International Journal of Serious Games

ISSN: 2384-8766
https://journal.seriousgamessociety.org/

Article

Development of Serious Games for Science Assessment Using Educational Design Research

Patrícia Christine Silva[1], Paulo Nuno Vicente[2] and Ana Valente Rodrigues[1]

[1]Research Centre on Didactics and Technology in the Education of Trainers, Department of Education and Psychology, Universidade de Aveiro, Aveiro, Portugal; [2]Department of Communication, Universidade Nova de Lisboa, Lisbon, Portugal
{Patrícia Christine Silva} christine.silva@ua.pt; {Paulo Nuno Vicente} pnvicente@fcsh.unl.pt; {Ana V. Rodrigues} arodrigues@ua.pt

Keywords:

Serious Games
Primary education
Science education
Learning assessment
Educational Design Research

Received: August 2024
Accepted: March 2025
Published: May 2025
DOI: 10.17083/ijsg.v12i2.854

Abstract

This article addresses concerns inherent to the process of assessing science learning, which is incompatible with the recommendations from an Inquiry-Based Science Education perspective. The challenges in achieving effective assessment intensify with evaluative practices primarily aimed at classification, focused on knowledge, and characterized by the inadequacy of instruments and assessment activities used in educational contexts. This article aims to present and describe the development cycle of four serious games designed for formative and summative assessment of science learning in primary education. The serious games were created through iterative cycles supported by the Educational Design Research (EDR) method, structured into two main stages: design and implementation. This process took place over two academic years in a primary school, with the participation of 12 classes. A total of 245 children participated in the first implementation cycle, and data from 158 children were collected in the second cycle through post-implementation survey responses about the serious games. The products presented here are four serious games that serve as legitimate, validated, and innovative assessment activities, fulfilling their purpose of evaluating science knowledge, skills, attitudes, and values while engaging the target audience. This study strengthens the theoretical field in this area of research, demonstrating the suitability of this methodological approach in designing such digital educational games.

1. Introduction

1.1 Assessment of and for Science Learning

The assessment of science learning remains a critical challenge in education, particularly in primary education, where traditional methods often fail to capture the complexity of students' cognitive, procedural, and attitudinal development. Conventional assessment practices, primarily reliant on summative tests, frequently emphasize rote memorization rather than fostering deeper conceptual understanding and inquiry-based learning. This misalignment is especially problematic within the framework of Inquiry-Based Science Education (IBSE), which advocates for active student engagement in scientific exploration. In light of these challenges, the integration of digital tools—particularly serious games (SGs)—has gained increasing attention as a means to enhance both formative and summative assessment. Serious games not only provide engaging and immersive learning experiences but also enable real-time data collection and adaptive feedback, making them a promising alternative for evaluating science learning in primary education. This study investigates the development and implementation of SGs as assessment tools, aiming to bridge the gap between assessment theory and practice while addressing the limitations of traditional evaluation methods.

Assessment in the school context is a concept that has evolved over time, characterized by its dynamic and constantly maturing nature, subject to reflection, dissemination, and consolidation. Initially, educational assessment was exclusively a means of determining the effects of the teaching process on students [1]. This conception, focused on assessment as measurement [2], had the exclusive aim of quantifying and measuring student knowledge [3], [4, 5], and was disconnected from the teaching process. Over time, the concept evolved and was reinterpreted as "description" [2]. This means that, in addition to the inherent qualification of assessment, a descriptive component was attributed to it [4, 5]. With this new perspective, a break from the conception of assessment as solely summative was established, marking the beginning of formative assessment. Aligned with principles that view assessment as an integral part of the teaching process, assessment emerged as a value judgment [2], which influences the teacher's decision-making for better student learning. Assessment as negotiation and construction fills the gaps of previous conceptions, becoming a process that regulates and defines the teaching process, actively involving the student [5, 6]. From this perspective, assessment becomes an ally to the teaching process, promoting more and better learning.

Assessment modalities can be defined according to the educational objective they report. These are classified into three types: diagnostic, formative, and summative assessment [7]. In Portugal, Decree-Law No. 65/2016 provides for these three modalities in the context of formal education. It defines diagnostic assessment as a means that "(...) facilitates the students' school integration, supports the definition of teaching strategies, and aids school and vocational guidance" (p.1125). Formative assessment allows for the management of "(...) pedagogical measures appropriate to the students' characteristics and the learning to be developed, using detailed information devices on performance" (p. 1125). Finally, summative assessment "takes place at the end of each school term and results in a decision at the end of the school year about the students' progression, retention, or reorientation of their educational path" (p.1125). Various authors [6 - 8] have made considerations about the characteristics of these assessment modalities, summarized in Table 1.

Table 1. Systematized Characteristics of Diagnostic, Formative, and Summative Assessment

Characteristics	Types of assessment		
	Diagnostic	**Formative**	**Summative**
Objective	Identify what students know before exploration	Monitor and improve learning	Summarize and identify what has been learned
Implications	Action planning	Action adjustment	Summary of action
Feedback Type	Descriptive and informative	Detailed, constructive, and continuous feedback	Quantitative feedback (grading)
Timing	Before starting the teaching process, topic approach	During the teaching process	After the teaching process

Frequency	Occurs either occasionally (beginning of the year/terms) or systematically (start of didactic exploration)	Systematic (follows the teaching-learning process)	Occasional (providing a status update at predetermined times)

Despite the evolution of the concept and the apparent stabilization of its definition, assessment continues to be a subject of debate due to its complexity. Nowadays, even the validity of associating assessment with learning is questioned [8]. Regardless of these discussions, combining these assessment modalities in the teaching process is recommended, leveraging the potential each offers.

The use of diverse and appropriate assessment instruments and activities in science education is crucial to meet the defined educational purposes [6, 9]. Various examples contrast with traditional techniques and are recommended:

- Portfolios: This assessment activity in science allows for a collection of evidence, with its potential depending on the systematic and timely guidance and feedback from the teacher [3].
- Mind or Concept Maps: These are generally used after the teaching process to stimulate the identification, definition, and relationship between concepts [10, 11], [12]. Their potential as a formative assessment tool is recognized [13], with the advantage of visually connecting concepts [10].
- Drawings: Children's drawings are a powerful assessment tool in science education [14-18]. Often used in the early years with young children who do not yet master writing [15], they provide more detailed information when combined with writing in summative assessments [14, 16] and allow young children to express their conceptions about the world around them in diagnostic assessments [18].
- Concept Cartoons: Introduced by Keogh and Naylor [19], these are considered innovative and continue to be used in science education to diagnose children's ideas and foster classroom discussions [17, 20, 21].
- Assessment Recording Instruments (rubrics, grids, checklists): These have been used by teacher-researchers in empirical studies in primary science education [22], [23], allowing systematic and individual evidence collection of practical science work.

Traditional assessment activities (e.g., tests) exhibit significant limitations, particularly their inability to evaluate competencies beyond factual knowledge, thereby fostering short-term memorization of canonical content [24], which contributes little, if anything, to meaningful learning. Such assessments are generally disconnected from students' realities, serving solely the purpose of evaluation rather than promoting further learning [25]. Moreover, traditional tests are often anxiety-inducing for students, adversely affecting their motivation [26]. Despite the broad array of suggested alternative assessment activities, assessment practices in Portugal remain predominantly traditional. This persistence is attributed to the inadequacy of assessment tools and the lack of periodic formative approaches [27–30]. Teachers' difficulties in effectively assessing practical science activities are frequently highlighted as a key challenge [31].

Given the urgent need to modernize science assessment practices in primary education, this study explores the development, implementation, and evaluation of SGs as assessment tools aligned with the principles of IBSE. Grounded in the EDR method, this research aims to design assessment activities that not only measure students' scientific knowledge but also capture their problem-solving skills, reasoning abilities, and engagement with scientific inquiry. By leveraging the interactive and adaptive nature of SGs, this study seeks to offer an alternative to conventional assessment methods, addressing their limitations in evaluating complex learning processes. Moreover, by integrating digital assessment tools into the curriculum, the study aspires to contribute to a broader shift in pedagogical practices, promoting assessment strategies that are both rigorous and engaging. Ultimately, this research not only provides empirical insights into the feasibility and effectiveness of SGs for assessment but also contributes to the theoretical discourse on innovative approaches to science education evaluation.

1.2 Serious Games as an Assessment Activity for Learning

The rapid evolution of digital technology has cemented its integration into the teaching process. Assessment has also incorporated this technology through SGs, which stand out as an innovative educational strategy, promoting motivation, greater student engagement, and facilitating more learning as an assessment activity [32-41]. According to some authors, this concept of "Game-Based Assessment," though promising, is still in its embryonic stages [42-44]. Thus, it is essential to empirically investigate its potential and scientifically demonstrate its validity as an assessment activity [41, 45].

Kato and Klerk [45] highlight several advantages and challenges associated with SGs. They point out the playful and interactive nature of these games compared to more traditional and tedious activities, and their ability to reduce children's anxiety by involving them in a narrative, a predictor of academic performance. SGs can be programmed to capture and record all student decisions by storing their actions, unlike analog activities [45, 46], which do not always allow tracking all processes leading to the final answer. Automatic feedback is also seen as an advantage compared to more traditional activities [46]. However, significant challenges remain for the effective integration of SGs as an assessment activity in educational settings [41]. Investment is necessary for SGs to be established and consolidated as valid assessment activities capable of fulfilling their educational objectives. The game's construction must be balanced not to hinder the educational focus [32, 46].

From the limited empirical research on SGs for learning assessment, some studies use quizzes for this purpose. For example, the use of Kahoot! has been commonly employed as a gamified assessment activity [35, 36, 47, 48, 49, 50, 51]. In the study by Iman, Ramli, and Saridewi [48], there was a positive appreciation for using these games, valuing the immediate feedback after answer validation, which enhances more and better learning. Students also refer to these activities as motivating and fun. Widyaningrum [51] reiterates previous conclusions and adds that their use benefits teachers' work, particularly by saving correction time, a feature highly valued by teachers in other studies [47].

Other SGs have been used as assessment activities. For example, the Virtual Cooking Setup [40] aims to assess the cognitive processes involved in multitasking and to determine if actions in the virtual task represent real-life actions. The results suggest the SG's effectiveness in achieving its objectives.

A SG was developed for preschool children with autism spectrum disorder, recreating everyday situations and challenging them to complete tasks [52]. This SG is identified as an innovative activity with several advantages: it allows monitoring learning, promotes a playful and relaxed assessment environment, and identifies ways to stimulate better learning.

Another study in Portugal found that using SGs as diagnostic, formative, and summative assessment activities in primary science education is promising in achieving its objectives [21]. The efficacy of assessment tests, concept cartoons, and SGs was compared, concluding that SGs promote greater autonomy and respect individual children's rhythms. Final questionnaires conducted with children indicated that SGs are the preferred assessment activity among participants.

In summary, the studies presented contribute to the growing body of evidence supporting the value of digital games in learning assessment, while also underscoring the need for further research into the specific potential of SGs within this context.

1.3 Experimental Science Teaching Program

With the development of the Experimental Science Education Program (PEEC), there is a proposal to contribute to the necessary renewal of science education in Portugal, based on the perspectives of IBSE and Science-Technology-Society (STS) orientation. PEEC consists of three interdependent components: curricular, activity-based, and assessment. All resources are intended to be available for free online on the official website. The PEEC curricular component includes a curriculum proposal for primary science education in Portugal, where specific objectives in knowledge, skills, and attitudes/values are organized into four areas: Biological Sciences, Physical Sciences, Earth Sciences, and Nature of science [53]. Regarding the curricular component, a set of activities was developed to facilitate its implementation [54]. Created teaching resources include static and dynamic infographics, analog classification games, card games, recorded interviews, record sheets, and more.

This article focuses on the development of the PEEC assessment component. Consistent with the other two components, the assessment component aims to evaluate science learning in terms of knowledge, skills, and attitudes/values. SGs are suggested as assessment activities as an alternative to traditional tests.

In this context, the expressed need for change in the assessment practices of science learning in primary education in Portugal, coupled with inspiration drawn from identified practical cases, led to the formulation of the guiding research question for this study: *How can serious games be developed for the assessment of and for primary school children's learning, consistent with the proposed experimental science teaching project?* Based on this research question, the study presents the development cycle (design, implementation, and evaluation) of four serious games (one for each grade level of primary education) for assessment purposes, employing the phases of EDR.

2. Methods and Material

This article aims to describe the development cycle of SGs for assessing science learning in primary education. As determined in other studies of this nature [55, 56], the EDR approach emerged as the most appropriate, given the defined characteristics and purposes. EDR unfolds in three essential phases: analysis, design and implementation, and evaluation. The article primarily focuses on the design, implementation, and evaluation phases, providing a brief description of the preliminary analysis phase.

2.1 Phase I - Analysis

The first phase of the EDR - analysis - identifies in-depth needs, gaps, and real problems in educational contexts [57, 58]. Detecting a problem is not a trivial task, and its thorough understanding is essential for a prolific resolution [59]. The problem must be researchable and solvable with a dual purpose: contributing to solving real problems in educational contexts and to the theoretical knowledge of the area in question [58, 59]. The research question defined for an EDR study was: "How to develop an assessment project for and of primary school children's learning consistent with the proposed experimental science teaching project?".

The problem addressed in this article results from identifying the need for a curricular restructuring in the natural sciences area in Portugal [60], which requires congruent proposals for assessment instruments and activities [61, 62]. A literature review revealed that, despite strong recommendations for learning-centered assessment, practice changes have been slow, particularly in primary science education [61, 63, 64]. Furthermore, as required by EDR research, understanding the problem also involved identifying and characterizing assessment practices in primary science education in Portugal. This analysis was conducted through the study of 119 Individual School Group Reports and three Annual Global Reports by the General

Inspection of Education and Science (IGEC) within the scope of experimental science teaching in primary education [65]. This study found that assessment practices in schools are decontextualized, do not promote child involvement in the process, are not transparent, and the instruments used are inadequate or even nonexistent for competency assessment, with practical work being almost excluded from the assessment component.

Given these findings, this phase of the EDR proposes guidelines for an educational solution, including forming a multidisciplinary team for diverse contributions and defining "critical friends" for designing science assessment activities for primary education, consistent with the already created curricular proposal.

2.2 Phase II - Design and Implementation

In Phase II, design and implementation, the still generic and abstract plans defined in Phase I take shape, precision, and support [57], with a detailed plan as schematized in Figure 1.

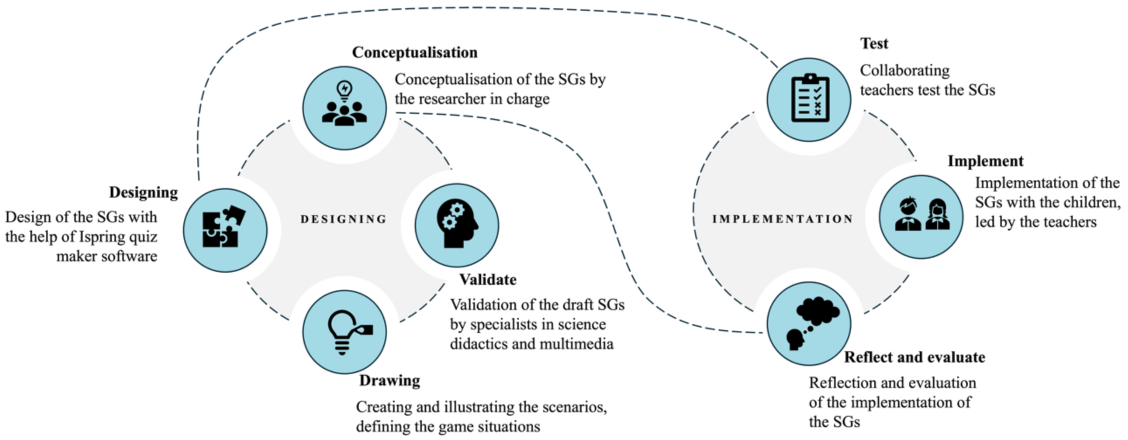

Figure 1. Outline of Tasks for Phase II - Design and Implementation

2.2.1 *Design Phase*

In this phase, the roles and functions of different members of the multidisciplinary team become evident [66]. Collaborating teachers act as “critical friends” [59], providing continuous and systematic feedback to refine and optimize the SGs. Testing and evaluating the SGs with the target audience, namely primary school children, allowed, in an initial phase (academic year 2020/2021), to test the feasibility of these games and, in a later phase (academic year 2021/2022), to ascertain the opinions of the end-users. Thus, the implementation of iterative cycles allowed stakeholders to voice their perspectives on the SGs created and tested [56], until expectations were met by all involved. This iterative and collaborative cycle ensured that the developed SGs were adapted to the real needs of the educational context, promoting an assessment that is not only more engaging but also more effective and aligned with the PEEC objectives.

It is common to integrate programmers into a multidisciplinary team for SG development [55, 67]. To compensate for limited research budgets, viable and effective software alternatives are often used [21, 68]. Nonetheless, mastering the chosen platform is essential to achieve the educational objectives of the SGs. In this case, licenses for Ispring Quiz Maker were acquired, a software already validated for assessment purposes [69-71].

In an EDR approach, for creating technological solutions, it is important to recognize the potential and identify technological possibilities capable of addressing the defined problem [72]. According to the taxonomy presented by De Lope and Medina-Medina [73], the categorization of these SGs is shown in Figure 2.

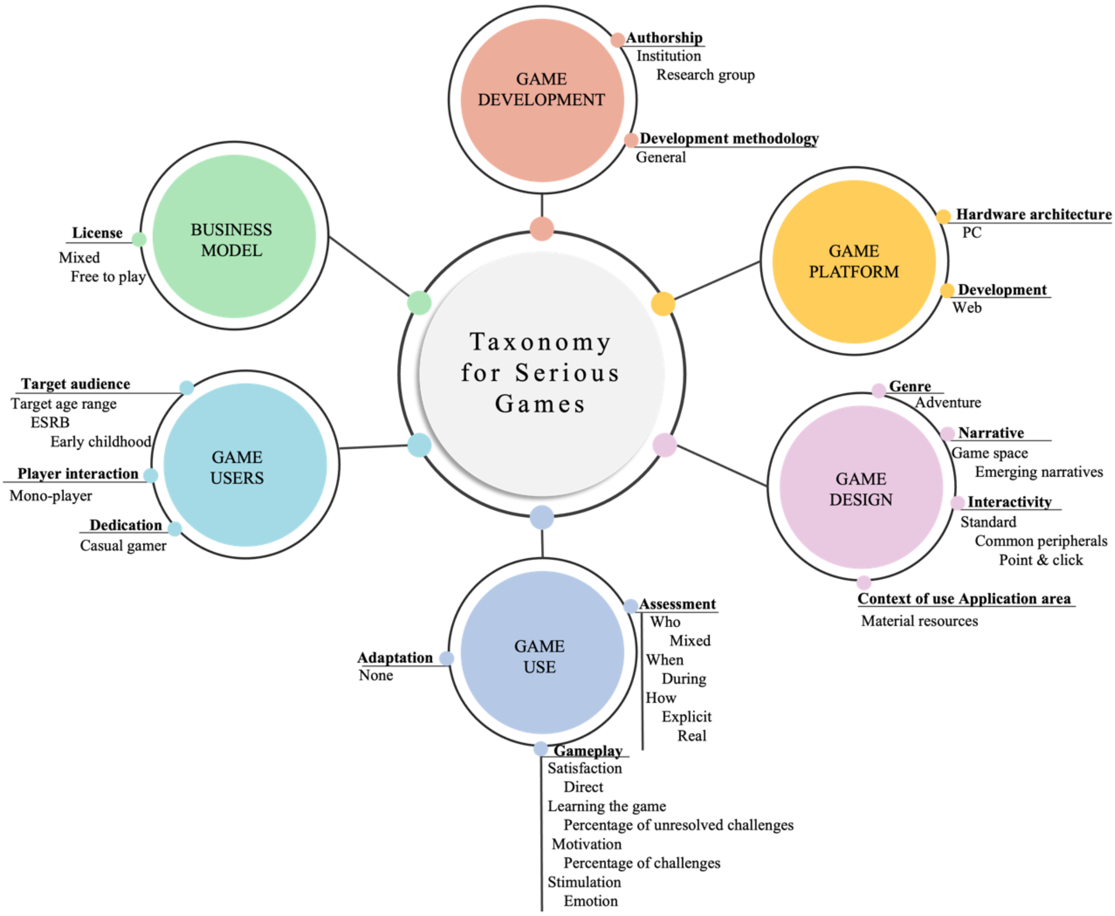

Figure 2. Categorization of SGs according to De Lope and Medina-Medina's Taxonomy [38]

In this regard, the design and construction of high-fidelity prototypes of the PEEC's serious games were supported by an in-depth theoretical foundation on the essential elements of SGs, as described below [56].

The narrative of a game constitutes one of the essential elements of SGs [73-76], capable of directly influencing gameplay by arousing curiosity, fostering motivation, stimulating engagement, and promoting more learning for players [74]. This element represents how the game's story is told and constructed through rule definition, message construction, inclusion of dialogues, and application of dynamics [73]. Following the recommendations of Janssen and collaborators [76] for narrative construction, attention was focused on the plot, characters, space, and time. The SGs share the same characters and time. The protagonists, Cien and Tista, are recurring in PEEC resources, especially in videos that contextualize all activities. Regarding time, all games occur during summer vacation, emphasizing a theme close and realistic to children's daily lives [75]. The narrative elements that distinguish these SGs are the settings and the problems the characters face throughout their journey (plot). SG 1 takes place in a campsite, SG 2 occurs on the grandparents' farm, SG 3 is set in a maritime context, and SG 4 happens on a motorhome trip. This type of narrative is characterized by the principle of "Narrative is everything" [73], where gameplay revolves around the characters' adventures and serves as a pretext for the progression of the SGs. The entire narrative is intrinsically related to the themes suggested in the PEEC curricular proposal for each grade level [53].

The feedback system in SGs stimulates player motivation, promoting and regulating their learning [75, 77, 78]. This consists of the information provided after the player's response. According to Johnson, Bailey, and Buskirk [78], feedback is characterized by type, timing, modality, and individual differences. Regarding type, it combines "corrective feedback" and "explanatory feedback." On the one hand, it indicates whether the player chose the correct, partially correct, or incorrect decision; on the other hand, it provides a brief explanation of why it is correct or not. "Explanatory feedback" can provide more learning than "corrective feedback," although it does not influence the player's motivation and enjoyment [77]. Timing-wise, it is immediate, appearing right after the student's response is validated. Regarding modality, it is presented in written form with visual support. Finally, concerning individual differences, only the age corresponding to the grade level is considered, with feedback for 6-7-year-olds having less text and shorter sentences, unlike games for 8-9-year-olds.

Some studies suggest that rewards enhance the player's experience [79]. In the SGs design, two types of "positive reinforcement" rewards were included [79]: those that interfere with gameplay (resources) and those that do not influence the narrative (scores) [80], where the player is rewarded whenever they get the answer right. Incorporating scores into SGs is also a way to ensure the creation of an exciting game [81]. On the other hand, as recommended, players are not penalized for incorrect answers [75].

One of the essential and valued aspects of SGs is the tracking and recording of player choices and paths, allowing the analysis and monitoring of learning [45, 75, 81]. Therefore, the possibility for the teacher to receive a detailed report of each player in their email account was incorporated.

The diversity of strategies and dynamics is also a relevant factor in SG development [75]. The PEEC SGs are characterized as "standard" and "point & click" [73]. Considering the software options used, "drag and drop," "hotspot," "matching," "numeric," "multiple choice," and "sequence" dynamics were incorporated. According to several authors, the diversity of dynamics is relevant to promoting player interest in the SG [67, 75], as it eliminates the predictability of subsequent dynamic types. Including audio for the instructions was deemed pertinent, anticipating possible reading difficulties that could affect gameplay. This aspect has been noted as highly relevant [55, 82].

Features from other SGs, such as the Two-tier Test [83], served as inspiration. Practically, the player is challenged to answer a question (first level). If they answer correctly, they collect the reward and move on to the second-level question. If the player does not answer correctly, they receive constructive feedback and can respond to another question. For SGs with learning assessment effects, it is relevant to include this dynamic to ensure formative assessment questions, stimulating more learning during assessment moments.

For each game, a set of learnings to be assessed was defined, facilitating its inclusion in the theme. An effort was made to organize and select learnings from Biological Sciences, Physical Sciences, and Earth Sciences, encompassing knowledge, skills, attitudes, and values.

After defining these game elements and the learnings to be assessed, the conceptualization and validation phase focused on developing the concept, mechanics, and narrative of the SGs. This step is usually carried out through a flowchart (Figure 3) to provide an overall view of the SGs' mechanics before their construction [55], allowing for the identification of potential usability problems, ensuring diverse dynamics and a logical and cohesive narrative.

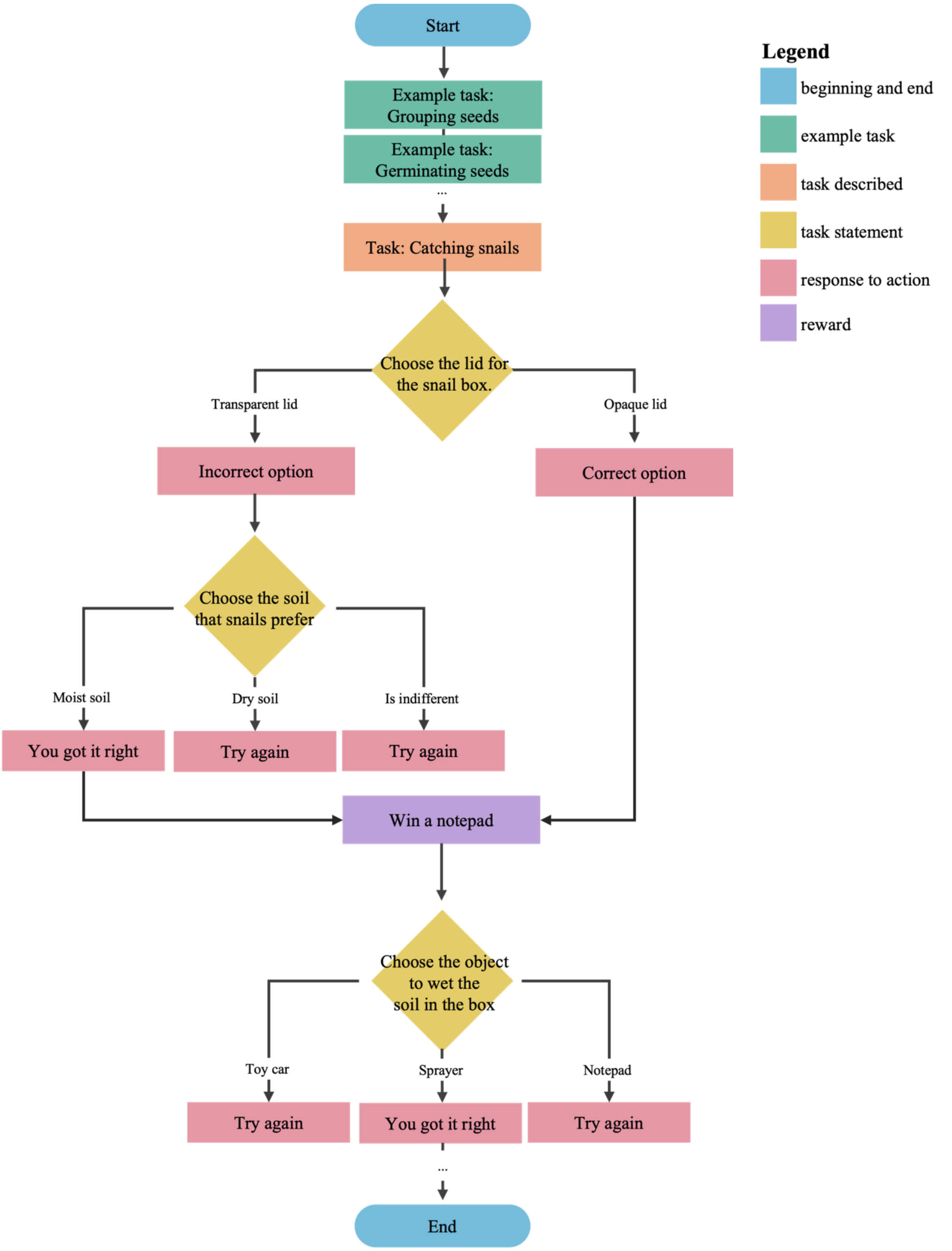

Figure 3. Initial gameplay flowchart

2.2.2 Implementation Phase of the SGs

In this study, as shown in Table 2, 245 children participated in the first implementation cycle and 158 in the second. The first cycle served as a pilot project, and data were only collected from the second cycle (2021/2022).

Table 2. Number of children participating in the study

Grade Level	Number of Children by Implementation Cycle	
	1st Cycle (2020/2021)	2nd Cycle (2021/2022)
1st Year	60	21
2nd Year	63	65
3rd Year	54	37
4th Year	68	35
Total	**245**	**158**

Each grade level had a supporting teacher responsible for implementation. In the second implementation cycle, data collection regarding the time taken to complete the SGs was recorded by the game software, while children’s emotional state and preferences between tests and SGs as assessment tools were collected through a questionnaire integrated into the game (Figure 4).

Figure 4. Screens from the game: final questions on children's satisfaction

In the days preceding the sessions, each supporting teacher tested the game corresponding to their grade level, providing feedback on minor errors, incorrect links, etc. This feedback allowed for final adjustments before implementation. Figure 5 illustrates feedback received regarding error detection in one of the SGs. For example, in the first conversation, the teacher identifies a coherence error where the game initially refers to a water gun as hydraulic and later as mechanical. In the second conversation, the teacher points out a technical error where one of the scenes displays a white screen. In the final conversation, the teacher suggests which game situations should be included.

.

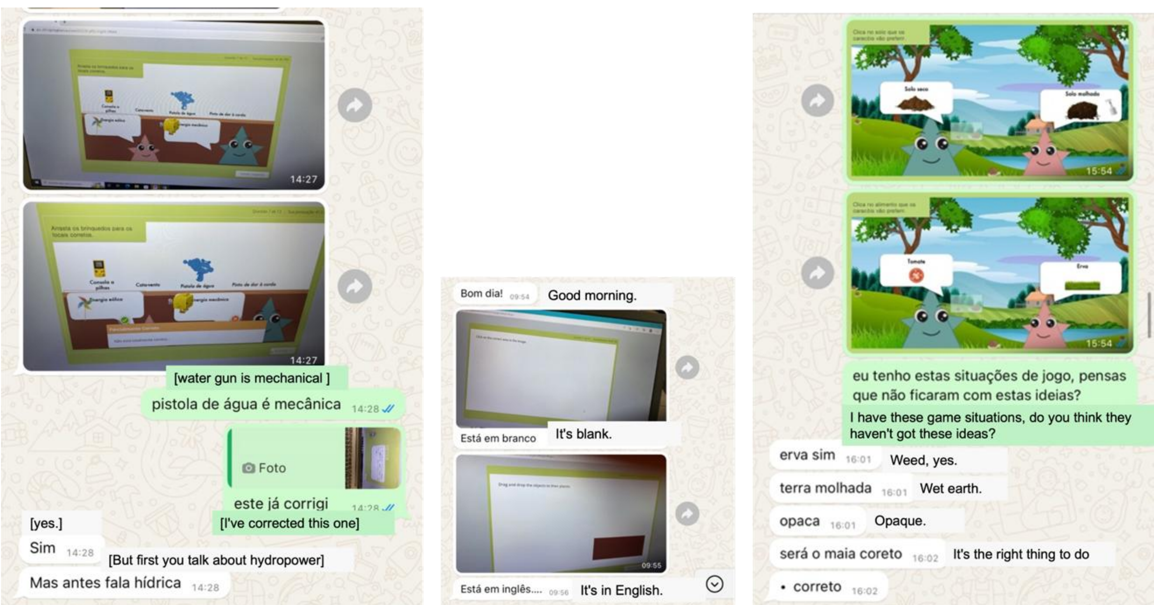

Figure 5. Feedback message from a teacher highlighting an error in one of the SGs and providing suggestions for correction.

At the end of two academic years (2020/2021 and 2021/2022), following the weekly implementation of the science activities from the PEEC, participating children were invited to play the SGs corresponding to their grade level as a means of learning assessment.

The SG sessions took place over a week in both implementation cycles. Collaborating teachers were responsible for their respective classes, guiding the sessions and addressing any questions or technical difficulties. Each teacher managed the organization of the sessions based on guidelines provided by the PEEC creator-researchers. In some cases, children worked in small groups in a computer-equipped room, while in others, sessions were conducted with the entire class. Regardless of the modality chosen by the teachers, each child had access to a mobile device, and the SGs were completed individually (Figure 6).

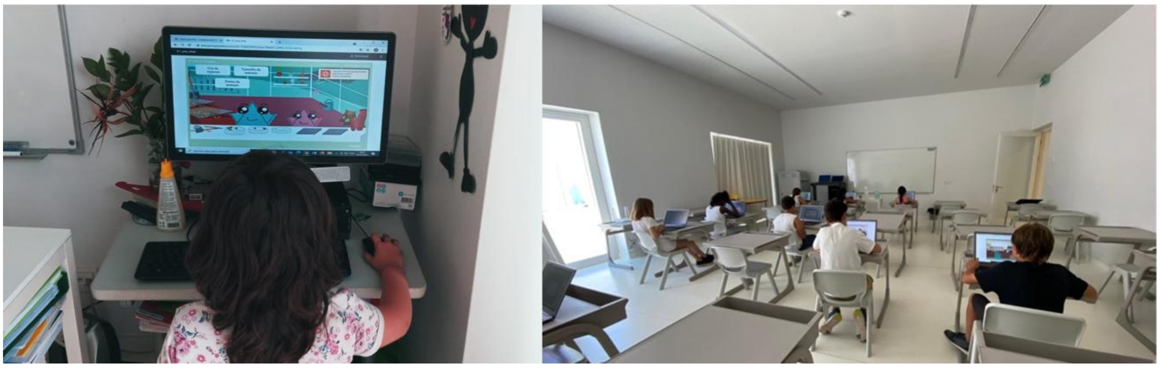

Figure 6. Implementation of SGs

Throughout the implementation process, the teachers reported their testing experiences, which allowed us to understand if the children were reacting positively to the SGs, as illustrated in Figure 7.

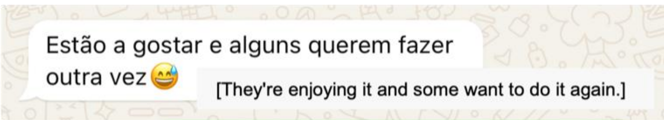

Figure 7. Message from a teacher about the implementation of SGs

3. Final Product

Four SGs were developed, one for each grade level in primary education, using an EDR approach. These games were designed to assess specific learning outcomes for each year, based on the PEEC curriculum. Given that they are intended for primary education, all games include audio to support children who have not yet mastered reading. These SGs are described below.

The player/child is invited to immerse themselves in a summer vacation narrative featuring two characters, Cien and Tista. The game designed for the 1st grade is titled *"Summer Holidays at the Campsite!"*; for the 3rd grade, *"Summer Holidays by the Sea!"*; and for the 4th grade, *"Summer Holidays on Safari!"*. The learning outcomes targeted in these games are detailed in the appendices.

Each game presents ten problem-based contexts that correspond to assessment scenarios evaluating knowledge, skills, and/or attitudes and values. To exemplify the created SGs, the situations from the 2nd-grade SG, titled *"Holidays at the Grandparents' Farm!"*, are presented. Table 3 provides an overview of the game scenarios, the tasks associated with each scenario, and the corresponding learning objectives.

Table 3. Description of the assessment game designed for 2nd-grade students (Holidays at the Grandparents' Farm!)

Game Situation	Task	Learning Objectives
A Greenhouse Confusion!	Click on the criterion used to group the seeds and drag the seeds according to the criterion.	Classify seeds based on color. Demonstrate consistency.
Extra Question	Click on the criterion corresponding to the size grouping.	Classify seeds based on size.
A Plant Challenge!	Click on the character that explains the essential conditions for a seed to germinate.	Understand that light does not influence seed germination and that water is essential for germination.
Extra Question	Click on the external factor that influences seed germination.	Recognize that the amount of water influences seed germination.
Desired Plant!	Click on the character explaining why pampas grass should not be taken home.	Understand that pampas grass is an invasive species in Portugal.
Extra Question	Click on the native plant species in Portugal.	Identify *Digitalis* as a native plant species in Portugal.
Snail, Snail, What Do You Like?	Drag the most appropriate lid [transparent or opaque] based on the snails' preferences.	Recognize that snails prefer dark environments in their habitat.
Extra Question	Click on the option indicating snails' preference regarding soil.	Understand that snails prefer moist environments.
Exploring Snails!	Click the button indicating the error in the snail anatomy drawing.	Identify that snail eyes are not located on their face.
Extra Question	Click the option indicating where snail eyes are located.	Understand that snail eyes are located on their upper tentacles.
Mushroom, Can I Eat You?	Click the character explaining whether all mushrooms are edible.	Understand that not all mushrooms are edible.
Extra Question	Click the option explaining how to prevent bee stings.	Learn that bee stings can be prevented by using a repellent spray.
Outdoor Cinema!	Drag the animals, arranging them from the most to the least voracious.	Interpret and analyze information on animal diets presented on a poster.
Extra Question	Click the animal that sleeps the least per day.	Interpret and analyze information about animals' sleep patterns.
Park Fun!	Drag the character to the swing seat requiring the least force to lift the sibling.	Understand that less effort is needed when force is applied farther from the fulcrum on a lever.
Extra Question	Click the object that does not include a pulley system.	Recognize that a two-pan balance does not contain pulleys.
Playing with Ramps!	Drag the ramp that allows the car to travel the greatest distance.	Understand that the steeper the inclined plane, the greater the distance an object travels.
Extra Question	Click the option describing the function of ramps.	Understand that ramps reduce the force needed to move an object.
Magnetic Gift!	Click on the materials attracted by a magnet.	Learn that aluminum, plastic, glass, and wood are not attracted to magnets, while nickel is.
Extra Question	Click the option confirming or refuting the statement "all metals are attracted to magnets."	Understand that not all metals are attracted to magnets.

The player accesses the SG and encounters the introduction screen displaying the game's title (Figure 8).

Figure 8. Introduction screen

When the player proceeds, they are asked to provide some identification information, such as their name, grade level, and the email of the recipient who will receive the game results, which, in a school context, is preferably the teacher's email. The gameplay experience begins with a progress map that situates the player but reveals little about the upcoming game situations, maintaining mystery and interest. The player then moves forward and is invited to solve various challenges.

In this case (Figure 9), the aim is to assess the ability to classify seeds through the situation "Greenhouse Confusion," which depicts the disorganization of seeds in the sisters' grandparents' greenhouse.

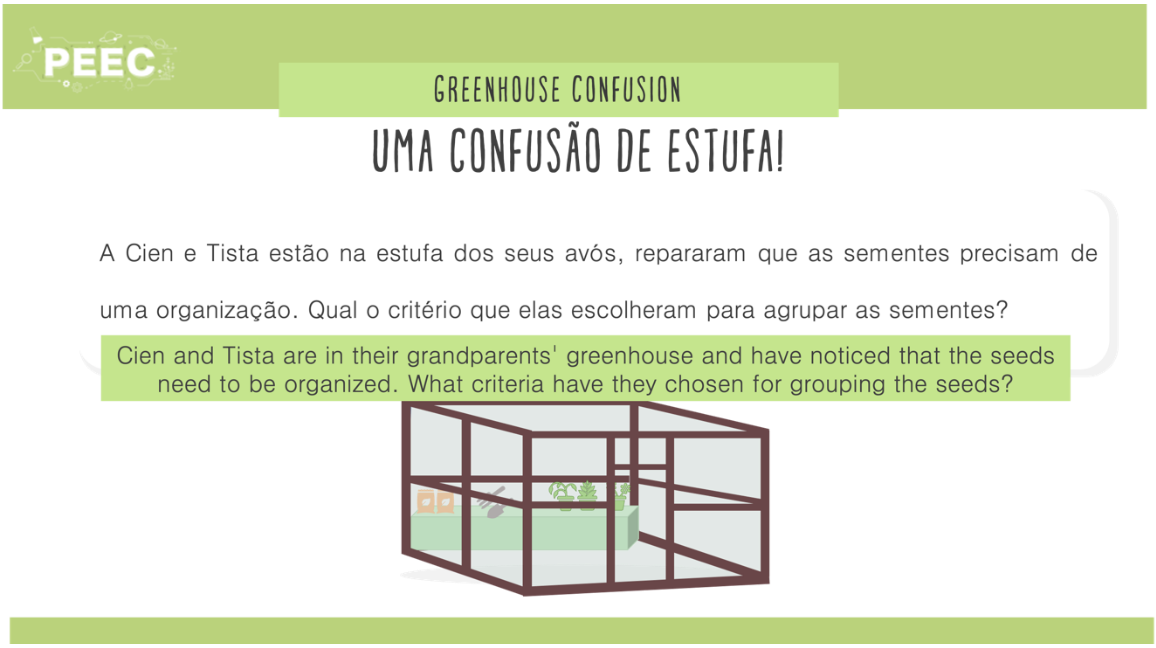

Figure 9. "Greenhouse Confusion" Scenario

The game scenario shows that the sisters have already organized the seeds and challenges the player to identify, by clicking, the criterion defined by the sisters for that grouping (Figure 10).

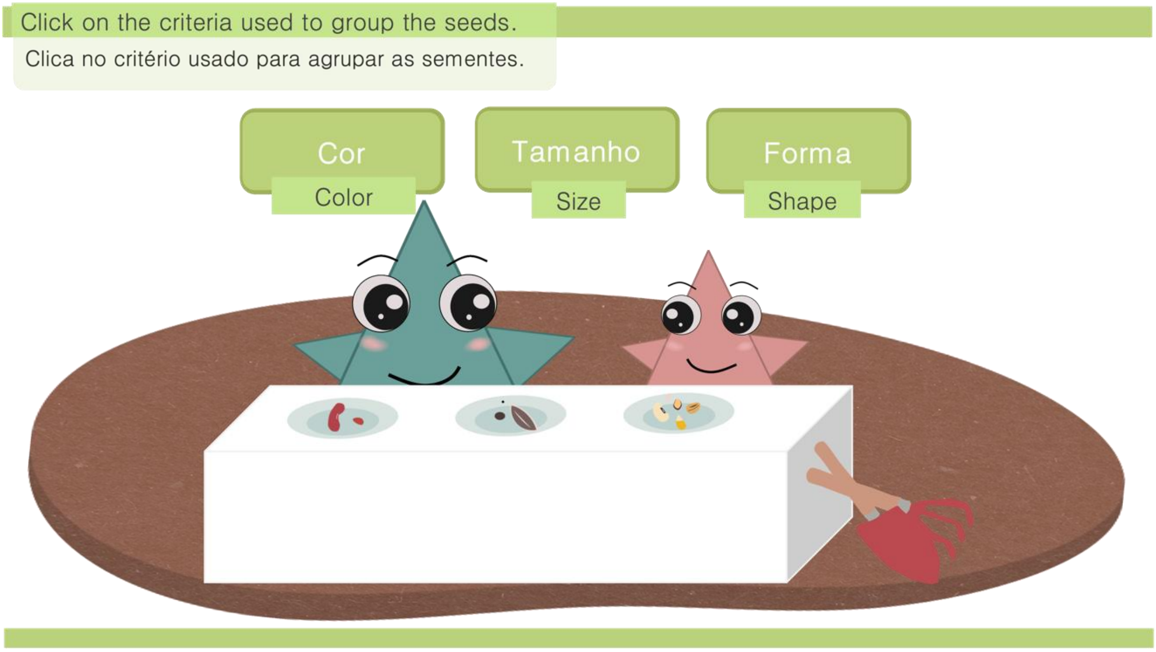

Figure 10. "Greenhouse Confusion" Game Scenario

Upon answering, feedback is provided indicating the correctness of the response, such as "Great, that's absolutely correct!", "Oh, that's not quite right," and "Oh, not this time." In addition to this feedback, informative feedback is given regardless of the quality of the response (Figure 11). If the player answers correctly, they earn 10 points.

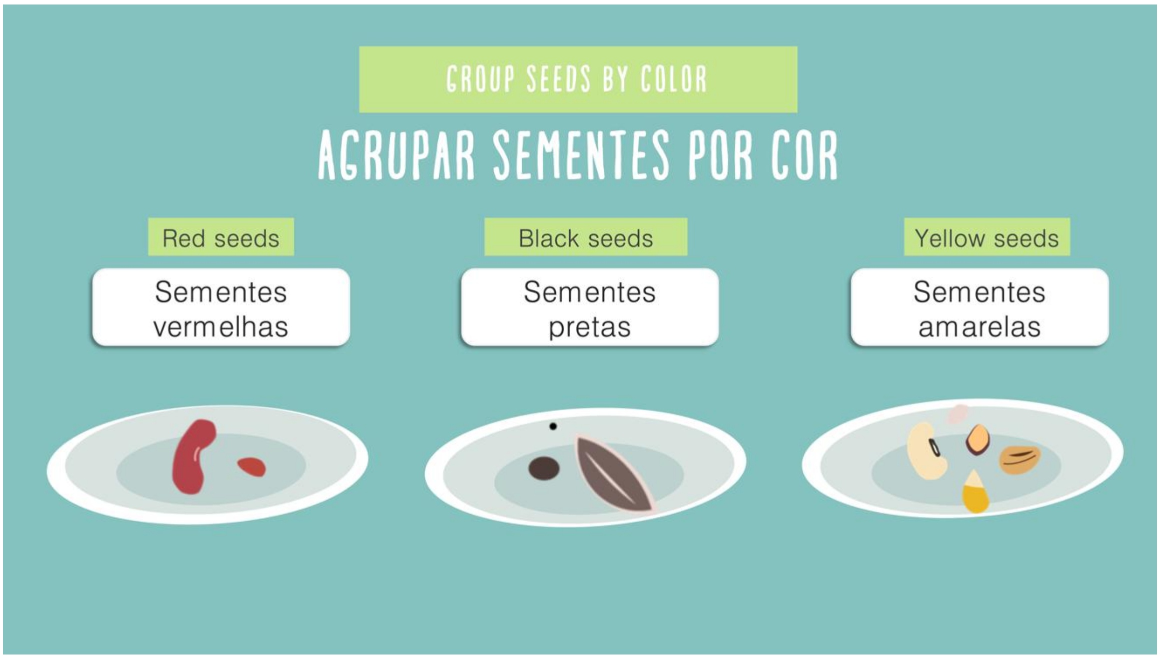

Figure 11. Informative Feedback on Seed Grouping

If the player does not answer correctly, a branched path is opened, and they are given an opportunity to respond to a quiz related to the assessment theme. In Figure 12, another possibility is shown for children to select the criterion for seed grouping. If they answer correctly, the player earns five points; otherwise, they lose two points with each incorrect attempt until they get it right.

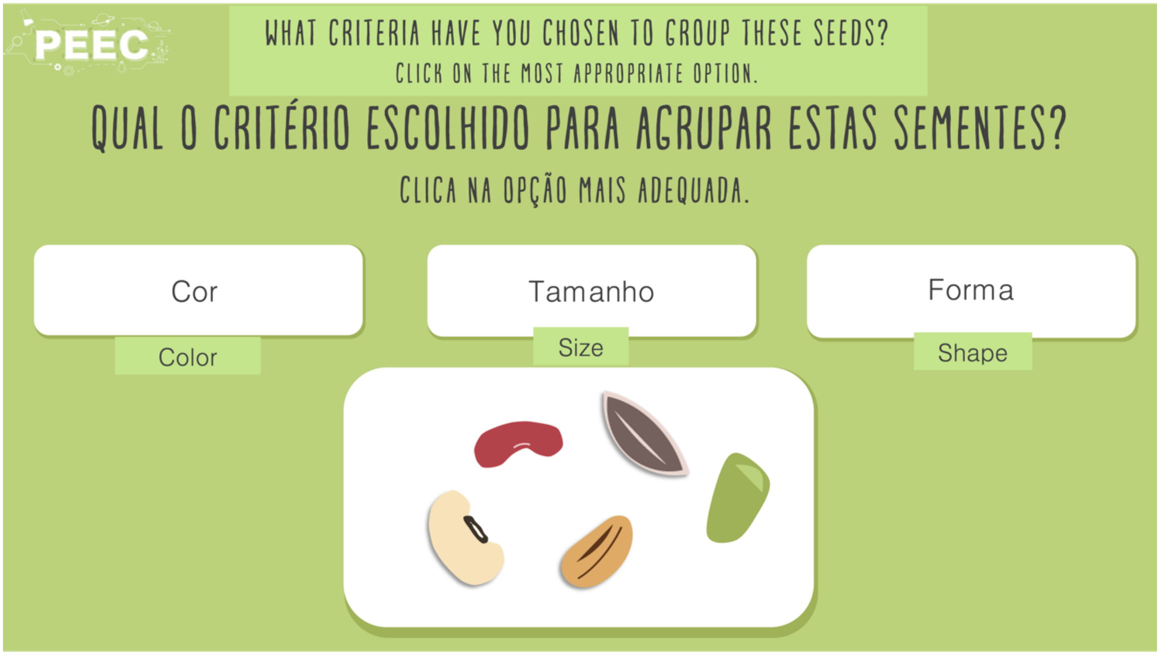

Figure 12. Game Situation: Extra Quiz “Grouping Seeds”

After correctly answering the question, in some cases, a reward is granted that will help progress to other parts of the game, as exemplified in Figure 13. The game is designed so that players earn various rewards. These artifacts will be necessary in the storyline, and the player will need them to advance in later situations.

Figure 13. Game situation: Prize and use of the prize "Vacation at Grandma's Farm"

The game concludes with a summary of the player’s journey, allowing players to revisit all their choices. The game summary is sent to the designated email provided at the start, detailing all responses, the chosen pathways, and total completion time (Figure 14).

Pergunta 1 Parcialmente Correto
Pontos: 2.5/5
Click on the correct area in the image.

Resposta do Usuário	Resposta Correta
✔ Gaivota	Gaivota
✘	Planta

Feedback: Não está totalmente correta...

Pergunta 2 Correto
Pontos: 5/5
Click on the correct area in the image.

Resposta do Usuário	Resposta Correta
✔ cenoura	cenoura

Feedback: Boa! Está correto!

Pergunta 3 Correto
Pontos: 5/5
Click on the correct area in the image.

Resposta do Usuário	Resposta Correta
✔ tampa opaca	tampa opaca

Feedback: Boa! Está correto!

Figure 14. Screenshot of the report received by email

4. Results

At the end of the experience, we gathered feedback from players about the SGs they had played. This evaluation assessed player experience, following approaches from similar studies [56, 84], and included comparative analysis with recommendations from the literature [85, 86].

To evaluate the player experience, data was collected on the time each child took to complete the SG, the emotions evoked during gameplay, and their preference between the SG and traditional assessment tests.

4.1 Player Experience and Learning Outcomes

The evaluation of the SGs focused on the player experience, measuring aspects such as completion time, emotional responses, and student preferences compared to traditional assessment methods. However, given that these SGs were designed not only as assessment tools but also as instruments for supporting formative learning, it is crucial to examine how these results relate to students' learning outcomes.

The observed positive emotional responses, particularly curiosity and engagement—suggest that the SGs fostered an affective learning environment, which is a critical factor in promoting motivation and sustained knowledge acquisition. Affective learning, characterized by students' emotional investment in the learning process, is known to enhance cognitive engagement and retention. The fact that most students expressed enjoyment and interest while playing the SGs suggests that the games successfully created a low-anxiety assessment environment, reducing the stress typically associated with traditional tests. This aligns with research indicating that when students are more engaged and emotionally invested in an activity, they are more likely to develop deeper conceptual understanding.

Moreover, the variation in completion times observed across different grade levels reinforces the notion that SGs can support differentiated learning by allowing students to progress at their own pace. Unlike rigid summative assessments, which often impose strict time constraints, the SGs provided a more flexible assessment structure, enabling students to explore and process information in a way that suited their individual learning needs. This self-paced approach is particularly beneficial for formative assessment, as it encourages reflection and iterative learning—key principles in IBSE.

The different completion times for the SGs, represented in Figure 15, demonstrate the varying needs of the children. For first graders, the minimum time to complete the SGs was approximately 5 minutes, while the maximum time was around 35 minutes, with a range of 30 minutes. The average completion time is about 20 minutes, and the median is around 21 minutes, suggesting a symmetrical distribution of the children's completion times.

For second graders, one child completed the SG in less than 3 minutes (minimum), while the child who took the longest completed it in 29 minutes, with a range of 27 minutes. The average completion time for the SGs is 15 minutes, and the median is 17 minutes. Like the first-grade results, the distribution is also symmetrical.

The range of completion times for third graders is over 45 minutes. The fastest child finished the SG in just over 6 minutes, while the one who took the longest finished in approximately 52 minutes. The average completion time is almost 18 minutes, and the median is around 16 minutes, again suggesting a symmetrical distribution, despite the evident different needs of the children in completing the SGs.

For fourth graders, the range is 24 minutes, with a minimum time slightly over 4 minutes and a maximum time of about 28 minutes. The average and median times are around 13 minutes, suggesting a symmetrical distribution. As observed in other grades, children need different amounts of time to complete the same assessment activity, respecting each one's pace.

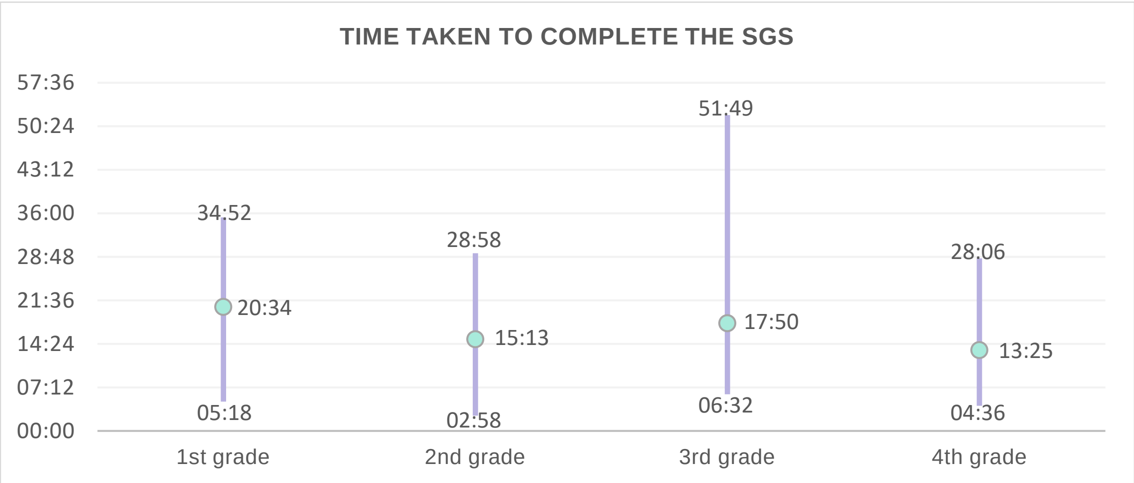

Figure 15. Minimum, maximum, and average time taken to complete the SGs by grade level

Overall, the emotional state of the children while playing the SGs was predominantly positive, as illustrated in Figure 16. The most highlighted feelings were surprise and curiosity, with frequent mentions of these two emotions during the SGs. A small percentage of children (almost 6%) reported feeling confused while playing the SGs.

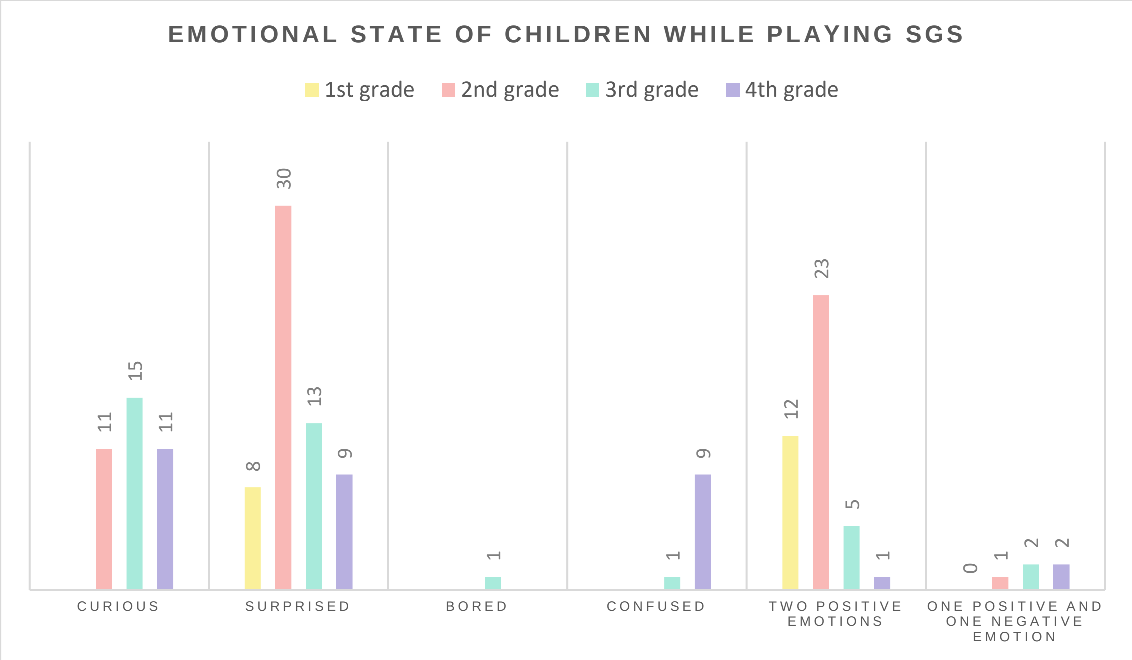

Figure 16. Emotional state of children while playing SGs by grade level

As shown in Figure 17, most children prefer digital games as an assessment activity. A small percentage indicated a preference for traditional tests. There are no significant differences across grade levels, suggesting that participating children generally prefer SGs over traditional assessment tests.

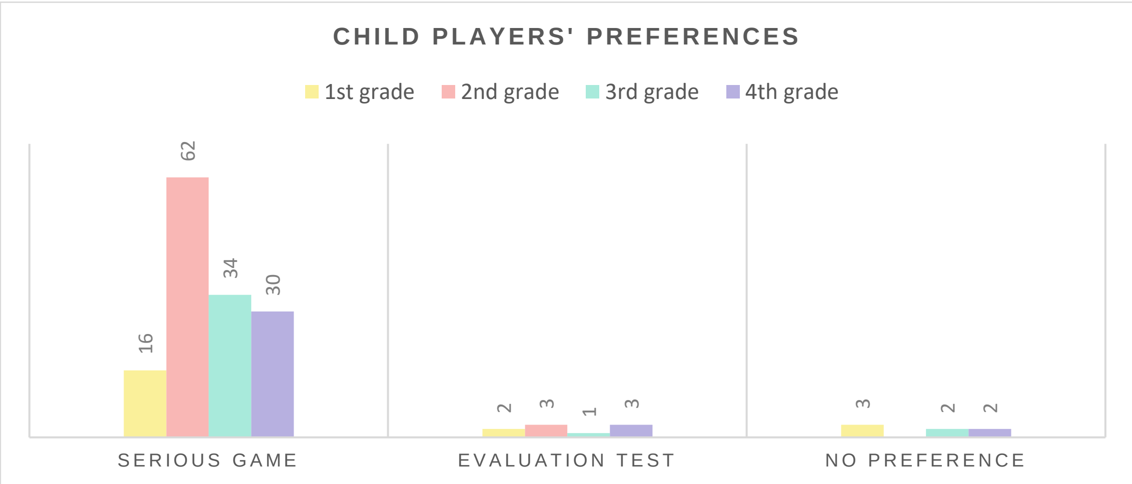

Figure 17. Children's preference by grade level

4.2 Evaluation of the Created SGs According to Literature Recommendations

Caserman and colleagues [87] propose a quality criterion framework for attractive and effective SGs, noting that quality assessment systems for SGs often lack a balance between the "serious" and "game" components. Thus, the SGs presented in this study were evaluated according to this quality framework, a process conducted by the game developers, who are researchers in the fields of science education and digital media.

The first criterion included the central elements for the serious part of the game, namely the existence of a characterization objective, the development of appropriate methods to achieve this objective, and the evaluation of quality (Table 4).

Table 4. Evaluation of SGs According to Quality Criteria for the Serious Part

	Quality criteria for the serious part by Caserman and colleagues	Applicability in PEEC Serious Games
CHARACTERIZING GOAL	Focus on the characterizing goal	
	Learning/training goal must remain in focus, for which a combination of physical and cognitive training can be beneficial	Not applicable. The SGs in question don't involve any physical activity on the part of the children, only cognitive activity.
	Support players to achieve the characterizing goal	The children are given some clues, by way of feedback, to help them fulfil the SGs in more difficult situations.
	Game elements should not interfere with the learning/training process	There is a balance between the elements of the game, allowing the educational focus not to be affected.
	Clear goals	
	Appropriate methods for the specific application area and target group	The statements are short and to the point, as audio support to overcome reading and interpretation difficulties.
	Goals are clear and appropriate so that players can work towards the characterizing goal	The statements are direct, always present and identify the challenge set to the children.
	Indispensability of the characterizing goal	
	Serious part must be mandatory	These are SGs intended to be used as a learning assessment activity in science.
	Characterizing goal must not be avoidable	They are presented as SGs for learning assessment purposes.
	Training and learning tasks should not be a hurdle	All the tasks don't materialise in one obstacle, so even if the student-player doesn't have the skills to play, they can move on to the next scenes.
METHODS	Correctness of the domain expert content	
	Avoid errors and ensure that the content is technically correct	The content of the SGs has been validated, as have all the illustrations used.
	Ensure correct technical language	The SGs were designed by three researchers and revised by ten primary school teachers.
	Remain neutral, especially on political and social issues	The content of the SGs is in the area of natural sciences and is scientifically validated and appropriate for all children in Portugal.
	Appropriate feedback on progress	
	Players should receive feedback on their performance and progress	A pop-up appears whenever the student-player validates their answer and the progress map allows the student-player to be situated.
	Visible and recognizable effects	The feedback given is represented by a color code and a message depending on the quality of the student-player's response.
	Provide simultaneous feedback (eg, visual, audio, haptic; multimodal feedback)	The feedback appears with a message that immediately indicates the quality of the answer and always provides an informative infographic relating to the learning being assessed.
	Appropriate rewards	
	Provide positive reinforcement and in-game awards	The SGs created feature scores and prizes whenever the student-player gets the game right.
QUALITY	Proof of effectiveness & sustainable effects	
	Prove that the characterizing goal is achieved	The effectiveness of SGs as an assessment activity was tested over two school years with 12 classes and their teachers.
	Learning/training effects need to be sustainable	The study involved 158 children and made it possible to ascertain the emotional state of the student-players, as well as their preferences.
	Awards and ratings	
	Game awards, professional and user ratings, recommendations by domain experts, game reviews, and number of players/downloads state the quality of the game	Not applicable. The SGs in question have educational characteristics for evaluation purposes, the purpose is not to achieve ratings of this nature.

In Table 5, the elements for appropriate game design and suitable interaction technology are presented.

Table 5. Evaluation of SGs According to Quality Criteria for the Game Part

Quality criteria for the game part by Caserman and colleagues		Applicability in PEEC Serious Games
ENJOYMENT	**Ensure player engagement and experience**	
	Ensure positive experience during playing	The participating children reported positive emotions about playing the SGs.
	Serious games should be engaging and enjoyable	The SGs created are colorful, with situations that are representative of children's daily lives.
	Provide an engaging experience for different player types	The situations created are real and/or realistic and each child can feel identified and involved in the narrative.
	Ensure flow	
	Keep a balance between a player's skills and challenge	The SGs present a set of problem situations that assess knowledge, skills, attitudes and values.
	Dynamically adapt the difficulty level depending on the current player's performance in the game	Not fully applicable. As the SGs are for assessing learning, a personalised path has been included when the child has not yet been able to answer correctly.
	Adapt to players to increase effectiveness (eg, motivate them to repeat the exercises continuously and regularly)	The SGs are available online and free to play whenever the children want.
	Increase complexity as the player gets better	Not applicable. Complexity in this sense is relative and the aim is to assess children's science learning.
	Provide varied gameplay	Different drag, select and click dynamics have been included.
	Establish an emotional connection	
	Allow emotions and arouse instinct	Rewards were included to keep the student-player interested and motivated in the SGs.
	Sense of control	
	Players should have control over their actions in the game	The student-player decides the actions in the SGs.
	Support social interactions	
	Provide different game modes (collaborative and competitive settings for players that perform better in groups)	Not applicable. These are individual SGs.
	Ensure immersive experience	
	Include multimodal sensory stimulations: visual, audio, haptics, smell	Not applicable.
	Ensure the sense of "being there"	Realistic settings and situations, the holiday contexts are reminiscent of the children's reality, which allows them to be immersed.
MEDIA PRESENTATION	**Attractive graphics**	
	Graphics must be appropriate for the game purpose, application area, and target group	Simple illustrations, colorful and rigorous in their representation to ensure an attractive interface for primary school children.
	Ensure clear interface without unnecessary information to not distract players from a specific task	All the elements we have included in the SGs are essentially fundamental to their playability.
	Appropriate sound	
	Include appropriate background music and sound effects	Not applicable. The SGs do not feature music because it is an evaluation activity.

Table 6 systematizes the criteria defined for the serious part and the game part, which must be strongly integrated and connected.

Table 6. Evaluation of SGs According to Quality criteria for balance between serious and game part

Quality criteria for the balance between serious and game part by Caserman and colleagues		Applicability in PEEC Serious Games
INTEGRATED SERIOUS PART WITH GAMEPLAY	Embedding characterizing goal into the gameplay	
	Integrate the characterizing goal into the gameplay	A simple, engaging and realistic narrative was defined.
	Learning/training tasks must be related to the game and should be connected to the game elements	The narrative commands the SGs and the tasks are intrinsically linked to it.
	Scientific foundation	
	Include interdisciplinary teams; game designer and domain experts should work together (also together with the target group)	The EDR method expects the involvement of several participants and in this case the team involved teachers, children, biologists, multimedia specialists and science didactics.
	Include state of the art in the relevant disciplines	The SGs are theoretically grounded, both in the field of science education and in the fundamental principles of SGs and learning assessment.
INTERACTION TECHNOLOGY	Appropriate interaction technology	
	Interaction technology must be suitable for the target group (ie, their physical and mental ability and game purpose)	The SGs have simple interactions, dynamics and mechanics suitable for the age level they are aimed at.
	Intuitive game mechanics and natural mapping	
	Provide tutorials for complex games; otherwise, players should discover the game mechanics themselves	Not applicable, the SGs have simple mechanics.
	Intuitive use of game controls (eg, the WASD keys to move and space bar to jump)	Simple click-and-drag tasks included.
	Enable natural mapping between technology and gameplay	SGs are intuitive in the tasks they include.
	No simplifying of the learning and/or training process due to technical features	
	Interaction technology must support players in achieving the characterizing goal	You don't need to be an advanced PC user to carry out the SGs' tasks.
	Ensure accurate tracking to prevent cheating in exergames	Not applicable to this type of SG.
	Avoid adverse effects	
	Low risk of accidents, injuries, or overload	The SGs are totally secure from that point of view.
	Avoid technical issues and ensure easy maintenance	The software used makes it possible, through encrypted code, to host the game on the website without any dependence on its subsistence, guaranteeing its sustainability.

5. Discussion

This study set out to answer the research question: *How can serious games be developed for the assessment of and for primary school children's learning, consistent with the proposed experimental science teaching project?* The findings demonstrate that SGs can serve as both assessment tools and learning facilitators when aligned with IBSE principles. The iterative EDR method used in this study enabled the development of SGs that not only evaluate scientific knowledge but also foster deeper learning through interactive and formative assessment mechanisms.

One of the key contributions of this research is the integration of formative assessment within a digital game-based environment. Unlike traditional tests, which often emphasize recall and classification, the SGs presented here promote assessment *for* learning by incorporating real-time feedback, interactive problem-solving, and a low-anxiety assessment context. The positive emotional responses reported by students—curiosity, engagement, and enjoyment—suggest that these SGs create a more motivating learning experience. This aligns with prior research indicating that affective learning plays a crucial role in cognitive engagement, concept retention, and overall academic achievement [55].

Moreover, the variation in completion times highlights an important advantage of digital game-based assessment: adaptability. While conventional tests impose uniform time constraints, SGs allow students to navigate assessment activities at their own pace, supporting differentiated learning. The ability of SGs to provide instant feedback further enhances their formative function, enabling students to adjust their understanding in real time—a key feature of effective learning-centered assessment [46, 48].

Teacher feedback further supports the claim that SGs contribute to learning. Informal observations indicated that students who engaged with SG-based assessment demonstrated improved recall and application of scientific concepts in subsequent class discussions. Additionally, the ability to track student decision-making within the SGs offers a more nuanced view of learning progression, allowing educators to identify patterns, misconceptions, and individual learning needs [45, 75, 81].

These findings reinforce the growing recognition that assessment should not merely be an endpoint in the learning process but rather an integral part of ongoing knowledge construction. The SGs developed in this study exemplify how assessment can be seamlessly embedded within engaging and interactive learning experiences, offering an alternative to traditional evaluation methods that often fail to capture the complexity of scientific understanding.

The EDR method has been used in various empirical studies proposing digital educational solutions [55, 56, 72, 88]. It is confirmed that this approach provides ideal conditions for the design of the assessment SGs produced. As in other studies, key characteristics for the success of this method are highlighted, including:

- Flexibility: It allowed for necessary adjustments throughout different phases [72], based on validations and evaluations by various stakeholders involved in the research during the two implementation cycles conducted.
- Iterative Cycles: In addition to the mentioned flexibility, the iterative nature of this method allowed for the gradual, conscious, and evidence-based design of SGs, supported by cycles of design, implementation, and validation-evaluation. This aspect is also noted and valued in some studies [56, 67]. This evaluation-validation, combined with implementation, enables systematic refinement of the SGs, with summative evaluation only occurring at the end of these final SGs. In this case, the first cycle involved developing a high-fidelity prototype of the SGs, aimed at testing their feasibility, appropriateness, and viability as an activity for assessing science learning for primary school children. Upon confirming their practicability with the target audience, the second implementation cycle assessed the satisfaction of the child-players through a final questionnaire.
- Multiple Perspectives: The critical and participatory input from those involved, from the design of the SGs to their final redesign, facilitated systematic and productive communication among stakeholders for necessary improvements and quick adjustments. Early involvement of the team responsible for both science education and multimedia was essential for identifying gaps and ensuring potential from the initial draft [67]. For instance, the study by Hart et al. [90] lists several changes based on feedback from various stakeholders during the SG design phase, demonstrating the importance of multiple perspectives for advancing the final development. In this particular study, the input from collaborating teachers enabled the identification of typographical errors, correction of response options, and detection of errors in the branching pathways. Beyond the feedback from the multidisciplinary team, external perspectives were also incorporated. This was achieved through the presentation of these SGs at a conference, where they were

subject to public analysis by specialists in science education. Their relevant feedback and significant reflections contributed to the final redesign of the games.

Regarding SGs as a research product, their evaluation followed approaches similar to those employed in other studies, focusing on the assessment of the player-child's experience [56, 84] and comparative analysis against recommendations in the literature [85, 86]. The collection of children's perceptions through the questionnaire embedded within the SGs allowed for the identification of their emotional state induced by gameplay, as well as their preference between SGs and traditional assessment tests. Consistent with findings in other studies [55], positive emotions related to gameplay and a preference for SGs over traditional tests were observed [21]. Through the two implementation cycles described in this article, situated within Phase II of the EDR framework, it can be affirmed that SGs serve as effective tools for assessing science learning. The main challenge in designing assessment SGs lies in achieving their educational purpose while simultaneously ensuring their playful nature [91]. To address this, the criteria proposed by Caserman and colleagues [87] were employed to evaluate the attractiveness and effectiveness of the SGs, considering the "serious part," the "game part," and the "serious-game part." As presented in the results section, the SGs largely satisfy these quality criteria. This framework has been highlighted in various studies: some have integrated it into their evaluations [92–95], others have expressed the intention to adopt these criteria for improved assessments in the future [96], while some have identified the lack of its use as a limitation for broader evaluation of their SGs [97]. Given the importance highlighted in studies on SG development, incorporating this component into the present study was deemed essential to evaluate the SGs across these three dimensions.

6. Limitations

This study contributes to the evolving discourse on digital game-based learning by demonstrating how SGs can function as innovative tools for science assessment in primary education. By leveraging the principles of IBSE and formative assessment, the developed SGs provide an engaging, interactive, and adaptive means of evaluating scientific knowledge, skills, and attitudes.

While the study highlights the promise of SGs as assessment tools, future research should further investigate their impact on long-term learning outcomes. Pre- and post-assessment comparisons, longitudinal studies, and expanded teacher feedback would provide additional insights into how SGs influence conceptual change and skill development over time.

Ultimately, this study affirms that digital game-based assessment, when thoughtfully designed, can transform how science learning is evaluated in primary education. By shifting from rigid, anxiety-inducing tests to interactive, formative, and student-centered assessment activities, SGs pave the way for a more meaningful and effective approach to evaluating and supporting science learning.

Similar to the results, the limitations of the study also reflect aspects of the EDR method approach and the process of developing SGs.

The hesitations in adopting the EDR approach, as outlined by Lehtonen [72], were addressed in this study. As in other projects, it is not always feasible to assemble a multidisciplinary team, and the lead researcher often has to assume more roles than anticipated [72]. When pursuing innovative digital technological solutions within an EDR approach, the lack of specialists can hinder the creation of stable prototypes [72]. Despite the limitations of the software, this study resulted in four original SGs for learning assessment, validated and effective in meeting both educational and playful objectives.

The inability to include a programmer in the SG development team, as achieved in other studies [55, 67], also emerged as a limitation. One way to address a limited budget in SG development is to use purpose-built software; however, this decision introduces a different set of limitations. The use of the Ispring Quiz Maker software is no exception. Although other studies describe it as “an excellent instrument for assessing pragmatic courses” [69, p. 94], certain shortcomings of the software were identified. Notably, it lacks features such as the integration of an interactive informational menu and a customizable progress bar. These limitations may stem from ambitions to use the software beyond its intended purposes—creating interactive quizzes rather than SGs. Consequently, the resulting games have a minimalist interface and offer limited options to the player, similar to the constraints observed in the game SimSustentabilidade [98].

Future research should explore adapting these SGs for mobile devices, a goal shared by other studies [99]. Additionally, the evaluation of these SGs should include other groups of children and teachers [55, 100, 101], addressing different aspects such as SG mechanics, technical challenges, and design elements.

7. Conclusions

The development of the SGs presented and described in this article aims to contribute to science assessment practices and learning in primary education in Portugal. To achieve this, an EDR method approach was adopted, which provided ideal characteristics for the design and implementation of the SGs, resulting in four tested and validated assessment games.

The SGs were:

- Designed according to theoretical assumptions about SGs and principles of science learning assessment;
- Categorized according to a SG framework;
- Evaluated based on quality criteria, ensuring a balance between educational and playful components;
- Tested and appreciated by the target audience.

From this validation, it was found that the SGs are suitable, evoke positive emotions in children during gameplay, respect each child's individual pace, and promote additional learning. The article discussed multiple opportunities for using SGs. This study confirmed that the presented SGs are feasible and appropriate as an innovative assessment activity for teaching science in primary education, aligning with the objectives of IBSE and STS, fulfilling both educational and playful goals.

Acknowledgments

Individuals' acknowledgements and funding sources. This work is financially supported by Fundação para a Ciência e Tecnologia I.P. with a PhD grant (SFRH/BD/143370/2019), through the European Social Fund and Human Capital Operational Program. The study was approved by the Comissão Permanente para os Assuntos de Investigação (CPAI) of Universidade de Aveiro (protocol code 11-CED/2021, 7 July 2021).

Conflicts of interest

The authors declare that there are no conflicts of interest.

References

[1] D. H. Jonassen, *Computadores, Ferramentas Cognitivas: Desenvolver o pensamento crítico nas escolas*. Porto: Porto Editora, 2007.
[2] D. Quinquer, "Evaluación como ayuda al aprendizaje," in *Modelos y enfoques sobre la evaluación: el modelo comunicativo*, Barcelona: Editorial GRAÓ, 2000, pp. 13–19.
[3] M. S. F. Chuieire, "Concepções sobre a Avaliação Escolar," *Estudos em Avaliação Educacional*, vol. 19, no. 39, pp. 49–64, Apr. 2008, doi: 10.18222/eae193920082469.
[4] S. Cristelo, "Avaliação das aprendizagens - um contributo para a melhoria das práticas de avaliação na sala de aula," *Psicologia.pt - O Portal dos Psicólogos.*, 2011. [Online]. Available: https://www.psicologia.pt/artigos/ver_artigo.php?codigo=A0605
[5] D. Fernandes, *Avaliação das aprendizagens: Uma agenda, muitos desafios*. Cacém: Texto Editores, 2004. [Online]. Available: http://hdl.handle.net/10451/5509
[6] W. Harlen, *Assessment & Inquiry-Based Science Education: Issues in Policy and Practice*, Derek Bell, Jens Dolin, Pierre Léna, Shelley Peers, Xavier Person, Patricia Rowell, Edith Saltiel. Trieste: Global Network of Science Academies & Science Education Programme, 2013.
[7] J. Lopes and H. S. Silva, *50 técnicas de avaliação formativa*. Lisboa: Lidel, 2020.
[8] W. Harlen, *Assessment, Standards and Quality of Learning in Primary Education*. York: Cambridge Primary Review Trust, 2014.
[9] G. T. L. Brown, "Is Assessment for Learning Really Assessment?," *Front. Educ.*, vol. 4, Jun. 2019, doi: 10.3389/feduc.2019.00064.
[10] R. P. de Ramos, V. A. Bagio, R. P. de Ramos, and V. A. Bagio, "Mapas conceituais no ensino de ciências: Uma estratégia potencialmente significativa para o processo didático," *Revista Exitus*, vol. 10, no. 1, Nov. 2020, doi: 10.24065/2237-9460.2020v10n1id1282.
[11] M. A. R. Primo, "El uso de mapas conceptuales como instrumento de evaluación del aprovechamiento en ciencias: lo que sabemos hasta ahora.," *Revista Electrónica de Investigación Educativa*, vol. 2, no. 1, pp. 29–52, 2000, [Online]. Available: https://redie.uabc.mx/redie/article/view/16
[12] T. Stoddart, R. Abrams, E. Gasper, and D. Canaday, "Concept maps as assessment in science inquiry learning - a report of methodology," *International Journal of Science Education*, vol. 22, no. 12, pp. 1221–1246, Jul. 2000, doi: 10.1080/095006900750036235.
[13] F. Choudhary and R. Bano, "Concept maps as an effective formative assessment tool in biology at secondary level," *Journal of Education and Educational Development*, vol. 9, no. 1, pp. 157–175, Jun. 2022, doi: 10.22555/joeed.v9i1.454.
[14] I. Nugraha, "The use of drawing as an alternative assessment tool in biology teaching," *J. Phys.: Conf. Ser.*, vol. 1013, no. 1, p. 012016, 2018, doi: 10.1088/1742-6596/1013/1/012016.
[15] M. C. de O. Reis, L. Jaguszevski, and L. de Freitas, "A utilização de desenhos como instrumento de avaliação de conhecimentos de ciências: o sistema solar," *Research, Society and Development*, vol. 9, no. 9, pp. e11996311–e11996311, Aug. 2020, doi: 10.33448/rsd-v9i9.6311.
[16] L. Horne, A. Manzanares, N. Atalan-Helicke, S. Vincent, S. W. Anderson, and C. L. Romulo, "An exploratory study of drawings as a tool to evaluate student understanding of the Food-Energy-Water (FEW) Nexus," *J Environ Stud Sci*, May 2024, doi: 10.1007/s13412-024-00929-x.
[17] C. Chin and L.-Y. Teou, "Formative assessment: Using concept cartoon, pupils' drawings, and group discussions to tackle children's ideas about biological inheritance," *Journal of Biological Education*, vol. 44, no. 3, pp. 108–115, Dec. 2010, doi: 10.1080/00219266.2010.9656206.
[18] H. B. Lobo and J. V. de S. Aguiar, "Ensino de ciências: o mundo vivido através dos desenhos," *Revista Valore*, vol. 6, pp. 848–860, Jul. 2021, doi: 10.22408/reva602021944848-860.
[19] B. Keogh and S. Naylor, "Concept cartoons, teaching and learning in science: an evaluation," *International Journal of Science Education*, vol. 21, no. 4, pp. 431–446, Apr. 1999, doi: 10.1080/095006999290642.
[20] I. A. Kusumaningrum, Ashadi, and N. Y. Indriyanti, "Concept cartoons for diagnosing student's misconceptions in the topic of buffers," *J. Phys.: Conf. Ser.*, vol. 1022, no. 1, p. 012036,

Sep. 2018, doi: 10.1088/1742-6596/1022/1/012036.
[21] P. C. B. da Silva, “Serious games: na avaliação de e para as aprendizagens em ciências,” *masterThesis*, Universidade de Aveiro, 2017. [Online]. Available: https://ria.ua.pt/handle/10773/22826
[22] J. Correia and C. Guerra, “‘Separa que o Planeta Repara’: proposta de uma sequência didática sobre separação de resíduos sólidos,” *Indagatio Didactica*, vol. 11, no. 4, pp. 301–323, Dec. 2019, doi: 10.34624/id.v11i4.10645.
[23] M. R. Clemente, R. M. Vieira, and F. Martins, “Educação para o Desenvolvimento Sustentável no 1º Ciclo do Ensino Básico – propostas didácticas no âmbito das ciências,” *Indagatio Didactica*, vol. 2, no. 1, pp. 5–42, Sep. 2010, doi: 10.34624/id.v2i1.4575.
[24] S. Zhu, Q. Guo, and H. H. Yang, “Beyond the Traditional: A Systematic Review of Digital Game-Based Assessment for Students’ Knowledge, Skills, and Affections,” *Sustainability*, vol. 15, no. 5, p. 4693, Mar. 2023, doi: 10.3390/su15054693.
[25] S. Dikli, “Assessment at a Distance: Traditional vs. Alternative Assessments,” *Turkish Online Journal of Educational Technology - TOJET*, vol. 2, no. 3, pp. 13–19, Jul. 2003, Accessed: Oct. 27, 2024. [Online]. Available: https://eric.ed.gov/?id=EJ1101956
[26] S. Laumer, A. Eckhardt, and T. Weitzel, “Online Gaming to Find a New Job – Examining Job Seekers’ Intention to Use Serious Games as a Self-Assessment Tool,” vol. 26, no. 3, pp. 218–240, Aug. 2012, doi: 10.2307/23279202.
[27] IGEC, “Gestão do Currículo: Ensino Experimental das Ciências - Relatório 2015,” Inspeção-Geral da Educação e Ciência, Lisboa, 2017. [Online]. Available: https://www.igec.mec.pt/upload/Relatorios/GCEE_2015_relatorio.pdf
[28] IGEC, “Gestão do Currículo: Ensino Experimental das Ciências - Relatório 2016,” Inspeção-Geral da Educação e Ciência, Lisboa, I/00235/DSAG/19, 2018. [Online]. Available: https://www.igec.mec.pt/upload/Relatorios/GC_EEC_2016_RELATORIO.pdf
[29] IGEC, “Gestão do Currículo: Ensino Experimental das Ciências - Relatório 2017,” Inspeção-Geral da Educação e Ciência, Lisboa, https://www.igec.mec.pt/upload/Relatorios/GC_EEC_2017_RELATORIO.pdf, 2019. [Online]. Available: https://www.igec.mec.pt/upload/Relatorios/GC_EEC_2017_RELATORIO.pdf
[30] IGEC, “Gestão do Currículo: Ensino Experimental das Ciências - Relatório 2018,” Inspeção-Geral da Educação e Ciência, Lisboa, I/01004/DSAG/22, 2020. [Online]. Available: https://www.igec.mec.pt/upload/Relatorios/GCEEC2018.pdf
[31] H. Oliveira and J. Bonito, “Practical work in science education: a systematic literature review,” *Front. Educ.*, vol. 8, Art. no. 1151641, May 2023, doi: 10.3389/feduc.2023.1151641.
[32] V. Verma, T. Baron, A. Bansal, and A. Amresh, “Emerging Practices in Game-Based Assessment,” in *Game-Based Assessment Revisited*, D. Ifenthaler and Y. J. Kim, Eds., Cham: Springer International Publishing, 2019, pp. 327–346. doi: 10.1007/978-3-030-15569-8_16.
[33] A. Burmich, O. Mashkin, and N. Stepanova, “The Role of Game-Based Assessment for the Enhancement of Learning,” in *The World of Games: Technologies for Experimenting, Thinking, Learning*, D. Bylieva and A. Nordmann, Eds., Cham: Springer Nature Switzerland, 2023, pp. 356–369. doi: 10.1007/978-3-031-48020-1_27.
[34] F. Danielli, K. T. D. de Freitas, and F. L. Cardoso, “Modelo de base teórica para criação de jogo digital pensado para o desenvolvimento infantil,” *Rev. Ibe. Est. Ed.*, vol. 18, no. 00, p. e023048, Aug. 2023, doi: 10.21723/riaee.v18i00.17204.
[35] N. N. Mdlalose, S. Ramaila, and U. Ramnarain, “Using Kahoot! As a Formative Assessment Tool in Science Teacher Education,” *International Journal of Higher Education*, vol. 11, no. 2, pp. 43–51, Jan. 2022, doi: 10.5430/ijhe.v11n2p43.
[36] S. N. Şad and N. Özer, “Using Kahoot! as a Gamified Formative Assessment Tool: A case study,” *IJARE*, vol. 5, no. 1–2, pp. 43–57, Dec. 2019, doi: 10.17985/ijare.645584.
[37] C. Udeozor, F. R. Abegão, and J. Glassey, “Measuring learning in digital games: Applying a game-based assessment framework,” *Brit J Educational Tech*, vol. 55, no. 3, pp. 957–991, 2023, doi: 10.1111/bjet.13407.
[38] S. de L. e Silva *et al.*, “Uma análise sobre a utilização de jogos didáticos como uma forma de avaliação mediadora para melhorar o ensino-aprendizagem de ciências,” *Caderno de ANAIS HOME*, 2023, [Online]. Available: https://homepublishing.com.br/index.php/cadernodeanais/article/view/1202

[39] S. Weber, M. Off, T. Hackenberg, M. Schumann, and F. Achtenhagen, “Serious Games as Assessment Tools: Visualizing Sustainable Creative Competence in the Field of Retail,” in *Game-based Learning Across the Disciplines*, C. Aprea and D. Ifenthaler, Eds., Cham: Springer International Publishing, 2021, pp. 109–140. doi: 10.1007/978-3-030-75142-5_6.
[40] V. Vallejo *et al.*, “Evaluation of a new serious game based multitasking assessment tool for cognition and activities of daily living: Comparison with a real cooking task,” *Computers in Human Behavior*, vol. 70, pp. 500–506, May 2017, doi: 10.1016/j.chb.2017.01.021.
[41] A. Oranje, B. Mislevy, M. I. Bauer, and G. T. Jackson, “Summative Game-Based Assessment,” in *Game-Based Assessment Revisited*, Cham: Springer International Publishing, 2019, pp. 37–65. [Online]. Available: https://doi.org/10.1007/978-3-030-15569-8_1
[42] A. Emerson, W. Min, R. Azevedo, and J. Lester, “Early prediction of student knowledge in game-based learning with distributed representations of assessment questions,” *British Journal of Educational Technology*, vol. 54, no. 1, pp. 40–57, Oct. 2022, doi: 10.1111/bjet.13281.
[43] Y. J. Kim, M. Knowles, J. Scianna, G. Lin, and J. A. Ruipérez-Valiente, “Learning analytics application to examine validity and generalizability of game-based assessment for spatial reasoning,” *British Journal of Educational Technology*, vol. 54, 2022, doi: 10.1111/bjet.13286.
[44] Y. J. Kim and D. Ifenthaler, “Game-Based Assessment: The Past Ten Years and Moving Forward,” in *Game-Based Assessment Revisited*, D. Ifenthaler and Y. J. Kim, Eds., Cham: Springer International Publishing, 2019, pp. 3–11. doi: 10.1007/978-3-030-15569-8_1.
[45] P. M. Kato and S. de Klerk, “Serious Games for Assessment: Welcome to the Jungle,” *Journal of Applied Testing Technology*, vol. 18, no. 1, pp. 1–6, Oct. 2017, [Online]. Available: https://jattjournal.net/index.php/atp/article/view/118669
[46] F. Bellotti, B. Kapralos, K. Lee, P. Moreno-Ger, and R. Berta, “Assessment in and of Serious Games: An Overview,” *Advances in Human-Computer Interaction*, vol. 2013, no. 1, p. 136864, Feb. 2013, doi: 10.1155/2013/136864.
[47] H. S. Çetin, “Implementation of the Digital Assessment Tool Kahoot in Elementary School,” *International Technology and Education Journal*, vol. 2, no. 1, pp. 9–20, Jun. 2018, [Online]. Available: https://dergipark.org.tr/en/pub/itej/issue/39211/461500
[48] N. Iman, M. Ramli, and N. Saridewi, “Kahoot as an Assessment Tools: Students’ Perception of Game-based Learning Platform,” *Jurnal Penelitian dan Pembelajaran IPA*, vol. 7, no. 2, pp. 245–259, Nov. 2021, doi: 10.30870/jppi.v7i2.8304.
[49] N. Mdlalose, S. Ramaila, and U. Ramnarain, “Using Kahoot! as a Formative Assessment Tool in Science Teacher Education,” *International Journal of Higher Education*, vol. 11, no. 2, pp. 43–51, Jan. 2022, doi: 10.5430/ijhe.v11n2p43.
[50] H. Muna, R. Setiyana, and F. Ismail, “A Game-based Assessment as A Formative Test in Academic Performance: Teachers’ Perspectives on Kahoot!,” *IJELR*, vol. 5, no. 2, pp. 180–189, Nov. 2023, doi: 10.35308/ijelr.v5i2.8661.
[51] B. W. Widyaningrum, “The Use of Kahoot! as Formative Assessment in Education,” *IJELS*, vol. 5, no. 2, pp. 75–84, Sep. 2019, doi: 10.24071/ijels.v5i2.2862.
[52] N. Tahiri and M. E. Alami, “A New Evaluation Technique Through Serious Games for Children with ASD,” *iJET*, vol. 15, no. 11, pp. 202–217, Jun. 2020, doi: 10.3991/ijet.v15i11.12843.
[53] P. C. Silva and A. V. Rodrigues, “Desenvolvimento de uma proposta curricular de ciências para os primeiros anos de escolaridade: Fundamentos, processo e produto,” *Revista on line de Política e Gestão Educacional*, p. e023072, 2023, doi: 10.22633/rpge.v27i00.18275.
[54] P. C. Silva and A. Rodrigues, “Ciclo de desenvolvimento das atividades do PEEC: análise, desenho, implementação, avaliação e redesenho,” in *Moving on Towards ‘New Normal’ In Education - ICEM2022 Conference: Book of Abstract*s, Santarém: IPSantarém, Oct. 2022, pp. 171–172. doi: 10.11144/Javeriana.M7-.
[55] P. Sá, P. C. Silva, J. Peixinho, A. Figueiras, and A. V. Rodrigues, “Sustainability at Play: Educational Design Research for the Development of a Digital Educational Resource for Primary Education,” *Social Sciences*, vol. 12, no. 7, p. 407, 2023, doi: 10.3390/socsci12070407.
[56] C. Boel, T. Rotsaert, M. Valcke, A. Vanhulsel, and T. Schellens, “Applying educational design research to develop a low-cost, mobile immersive virtual reality serious game teaching safety in secondary vocational education,” *Educ Inf Technol*, vol. 29, no. 7, pp. 8609–8646, May

2024, doi: 10.1007/s10639-023-12126-4.
[57] S. McKenney and T. C. Reeves, “Educational design research: Portraying, conducting, and enhancing productive scholarship,” *Med Educ*, vol. 55, no. 1, pp. 82–92, Jan. 2020, doi: 10.1111/medu.14280.
[58] W. Sandoval, “Educational design research in the 21st century,” in *Handbook of Design in Educational Technology*, R. Luckin, S. Puntambekar, P. Goodyear, J. Underwood, and N. Winters, Eds., New York: Routledge, 2013, pp. 388–396.
[59] S. McKenney and T. Reeves, *Conducting Educational Design Research*. London: Routledge, 2019. [Online]. Available: https://www.routledge.com/Conducting-Educational-Design-Research/McKenney-Reeves/p/book/9781138095564
[60] P. C. Silva, A. V. Rodrigues, and P. N. Vicente, “Currículo de ciências para o ensino primário: Uma análise comparativa entre Portugal, Inglaterra, Estados Unidos, Austrália e Singapura,” *Education Policy Analysis Archives*, vol. 31, 2023, doi: 10.14507/epaa.31.8192.
[61] C. Galvão, P. Reis, A. Freire, and T. Oliveira, *Avaliação de competências em ciências: sugestões para os professores dos ensinos básico e secundário*. Porto: ASA Editores, 2006.
[62] E. W. Krüger and R. I. M. Uhmann, “Avaliação da aprendizagem em foco no currículo do ensino de ciências,” *Revista Electrónica de Investigación en Educación en Ciencias*, vol. 17, no. 1, pp. 50–65, Jul. 2022, doi: https://doi.org/10.54343/reiec.v17i1.301.
[63] L. A. e L. do Nascimento and G. Rôças, “Portfólio: uma opção de avaliação integrada para o ensino de Ciências,” *Estud. Aval. Educ.*, vol. 26, no. 63, pp. 742–767, Dec. 2015, doi: 10.18222/eae.v0ix.3209.
[64] J. L. P. B. Silva and E. F. de Moradillo, “Avaliação, ensino e aprendizagem de ciências,” *Ensaio Pesquisa em Educação em Ciências*, vol. 4, no. 1, pp. 25–36, Jul. 2002, [Online]. Available: https://periodicos.ufmg.br/index.php/ensaio/article/view/10036
[65] P. Silva, A. Rodrigues, and P. Vicente, “Práticas de Ensino Experimental das Ciências no 1.o Ciclo do Ensino Básico em Portugal: uma análise dos relatórios da Inspeção-Geral da Educação e Ciência,” *Praxis Educativa*, vol. 18, no. 1, pp. 1–22, 2023, doi: 10.5212/PraxEduc.v.18.21358.062.
[66] T. C. Reeves, S. McKenney, and J. Herrington, “Publishing and perishing: The critical importance of educational design research,” *Australasian Journal of Educational Technology*, vol. 27, no. 1, Art. no. 1, Mar. 2011, doi: 10.14742/ajet.982.
[67] A. Cardoso, J. Souza, R. P. Rodrigues, and N. Palma, “Ciclo de desenvolvimento de um recurso educativo digital para o 1.º Ciclo do Ensino Básico,” *1*, vol. 13, no. 5, pp. 125–142, Dec. 2021, doi: 10.34624/id.v13i5.27175.
[68] N. Lopes and L. Amante, “Conceção e desenvolvimento de jogos digitais simples por professores e técnicos socais. Estado da arte e estudo empírico de um modelo de formação,” in *Educação e Tecnologias Web: Contributos de pesquisa Luso Brasileiros*, Torres, P.L. & L. Amante., Curitiba: Editora Appris, 2019. [Online]. Available: http://hdl.handle.net/10400.2/9448
[69] H. R. Dinia, C. Q. Lubis, M. Zawawi, U. Machmudah, M. Huda, and M. H. A. Lubis, “Developing of Ispring Quiz Maker as an Evaluation Media in the Pragmatics Course,” *Izdihar*, vol. 7, no. 1, pp. 83–98, Apr. 2024, doi: 10.22219/jiz.v7i1.31373.
[70] D. Mayuri, P. Aswirna, and H. Hurriyah, “Development of a quizmaker ispring approach assessment instrument to measure understanding of student concepts,” *Jurnal Inovasi Pendidikan IPA*, vol. 7, no. 2, pp. 213–225, 2021, doi: 10.21831/jipi.v7i2.30847.
[71] P. I. V. D. Radjibu, H. Kuswanto, and Sugiharto, “Analysis of critical thinking skills and scientific communication of students for SHM concepts assisted by Ispring quiz maker test instrument,” *Journal of Physics: Conference Series*, vol. 1440, no. 1, p. 012054, 2020, doi: 10.1088/1742-6596/1440/1/012054.
[72] D. Lehtonen, “Constructing a design framework and design methodology from educational design research on real-world educational technology development,” *EDeR*, vol. 5, no. 2, Nov. 2021, doi: 10.15460/eder.5.2.1680.
[73] R. P. De Lope and N. Medina-Medina, “A Comprehensive Taxonomy for Serious Games,” *Journal of Educational Computing Research*, vol. 55, no. 5, pp. 629–672, Dec. 2016, doi: 10.1177/0735633116681301.
[74] E. Naul and M. Liu, “Why Story Matters: A Review of Narrative in Serious Games,” *Journal of Educational Computing Research*, vol. 58, no. 3, pp. 687–707, Jul. 2020, doi:

10.1177/0735633119859904.
[75] W. S. Pereira, G. Cysneiros, and Y. P. C. Aguiar, "Diretrizes para o Desenvolvimento de Serious Games: Um Mapeamento Sistemático da Literatura," in *Anais do XXX Simpósio Brasileiro de Informática na Educação (SBIE 2019)*, Brasília: Sociedade Brasileira de Computação (SBC), 2019, pp. 714–722. doi: https://doi.org/10.5753/cbie.sbie.2019.714.
[76] F. L. V. Janssen, B. L. Xaviera, V. J. Serva, and Santos, "Investigando a Elaboração de Narrativas em Jogos Sérios," in *SBC – Proceedings of SBGames 2019*, Rio Janeiro, Brazil, 2019, pp. 1287–1295.
[77] Y. Kahyaoğlu Erdoğmuş and A. A. Kurt, "Digital game-based learning: Pedagogical agent and feedback types on achievement, flow experience, and cognitive load," *Education and Information Technologies*, vol. 29, pp. 12943–12968, Dec. 2023, doi: 10.1007/s10639-023-12368-2.
[78] C. I. Johnson, S. K. T. Bailey, and W. L. Van Buskirk, "Designing Effective Feedback Messages in Serious Games and Simulations: A Research Review," in *Instructional Techniques to Facilitate Learning and Motivation of Serious Games*, P. Wouters and H. van Oostendorp, Eds., Cham: Springer International Publishing, 2017, pp. 119–140. doi: 10.1007/978-3-319-39298-1_7.
[79] Y. A. Sekhavat, "A Rewarding Framework for Multiplayer Serious Games based on Competitive Reinforcement," in *2019 IEEE 7th International Conference on Serious Games and Applications for Health (SeGAH)*, 2019, pp. 1–8. doi: 10.1109/SeGAH.2019.8882473.
[80] H. Wang and C.-T. Sun, "Game reward systems: Gaming experiences and social meanings," in *Proceedings of DiGRA 2011 Conference: Think Design Play*, Jan. 2011. [Online]. Available: https://dl.digra.org/index.php/dl/article/view/594
[81] D. A. M. Darwesh, "Concepts Of Serious Game In Education," *International Journal of Engineering and Computer Science*, vol. 4, no. 12, 2015, [Online]. Available: http://www.ijecs.in/index.php/ijecs/article/view/2781
[82] B. M. F. de Souza, N. B. Xavier, E. S. R. da Silva, and M. V. de Oliveira, "U.M.E: A game to raise environmental awareness in children from the 2nd to the 5th year," *Brazilian Journal of Development*, vol. 6, no. 7, pp. 43823–43828, Jul. 2020, doi: 10.34117/bjdv6n7-121.
[83] H.-C. Chu and S.-C. Chang, "Developing an educational computer game for migratory bird identification based on a two-tier test approach," *Education Tech Research Dev*, vol. 62, no. 2, pp. 147–161, Nov. 2014, doi: 10.1007/s11423-013-9323-4.
[84] D. P. Paz, E. Pontarolo, and N. A. de Mello, "Avaliação do Kahoot como recurso para práticas de Educação Ambiental e em Solos," *TEyET*, no. 34, p. e15, Jun. 2023, doi: 10.24215/18509959.34.e15.
[85] J. Moitinho and M. F. Barreto, "O jogo digital como estratégia de Educação Alimentar e Nutricional para o público infantil," *Rev. Entreideias*, vol. 12, no. 02, Sep. 2023, doi: 10.9771/re.v13i02.48296.
[86] C. S. Fernandes, M. J. Campos, M. T. Moreira, A. Lima, S. Ferreira, and M. M. Martins, "Development and Validation of the Serious Educational Game in Nursing Appraisal Scale," *Nurs Rep*, vol. 14, no. 2, pp. 1148–1157, May 2024, doi: 10.3390/nursrep14020087.
[87] P. Caserman *et al.*, "Quality Criteria for Serious Games: Serious Part, Game Part, and Balance," *JMIR Serious Games*, vol. 8, no. 3, p. e19037, Jul. 2020, doi: 10.2196/19037.
[88] E. Turucz *et al.*, "Educational design research, an innovative method to investigate virtual reality applications in disaster management training programs," *Acta Marisiensis - Seria Medica*, vol. 67, no. 1, Art. no. 1, 2021, doi: 10.2478/amma-2021-0003.
[89] R. F. V. Tavares, "Aplicação móvel para educação em ciências integrando um framework de exploração de dados educacionais," doctoralThesis, Universidade de Aveiro, Aveiro, 2021. [Online]. Available: https://ria.ua.pt/handle/10773/31037
[90] S. Hart, A. Margheri, F. Paci, and V. Sassone, "Riskio: A Serious Game for Cyber Security Awareness and Education," *Computers & Security*, vol. 95, p. 101827, 2020, doi: 10.1016/j.cose.2020.101827.
[91] M. S. Siedler, R. C. Cardoso, L. A. D. S. Pinto, T. A. Tavares, and T. T. Primo, "Entretenimento e Aprendizagem: os desafios da criação de Serious Games educacionais," in *Workshop de Informática na Escola (WIE)*, SBC, 2023, pp. 496–507. doi: 10.5753/wie.2023.234887.

[92] L. Scienza and L. P. F. da Silva, "Game Design de 'Trate Bem!' e 'Fables Maker': Desenvolvimento e Testagem de Jogos Sérios," in *Anais do XV Seminário Jogos Eletrônicos, Educação e Comunicação*, Bahia, 2024, pp. 207–216. doi: 10.5281/zenodo.11271561.
[93] J. Gaspar *et al.*, "O Jogo Sério 'COVID-19 - Você sabia?' Descrição dos acessos ao jogo após 30 meses de implementação," *Anais do Computer on the Beach*, vol. 14, pp. 515–517, May 2023, doi: 10.14210/cotb.v14.p515-517.
[94] M. Karthan, P. Romahn, R. Pryss, diger, H. Kestler, and J. Schobel, "Collecting Data from Senior Citizens Using Serious Games," in *Advances in Informatics, Management and Technology in Healthcare*, IOS Press, 2022, pp. 234–237. doi: 10.3233/SHTI220705.
[95] J. D. S. Gaspar *et al.*, "A Mobile Serious Game About the Pandemic (COVID-19 - Did You Know?): Design and Evaluation Study," *JMIR Serious Games*, vol. 8, no. 4, p. e25226, Dec. 2020, doi: 10.2196/25226.
[96] C. Papoutsi, A. Drigas, C. Skianis, C. Skordoulis, and M. Pappas, "'The Park of Emotions': A Serious Game for the Development of Emotional Intelligence in Children," *Appl. Sci.*, vol. 14, no. 14, Art. no. 6067, Jul. 2024, doi: 10.3390/app14146067.
[97] A. Gordillo, D. López-Fernández, and J. Mayor, "Examining and Comparing the Effectiveness of Virtual Reality Serious Games and LEGO Serious Play for Learning Scrum," *Appl. Sci.*, vol. 14, no. 2, Art. no. 830, Jan. 2024, doi: 10.3390/app14020830.
[98] A. Jesus, I. Silveira, M. Araújo, and P. X. da Penha, "Simsustentabilidade: um jogo digital de estratégia para educação ambiental," *ENCITEC*, vol. 11, no. 3, Art. no. 3, Nov. 2021, doi: 10.31512/encitec.v11i3.40.
[99] G. M. dos S. A. Mol and P. V. R. de Carvalho, "Desenvolvimento de um jogo digital educacional de acordo com critérios de game design e pedagógico," *RECITE*, vol. 7, no. 1, pp. 147–167, Jun. 2022, doi: 10.29327/2283237.7.1-10.
[100] S. Sukirman, L. F. M. Ibharim, C. S. Said, and B. Murtiyasa, "Development and Usability Testing of a Virtual Reality Game for Learning Computational Thinking," *IJSG*, vol. 11, no. 3, Art. no. 3, Aug. 2024, doi: 10.17083/ijsg.v11i3.670.
[101] C. Reina-Reina, E. Antón, and J. A. Du´ñabeitia, "Impact of a cognitive training on reading of 6-year-old children," *IJSG*, vol. 11, no. 3, Art. no. 3, Sep. 2024, doi: 10.17083/ijsg.v11i3.754.

Appendixes

Description of the assessment SGs for the 1st, 3rd, and 4th grades.

Game Situation 1st grade - Summer Holidays at the Campsite!	Task	Learning Objectives
Boat Race!	Order the steps and identify permeable materials.	Plan an experiment.
Extra Question	Click on the instrument needed for the action "add water using a dropper."	A dropper is required for the action "add water using a dropper."
Loading the Boat!	Click on the character who is correct about the paper boat's load.	A boat has a maximum load capacity; once exceeded, it sinks.
Extra Question	Click on the option describing the boat's behavior under load.	Draw conclusions about the boat's behavior when carrying a load.
Mysterious Footprints!	Click on the animal corresponding to the footprints.	Woodpeckers have opposing toes that help them climb.
Extra Question	Click on the feather that corresponds to the woodpecker.	Woodpeckers have black and white feathers.
A Scientist?	Click on the character indicating where scientists work.	Scientists can also work in mountains.
Extra Question	Click on the image showing an area not studied by biologists.	Biologists do not study planets.
Cork, Where Do You Come From?	Click on the character naming the tree that produces cork.	Cork comes from the cork oak tree.

Game Situation 1st grade - Summer Holidays at the Campsite!	Task	Learning Objectives
Extra Question	Click on the material of mineral origin.	Aluminum is of mineral origin.
So Windy!	Drag the windsock corresponding to the day with the highest wind intensity.	Interpret wind intensity information using the windsock.
Extra Question	Click on the function of an anemometer.	An anemometer measures wind speed.
A Living Collection!	Click on the living beings at the beach.	Beach plants (ice plant), starfish, crabs, and seagulls are living beings.
Extra Question	Click on the image of a non-living being.	Robots are non-living beings.
A Fresh Snack!	Drag the plant part corresponding to the root.	A carrot is a root.
Extra Question	Click on the group of plant parts that are fruits.	Banana, apple, and pumpkin are fruits.
Waste Sorting!	Select and group materials into their correct recycling bins.	Demonstrate consistency in recycling.
Extra Question	Click on the object that should not go in the yellow bin.	Metal cutlery should not be placed in the yellow bin.
Shadow Theater!	Drag the translucent material to the box.	A translucent object partially lets light through.
Extra Question	Click on the material that lets light pass through.	Glass is a material that typically allows light to pass through completely.

Game Situation 3rd grade - Summer Holidays by the Sea!	Task	Learning Objectives
Sweet Breakfast!	Click on the character identifying the glass where chocolate dissolves fastest.	Chocolate dissolves faster in the glass with the highest temperature.
Extra Question	Click on the factor that does not influence dissolution time.	The type of glass does not influence the dissolution time of sugar.
A Light Ball!	Click on the character who is correct about whether air has mass.	Air has mass.
Extra Question	Click on the option identifying the relationship between balloons of different sizes.	A balloon with more volume (more air) is heavier.
A Melting Drink!	Drag the drink with and without ice based on the ice state for each character.	Crushed ice melts faster.
Extra Question	Click on the reason why the ice cream melted.	The ice cream melted because the ambient temperature increased.
Sandcastles	Click on the character identifying the permeability capacity of the sand.	Sandy soil is permeable.
Extra Question	Click on the option describing a permeable soil.	Permeable soil allows excess water to pass through.
A Ride in a Hot Air Balloon!	Drag the words describing how a hot air balloon works.	A hot air balloon rises because hot air is less dense. It descends because cold air is denser.
Extra Question	Click on the option identifying what happens when the hot air balloon descends.	A hot air balloon descends because cold air is denser.
A Shadow Game!	Drag the flashlight to the spot where the shadow is largest.	The closer the light source is to the object, the larger the shadow.
Extra Question	Click on the image representing the bear's shadow based on the flashlight's position.	A shadow always forms on the side opposite to the light source.
Interaction on the Ocean Floor!	Click on the character identifying the interaction between the anemone and clownfish.	The clownfish uses the anemone as a home, providing a hiding place from predators.
Extra Question	Click on the parasite that infects humans.	Roundworms infest human intestines.
Helping Parents!	Click on the character describing how dolphin parents assist their young.	Dolphins help their young by guiding them to the water's surface to breathe.
Extra Question	Click on the animal that brings food to its young.	Seagulls bring food to their young.

Game Situation 3rd grade - Summer Holidays by the Sea!	Task	Learning Objectives
So Many Eggs! Whose Are They?	Drag and match the eggs to the animals they belong to.	Octopus eggs are elongated, shark eggs are dark and capsule-like, and turtle eggs are round.
Extra Question	Click on the animal that is viviparous.	The whale is a viviparous animal.
Shark Pamphlet!	Click on the information matching the poster.	Analyze and interpret information about sharks.
Extra Question	Click on the option corresponding to information about the turtle.	Analyze and interpret information about turtle eggs.

Game Situation 4th grade - Summer Holidays on Safari	Task	Learning Objectives
An Injury!	Click on the most correct option and drag the corresponding X-ray to the injury.	A sprain occurs when the ligaments connecting bone to bone are torn.
Extra Question	Click on the instrument that amplifies the sound of the human body.	A stethoscope amplifies the sounds of the human body.
Preparing for the Safari!	Click on the character identifying the quickest way to cool drinks.	Water at a higher temperature freezes faster than water at a lower temperature.
Extra Question	Click on the state of ice division that cools water faster.	Crushed ice cools water faster.
The Lost Lunchbox!	Click on the character identifying the material that can slow the melting of ice.	Wool can be used as an insulator to slow the melting of ice.
Extra Question	Click on the state of ice division that melts faster.	The smaller the ice cube’s division, the faster it melts.
Sustainable Consumption!	Drag the most sustainable apples to the shopping basket.	Locally produced, unpackaged apples are the most sustainable option.
Extra Question	Click on the most sustainable bag.	A fabric bag is the most sustainable option.
An Afternoon on Safari!	Drag savannah animals to construct a food chain.	Grass is a producer, gazelles are primary consumers, and lions are secondary consumers.
Extra Question	Click on the herbivore’s role in the food chain.	Herbivores are primary consumers in the food chain.
Rhino in Sight!	Click on the character identifying why the black rhino is endangered.	The black rhino is endangered due to illegal poaching.
Extra Question	Click on the main reason for poaching.	Poaching occurs because the materials from these animals are rare and valuable.
Careful What You Catch!	Drag the gilthead seabreams longer than 19 cm.	Measure the length of gilthead seabreams.
Extra Question	Click on the reason why gilthead seabreams under 19 cm cannot be caught.	Gilthead seabreams under 19 cm cannot be caught because they have not yet reached reproductive maturity.
A Polluted Ocean!	Click on the character identifying the effect of excess carbon dioxide on fish.	Excess carbon dioxide in the water affects fish respiration.
Extra Question	Click on the option describing what happens to plastics in oceans.	Plastics in oceans fragment into very small pieces.
Where Do I Use Water?	Click on activities where water is used indirectly.	Activities like playing sports, flying a kite, and shopping consume water indirectly.
Extra Question	Click on the truthfulness of the statement “Producing food uses water and other energy types.”	Producing food also consumes water and other forms of energy.
Animal Collection!	Group animals by dragging them into categories: mammals, fish, and reptiles.	Animals are classified into groups: mammals (lion, giraffe), fish (gilthead seabream), and reptiles (turtle, crocodile).
Extra Question	Click on the option describing the characteristics of reptiles.	Reptiles are vertebrates, oviparous, have scaly skin, and are cold-blooded.