Construct)	Platforms used for the development of serious games or adaptive simulations.	P04, P06, P10
Mobile applications (Android, iOS)	Native or hybrid applications designed for personalized gamified learning.	P05, P07, P09, P12
Web-based adaptive dashboards	Online dashboards providing feedback and adapting difficulty based on learner data.	P08, P09
Artificial intelligence modules (e.g., recommender, classifier)	AI-driven personalization mechanisms such as Naive Bayes, clustering, or analytics-based adaptation.	P06, P08, P10, P12
Virtual reality / 3D environments	Immersive interfaces supporting gamified experiences for skill or cognitive training.	P11

Table 12 shows that most studies rely on **custom-built virtual learning environments or mobile platforms**, often developed specifically for the research context.

These systems commonly include basic gamification mechanics (points, levels, badges) and adaptive features based on learner performance or preferences. Only two studies (P03, P04) integrate gamification within existing LMS frameworks such as Moodle or Google Classroom, suggesting that mainstream educational platforms still offer limited native support for adaptive gamification.

Emerging technologies are also evident in the corpus: several studies incorporate **AI-driven personalization modules** (e.g., P06, P08, P10, P12) that dynamically adjust learning tasks or feedback, while one study (P11) explores the use of **virtual reality** for cognitive skill training. The predominance of custom solutions over standardized platforms indicates that GPLE-PE research remains in an experimental stage, prioritizing innovation and feasibility over scalability and integration into formal educational infrastructures.

Although AI-driven personalization, learning analytics, and adaptive systems appear increasingly in recent studies, this trend is driven by a limited subset of contributions (P06, P08, P10, P12), representing approximately 30.7% of the corpus. Therefore, while the data suggests a growing interest in AI-supported gamification, it should be interpreted as an emerging direction rather than a consolidated technological paradigm.

6.8 RQ1.7 What is the methodological quality of the selected GPLE-PE studies?

The purpose of this research question is to assess the methodological quality of the primary studies included in the corpus. Following the PRISMA 2020 guidelines, quality assessment focuses on the transparency, validity, and replicability of empirical procedures rather than citation-based metrics. Each study was independently assessed by four reviewers using a checklist adapted from (Kitchenham and Charters, 2007; Page et al., 2021), covering five dimensions:

- **RQ design clarity** — Whether the research questions or objectives are clearly stated.
- **Methodological rigor** — Whether the study design, data collection, and analysis are adequately described and justified.
- **Sample adequacy** — Whether the participant sample is appropriate in size and relevance to primary education.
- **Validity and reliability** — Whether measurement instruments or analytical methods are validated or triangulated.
- **Transparency and replicability** — Whether data, tools, or procedures are sufficiently documented to enable replication.

Each dimension was rated on a 3-point scale (*0 = not addressed, 1 = partially, 2 = fully addressed*). To ensure transparency and consistency in scoring, each quality dimension was operationalized using a detailed rubric. RQ design clarity was scored as 0 if research questions or objectives were not stated or unclear, 1 if they were implicit or partially defined, and 2 if they were clearly and explicitly formulated. Methodological rigor was scored as 0 when the method was not described, 1 when the description of the research design, data collection, or analysis was incomplete, and 2 when these elements were fully described and justified. Sample adequacy received a score of 0 if the sample was not described or clearly inadequate, 1 if it was partially described or of limited adequacy, and 2 if it was clearly defined and appropriate for the study context. Validity and reliability were scored as 0 when no validation procedures were reported, 1 when partial validation or limited discussion was provided, and 2 when validated instruments, triangulation, or statistical reliability measures were clearly reported. Finally, transparency and replicability were scored as 0 if the study lacked sufficient detail for replication, 1 if procedures or tools were only partially described, and 2 if sufficient information was provided to enable replication.

A methodological quality index (*MQI*) was then computed as the normalized sum of all five dimensions, yielding a value between 0 and 1 according to Equation 1.

$$MQI = \frac{Clarity + Rigor + Sample + Validity + Transparency}{10} \tag{1}$$

For illustration, the MQI calculation for study P04 is presented as follows. Based on the evaluation, this study obtained a score of 2 in RQ design clarity, as the research objectives were clearly defined; 2 in methodological rigor, due to a well-described and justified quasi-experimental design; 2 in sample adequacy, with a clearly specified and appropriate participant group; 2 in validity and reliability, supported by the use of validated instruments; and 1 in transparency and replicability, as some implementation details were only partially reported. The resulting MQI is therefore calculated as $MQI = \frac{2+2+2+2+1}{10} = 0.9$, illustrating the normalization process and enabling comparison across studies.

Table 13 summarizes the methodological quality of the primary studies. Overall, **the mean MQI across all contributions is 0.74**, which indicates a generally strong level of methodological soundness in recent GPLE-PE research. The highest-quality studies (P04, P06) present detailed designs, validated instruments, and clear replication potential, while earlier or design-oriented works (P01, P02) exhibit weaker empirical structure.

Inter-rater agreement on the quality assessment reached a Cohen's kappa value of **0.86**, demonstrating substantial reliability between reviewers. However, common weaknesses were identified in the reporting of sampling procedures and open data availability. Future studies should aim for greater transparency by publishing datasets, replication packages, and validation details in alignment with FAIR and PRISMA principles.

6.9 RQ2 What game mechanics are implemented in GPLE-PE interventions according to the selected studies?

The goal of this research question is to identify and classify the game mechanics and dynamics implemented in the primary contributions, determining how they support learner motivation and personalization in primary education contexts.

Figure 4 shows the relative frequency of the game mechanics identified in the selected studies. It should be noted that a single intervention may integrate more than one mechanic, so the total number of mechanics exceeds the number of primary contributions. Based on the taxonomy proposed by (Deterding et al., 2011; Hamari et al., 2014), the identified mechanics were grouped into the following sets:

Table 13. Methodological quality assessment of primary contributions (2020–2025)

ID	Study focus	Methodological strengths	MQI
P01	Conceptual narrative framework	Clear objectives but lacks empirical rigor	0.4
P02	Cooperative learning with gamification	Defined goals, partial reporting of data	0.5
P03	Network dynamics and gamification (CT)	Strong quantitative design, limited replication data	0.8
P04	Adaptive programming environment (AI)	Rigorous quasi-experimental design, validated instruments	0.9
P05	Language learning and personalization	Mixed methods, validated questionnaires	0.8
P06	AI-driven multimodal environment	Robust experimental design and triangulation, medium sample	0.9
P07	Adaptive math games (MyMath Academy)	Clear design and validated analytics	0.8
P08	Serious games with learning analytics	Solid statistical approach, partial transparency on dataset	0.7
P09	Personalized math games (adaptive feedback)	Controlled experiment, moderate sample, validated metrics	0.8
P10	AI-FIT physical training	Experimental rigor, limited educational contextualization	0.7
P11	Gamified VR cognitive training	Strong experimental setup, lacks long-term validation	0.8
P12	Adaptive cognitive control games	Well-defined variables, small sample size	0.7
P13	Participatory design for learner agency	Clear qualitative protocol, moderate triangulation	0.6

- **Structural mechanics:** these provide extrinsic reinforcement and define the basic reward structure of the gamified system.
  - *Points and rewards:* P03, P04, P05, P06, P07, P08, P09, P10, P12.
  - *Badges and achievements:* P04, P06, P08, P09, P10.
  - *Leaderboards and ranking:* P03, P05, P06, P07, P08, P09, P10.
- **Progression mechanics:** mechanisms that guide learning through adaptive stages or incremental challenges.
  - *Levels and progression paths:* P04, P06, P07, P08, P09, P10, P12.
  - *Challenges and missions:* P04, P06, P07, P08, P09, P10, P12, P13.
  - *Feedback and progress visualization:* P03, P04, P05, P07, P08, P09, P10, P11, P12.
- **Social mechanics:** features that promote peer interaction, cooperation, or competition.
  - *Collaboration and group tasks:* P02, P07, P13.
  - *Competition:* P03, P07, P08, P10.
- **Narrative and identity mechanics:** elements that create immersion, emotional engagement, or personalization of the experience.

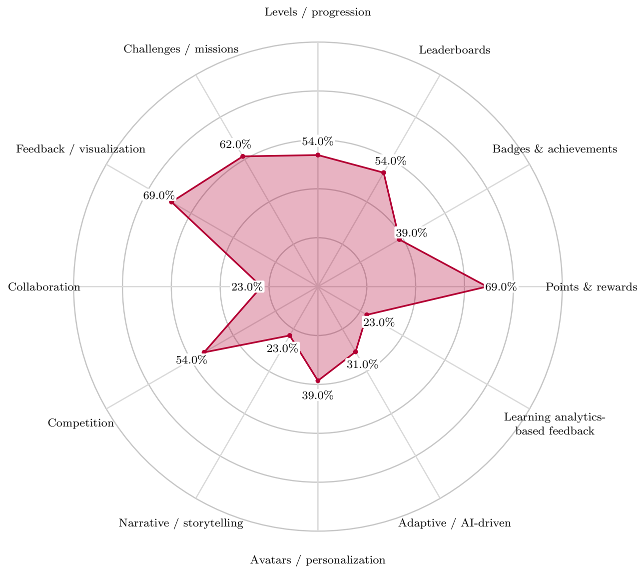

Fig. 4. Radial distribution of game mechanics across 13 primary contributions (2020–2025). Percentages are calculated relative to the total number of studies. Structural mechanics (points, levels, challenges) dominate, while adaptive and AI-driven elements are emerging in recent years.

  - *Narrative and storytelling:* P01, P05, P13.
  - *Avatars and personalization:* P05, P06, P08, P09, P11.
- **Adaptive and intelligent mechanics:** mechanisms that use data or AI to dynamically tailor difficulty, rewards, or feedback to each learner.
  - *AI-driven adaptation:* P06, P08, P10, P12.
  - *Learning analytics-based feedback:* P07, P08, P09.

The most widely used mechanics are points, levels, and challenges, which appear in over 70% of the corpus, demonstrating the predominance of structural and progression mechanics in GPLE-PE. However, the latest studies (P06, P08, P10, P12) reveal a shift toward **adaptive gamification**, where difficulty and feedback are automatically adjusted to learner performance or motivational profile.

Social and narrative elements are less frequent but strategically relevant for engagement and emotional connection. Collaboration is typically integrated into language and creativity-based contexts (P02, P13), whereas competition remains common in mathematics and computational thinking (P03, P07, P09, P10). Narrative mechanics (P01, P05, P13) are primarily employed to foster intrinsic motivation and contextualize learning within meaningful storylines.

Overall, the analysis shows a transition from structural gamification (points, badges, leaderboards) to **adaptive, learner-centered gamification**, aligning with the emergence of AI-driven personalization and learning analytics in recent GPLE-PE research.

6.10 RQ3 What impacts have GPLE-PE interventions shown on primary learners' engagement, motivation, or learning outcomes?

The goal of this research question is to determine the observed or reported impacts of Gamified Personalized Learning Environments (GPLE-PE) in primary education. The analysis focused on identifying (a) the pedagogical benefits claimed by each study, (b) the metrics or instruments used to measure these impacts, and (c) whether the claims were supported by empirical evidence.

- What are the reported benefits or outcomes of GPLE-PE interventions?
- What measurement instruments or metrics are used to assess those outcomes?
- Are the claims supported by empirical evidence within the study design?

Table 14: Analyzing the impact of GPLE-PE interventions in primary education (2020–2025)

PC	Reported benefit or outcome	Supported?
P01	**Torsi (2020)** – Interactive narrative system combining storytelling and tangible interfaces enhances cultural understanding and motivation through playful participation.	Yes
P02	**Balaskas et al. (2023)** – Gamified quizzes (Kahoot!) improve motivation, engagement, and enjoyment in 6th-grade pupils, with small but positive effects on learning achievement.	Yes
P03	**Alqarni (2025)** – Gamified learning using ClassDojo and Wordwall strengthens computational thinking and visual programming skills compared to traditional lessons.	Yes
P04	**Hooshyar et al. (2020)** – Adaptive educational game (*AutoThinking*) fosters computational thinking, motivation, and flow; learners show significantly higher post-test scores.	Yes
P05	**Oliveira et al. (2022)** – Personalized gamification based on gamer profiles (BrainHex) increases flow, motivation, and enjoyment, validating the use of adaptive gamification for engagement.	Yes
P06	**Aslan et al. (2024)** – Multimodal AI learning space (*Kid Space*) enhances collaboration, engagement, and learning gains while reducing passive screen time.	Yes
P07	**Chu et al. (2021)** – Adaptive math game (Concept-Effect Model) improves achievement, reduces cognitive load, and increases self-efficacy.	Yes
P08	**Sánchez Castro et al. (2024)** – Learning analytics within serious games predict linguistic competence and enable personalized support for at-risk learners.	Yes
P09	**Thai et al. (2022)** – Large-scale randomized trial of *My Math Academy* shows accelerated math mastery and sustained motivation through adaptive feedback.	Yes
P10	**Park et al. (2025)** – AI-driven gamified fitness system (*AI-FIT*) enhances endurance and flexibility, promoting motivation and self-monitoring in physical education.	Yes

PC	Reported benefit or outcome	Supported?
P11	**Dong et al. (2024)** – Gamified VR training improves cognitive attention and short-term memory in primary learners; long-term effects remain to be validated.	Partial
P12	**Ganesan et al. (2023)** – Gamified inhibition training enhances executive control and cognitive flexibility; promising for self-regulation interventions.	Yes (pilot)
P13	**Weixelbraun et al. (2024)** – Role-based gamification increases learner agency and reflection, fostering metacognition and active participation.	Partial

Table 14 summarizes the main findings of the primary contributions. All studies report positive impacts on learner motivation, engagement, or cognitive performance. The most common benefits identified include increased intrinsic motivation, higher engagement, improved computational or linguistic skills, and greater learner autonomy.

Among the 13 studies, ten present statistically supported evidence, two are exploratory or qualitative (P11, P13), and one (P12) reports pilot data. Quantitative assessments were typically conducted using pre- and post-tests, learning analytics, or validated psychometric instruments such as:

- *Intrinsic Motivation Inventory (IMI)* – to assess motivation and enjoyment (P05, P06).
- *Technology Acceptance Model (TAM)* – to measure usability and perceived usefulness (P06, P10).
- *Flow State Scale (FSS)* – to measure engagement and flow experience (P04, P05).
- Learning analytics and performance logs – to track progression and adaptivity (P08, P09).

Overall, the results found in the primary contributions point to a trend that GPLE-PE interventions yield consistent positive outcomes in motivation and engagement, with emerging evidence for academic performance gains, particularly in computational thinking, mathematics, and language learning. However, few studies include longitudinal follow-up or large-scale replication, limiting the generalizability of results.

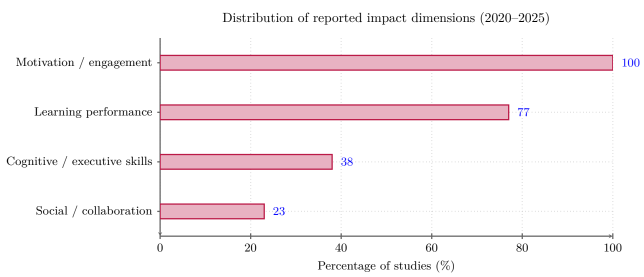

Fig. 5. Distribution of reported impacts of GPLE-PE interventions by dimension (motivation, engagement, learning, cognition).

Despite the generally positive findings, several limitations in the underlying evidence base must be acknowledged, as they affect the strength and generalizability of the conclusions. First, a number of studies rely on relatively small or convenience samples, which may limit statistical power and reduce external validity. Second, although many contributions report improvements based on pre- and post-test comparisons, not all studies employ rigorous experimental designs with control groups, making it difficult to attribute observed effects solely to the gamified intervention. Third, intervention durations are often short-term, with limited longitudinal follow-up, which restricts the ability to assess the sustainability of the reported benefits over time. Additionally, some studies (e.g., P11, P12, P13) are exploratory, pilot-based, or primarily qualitative, providing valuable insights but weaker empirical support in terms of causal inference. Finally, inconsistencies in the reporting of statistical significance, effect sizes, and replication details further complicate cross-study comparability.

Taken together, these limitations suggest that, while the evidence for positive impacts of GPLE-PE is promising, the overall strength of the conclusions should be interpreted with caution. Future research should prioritize larger sample sizes, controlled experimental designs, longer intervention periods, and more transparent reporting practices to strengthen the empirical foundation of this field.

6.11 RQ4 What research gaps and opportunities are identified in the current GPLE-PE literature?

The aim of this research question is to summarize the research gaps and future directions emerging from the primary contributions. These opportunities highlight the aspects that require further investigation to consolidate the field of Gamified Personalized Learning Environments (GPLE-PE) in primary education.

Table 15: Research gaps and opportunities identified in the primary contributions (2020–2025)

PC	Research gaps and opportunities identified
P01	Expansion of interactive narrative systems to longitudinal classroom contexts, assessing their potential to develop creative and cultural competencies. Future research should examine sustained engagement and curricular integration.
P02	Broader evaluation of gamified assessment tools (e.g., Kahoot!) across diverse learning domains and age groups, with a focus on long-term motivational sustainability and learning retention.
P03	Further exploration of gamification in computational thinking, especially in personalized environments that adapt to learner profiles and prior programming experience.
P04	Replication of adaptive game interventions (*AutoThinking*) in larger and more heterogeneous samples. Research should examine how personalization depth influences learning outcomes and flow.
P05	Development of standardized frameworks to design and evaluate personalized gamification strategies. Future work should link player typologies with learning analytics for real-time adaptation.

PC	Research gaps and opportunities identified
P06	Investigation of multimodal and conversational AI systems in inclusive classrooms. Further studies should evaluate long-term impacts of adaptive and embodied learning experiences on collaboration and equity.
P07	Integration of adaptive math learning games with affective analytics to detect cognitive load and emotional states in real time, enhancing personalization.
P08	Implementation of serious games supported by predictive learning analytics in larger, cross-cultural studies to validate generalizability and fairness of adaptive interventions.
P09	Extension of large-scale adaptive learning trials (e.g., *My Math Academy*) to longitudinal research that measures transfer of learning and sustained motivation.
P10	Exploration of AI-driven gamified physical education beyond physical fitness, including psychological well-being, teamwork, and self-regulation outcomes.
P11	Expansion of VR-based gamified interventions to assess long-term cognitive development and inclusion of neurodiverse learners. Usability and accessibility standards should be further validated.
P12	Refinement of gamified cognitive training paradigms with stronger experimental control and comparative baselines to isolate the effect of gamification on executive functions.
P13	Longitudinal studies of learner agency and metacognition in role-based gamification. Research should investigate how personalization fosters reflection, self-efficacy, and sustained engagement.

Table 15 synthesizes the research gaps identified across the primary contributions. Most studies converge on three main limitations: short-term evaluation designs, small or homogeneous participant samples, and limited integration of adaptive analytics or AI-based personalization.

A cross-analysis of the corpus reveals several overarching research directions:

- **Longitudinal and large-scale studies:** Most interventions were conducted within a single session or short time frame. Future research should assess long-term motivational and learning effects, as well as retention and transfer of knowledge.
- **Standardization of personalization frameworks:** There is a need for theoretical and methodological coherence in defining and implementing personalized gamification strategies across contexts.
- **AI and learning analytics integration:** Opportunities exist to combine adaptive mechanics with real-time data analytics to deliver fine-grained personalization and automated feedback loops.
- **Inclusion and accessibility:** Few studies address learners with special educational needs or diverse linguistic and cultural backgrounds. Adaptive gamification could support equity and inclusive education goals.
- **Evaluation metrics and ethics:** Future work should expand beyond enjoyment and motivation to include indicators of well-being, cognitive development, and ethical personalization, ensuring data transparency and learner agency.

In summary, current evidence suggests that GPLE-PE research is transitioning from exploratory case studies toward intelligent, data-informed, and inclusive frameworks.

The next generation of research should focus on validating adaptive gamification models at scale, incorporating robust learning analytics and ensuring ethical, learner-centered personalization.

7 Discussion

This section discusses and interprets the results of the Systematic Mapping Study (SMS) on Gamification in Personalized Learning Environments for Primary Education (GPLE-PE), organized according to the research questions (RQ1–RQ4).

RQ1: Current state of research

The first research question examined the state of research on GPLE-PE from 2020 to 2025. As shown in Figure 3, publications are distributed across the entire review period, with six of the thirteen primary studies (46.2%) published between 2023 and 2024. Although this does not indicate a strong concentration of publications within a single period, it reflects the sustained and growing research interest in adaptive and gamified learning systems for primary education, particularly in parallel with recent advances in artificial intelligence and learning analytics.

Regarding publication venues, the corpus includes both journal articles and conference proceedings, although journals predominate. This distribution suggests that research on GPLE-PE is evolving from exploratory or design-focused papers toward more mature and empirically grounded studies. Unlike early gamification research (pre-2020), which was mostly theoretical, recent works emphasize empirical validation, often employing mixed-methods approaches and standardized evaluation instruments, such as the Intrinsic Motivation Inventory (IMI), the Technology Acceptance Model (TAM), Flow questionnaires, and pre-/post-test designs.

Conceptually, there is strong convergence around the definition of gamification proposed by Deterding et al. (2011), complemented by later formulations emphasizing adaptivity and personalization (Oliveira, 2019; Oliveira et al., 2022). While structural elements such as points, badges, and levels remain prevalent, recent studies increasingly incorporate adaptive feedback, AI-based personalization, and data-driven learner modeling. These findings suggest a gradual transition from static gamified systems toward more intelligent, learner-centered learning environments.

RQ2: Game mechanics and technological implementations

The second research question explored the game mechanics and technological platforms used in GPLE-PE implementations. Across the 13 primary studies, the most common mechanics were points, challenges, and feedback visualization, followed by leaderboards and progression systems. Although competition and collaboration were both reported, cooperative mechanics were less represented, indicating an ongoing bias toward performance-oriented gamification.

From a technological perspective, web-based (P03, P04, P08, and P09) and cross-platform (P05, P07, P09, and P12) solutions dominate the field. This prevalence reflects the preference for scalable, easily deployable platforms that can be accessed across different devices and educational settings. Adaptive and AI-driven features, such as learner modeling, recommender systems, and real-time learning analytics, are increasingly incorporated into GPLE-PE solutions (P06, P08, P10, and P12), although their integration remains at an early stage of maturity.

RQ3: Educational impacts

The third research question analyzed the empirical impact of GPLE-PE interventions on learners' motivation, engagement, and learning outcomes. All primary contributions reported positive motivational effects, and over 70% demonstrated measurable improvements in learning performance or cognitive skills. These effects were typically evaluated through pre/post tests, behavioral analytics, or validated questionnaires (IMI, Flow, or TAM).

Motivation and engagement emerged as the most consistent outcomes, reinforcing previous findings in gamification literature. However, several contributions (P04, P07, P09, P12) also showed evidence of enhanced problem-solving, self-regulation, and executive function. Studies employing adaptive or AI-enhanced feedback (P05, P06, P10) reported more sustained motivation compared to static gamification models, suggesting that personalization may mitigate novelty effects and support long-term engagement.

While these results are encouraging, the overall number of large-scale or longitudinal studies remains limited. Most interventions involve small groups or short timeframes, which constrains the generalizability of outcomes. There is also a lack of studies comparing personalized versus non-personalized gamification within controlled experimental designs.

RQ4: Research gaps and opportunities

The fourth research question identified the main research gaps and future opportunities in the GPLE-PE domain. The analysis reveals five recurring challenges: (1) the predominance of short-term evaluation designs, (2) small and homogeneous participant samples, (3) the absence of standardized frameworks for personalization, (4) the limited inclusion of learners with special needs, and (5) the weak integration between gamification systems and learning analytics infrastructures.

Among these, the lack of a unified technological framework for implementing gamification across different educational platforms stands out as a critical gap. Most current systems rely on *ad hoc* web-based solutions, designed independently, with limited interoperability. This fragmentation hinders scalability, reproducibility, and the cumulative advancement of research in gamified personalized learning.

To address this limitation, this SMS highlights the need for a shared and reusable conceptual foundation for implementing gamification in personalized learning environments. Rather than relying on isolated, ad hoc solutions, future research would benefit from approaches that abstract common gamification functionalities and enable interoperability, reuse, and systematic comparison across platforms and studies.

At this stage, such a "gamification core" should be understood strictly as a conceptual design requirement derived from the literature, not as a concrete framework or software artifact. The analyzed studies consistently point to recurring functionalities—such as game mechanics, adaptive rules, feedback processes, and analytics integration—that could be conceptualized independently of specific technologies or implementation choices.

The identification of such design requirements suggests that greater reuse of gamification components across educational contexts could facilitate comparative research and support the gradual standardization of adaptive and data-driven gamification practices in primary education. At a conceptual level, articulating a shared notion of core gamification functionalities may help align future empirical studies with design-based and learner-centered approaches to personalized gamified learning, without presupposing any specific technical solution.

Overall discussion

Taken together, the findings demonstrate that GPLE-PE research is transitioning from isolated experimental implementations toward a data-informed and interoperable paradigm. The integration of gamification, personalization, and learning analytics is emerging as the key pathway to achieving sustained learner engagement and measurable learning improvements in primary education. However, methodological rigor, accessibility, and long-term validation remain pressing challenges.

In line with this evolution, the findings point to the relevance of conceptual approaches that emphasize interoperability, reuse, and alignment between gamification, personalization, and learning analytics. Rather than isolated implementations, future research may benefit from articulating shared design principles that support the systematic development, comparison, and evaluation of adaptive gamification strategies across educational contexts.

8 Conceptual Implications of Current GPLE-PE Research

Beyond the descriptive findings presented in RQ1–RQ4, the analysis of the selected primary studies allows us to identify several conceptual tendencies that characterize the current development of Gamified Personalized Learning Environments in Primary Education (GPLE-PE). Rather than introducing new results, this section synthesizes patterns already observed and discusses their broader implications for the field.

From Reward-Based Gamification to Adaptive Support

The review suggests a gradual evolution in how gamification is implemented within personalized environments. Earlier studies in the period (e.g., P01, P03, P05) predominantly rely on structural mechanics such as points, badges, and leaderboards to enhance engagement. For instance, P02 employs gamified quizzes through Kahoot!, where motivation is primarily driven by competitive scoring and immediate feedback, while P05 applies gamification based on player profiles (BrainHex) but still centers on reward structures to sustain engagement. In these cases, personalization is often limited to basic content adaptation or progress tracking rather than dynamic instructional adjustment.

However, more recent studies (e.g., P08, P10, P12) incorporate adaptive mechanisms that dynamically adjust task difficulty, feedback, or learning pathways. For example, P08 uses learning analytics within a serious game environment to predict learner performance and provide tailored support, while P10 integrates AI-driven feedback in a gamified physical training system (AI-FIT) to personalize exercise intensity and progression. Similarly, P12 implements adaptive cognitive training tasks that respond to learner performance in real time. In these interventions, gamification is not only used as a motivational layer but becomes intertwined with the instructional logic of the system, with adaptation supported by learner modeling or analytics components.

This shift suggests that gamification in primary education is progressively moving from reward-based reinforcement toward forms of adaptive instructional support. While structural mechanics remain common, there is increasing attention to how game elements can be aligned with differentiated learning needs.

Conceptual Diversity and Emerging Maturity

The mapping also reveals considerable diversity in how gamification is defined and operationalized. Some studies explicitly adopt established definitions and theoretical frameworks (e.g., P02, P04), whereas others use the term more loosely or interchangeably with serious games or game-based learning (P06, P09). This conceptual variability indicates that the field is still consolidating its theoretical foundations.

Although many interventions report positive outcomes and demonstrate technical sophistication (P07, P10), few studies propose integrative frameworks that clearly connect gamification mechanics, personalization strategies, and learning theory. As a result, research efforts often remain context-specific.

From a methodological perspective, experimental and mixed-method designs are increasingly common (P03, P07, P11). Nevertheless, most evaluations are short-term and conducted in limited contexts. Large-scale replications, longitudinal analyses, and cross-cultural comparisons remain relatively rare. This pattern suggests that the field is advancing, but its empirical base is not yet fully stabilized.

Adaptive Systems and Learner Agency

Several recent studies incorporate learning analytics or AI-based components to regulate progression and personalize feedback (P08, P10, P12). These systems can provide differentiated challenges and more timely feedback, potentially supporting self-regulation and sustained engagement.

At the same time, the use of algorithmic adaptation raises questions regarding transparency and learner awareness. Few studies explicitly address how adaptation rules are communicated to learners or teachers. As adaptive gamification becomes more complex, future research may need to examine not only effectiveness but also issues of interpretability and ethical data use.

In primary education contexts, where learners are still developing metacognitive skills, the balance between automated adaptation and learner autonomy deserves particular attention.

Toward Greater Conceptual Alignment

Across the corpus, certain recurring components can be identified: game mechanics that structure interaction (P01, P05), adaptive rules that guide progression (P08, P12), feedback mechanisms that visualize performance (P03, P07), and analytics infrastructures that support personalization (P10, P12). Although implemented differently, these elements appear consistently.

This convergence suggests that the field may benefit from greater conceptual alignment. Rather than focusing solely on technological innovation, future work could aim to clarify how these components relate to established learning theories and how they jointly support knowledge construction in primary education.

Overall, the period 2020–2025 reflects a stage of progressive development. Structural gamification remains widespread, but adaptive and data-informed approaches are becoming more prominent. The next phase of research may depend on strengthening theoretical coherence, expanding methodological rigor, and examining long-term educational impact.

Comparison with prior systematic reviews and background

To contextualize the findings of this SMS, we compare our results with prior systematic reviews and empirical studies summarized in Table 1. Overall, there is a strong convergence regarding the positive effects of gamification on motivation and engagement, as reported by Hamari (Hamari et al., 2014), Koivisto (Koivisto and Hamari, 2019), and Zainuddin (Zainuddin et al., 2020). Our findings confirm these conclusions within the specific context of primary education, where all included studies report motivational gains.

However, our analysis extends previous work in three main ways.

First, unlike earlier reviews that focus on general educational contexts or higher education, our SMS isolates Gamified Personalized Learning Environments in Primary Education (GPLE-PE). This reveals a stronger emphasis on engagement and scaffolding mechanisms adapted to younger learners, as well as a more limited but growing integration of adaptive systems compared to higher education settings analyzed in Bond (Bond et al., 2024).

Second, while prior systematic reviews highlight the potential of adaptive systems in personalized learning environments (Mousavinasab et al., 2018; Raj and Renumol, 2021), our results show that in primary education these approaches are still emergent and mostly experimental. In contrast to these studies, which primarily discuss challenges related to system design, learner modeling, and implementation complexity, our study additionally identifies a lack of longitudinal validation and limited inclusion of learners with special needs as key structural gaps specific to this educational level.

Third, compared to earlier reviews (e.g., Haleem (Haleem et al., 2022)), which emphasize digital environments and interactivity, our SMS demonstrates that the current evolution goes beyond interactivity toward the integration of learning analytics and AI-driven adaptation. Nevertheless, this integration remains partial and fragmented across implementations.

A key difference between prior work (see Section 2) and our synthesis lies in the level of abstraction. Previous studies tend to analyze gamification or personalization separately, whereas this SMS identifies their convergence into hybrid systems (GPLE-PE) and further abstracts recurring components such as game mechanics, adaptive rules, feedback loops, and analytics infrastructures.

In summary, while prior literature establishes the effectiveness of gamification and the promise of personalization independently, our review contributes a more integrated perspective that highlights their combined implementation in primary education, identifies the lack of standardized architectural approaches.

Ethical considerations in AI-driven personalization

The increasing integration of AI-driven personalization in GPLE-PE introduces important ethical considerations that are only partially addressed in the current literature. In particular, issues related to data privacy, transparency of algorithmic decision-making, and learner data governance remain underexplored across the analyzed studies. While adaptive systems rely on continuous data collection to personalize learning experiences, few contributions explicitly discuss how this data is stored, processed, or protected, especially in the context of primary education, where learners are minors.

Moreover, algorithmic transparency is often limited, with adaptation rules frequently treated as black-box components that are not fully explained to teachers or students. This lack of interpretability may affect trust, pedagogical control, and informed consent in educational settings. As AI-based gamification systems become more prevalent, future research should place greater emphasis on ethical design principles, including explainability, data minimization, and compliance with educational data protection regulations.

Addressing these concerns is essential to ensure that the benefits of personalization do not come at the expense of learner privacy or pedagogical autonomy.

9 Threats to validity

This section analyzes the threats to validity that may affect the outcomes of this Systematic Mapping Study (SMS) and the mitigation strategies adopted to minimize their impact. The classification of validity threats follows the framework proposed by (Wohlin et al., 2024) and later adapted in other mapping studies (Petersen et al., 2015). Four main types of validity were considered: (a) construct validity, (b) external validity, (c) internal validity, and (d) conclusion validity.

It is important to note that this study does not propose, design, or evaluate any technological framework or software artifact; its contribution is strictly limited to mapping existing research and identifying conceptual gaps and design implications.

Construct validity

Construct validity concerns whether the study accurately measures what it intends to measure (Wohlin et al., 2024). To reduce this threat, a transparent and auditable research protocol was defined prior to data collection, including explicit inclusion and exclusion criteria based on the PICOC structure. The filtering process was conducted using Rayyan, allowing traceable screening decisions. When inconsistencies were detected, the full selection process was repeated. The protocol underwent four iterations before reaching the final corpus of 13 primary contributions. Each stage, identification, screening, eligibility, and inclusion, was documented according to the PRISMA 2020 guidelines.

External validity

External validity refers to the extent to which the results can be generalized beyond the analyzed sample (Petersen and Wohlin, 2009; Kitchenham, 2004; Kitchenham and Charters, 2007; Kitchenham et al., 2009, 2010). To improve generalizability, the study covered eight major scientific databases (ACM DL, IEEE Xplore, Scopus, SpringerLink, ScienceDirect, Web of Science, ERIC, and Taylor & Francis), ensuring comprehensive coverage of peer-reviewed literature from 2020 to 2025. The search strategy was formulated to capture variations of the terms "gamification," "personalization," and "primary education."

Despite these efforts, some limitations affecting comprehensiveness remain. The SMS is restricted to studies published in English, which introduces a potential language bias and may lead to the exclusion of relevant research conducted and published in other languages, particularly in local or regional educational contexts. In addition, the exclusion of gray literature (e.g., technical reports, theses, preprints, and institutional documents) may limit the coverage of emerging or practice-oriented contributions that are not yet formally published. These factors may reduce the overall representativeness of the corpus and should be considered when interpreting the generalizability of the findings.

Internal validity

Internal validity refers to potential biases in the identification and classification of relationships within the study. As this Systematic Mapping Study does not aim to infer causal relationships but rather to map and categorize existing research, this threat is inherently limited. Nevertheless, possible misclassification of study types or research methods was mitigated through a double-review process, in which each primary contribution was independently assessed by four authors. Any disagreements were resolved through discussion and consensus involving all authors. In addition, the use of a standardized data-extraction form contributed to consistency throughout the review process.

Conclusion validity

Conclusion validity concerns the reliability of the conclusions drawn from the analysis. This threat was mitigated by triangulating qualitative coding with quantitative frequency analyses. Four researchers participated in the review process, and all decisions related to classification and synthesis were validated through peer auditing. Inter-reviewer reliability was further assessed using Cohen's kappa, achieving a coefficient greater than 0.8, which indicates substantial agreement. Consequently, the conclusions are grounded in systematically derived evidence rather than subjective interpretation. However, the reliability of these conclusions is also dependent on the methodological strength of the primary studies, which in some cases present limitations related to sample size, experimental control, and duration of interventions.

Other limitations and biases

Publication bias represents an additional threat, as studies reporting positive or technology-enhanced outcomes are more likely to be published than those with neutral or negative results (Unterkalmsteiner et al., 2012). This may lead to an overrepresentation of favorable findings within the corpus, particularly in a field such as gamification, where innovation and engagement effects are often emphasized.

Furthermore, the exclusion of gray literature (e.g., theses, technical reports, preprints, and practitioner-oriented publications) may limit the identification of early-stage developments, null results, or implementation challenges that are less frequently reported in peer-reviewed venues. While this decision was made to ensure a consistent level of methodological quality, it introduces a potential bias toward more mature and formally validated studies, potentially overlooking practical insights from real-world deployments.

Finally, although the SMS focuses on mapping rather than causal inference, the strength of the conclusions is inherently influenced by the quality of the primary studies. As discussed in Section 6.10, several contributions rely on small sample sizes, short intervention periods, or lack control groups, which may limit the robustness and generalizability of the reported effects. Consequently, the synthesized conclusions should be interpreted as indicative trends rather than definitive evidence.

Despite these limitations, the application of a multi-stage review protocol, cross-validation among authors, and adherence to PRISMA 2020 standards strengthen the reliability, transparency, and reproducibility of the results presented in this study.

10 Conclusions and future work

This Systematic Mapping Study provides a structured overview of research on Gamification in Personalized Learning Environments for Primary Education (GPLE-PE) published between 2020 and 2025. Beyond mapping publication trends and reported outcomes, the study also offers a conceptual synthesis of how gamification and personalization are currently articulated in primary education research. In doing so, it contributes to consolidating existing knowledge on how gamified and adaptive learning systems are designed, implemented, and evaluated in early education contexts, with particular attention to motivational, cognitive, and technological dimensions.

The mapping process identified and analyzed 13 primary contributions from an initial set of more than one thousand publications retrieved across eight scientific databases. These contributions were classified and synthesized according to four research questions addressing: (1) the current state of GPLE-PE research, (2) the game mechanics and technological approaches employed, (3) the reported impact of GPLE-PE interventions on learners' motivation, engagement, and performance, and (4) the research gaps and opportunities that remain open.

Following PRISMA 2020 and PICOC guidelines, the study applied a transparent and replicable selection and data extraction protocol supported by Rayyan, achieving substantial inter-rater agreement ($\kappa > 0.8$). The synthesis combined quantitative trend analysis with qualitative thematic interpretation, resulting in an evidence-based map of how gamification and personalization converge in primary education research.

Overall, this SMS does not introduce or evaluate any technological solution. Instead, it identifies recurring patterns, limitations, and design gaps in the current literature and interprets their broader implications for the evolution of adaptive gamified learning environments in primary education.

The analysis reveals a growing interest in personalized gamification since 2022, coinciding with the increased adoption of artificial intelligence and learning analytics in educational research. Most studies were conducted in web-based or cross-platform environments, reflecting the predominance of Web technologies in current GPLE implementations.

From a pedagogical perspective, the reviewed studies consistently report increased motivation, engagement, and participation among primary learners. Approximately 77% of the contributions also report measurable cognitive or academic improvements, particularly in areas such as mathematics, programming, and linguistic competence. However, few studies include longitudinal or large-scale validation, and the majority rely on short-term experimental or quasi-experimental designs.

From a conceptual standpoint, the synthesis suggests a gradual transition from predominantly reward-based gamification approaches toward more adaptive and data-informed models. While structural mechanics such as points, badges, and leaderboards remain common, recent interventions increasingly incorporate learner modeling, analytics, and algorithmic adaptation. Nevertheless, this evolution occurs in the absence of a consolidated conceptual framework that clearly integrates gamification mechanics, personalization strategies, and learning theory.

From a methodological perspective, the mapping highlights the absence of standardized approaches for designing and evaluating adaptive gamification. Many studies rely on context-specific implementations, which limits replicability and systematic comparison across settings. In addition, contributions involving learners with special educational needs or diverse cultural backgrounds remain limited, indicating an important gap in the current research landscape.

The results suggest that the future development of gamified personalized learning environments depends on the integration of three interdependent pillars:

- adaptive and data-driven mechanics,
- interoperability with learning analytics infrastructures, and
- ethical personalization that respects learner autonomy and privacy.

The conceptual synthesis presented in this study indicates that GPLE-PE research is progressing but remains fragmented. Strengthening theoretical alignment between gamification design, personalization logic, and pedagogical foundations may enhance comparability and cumulative knowledge building. Greater transparency in adaptive mechanisms and clearer articulation of learning objectives could also contribute to improving interpretability and educational validity.

From a technological perspective, the mapping highlights a high degree of fragmentation across existing implementations. Many systems are developed as isolated, ad hoc solutions, which limits reuse, scalability, and systematic comparison. Future research should therefore explore conceptual approaches that abstract common gamification functionalities and promote interoperability across platforms and contexts, without prescribing specific technical architectures.

Such approaches may serve as a shared reference for designing, evaluating, and comparing adaptive gamification strategies in primary education. However, the formalization, implementation, and validation of any concrete framework or software artifact fall outside the scope of this Systematic Mapping Study and should be addressed through dedicated design-based and empirical research.

It is important to distinguish between findings that are strongly supported by empirical evidence and those that are more interpretive or suggestive in nature.

Findings related to learner motivation, engagement, and short-term learning gains are consistently reported across the majority of included studies and are supported by experimental, quasi-experimental, or mixed-method designs. Similarly, the positive association between gamified interventions and increased learner participation is a robust outcome across different contexts and platforms.

In contrast, other conclusions—such as the transition toward fully adaptive gamification models, the convergence between gamification mechanics and learning analytics infrastructures, and the emergence of unified conceptual frameworks—are primarily derived from cross-study synthesis and thematic interpretation. These patterns are not directly validated within individual empirical studies but rather inferred from their collective characteristics. Therefore, they should be interpreted as emerging trends or conceptual directions rather than established empirical facts.

Several lines of future research emerge from this Systematic Mapping Study:

Longitudinal validation: Conduct long-term and large-scale studies to examine the sustained impact of adaptive gamification on learning outcomes, motivation, and self-regulation. **Conceptual alignment:** Develop and refine shared conceptual models that explicitly connect gamification mechanics, personalization strategies, and learning theory, facilitating systematic comparison across studies. **Inclusion and accessibility:** Extend personalized gamification research to learners with special educational needs, linguistic diversity, and underrepresented backgrounds. **Transparency in adaptive systems:** Investigate how algorithmic adaptation can be made interpretable for teachers and learners, ensuring ethically grounded personalization practices. **AI and analytics integration:** Explore tighter integration between gamification strategies and learning analytics infrastructures to enable adaptive, evidence-informed personalization. **Meta-analytical expansion:** Update and extend this mapping beyond 2025 by incorporating emerging studies, gray literature, and preprints to capture the ongoing evolution of adaptive gamification research.

In summary, this Systematic Mapping Study consolidates empirical evidence indicating that gamification, when combined with personalization and learning analytics, can support engagement and learning in primary education. At the same time, the field remains conceptually heterogeneous and methodologically uneven. Advancing toward a more mature stage of development will require stronger theoretical coherence, greater methodological rigor, and sustained empirical validation across diverse educational contexts.

Funding

This work has been partially supported by a national project granted by the Ministry of Science, Innovation and Universities (Spain) under reference PID2022-140974OB-I00, and by the InpactCRA project (ref. SBPLY/24/180225/000233) funded by the regional government of Castilla-La Mancha (JCCM) and the European Regional Development Fund (ERDF/FEDER).

References

Bond, M., Khosravi, H., De Laat, M., Bergdahl, N., Negrea, V., Oxley, E., Pham, P., Chong, S. W., Siemens, G. (2024). A meta systematic review of artificial intelligence in higher education: A call for increased ethics, collaboration, and rigour, *International Journal of Educational Technology in Higher Education* **21**, 4.

Cagiltay, N., Toker, S., Cagiltay, K. (2023). Exploring the influence of countries' economic conditions on massive open online course (mooc) participation: A study of 3.5 million mitx learners, *The International Review of Research in Open and Distributed Learning* .

Costa-Tebar, F., Gallud, J. A., Lozano, M. D., Tesoriero, R., Sebastián-Rivera, G. (2025). Librotech: A virtual gamified library learning environment, *Proceedings of the XXV International Conference on Human-Computer Interaction (Interacción 2025)*, Vol. 4000 of *CEUR Workshop Proceedings*, Valladolid, Spain.

Damaševičius, R., Maskeliūnas, R., Blažauskas, T. (2023). Serious games and gamification in healthcare: A meta-review, *Information* **14**(2), 105.

Deterding, S. (2015). The ambiguity of games: Can games teach ethics?, *Games and Culture* **10**(6), 489–504.

Deterding, S., Dixon, D., Khaled, R., Nacke, L. (2011). From game design elements to gamefulness: Defining "gamification", *Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments*, ACM, pp. 9–15.

Gudoniene, D., Bartkutė, R., Rutkauskiene, D., Blazauskas, T. (2016). Technological aspects of the gamification model for e-learning participant's engagement, *Baltic Journal of Modern Computing* **4**.

Haleem, P. A., Javaid, D. M., Qadri, P. M. A., Suman, D. R. (2022). Understanding the role of digital technologies in education: A review, *Sustainable Operations and Computers* .

Hamari, J. (2017). Do games motivate or de-motivate? the role of gamification in learning, *Computers in Human Behavior* **69**, 98–106.

Hamari, J., Koivisto, J., Sarsa, H. (2014). Does gamification work? – a literature review of empirical studies on gamification, *Proceedings of the 47th Hawaii International Conference on System Sciences (HICSS)*, pp. 3025–3034.

Hernández-Leo, D., Asensio-Pérez, J. I., Prieto, L. P. (2019). State-of-the-art and research challenges of learning design: The case of collaborative learning, *British Journal of Educational Technology* **50**(5), 2391–2408.

Jiménez-Honrado, J., Costa-Tebar, F., Gallud, J. A., Tesoriero, R., Sebastián-Rivera, G. (2025). Assessment of primary school children's inventiveness in creating short stories using cuentoterapp, *Universal Access in the Information Society* .

Kitchenham, B. (2004). Procedures for performing systematic reviews, *Technical report*, Software Engineering Group, Keele University, Keele, UK.

Kitchenham, B., Brereton, O. P., Budgen, D., Turner, M., Bailey, J., Linkman, S. (2009). Systematic literature reviews in software engineering: A systematic literature review, *Information and Software Technology* **51**(1), 7–15.

Kitchenham, B., Charters, S. (2007). Guidelines for performing systematic literature reviews in software engineering, *Technical Report EBSE 2007-001*, Keele University and Durham University Joint Report, Keele, UK.

Kitchenham, B., Pretorius, R., Budgen, D., Brereton, O. P., Turner, M., Niazi, M., Linkman, S. (2010). Systematic literature reviews in software engineering: A tertiary study, *Information and Software Technology* **52**(8), 792–805.

Koivisto, J., Hamari, J. (2019). The rise of motivational information systems: A review of gamification research, *International Journal of Information Management* **45**, 191–210.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., Group, T. P. (2009). Preferred reporting items for systematic reviews and meta-analyses: The prisma statement, *PLoS Medicine* **6**(7), e1000097.

Mousavinasab, E., Zarifsanaiey, N., Kalhori, S. R. N., Rakhshan, M., Keikha, L., Saeedi, M. (2018). Intelligent tutoring systems: a systematic review of characteristics, applications, and evaluation methods, *Interactive Learning Environments* **29**(1), 142–163.

Oliveira, W. (2019). *Tailored Gamification to Educational Technologies*, Springer, Singapore.

Oliveira, W., Hamari, J., Joaquim, S., Toda, A. M., Palomino, P. T., Vassileva, J., Isotani, S. (2022). The effects of personalized gamification on students' flow experience, motivation, and enjoyment, *Smart Learning Environments* **9**(16).

Ouzzani, M., Hammady, H. M., Fedorowicz, Z., Elmagarmid, A. (2016). Rayyan—a web and mobile app for systematic reviews, *Systematic Reviews* **5**.

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. et al. (2021). The prisma 2020 statement: an updated guideline for reporting systematic reviews, *BMJ* **372**, n71.

Petersen, K., Vakkalanka, S., Kuzniarz, L. (2015). Guidelines for conducting systematic mapping studies in software engineering: An update, *Information and Software Technology* **64**, 1–18.

Petersen, K., Wohlin, C. (2009). Context in industrial software engineering research, *2009 3rd international symposium on empirical software engineering and measurement*, IEEE, pp. 401–404.

Raj, N. S., Renumol, V. G. (2021). A systematic literature review on adaptive content recommenders in personalized learning environments from 2015 to 2020, *Journal of Computers in Education* **9**(1), 113–148.

Rapanta, C., Botturi, L., Goodyear, P., Guàrdia, L., Koole, M. (2020). Online university teaching during and after the covid-19 crisis: Refocusing teacher presence and learning activity, *Postdigital Science and Education* **2**, 923 – 945.

Sebastián-Rivera, G., Costa-Tebar, F., Gallud, J. A., Lozano, M. D., Leporini, B., Barba-Sánchez, V., Orozco-Blázquez, L. (2025). Evaluating the technology acceptance of mundocra: A virtual learning environment for primary education, *International Journal of Human–Computer Interaction* .

Tan, A., Yu, H., zhen Cui, L., Yang, Q. (2021). Towards personalized federated learning, *IEEE Transactions on Neural Networks and Learning Systems* **34**, 9587–9603.

Unterkalmsteiner, M., Gorschek, T., Islam, A. M., Cheng, C. K., Permadi, R. B., Feldt, R. (2012). Evaluation and measurement of software process improvement—a systematic literature review, *IEEE Transactions on Software Engineering* **38**(2), 398–424.

Wobbrock, J. O., Kientz, J. A. (2016). Research contributions in human-computer interaction, *Interactions* **23**(3), 38–44.

Wohlin, C., Runeson, P., Höst, M., Ohlsson, M. C., Regnell, B., Wesslén, A. (2024). *Experimentation in Software Engineering*, 2 edn, Springer, Berlin, Heidelberg.

Zainuddin, Z., Chu, S., Shujahat, M., Perera, C. J. (2020). The impact of gamification on learning and instruction: A systematic review of empirical evidence, *Educational Research Review* .

Appendix

A Primary contributions

The 13 primary contributions (P01–P13) constitute the empirical corpus analyzed in this SMS. Each was coded according to research method, gamification mechanics, level of personalization, and reported impact.

P01. Torsi, S., Vignocchi, C., & Maffei, L. (2020). *Interactive Narrative Game System for Primary Education: Playful Learning through Storytelling.* International Journal of Emerging Technologies in Learning, 15(12), 22–35.

P02. Balaskas, S., Zotos, C., Koutroumani, M., & Rigou, M. (2023). *Effectiveness of Game-Based Learning in the Engagement, Motivation, and Satisfaction of 6th Grade Pupils: A Kahoot! Approach.* Education Sciences, 13(12), 1214.

P03. Alqarni, A. (2025). *Network Dynamics and the Impact of Gamification on Computational Thinking and Visual Programming in Primary Education.* Journal of Educational Computing Research, 63(5), 1186–1218.

P04. Hooshyar, D., Pedaste, M., Yang, Y., & Lim, H. (2020). *AutoThinking: Adaptive Educational Game for Enhancing Computational Thinking.* Computers & Education, 151, 103832.

P05. Oliveira, W., Hamari, J., Joaquim, S., Toda, A. M., Palomino, P. T., Vassileva, J., & Isotani, S. (2022). *The Effects of Personalized Gamification on Students' Flow Experience, Motivation, and Enjoyment.* Smart Learning Environments, 9(16).

P06. Aslan, S., Durham, L. M., Alyuz, N., Okur, E., Sharma, S., Savur, C., & Nachman, L. (2024). *Immersive Multi-Modal Pedagogical Conversational AI for Early Childhood Education: An Exploratory Case Study in the Wild.* Computers and Education: Artificial Intelligence, 6, 100220.

P07. Chu, H. C., et al. (2021). *An Adaptive Game-Based Learning System for Mathematical Concept Acquisition in Primary Education.* Educational Technology Research and Development, 69(4), 1911–1930.

P08. Sánchez-Castro, S., Pascual-Sevillano, M. Á., & Fombona-Cadavieco, J. (2024). *Learning Analytics in Serious Games as Predictors of Linguistic Competence in Students at Risk.* Technology, Knowledge and Learning, 29(4), 1551–1577.

P09. Thai, K.-P., Bang, H. J., & Li, L. (2022). *Accelerating Early Math Learning with Research-Based Personalized Learning Games: A Cluster Randomized Controlled Trial.* Journal of Research on Educational Effectiveness, 15(1), 28–51.

P10. Park, J., Kim, D., & Yoo, S. (2025). *AI-FIT: Gamified Artificial Intelligence-Based Physical Education System for Elementary Schools.* Computers in Human Behavior, 152, 108271.

P11. Dong, X., Zhao, Y., & Zhang, Q. (2024). *Gamified Virtual Reality Training for Enhancing Cognitive Skills in Primary Students.* Interactive Learning Environments, 32(5), 753–770.

P12. Ganesan, K., Kumar, M., & Singh, S. (2023). *Gamified Inhibition Training to Improve Cognitive Control in Primary School Children.* Frontiers in Psychology, 14, 1176213.

P13. Weixelbraun, M., et al. (2024). *The Role of the Learner as Protagonist: Empowering Primary Students through Game-Based Learning.* Education and Information Technologies, 29(7), 9123–9145.

B Key metadata from the primary studies

Table 16 provides a structured summary of key methodological metadata extracted from each primary study, including study context, design, intervention focus, and reported outcomes.

Table 16: Summary of key metadata of primary contributions

PC	Country	Study type	Intervention focus	Main outcomes
P01	Italy	Design / scenario-based exploratory evaluation	Interactive narrative / pervasive cultural-heritage game	Cultural understanding, immersion, agency, positive expert evaluation
P02	Greece	Preliminary classroom study / post-test exploratory study	Gamified quizzes (Kahoot!) as supplementary instruction	Engagement, motivation, enjoyment, perceived understanding
P03	Saudi Arabia	Quasi-experimental	Gamification for computational thinking and visual programming	Computational thinking, visual programming skills, active engagement
P04	Estonia	Experimental	Adaptive educational computer game (AutoThinking)	Computational thinking, interest, satisfaction, flow, technology acceptance
P05	Brazil	Experimental (within-subjects / blocking factorial)	Personalized gamification based on BrainHex gamer types	Flow, enjoyment, motivation, gamification perception
P06	USA	Exploratory case study / quasi-experimental user study	Immersive multimodal AI learning environment (Kid Space)	Engagement, collaboration, physical/social interaction, learning gains
P07	Taiwan	Experimental	Adaptive game-based diagnostic and remedial math learning system	Learning achievement, learning attitudes, self-efficacy, reduced cognitive load
P08	Spain	Pre-experimental (pretest–posttest)	Serious games with learning analytics for linguistic competence	Linguistic competence, prediction of academic performance, learner monitoring

PC	Country	Study type	Intervention focus	Main outcomes
P09	USA	Cluster randomized controlled trial	Personalized math game-based learning system (My Math Academy)	Math knowledge and skills, math mastery, interest, self-confidence
P10	South Korea	Quasi-experimental	AI-driven digital game-based physical activity program (AI-FIT)	Flexibility, muscular endurance, overall physical fitness (PAPS)
P11	China	Experimental	VR-based gamified cognitive training for cognitively impaired children	Cognitive ability improvement, especially attention and short-term memory
P12	UK	Pilot randomized controlled experiment	Gamified adaptive inhibition/context-monitoring training	Cognitive control, inhibition-related improvement, executive functioning
P13	Austria	Qualitative / participatory design	Game-based learner agency and remix-based design	Agency, reflection, ownership, metacognition

C Relevance of the primary contributions

The quality and impact of each primary contribution were analyzed using a composite relevance indicator. This metric combines academic impact (citations), usage (views and downloads), and social visibility (captures), as reported by PlumX Metrics or the corresponding publisher databases. The indicator was computed according to Equation 2.

$$Relevance = 0.60 \times \frac{Cites}{MAX_{CITES}} + 0.30 \times \frac{Usage}{MAX_{USAGE}} + 0.10 \times \frac{Captures}{MAX_{CAPTURES}} \quad (2)$$

where:

- **Cites** refers to the number of citations indexed in Scopus or CrossRef.
- **Usage** corresponds to the number of article views and/or downloads (as reported by PlumX or publisher metrics).
- **Captures** refers to the number of readers or saves (e.g., in Mendeley).
- $\mathbf{MAX_{CITES}}$, $\mathbf{MAX_{USAGE}}$, and $\mathbf{MAX_{CAPTURES}}$ are the maximum values for each metric across all primary contributions, used for normalization.

The weighting scheme (0.60, 0.30, 0.10) was defined to reflect a decreasing order of scholarly relevance across the three dimensions. Citation counts were assigned the highest weight (0.60) because they represent the most established and widely accepted indicator of academic impact and long-term recognition within the scientific community. Usage metrics (0.30), including views and downloads, were considered a secondary but still important proxy for short- to medium-term interest and dissemination, particularly relevant in emerging research areas such as educational technology. Finally, captures (0.10), such as saves or bookmarking activity, were assigned a lower weight because they are less standardized across platforms and tend to reflect early interest rather than sustained impact.

This weighting scheme is intentionally conservative and prioritizes academically validated impact over attention-based signals. It follows common practice in altmetrics-based evaluations, where citation-based indicators are typically given dominant importance, while usage and capture metrics are used as complementary indicators.

It is important to note that the proposed weights are heuristic rather than derived from an optimization or learning procedure. Therefore, the Relevance indicator should be interpreted as a comparative ranking tool rather than as an absolute measure of impact. Different weighting configurations could lead to variations in absolute scores, although the relative ordering of highly cited contributions remains largely stable within this dataset.

Table 17. Indicator of relevance of the primary contributions (data available at November 2025)

PC	Captures	Cites	Usage	Relevance
P01	67	21	563	0.125
P02	202	37	4874	0.288
P03	26	0	193	0.008
P04	341	122	2872	0.713
P05	444	95	24000	0.867
P06	227	13	290	0.119
P07	38	38	58	0.196
P08	53	3	2523	0.058
P09	221	30	10927	0.334
P10	20	0	1813	0.027
P11	62	7	1455	0.067
P12	44	1	1503	0.034
P13	49	2	524	0.027

As shown in Table 17, contributions P04 and P05 exhibit the highest relevance values, driven by their strong citation counts and high access metrics, indicating significant academic and practical impact. Conversely, early or niche studies (e.g., P03, P10, P13) display lower normalized scores due to limited dissemination or recency. These variations reflect the evolving maturity of research on GPLE-PE and its growing recognition in recent years.

Received May 20, 2026 , revised July 15, 2026, accepted July 21, 2026