Multimodal Technologies and Interaction

Special Issue Reprint

Video Games

Learning, Emotions, and Motivation

Edited by
Ana Manzano-León and José M. Rodríguez Ferrer

mdpi.com/journal/mti

MDPI

Video Games: Learning, Emotions, and Motivation

Video Games: Learning, Emotions, and Motivation

Guest Editors

Ana Manzano-León
José M. Rodríguez Ferrer

MDPI

Basel • Beijing • Wuhan • Barcelona • Belgrade • Novi Sad • Cluj • Manchester

Guest Editors

Ana Manzano-León
Department of Psychology
University of Almería
Almería
Spain

José M. Rodríguez Ferrer
Department of Health
Sciences
University of Burgos
Burgos
Spain

Editorial Office
MDPI AG
Grosspeteranlage 5
4052 Basel, Switzerland

This is a reprint of the Special Issue, published open access by the journal *Multimodal Technologies and Interaction* (ISSN 2414-4088), freely accessible at: https://www.mdpi.com/journal/mti/special_issues/8F37PJ5A91.

For citation purposes, cite each article independently as indicated on the article page online and as indicated below:

Lastname, A.A.; Lastname, B.B. Article Title. *Journal Name* **Year**, *Volume Number*, Page Range.

ISBN 978-3-7258-7919-9 (Hbk)
ISBN 978-3-7258-7920-5 (PDF)
https://doi.org/10.3390/books978-3-7258-7920-5

© 2026 by the authors. Articles in this reprint are Open Access and distributed under the Creative Commons Attribution (CC BY) license. The reprint as a whole is distributed by MDPI under the terms and conditions of the Creative Commons Attribution-NonCommercial-NoDerivs (CC BY-NC-ND) license (https://creativecommons.org/licenses/by-nc-nd/4.0/).

Contents

About the Editors . . . . . **vii**

Alfonso Lago-Ferreiro, María Ángeles Gómez-González and José Carlos López-Ardao
A New Serious Game (*e-SoundWay*) for Learning English Phonetics
Reprinted from: *Multimodal Technol. Interact.* **2025**, *9*, 54,
https://doi.org/10.3390/mti9060054 . . . . . **1**

Noha Badkook, Doaa Sinnari and Abeer Almakky
Developing Comprehensive e-Game Design Guidelines to Support Children with Language Delay: A Step-by-Step Approach with Initial Validation
Reprinted from: *Multimodal Technol. Interact.* **2025**, *9*, 68,
https://doi.org/10.3390/mti9070068 . . . . . **28**

Juan Pedro Tacoronte-Sosa and María Ángeles Peña-Hita
Design and Evaluation of a Serious Game Prototype to Stimulate Pre-Reading Fluency Processes in Paediatric Hospital Classrooms
Reprinted from: *Multimodal Technol. Interact.* **2025**, *9*, 90,
https://doi.org/10.3390/mti9090090 . . . . . **55**

Yosra Achab-Moukram, Paula Rodríguez-Rivera, José Manuel Aguilar-Parra, José M. Rodríguez-Ferrer, Javier Rodríguez-Moreno and Ana Manzano-León
Evaluating a UDL-Based Digital Escape Room for Bullying Awareness in Secondary Education
Reprinted from: *Multimodal Technol. Interact.* **2025**, *9*, 55,
https://doi.org/10.3390/mti9060055 . . . . . **97**

Martin Hundrup, Jessi Holte, Ciara Bordeaux, Emma Ferguson, Joscelyn Coad, Terence Soule and Tyler Bland
Space Medicine Meets Serious Games: Boosting Engagement with the Medimon Creature Collector
Reprinted from: *Multimodal Technol. Interact.* **2025**, *9*, 80,
https://doi.org/10.3390/mti9080080 . . . . . **112**

Ioannis Sarlis, Dimitrios Kotsifakos and Christos Douligeris
Evaluating Educational Game Design Through Human–Machine Pair Inspection: Case Studies in Adaptive Learning Environments
Reprinted from: *Multimodal Technol. Interact.* **2025**, *9*, 92,
https://doi.org/10.3390/mti9090092 . . . . . **129**

Sijie Sun
Designing Gamified Intergenerational Reverse Mentorship Based on Cognitive Aging Theory
Reprinted from: *Multimodal Technol. Interact.* **2025**, *9*, 64,
https://doi.org/10.3390/mti9060064 . . . . . **162**

Raul Brumar, Stelian Nicola and Horia Ciocârlie
A Digital Model-Based Serious Game for PID-Controller Education: One-Axis Drone Model, Analytics, and Student Study
Reprinted from: *Multimodal Technol. Interact.* **2025**, *9*, 111,
https://doi.org/10.3390/mti9110111 . . . . . **190**

About the Editors

Ana Manzano-León

Ana Manzano-León is an Assistant Professor in the Department of Psychology at the University of Almería, Spain, where she is also a member of the HUM-878 Research Team at the Health Research Centre. Her research focuses on game-based learning, gamification, and playful strategies in education, with particular emphasis on student motivation, emotional engagement, and inclusive educational practices. She has led and participated in multiple research and teaching innovation projects, including studies on the use of serious games and gamification with students with dyslexia, autism spectrum disorder, and other special educational needs. She is the author of widely cited systematic reviews on educational gamification and has published extensively in international peer-reviewed journals on the role of video games and playful methodologies in formal and non-formal learning contexts.

José M. Rodríguez Ferrer

José M. Rodríguez Ferrer is an Assistant Professor affiliated with the University of Burgos, specializing in game-based learning, gamification, and educational innovation across multiple levels of schooling. His work examines the impact of playful methodologies on academic performance, reading processes, motivation, and social inclusion, with studies spanning primary, secondary, and higher education. He has collaborated on numerous research projects addressing the use of digital and analog games in diverse educational contexts, including interventions with students with dyslexia and language delays, and students from socially deprived communities. His publications appear in leading international journals and contribute to advancing evidence-based approaches to the design and evaluation of gamified learning environments.

Multimodal Technologies and Interaction

MDPI

Article

A New Serious Game (*e-SoundWay*) for Learning English Phonetics

Alfonso Lago-Ferreiro [1], María Ángeles Gómez-González [2] and José Carlos López-Ardao [3,*]

[1] Department of Electronic Technology, Universidad de Vigo, 36310 Vigo, Spain; alago@uvigo.gal
[2] Department of English and German Philology, Universidad de Santiago de Compostela, 15782 Santiago de Compostela, Spain; mdelosangeles.gomez@usc.es
[3] atlanTTic Research Center for Telecommunications Technology, Universidad de Vigo, 36310 Vigo, Spain
* Correspondence: jardao@det.uvigo.es; Tel.: +34-986-81-21-76

Abstract: This paper presents the design and evaluation of *e-SoundWay*, a cross-platform serious game developed to improve English phonetic competence through a multimodal and narrative-driven approach. While the platform is specifically tailored to meet the needs of Spanish-speaking learners, it is adaptable for a wider range of English as a Foreign Language (EFL) users. *e-SoundWay* offers over 600 interactive multimedia minigames that target three core competencies: perception, production, and transcription. Learners progress along a gamified version of the *Camino de Santiago*, interacting with characters representing diverse English accents. A mixed-methods evaluation combining pre- and post-tests with a user experience questionnaire revealed statistically significant improvements across all domains, particularly in perception. Reduced post-test variability indicated more equitable learning outcomes. User satisfaction was high, with 64% of participants reporting satisfaction with their phonetic progress and 91% stating they would recommend the platform. These findings highlight the educational effectiveness, accessibility, and motivational value of *e-SoundWay*, reinforcing the role of serious games and multimodal technologies in delivering inclusive and engaging pronunciation instruction.

Keywords: English pronunciation; serious games; multimodal learning; computer_assisted language learning (CALL); computer-assisted pronunciation training (CAPT); phonetic competence; gamified language education

1. Introduction

The ability to communicate effectively in English has become a fundamental skill in today's globalized world, where English functions as a *lingua franca* across academic, professional, and social domains. With more than 1.35 billion people using *English as an international, second or foreign language* (*EIL/ESL/EFL*), linguistic competence is no longer optional—it is strategic [1,2]. The growth of digital communication, global mobility, and intercultural interaction has led to the proliferation of English varieties and communicative norms, reinforcing the need for learners to navigate diverse linguistic contexts [3,4].

Within this evolving landscape, communicative competence must encompass not only grammatical and sociolinguistic knowledge but also phonetic awareness. This includes the ability to produce intelligible speech, accurately control rhythm and intonation, and recognize different English accents [5,6]. However, international assessments such as PISA and the European Survey on Language Competences consistently identify *pronunciation* as a major challenge for Spanish-speaking EFL learners, especially at the secondary and tertiary lev-

els. These difficulties are often linked to cross-linguistic interference, phoneme-grapheme mismatches, and the persistence of outdated or teacher-centered methodologies [7,8].

In response to these challenges, this study introduces *e-SoundWay*, a cross-platform serious game aimed at developing English phonetic competence through an *innovative, learner-centered, and gamified approach*. While specifically designed with Spanish-speaking learners in mind, the platform's adaptability makes it relevant for a broader EFL population. *e-SoundWay* leverages the power of *mobile learning*, *interactive multimedia*, and *narrative-driven design* to support lifelong learning and promote inclusive, accessible phonetics instruction.

Co-developed by the Electronic Equipment Engineering Group at the University of Vigo and the Scimitar linguistics team of the Discourse and Identity Group at the University of Santiago de Compostela, the tool reflects a cross-disciplinary partnership that integrates technological innovation with pedagogical and linguistic expertise. As will be further detailed in Section 2, the platform includes over *600 interactive minigames*, targeting three core phonetic skills: (1) *perception*—the ability to distinguish phonemes, allophones, and accentual variation; (2) *production*—accurate articulation of English speech sounds; and (3) *transcription*—mastery of the International Phonetic Alphabet (IPA) for decoding and encoding speech. A custom virtual keyboard, "TecladoFonetico," featuring 51 IPA symbols, supports transcription tasks [9].

The pedagogical design of *e-SoundWay* is grounded in *blended and mobile learning paradigms* that combine synchronous and asynchronous instruction [10,11]. Drawing from principles of *Computer-Assisted Language Learning (CALL)* and *Mobile-Assisted Language Learning (MALL)*, the platform fosters learner autonomy, multimodal engagement, and digital literacy [12,13]. Importantly, it also promotes cultural awareness through a storyline centered on the *Camino de Santiago*, which introduces learners to diverse English accents embedded in rich communicative contexts.

Although research has consistently highlighted the benefits of *Computer-Assisted Pronunciation Training (CAPT)* [14,15], many existing tools remain limited in scope. Common criticisms include the lack of personalized feedback, overly simplified phonological models, and assumptions about user proficiency that do not align with real learner needs [16,17]. Moreover, issues such as low digital literacy and insufficient scaffolding contribute to high dropout rates in online phonetics instruction [18].

This study addresses these limitations by evaluating the pedagogical impact of *e-SoundWay* as a next-generation CAPT solution. Section 2 presents the research design, platform features, as well as the guiding research questions. Sections 3 and 4, in turn, detail and discuss the results, while Section 5 draws together key conclusions and implications for future research and educational practice.

2. Materials and Methods

2.1. Tools for Game Creation

The serious game *e-SoundWay* was developed to enhance learner motivation and engagement in acquiring English phonetics, with a particular focus on segmental and suprasegmental pronunciation. Grounded in the EPSS method [19], which combines perceptual training, articulatory practice, and phonological transcription through task-based, multimodal instruction, the platform is purposefully designed for seamless integration into university-level English Phonetics and Phonology courses. At the same time, its user-friendly interface and autonomous learning features make it equally accessible to independent learners, educators, and professionals aiming to strengthen their phonetic awareness and English pronunciation skills in a flexible, engaging format. The method has

been shown to improve learners' intelligibility, self-monitoring, and transcription accuracy in empirical studies conducted in higher education contexts [20,21]. In particular, the authors in [20] report an empirical evaluation involving over 200 Spanish-speaking university students, demonstrating significant improvements in segmental accuracy, phonemic transcription, and learner motivation following the use of EPSS-based multimedia materials. These findings are further supported by the authors' longitudinal study [21], confirming statistically significant gains in pronunciation accuracy, transcription fluency, and error detection skills among students using the EPSS Multimedia Lab. Together, these studies confirm the value of EPSS-informed digital tools for phonetics instruction and validate their adaptation to serious game formats such as *e-SoundWay*. Designed as a complementary digital resource to traditional classroom instruction, the game promotes self-directed, gamified learning that supports phonological awareness and active learner participation.

The *e-SoundWay* design philosophy is also grounded in research demonstrating the motivational and pedagogical effectiveness of serious games in second-language learning environments [22–35]. Meta-analyses such as those by Girard et al. [22] and Subhash and Cudney [23] provide compelling evidence for the effectiveness of serious games and gamified learning in higher education. Girard et al. [22] demonstrate that serious games not only improve motivation and engagement but also lead to significant gains in knowledge acquisition and skill development, outcomes that are particularly relevant for pronunciation training, where repetitive, focused practice is essential. Similarly, Subhash and Cudney [23] identify critical gamification elements, such as feedback, rewards, progress tracking, and learner autonomy, as key drivers of improved academic performance and sustained engagement. These findings underscore the pedagogical value of integrating serious games into EFL learning environments, where they can provide personalized, interactive, and motivating contexts for developing phonological competence and pronunciation accuracy.

The application was developed using the Unity engine [36], chosen for its robust multimedia capabilities and broad compatibility across operating systems (Windows, macOS, Linux) and mobile platforms (Android, iOS). Unity's architecture enables the creation of immersive 2D environments through real-time rendering, including dynamic shadows, relief mapping, and reflections. Game logic is programmed using MonoDevelop, an open-source IDE that supports C# and other .NET languages, allowing for precise scripting and control of visual and interactive components. The development environment features four primary components:

1. **Hierarchical Tree**: Enables the creation of screen elements and their interdependencies, including layered objects such as buttons, text, and images.
2. **Project Window**: Hosts the organizational structure of assets and code modules used throughout the application.
3. **Scene and Game View**: Provides a live preview of the interface, including audiovisual materials and user interactions.
4. **Inspector and Console**: Allows developers to configure each element's properties (e.g., size, position, behavior) and troubleshoot runtime issues.

Scene resources are unloaded between stages to optimize memory usage, except for background music, which uses a Singleton pattern to maintain continuity. Care is taken to avoid redundant Singleton instances that could degrade performance.

Additionally, a core functionality of the game is its integration of offline speech recognition, implemented via the *Recognissimo Module* [37]. This system converts learner speech into text, enabling real-time comparison with target phonemes. All voice processing routines are implemented as asynchronous co-routines that run in parallel with the main gameplay. This design ensures that speech recognition tasks, such as audio capture,

preprocessing, and text conversion, do not interfere with interactive gameplay elements. To maintain performance stability, these routines allocate system resources dynamically and introduce controlled delays to accommodate real-time sentence processing without compromising user experience. All mini-games that require speech recognition include specialized code blocks configured to meet the technical and pedagogical requirements of each task. These blocks handle input signal cleaning, timing control, and target-specific evaluation procedures to ensure consistent, accurate performance across diverse game types as follows:

- **Language Model Provider:** References the target language model used to evaluate user speech input (see Figure 1). It specifies the location of the language-specific dictionaries and phonetic resources that support the speech recognition process.
- **Audio fragment as voice source:** Handles incoming audio clips, including cleaning, segmentation, and preparation for recognition tasks (Figure 2). It is specifically used in gameplay sections that require temporary audio file creation for comparing user-generated speech with predefined reference samples. By preprocessing the audio input, this module improves the accuracy and reliability of the speech recognition system.
- **Speech Recognizer:** Executes the speech recognition process by integrating various program elements (Figure 3). It utilizes the target language configuration defined in the Language Model Provider and accepts input either directly from a microphone or from pre-recorded audio clips. Target words are dynamically assigned to guide the evaluation process. A dedicated function supports control and comparison algorithms, allowing the system to generate reference lists and assess learner output through tolerance thresholds and approximation-based matching.
- **External speech recognizer:** Manages the execution of the speech recognition program based on the mini-game type selected (Figure 4). Within the development environment, a Boolean selector is used to configure the input format: short recordings for *Listen and Repeat* tasks (RECORDING1), medium-length recordings for *Minimal Pairs* (RECORDING2), and extended recordings for *Connected Speech* exercises (RECORDING3). This configuration ensures that the recognition system adapts accurately to the phonetic and temporal complexity of each task type.

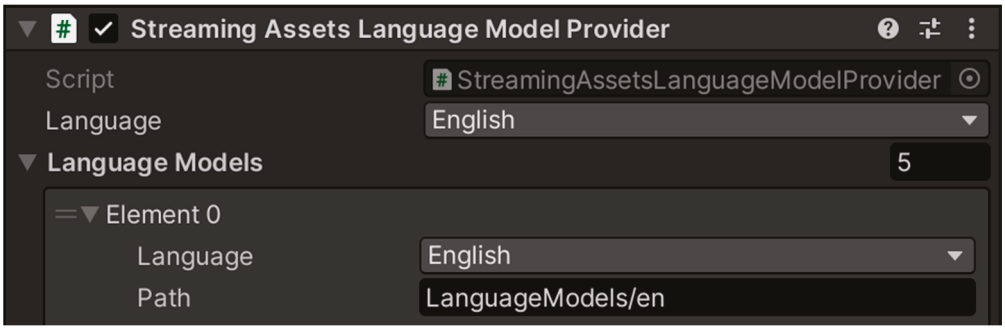


Figure 1. Language provider configuration panel.

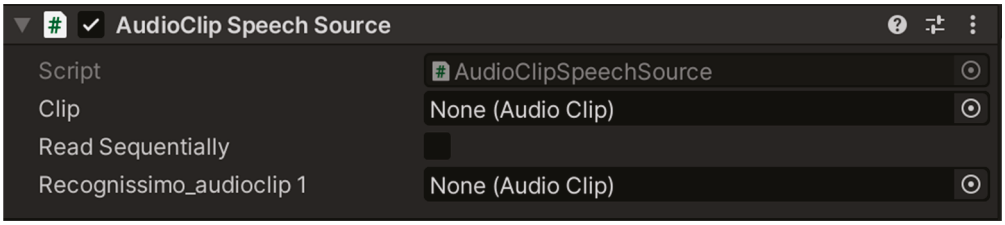


Figure 2. Script configuration audio snippet as voice source.

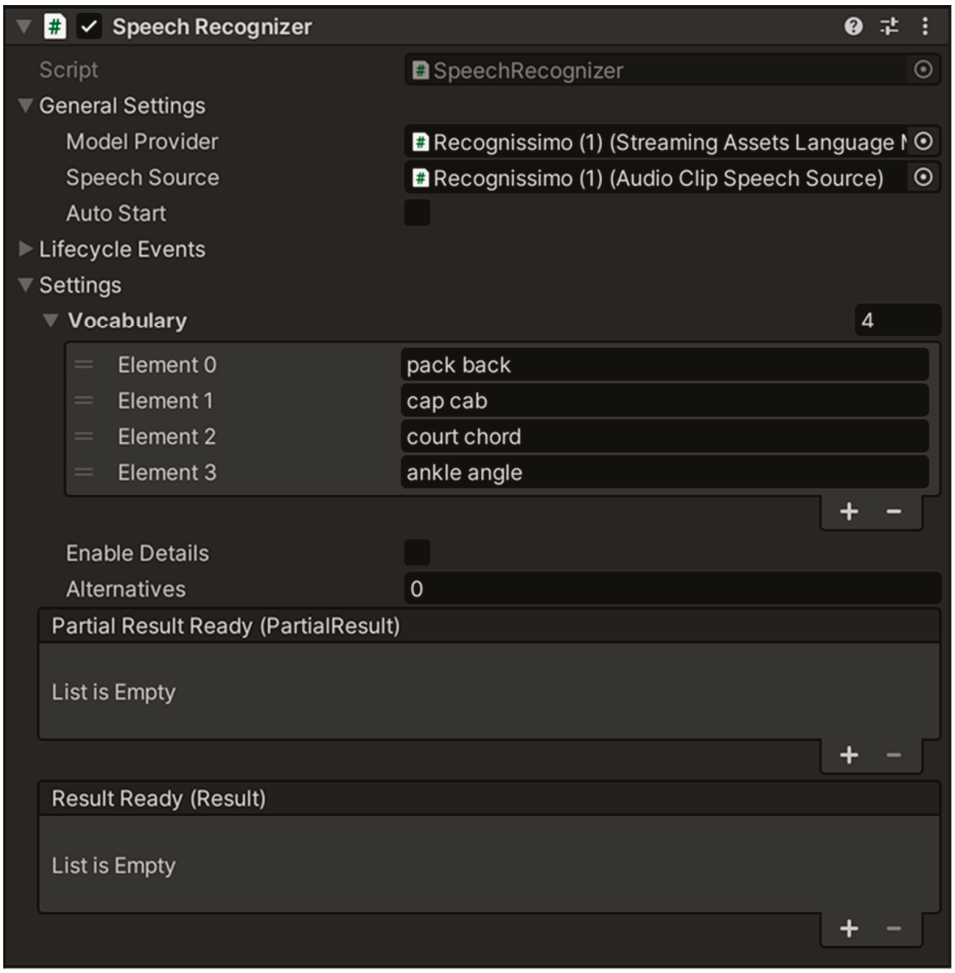


Figure 3. Speech recognizer configuration block.

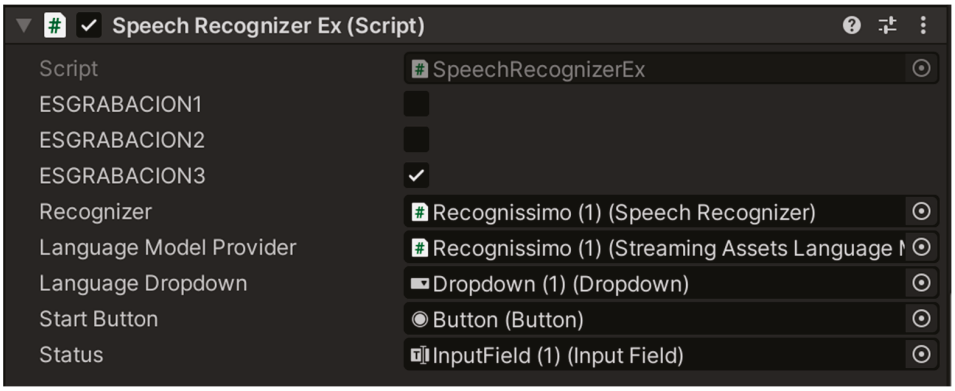


Figure 4. Script created for external speech recognition control.

2.2. General Structure and Game Contents

e-SoundWay delivers its phonetic content through 12 carefully designed challenge types, implemented as multimedia mini-games that develop learners' perception, production, and transcription skills. Each mini-game aligns auditory input with visual and textual elements, creating multisensory reinforcement and fostering phonological mapping. Drawing on Mayer's multimedia learning theory [4] and informed by Derwing and Munro's principles for effective pronunciation instruction [6], an approach that enhances both phoneme recognition and segmental intelligibility.

The tool adopts a microlearning framework [2], segmenting instruction into targeted, manageable "phonetic pills" that promote focused practice and cognitive retention. Empirical studies have shown that modular, feedback-driven phonetics tasks, particularly when implemented in gamified digital environments, boost learner motivation and improve long-term outcomes [13,16]. To support accessibility and reinforce phonological processing, the platform incorporates visual stimuli based on *ARASAAC (Aragonese Portal of Augmentative and Alternative Communication) pictograms*, which, as demonstrated by Allen et al. [38], facilitate sound-word mapping and phonemic awareness development, particularly for learners with additional processing needs.

High-quality audio recordings, produced at University College London's Speech, Hearing & Phonetic Sciences division, ensure articulatory accuracy and pedagogical consistency. The musical soundtrack, featuring Galician groups *Luar na Lubre* and *Resonet*, provides a culturally immersive experience that enriches the phonetic input. This aligns with findings by Lems [39], who emphasizes that music rooted in learners' cultural backgrounds fosters rhythmic sensitivity, enhances prosodic modeling, and increases engagement, key components in pronunciation training.

The user interface comprises a home screen, a progress-tracking dashboard, challenge-specific instructions, the mini-games themselves, and a Library of supplementary didactic materials. Although the total number of challenge types remains fixed at 12, the user's interaction determines the order, frequency, and depth of engagement with each module. These design choices reflect learner-centered principles in CAPT [24] and serious game research [32,33,35], which emphasize scaffolding, adaptive feedback, learner autonomy, and sustained engagement.

Upon launching the application, users enter the general interface via the home screen. From there, they can access progress metrics, browse categorized mini-games and stage information, or consult the Library for additional learning materials. Figure 5 illustrates the overall game flow. While the number of challenge stages is generally stable, it may vary slightly based on task configuration but never exceeds 12.

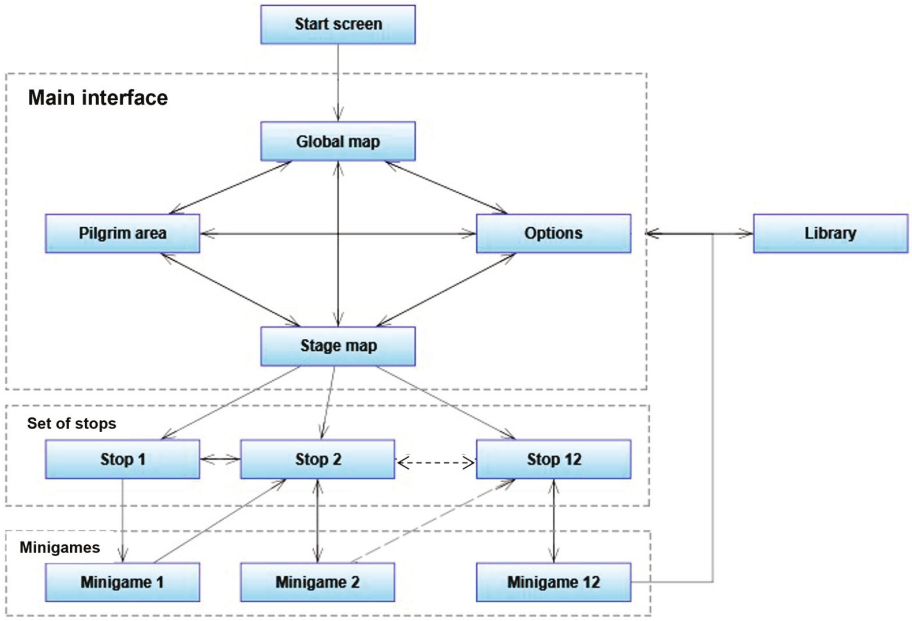


Figure 5. Connection between game scenes. **NOTE:** The dotted dashed line between stop 2 and stop 12, as well as Mini-game 2 to stop 12, refers to the pattern that repeats between stop 2 and 12 for stops and mini-games 2 to 11.

2.2.1. Home Screen

The e-SoundWay platform incorporates a user authentication system to personalize the learning experience and ensure data privacy. On the home screen, users are presented with two entry options: creating a new user profile or logging in using a previously registered username and password. This structure allows for individualized learning trajectories while safeguarding player data against unauthorized access, a critical feature in educational technologies where progress tracking and personalization are essential components of learning effectiveness [1,5].

When a new user initiates the registration process, the interface prompts for a username and password. The system then verifies whether all required fields are completed. If the data are missing, the profile cannot be created. Upon successful validation, the program checks for redundancy by comparing the input to existing entries. If the user is unique, the system creates a dedicated data table on the local device. This table remains persistent and dynamically records the player's progress throughout the game. The back-end structure also maintains a global registry to manage all users. This includes an incremental counter tracking the number of players created, and a reference map linking each username to a unique creation identifier. Such functionality supports both administrative transparency and pedagogical scalability, features increasingly emphasized in gamified learning environments [27,33]. In cases where a user attempts to log in with an existing profile, the system authenticates their identity and retrieves the corresponding data table, enabling a seamless continuation of the game and individualized pronunciation training.

This model is consistent with best practices in CALL and CAPT, where progress tracking and adaptive content delivery are recognized as key factors in maintaining learner engagement and ensuring pedagogical alignment with prior performance [24]. Additionally, studies in game-based learning environments highlight that personalization and ownership over learning data contribute significantly to learner autonomy and sustained motivation [33,34].

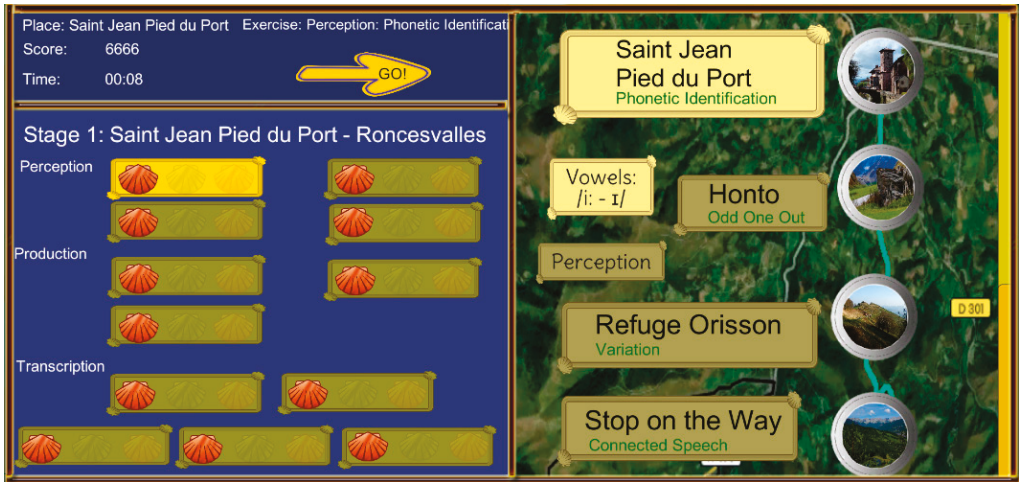
2.2.2. Main Interface

The structural design of *e-SoundWay* draws on the *roadmap game metaphor*, which, as noted by Peterson and Jabbari [40], leverages spatially oriented progression systems to enhance learner motivation, engagement, and perceived competence in EFL contexts, especially when embedded within meaningful cultural narratives. Accordingly, the user interface consists of four core interactive modules: the *Global Map*, the *Stage Map*, the *Options Panel*, and the *Pilgrim Zone*. These modules operate within a unified scene and are supported by a set of initialization routines responsible for rendering information dynamically on screen. Their integration ensures consistency in user interaction, performance tracking, and pedagogical feedback across the platform. A top navigation bar enables users to switch seamlessly between modules via button-based actuators (Figure 6). These buttons allow users to exit the game, access the current stage map (based on a stored stage ID), explore the global route ("The Way"), consult the Pilgrim Zone for individual learning analytics, retrieve additional resources from the Library, or configure settings through the Options panel. The interface adheres to principles of minimalist design in CALL environments, where streamlined navigation, visual clarity, and functional efficiency are essential to sustaining learner engagement and facilitating effective interaction [24,33].



Figure 6. Menu in main interface.

- **Stage Map:** Displays an image of a real section of the Camino de Santiago with interactive stop markers placed along the route (Figure 7). On the left side of the interface, a static frame presents information on player progress, including stop descriptions and a visual record of earned medals. Scallop shells, chosen for their cultural association with the Camino, represent achievement tiers: bronze, silver, or gold, awarded based on learner performance. The use of culturally grounded symbols aligns with findings by Lems [39], who emphasizes that familiar cultural cues increase emotional investment and deepen engagement with learning tasks. Each stage map includes interactive buttons programmed to activate animations, illuminate earned medals, and display stop-specific data such as time spent and phonetic scores. A "GO!" button becomes available once a stop is activated, allowing learners to continue their journey. Across the platform, 34 stage maps provide unique content and visual references while maintaining a consistent interaction logic.



(**a**) (**b**)

Figure 7. Stage map object image: (**a**) Stage of the Way of St. James that is being carried out and GO! arrow to advance. Medals achieved at that stage in the different skills; (**b**) stops where the student is in the stage and skills to be worked on.

- The **Global Map ("The Way"):** includes hidden actuators that, once triggered, load associated stage data and dynamically generate the relevant interface elements. This modular design supports data-driven content delivery and simplifies stage transitions, enhancing both navigational clarity and scalability [35].
- The **Pilgrim Zone:** Includes a personalized learning dashboard (see Figure 8) that provides learners with real-time feedback on their current stage, accumulated scores, unlocked content, and overall progress across three core skill areas: perception, production, and transcription. These competencies are visually represented in a *dynamic 3D pyramid*, which reflects the learner's advancement along each dimension. Additional features, such as an experience bar and a usage time clock, are incorporated to promote sustained engagement, metacognitive awareness, and effective time management, in accordance with established principles of serious game design [33,34]. Consistent with other EFL-focused serious games [23,24], *e-SoundWay* adopts a scaffolded,

performance-based learning path grounded in the EPSS model. This structure ensures a gradual increase in phonetic and phonological complexity, thereby supporting effective and individualized pronunciation development.

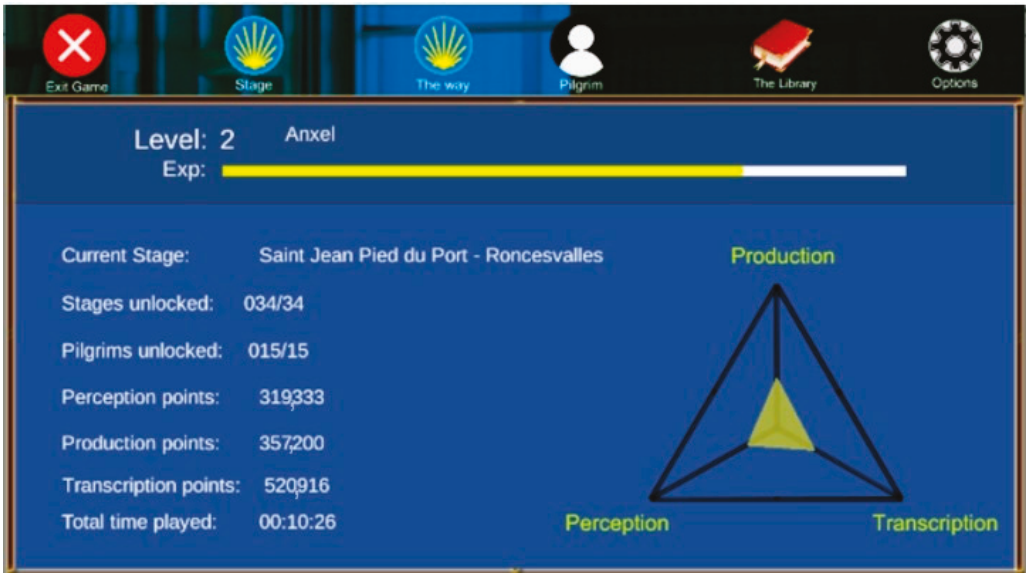


Figure 8. Screenshot Pilgrim's Zone.

- The **Options Panel:** Provides basic configuration tools, such as adjusting game audio levels and resetting progress. Selecting the reset option deletes all persistent data tied to the player profile, restoring the game to its default state. This functionality supports both usability testing and learner control, two features highlighted ineffective CALL platform design [24,27].

2.2.3. Stop Screens

Stop screens serve as intermediary nodes between the main interface and the minigames, offering users contextual information and navigational control within each stage. These screens are composed of a flexible set of buttons, the number and arrangement of which vary depending on the volume of content and interaction options associated with each stop. The interface supports up to nine buttons per screen. At the top, one or two arrow-shaped icons, visually inspired by the Camino de Santiago, enable learners to move sequentially between stops without returning to the main interface. This design facilitates fluid navigation and maintains cognitive continuity within stages. Additional top-level controls include sound settings and a return button to the main interface. Centrally positioned buttons (typically between two and five) provide access to individual informational modules, while a dedicated final button always launches the phonetic minigame associated with the current stop. This consistent layout reinforces procedural fluency and supports repeated exposure, two key principles in the design of educational games for language learning [33,34].

When learners interact with any of the informational buttons, a modal window appears containing written explanations and corresponding audio narration (Figure 9b). These dual-channel elements promote phonological awareness through synchronized verbal and visual input, a technique supported by multimedia learning theory [4] and studies on pronunciation-focused CALL environments [24]. Additionally, some informational windows include interactive hyperlinks, open in the default system browser and highlighted in blue, which redirect learners to external resources such as the EPSS Multimedia Lab

website (Figure 9b), thereby bridging in-game content with broader phonetic instruction. The structure of these stop screens aligns with scaffolding principles in serious game design, providing layered access to core content while encouraging learner autonomy through selective exploration.

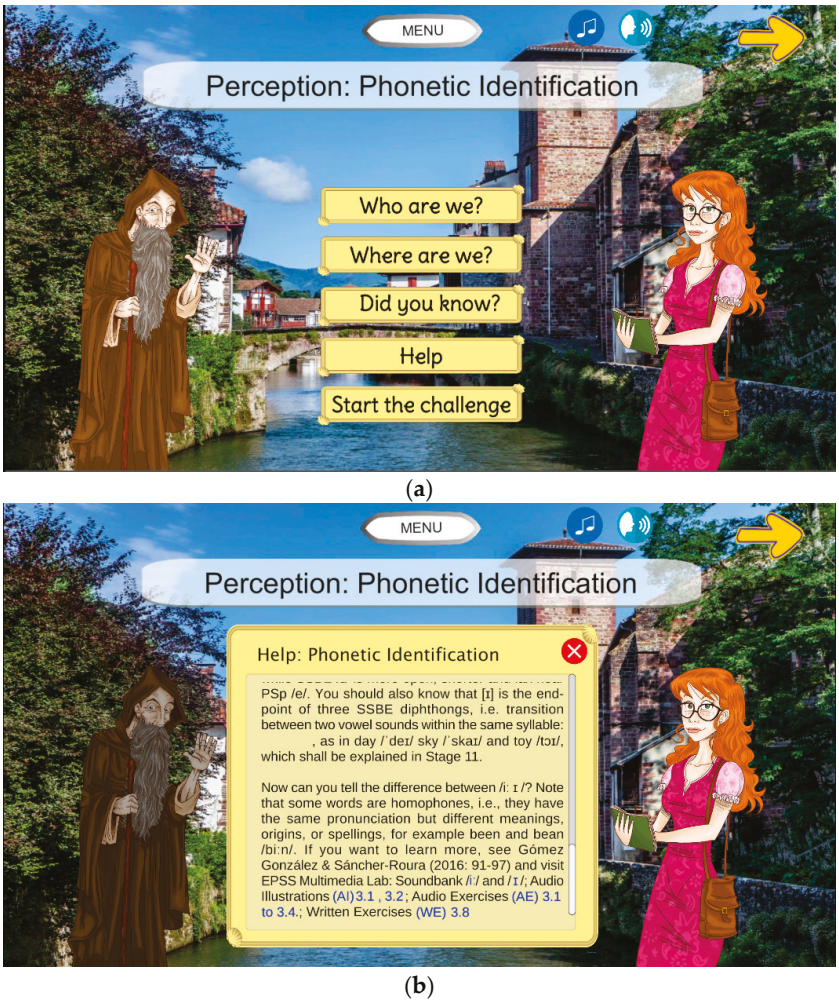


Figure 9. (**a**) Capture of a stop; (**b**) example of a displayed information element.

2.2.4. The Library as a Multimodal Resource

The Library in *e-SoundWay* offers a dedicated space for extended, multimodal learning that reinforces the game's phonetic training goals. It provides curated content that includes historical insights into pilgrimage traditions, cultural references linked to the *Camino de Santiago*, and phonetic resources anchored in the EPSS framework. Accessible through hyperlinked audiovisual modules, the Library connects learners to the EPSS Multimedia Lab, allowing them to deepen their understanding of phonetic theory and pronunciation practice beyond the game itself (Figure 10).

By bridging in-game tasks with evidence-based content and reflective learning, the Library significantly expands the pedagogical scope of CAPT. It unites task-based learning with structured phonetic theory, transforming *e-SoundWay* into a dynamic, multimodal environment that supports both formal instruction and autonomous learning. This integration fosters learner autonomy, metacognitive strategy development, and sustained motivation, as emphasized in research on content-rich CALL environments [16,20,22]. Moreover, by

embedding culturally and linguistically relevant resources, the Library enhances emotional engagement and contextual relevance [7,39].

Technically, the Library consists of four modular, dynamically loaded objects, each combining text, images, and interactive components that activate only when selected. This design prevents scene conflicts, ensures efficient memory use, and aligns with best practices in mobile-assisted language learning, where modularity and responsiveness are essential [15,24].



Figure 10. Screenshots of the Library.

2.2.5. Scoring System and Game Progress

The scoring architecture in *e-SoundWay* is structured around two layers of scripting: a resident script specific to each mini-game type, and a central script that aggregates scores and calculates overall player experience. This layered design enables both localized feedback and global performance tracking, key features in formative assessment frameworks for language learning [33,34].

Each mini-game generates an individualized score based on the player's performance, which is then evaluated against a success rate (%success). This rate is calculated using

Equation (1) for most mini-game types, while Equation (2) applies to the final two game categories, reflecting their increased complexity.

$$\begin{gathered}\%\text{success} = \text{if (success} \geq \text{mistakes)} \rightarrow (\text{success} - \text{mistakes})/\text{number of questions} \\ \text{if (success} < \text{mistakes)} \rightarrow 0\end{gathered} \tag{1}$$

$$\%\text{success} = \text{success}/\text{number of questions} \tag{2}$$

Upon completion of a mini-game, the player receives a final score calculated by multiplying the %success value by the game's maximum point allocation (Equation (3)). This result is visually displayed at the end of the activity, providing immediate performance feedback (Figure 11).

$$\text{Score} = \%\text{success} \times 10{,}000 \tag{3}$$



Figure 11. Sample score and medal on completion of mini-game.

In addition to numeric scores, the game employs a symbolic reward system in the form of medals, bronze, silver, and gold scallop shells, symbols of the Jacobean route, assigned according to performance thresholds. Table 1 illustrates the correlation between %success, accumulated points, and the corresponding medal. The use of culturally symbolic visual rewards, such as the scallop shell from the Camino de Santiago, enhances learner engagement and reinforces thematic continuity [39,40].

Table 1. Scored required to earn *e-SoundWay* medals.

% Success	Points	Scallop Shell Medals
$0\% < X < 50\%$	3	None
$50\% \leq X < 75\%$	5	Bronze
$75\% \leq X < 90\%$	7	Silver
$90\% \leq X \leq 100\%$	10	Gold

Cumulative scores are displayed in the Pilgrim Zone, where learners can monitor their overall progress. These scores are calculated by summing the total points earned across mini-games. To produce the progress graph shown in Figure 8, the total achieved scores are compared against the theoretical maximum, yielding a normalized success rate

between 0 and 1, as defined in Equation (4). This visual metric reinforces metacognitive reflection, allowing learners to assess their performance trajectory over time [16,24]. To ensure pedagogical consistency and minimum competence acquisition, progression within the game is gated. Learners must earn at least a bronze medal in each stage to unlock the subsequent one. This design reflects the principles of competence-based advancement and mastery learning [5], ensuring that players engage meaningfully with each stage's phonetic objectives before continuing.

$$\%\text{skill} = \sum (\%\text{success at that mini-game of that skill})/\text{number mini-games of that skill played} \quad (4)$$

2.2.6. Structure of the Mini-Games

The e-SoundWay tool consists of 614 mini-games divided into 3 categories: Perception games (266 mini-games)), Production Games (154 mini-games), and Transcription (194 mini-games).

Perception

In second language phonetics, *perception* refers to the learner's ability to detect, differentiate, and categorize speech sounds, particularly phonemic contrasts that may not exist in their first language [6,24]. Among perception-focused activities, *identification* and *discrimination* tasks play a central role, as noted by Strange and Shafer [41], Bradlow et al. [42], or Lengeris and Hazan [43], among others, who emphasize that such tasks are essential for "re-educating selective perception" in adult L2 learners, leading to measurable gains in both perceptual acuity and production accuracy, which supports the integration of structured perception training into digital pronunciation tools. *e-SoundWay* includes four perception mini-games that reflect this research-based approach:

- **Identification:** Learners recognize and label a sound by filling in missing phonemes in partially displayed words using audio cues and the virtual phonetic keyboard (Figure 12a). This reinforces sound-to-symbol mapping and contrastive listening [19,24].
- **Odd one out:** Players listen to four items and select the phonologically distinct one (A-B-A-A, B-A-A-A) (Figure 12b). Focusing on phonetic discrimination, i.e., comparing sounds and deciding whether they are the same or different, these activities build categorical perception and raise awareness of subtle L2 contrasts, often without requiring explicit labeling [6,42].
- **Variation:** These mini-games train learners to recognize phonetic variants, both context-driven and accent-based, by exposing them to multiple realizations of target phonemes (Figure 12c). This fosters flexible listening and perceptual tolerance, aligning with research that stresses the need to prepare learners for phonetic diversity in real-world speech (Setter & Jenkins [44]; Munro & Derwing [45]) and supports the development of robust phonological categories and speech convergence (Bradlow et al. [42]).
- **Connected speech.** Learners identify target sounds embedded in naturalistic text. By interacting with audio and receiving visual feedback, they develop an awareness of linking, assimilation, and reduction (Figure 12d). Strange and Shafer [41] and Wang and Munro [46] both advocate for such training to improve real-world listening comprehension.

These perception tasks reflect the EPSS method's emphasis on input-rich, feedback-oriented learning [20,21]. By embedding these tasks within a gamified progression model, e-SoundWay promotes sustained attention, repeat exposure, and active engagement—factors shown to accelerate phonological development [32,33].



Figure 12. Screenshot of perception mini-game: (**a**) identification; (**b**) odd one out; (**c**) variation; (**d**) connected speech.

Production

Production alludes to the learner's ability to articulate speech sounds accurately and intelligibly, involving the motor execution of phonological representations. Accurate production is a crucial component of oral intelligibility and is deeply intertwined with perceptual learning, as learners must first develop the ability to hear contrasts before they can reliably reproduce them [41].

Production alludes to the learner's ability to articulate speech sounds accurately and intelligibly, involving the motor execution of phonological representations. Accurate production is a crucial component of oral intelligibility and is deeply intertwined with perceptual learning, as learners must first develop the ability to hear contrasts before they can reliably reproduce them [41]. The production module in e-SoundWay aims to strengthen these articulatory skills through interactive mini-games of three different kinds that provide structured, feedback-rich output practice, each targeting different aspects of segmental and suprasegmental production:

- **Listen and Repeat** tasks are rooted in traditional mimicry-drill techniques, where learners attempt to replicate the target sound after auditory exposure (Figure 13a). Wang and Munro [46], among others, demonstrated that, despite their simplicity, such drills remain effective for developing articulatory precision, particularly when paired with immediate feedback and visual reinforcement. It is therefore assumed that computer-assisted listen-and-repeat tasks will improve sound contrast production and lead to measurable gains in intelligibility and learner confidence.
- **Minimal pairs:** Provide contrastive practice with phonemes that are especially challenging for EFL learners. By requiring learners to differentiate and articulate pairs of words that differ in the pronunciation of one, often challenging phoneme, these tasks promote fine-grained phonological awareness and motor control (Figure 13b), as

previously found by authors like Bradlow et al. [42] or Lengeris and Hazan [43]. They explained that explicit minimal pair training leads to improvements in both perception and production, especially when it incorporates high-variability input across speakers and contexts.

- **Connected speech** mini-games focus on the natural flow of spoken English, requiring learners to articulate word sequences that exhibit phonetic processes such as assimilation, linking, and elision (Figure 13c). These features often contribute to reduced intelligibility among EFL speakers who have primarily trained on isolated words or canonical forms. Studies such as those by Jenkins [44] emphasize that connected speech awareness is critical for developing rhythm, fluency, and listener-oriented intelligibility. Training learners to perceive and produce connected speech patterns equips them with the tools to understand rapid, native-like speech and to improve their own prosodic delivery in extended utterances. Incorporating connected speech into production tasks also supports the communicative goal of pronunciation instruction, which is not only accuracy but also comprehensibility in real-world interaction [45].

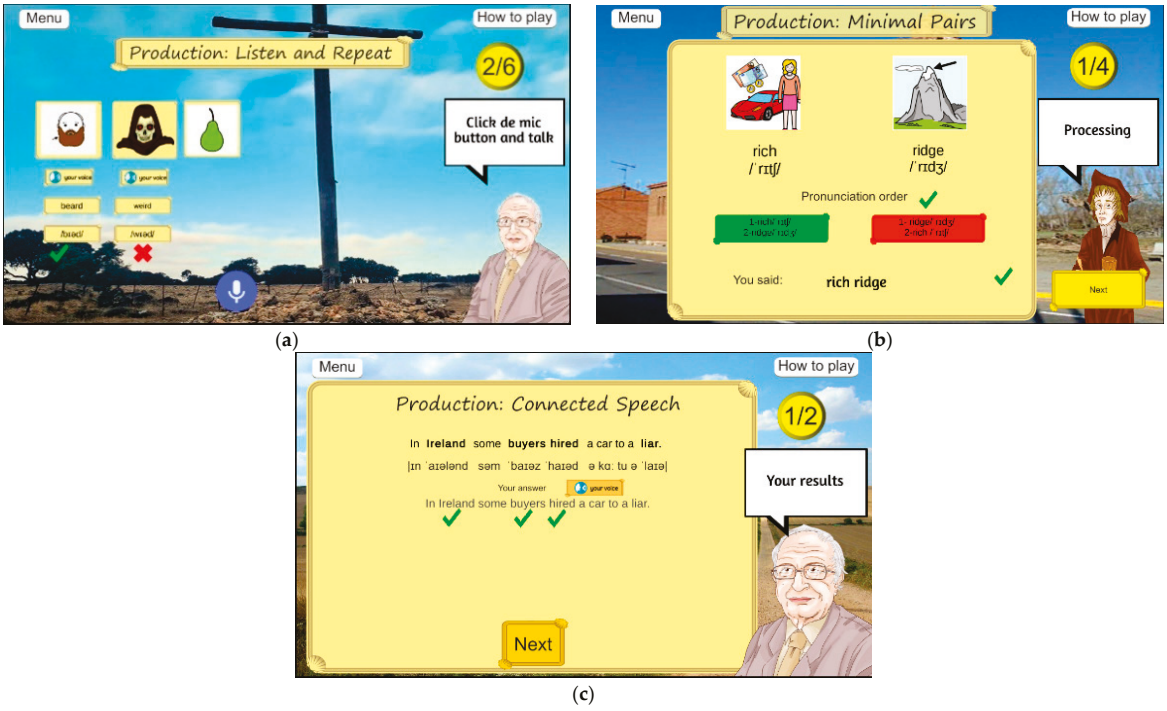


Figure 13. Screenshot of production mini-game: (**a**) listen and repeat; (**b**) minimal pairs; (**c**) connected speech.

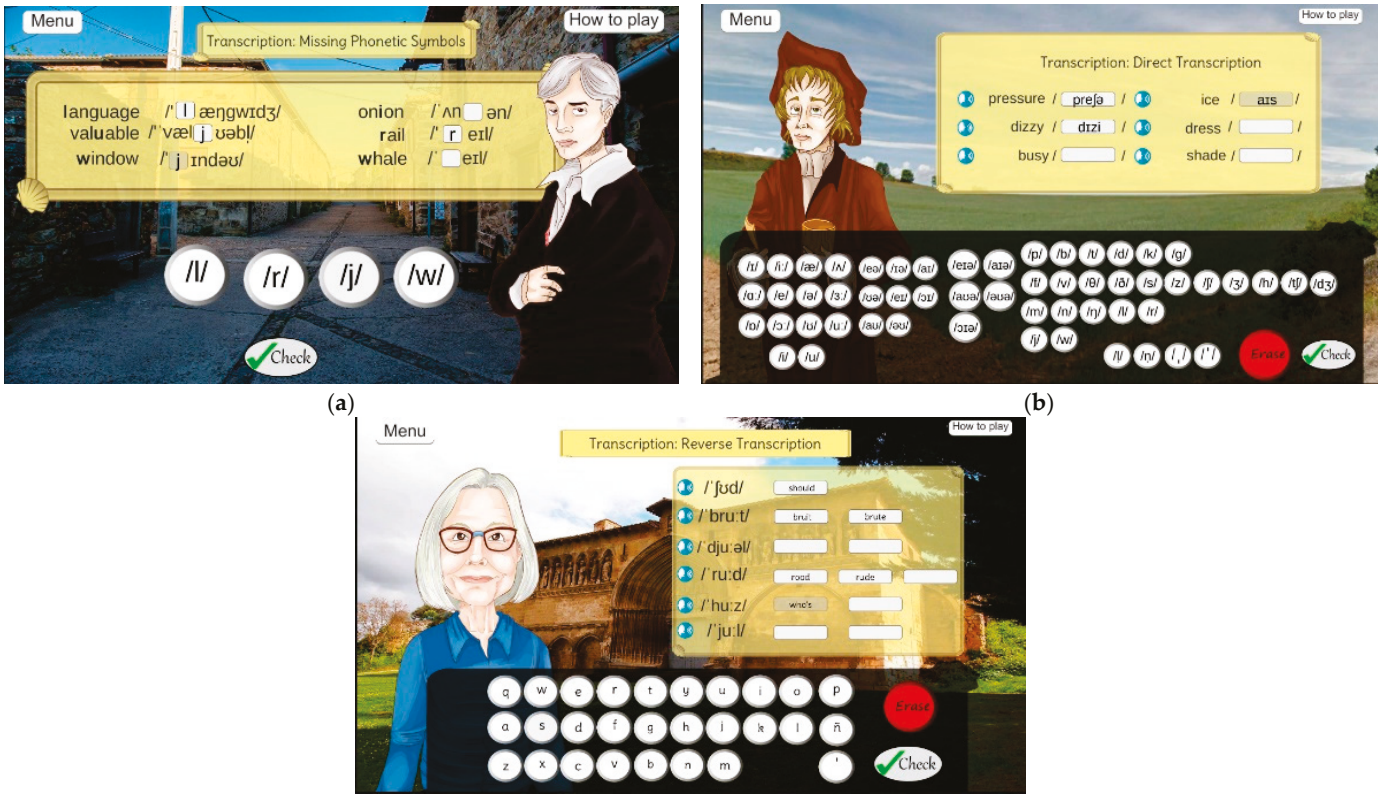
As with the perception module, the production mini-games in *e-SoundWay* are structured in accordance with the EPSS framework [20,21], which advocates for the progressive layering of phonetic skills, beginning with perception, followed by production, and culminating in transcription. The production tasks are not only gamified to enhance learner motivation but are also carefully sequenced to allow for repeated, scaffolded practice with increasing levels of complexity. This pedagogical design fosters the internalization of articulatory patterns while simultaneously reducing the anxiety often associated with pronunciation training in traditional classroom contexts.

Transcription

Transcription refers to the ability to decode or encode the sounds of spoken language using phonemic symbols or corresponding graphemes. It plays a critical mediating role between perception and production by encouraging learners to attend closely to sound-symbol correspondences, segmental detail, and phonological structure. As Strange and Shafer [41] emphasize, transcription promotes the re-education of perceptual selectivity and the formation of stable phonemic categories in adult learners. It also reinforces the articulatory and acoustic properties of L2 phonemes through metacognitive engagement with speech. As already noted, drawing on the EPSS framework [20,21], the *e-SoundWay* transcription module targets the third and final skill in the perceptual-articulatory sequence. It consists of five mini-games that train learners in both phoneme-to-grapheme and grapheme-to-phoneme conversion. Four types of transcription games are designed not only to reinforce previous perceptual and production learning but also to enhance learners' ability to decode a range of phonological inputs, including those from different English accents:

- **Missing symbols:** In these mini-games, learners are required to fill in missing phonemes within a set of words displayed on the screen in order to reinforce sound-symbol associations and test learners' segmental decoding accuracy [46]. Unlike in identification tasks, here learners must select phonemic symbols from a virtual IPA keyboard to complete the word forms. The player can overwrite previously placed entries before finalizing their response using the "Check" button (Figure 14a).
- **Direct transcription** mini-games require learners to transcribe full words from graphemes into phonemic script, in line with Wang and Munro's [45] suggestion that transcription-based tasks improve learners' sensitivity to non-native vowel contrasts and provide insight into their own production limitations. Audio support is provided for each word, and users interact with the IPA keyboard to generate phonological representations. The format mirrors authentic tasks used in pronunciation instruction to strengthen learners' understanding of English spelling-pronunciation irregularities (Figure 14b).
- **Reverse transcription:** Learners are given phonemic transcriptions and must produce the corresponding graphemic forms. The game employs a QWERTY keyboard and allows for multiple correct spellings, as one phoneme string may correspond to several orthographic variants. The evaluation system accounts for such variability, penalizing only duplicated or incorrect entries while accommodating lexical diversity (Figure 14c). As in the previous case, these minigames intend to enhance learners' decoding strategies and are especially relevant for those who struggle with the irregularities of English spelling conventions.
- **Crossword.** In these interactive puzzles, players complete a crossword grid using either phonemes or graphemes, depending on the given clues. The clues may include audio, IPA symbols, or graphemic hints, and the keyboard script dynamically updates each entry and score. Evaluation is based on the correctness of full entries, though the task avoids penalizing partial errors. This game helps consolidate learners' segmental and suprasegmental awareness across multiple word forms (Figure 15a). as suggested by authors like Setter and Jenkins [44], who note that explicit transcription activities enhance learners' ability to monitor their own pronunciation more effectively over time.
- **Word puzzle:** In these mini-games, players must locate words within a grid using graphemes, phonemes, or only audio clues. Increased difficulty is introduced in stages involving accent confrontation, where audio becomes the primary input. The grid is

dynamically generated using stored lexical data, and correct matches are highlighted visually and aurally (Figure 15b). This task promotes orthographic and phonemic scanning skills, crucial for real-time decoding and listening comprehension [41,44].



(**a**) (**b**)

(**c**)

Figure 14. Screenshot of transcription mini-game: (**a**) missing symbols; (**b**) direct transcriptions; (**c**) reverse transcription.



(**a**) (**b**)

Figure 15. A Transcription mini-game: (**a**) crossword; (**b**) world-puzzle.

Together, these transcription mini-games bridge the gap between perceptual training and productive control and writing (i.e., reinforcing associative mappings across sounds-spellings-IPA symbols) by enabling learners to analyze the structure of words in both auditory and symbolic forms. By developing explicit knowledge of the phonological system through repeated, interactive transcription practice, learners can enhance both their intelligibility and their phonological autonomy in real-world communication.

2.3. Methodology

2.3.1. Research Questions

This study was guided by two main research questions to evaluate the instructional effectiveness and learner reception of the platform:

1. To what extent does the use of *e-SoundWay* improve Spanish-speaking university EFL learners' performance in phonetic perception, production, and transcription as measured by pre- and post-intervention tests?
2. How do learners perceive the pedagogical effectiveness, usability, and motivational value of the *e-SoundWay* platform according to post-intervention user experience questionnaires?

2.3.2. Research Design

The experiment employs a *mixed-methods design* to evaluate the pedagogical effectiveness and user experience of *e-SoundWay*. Quantitative data were collected via pre- and post-tests designed to assess learners' phonetic perception, production, and transcription skills, following the EPSS model [20,21]. Qualitative and attitudinal data were gathered through post-intervention user experience questionnaires. Studies by Subhash and Cudney [33], Girard et al. [32], and Chen et al. [16] underscore that learner motivation, engagement, and perceived usability are key factors influencing the success of gamified educational tools in EFL contexts. Gathering learner perspectives not only offers insights into the affective dimensions of the learning experience but also supports iterative improvements in tool design and pedagogical alignment.

2.3.3. Participants

Participants consisted of Spanish-speaking university students (N = 33) enrolled in a first-year undergraduate course in English Phonetics and Phonology at the University of Santiago de Compostela in the academic year 2023/2024, which followed the EPSS method [19]. All participants had an upper-intermediate level of English (B2, CEFR) and limited prior exposure to phonetic transcription or pronunciation-focused training.

As the primary aim of this study was to evaluate the effectiveness of the *e-SoundWay* platform through within-subject comparison, no control group was included. This approach is consistent with similar studies in CALL and CAPT contexts (e.g., [16,45]), where pre- and post-intervention testing has been used effectively to assess learning gains among target user populations. Participation was voluntary, and all students provided informed consent.

2.3.4. Player Progression and Level System

The e-SoundWay platform features a gamified leveling system designed to monitor learner progress and foster sustained engagement across the 34-stage Camino itinerary. Learners begin at Level 1 and advance by completing tasks, achieving accuracy in phonetic transcription, and performing successfully in pronunciation and perception challenges and minigames. Progression is driven by the accumulation of points, which unlock higher levels characterized by increased linguistic complexity and greater exposure to a variety of English accents.

Learners' development in perception, production, and transcription skills is visually represented through a dynamic 3D pyramid (see Section 2.2.2, Pilgrim Zone, Figure 8), based on the scoring system detailed in Section 2.2.5. This multimodal representation allows for continuous monitoring of individual progress across the three core phonetic competencies. Beyond its motivational function, the leveling system acts as a formative assessment tool, providing learners with immediate, visually accessible feedback on their evolving proficiency in specific phonetic categories. Moreover, it supports phonological

development through a scaffolded learning trajectory that incrementally increases both cognitive and articulatory demands.

2.3.5. Instruments

Three types of diagnostic pre- and post-tests were designed, each corresponding to a specific mini-game category—perception, production, and transcription—to assess learners' phonetic abilities, whose relevance for EFL phonetic training has been previously mentioned. In addition, a post-intervention questionnaire comprising 12 items was administered to collect data on learner profiles, perceived usability, motivation, task difficulty, and perceived learning outcomes. The questionnaire items were organized into two categories, as shown in Tables 2 and 3. This research design draws on previous studies (e.g., Chen et al. [16]; Girard et al. [32]; Cudney [33]) that highlight the importance of learner perception data in evaluating digital learning environments. These studies advocate for the integration of user-centered feedback mechanisms to better capture affective and cognitive engagement and to inform iterative design improvements.

Table 2. Student profile.

Number	Question
1	What is your occupation?
2	How many hours a day are you online?
3	How long do you devote to playing games online?
4	What kind of games do you play?

Table 3. Student Satisfaction.

Number	Question
5	Are the instructions to play clear enough?
6	Do you find the interactivity of the game is user-friendly?
7	Do you find it challenging enough?
8	Does *e-SoundWay* feel like a game?
9	Which was the most difficult game section?
10	Do you think you have improved your level in English Phonetics and Phonology?
11	How satisfied are you with *e-SoundWay* as a whole?
12	Would you recommend this game to others in your situation?

Questions 6, 7, 8, 10, and 11 were answered by the students on a Likert scale from 1 to 5, where 1 is a very poor rating and 5 is an excellent rating. The students' results are shown in Section 3.2.

2.3.6. Procedure

The experiment was conducted over a four-month period, consisting of one training session per week. In parallel with their EPSS-based formal instruction offered in class, learners used *e-SoundWay* independently, engaging in self-paced sessions outside the classroom on their own devices. This dual approach allowed them to consolidate course content through gamified, task-based practice aligned with perceptual training, articulatory exercises, and phonological transcription.

In the first week, participants completed a 50 min pre-test comprising 15 mini-games, five targeting each of the three core phonetic competencies: perception, production, and transcription. From weeks 2 to 15, learners accessed the platform autonomously once per week, progressing through scaffolded pronunciation tasks embedded in the game's

storyline. No direct teacher intervention was provided during the training period, ensuring that learners' progress was driven by interaction with the platform's multimodal feedback and learning tools.

In the final week, a post-test mirroring the structure of the pre-test was administered, followed by a user experience questionnaire. While the post-test assessed the same three phonetic competencies, perception, production, and transcription, it included a different set of mini-games. This approach ensured that learners were not simply recalling previous task content but were instead applying acquired skills to novel items. Using parallel but distinct tasks is not only consistent with sound assessment practices in language learning research, as it minimizes memory effects and enhances construct validity [4,14,24]. It is also grounded on principles of authentic assessment increasingly relevant to contemporary language education, where learners demonstrate skill transfer in varied contexts so that learning outcomes can be evaluated in conditions that reflect real-world blended and independent learning contexts [10,27].

2.3.7. Data Analysis

To assess the effectiveness of *e-SoundWay*, a comprehensive set of statistical analyses was conducted. Paired-sample t-tests compared pre- and post-test scores across the three phonetic competencies—perception, production, and transcription—while Cohen's d was calculated to determine the magnitude of observed effects. To evaluate the consistency of learning outcomes, standard deviation values were analyzed across test phases.

User Experience Questionnaire responses were also examined using frequency distributions and additional tests to identify patterns related to satisfaction, motivation, and perceived learning. Specifically, a chi-square test of independence was applied to determine whether prior exposure to serious games influenced usability perceptions, while a one-way ANOVA explored the effect of time spent on online gaming on learners' perceived levels of challenge and engagement. Perceived difficulty across the three phonetic domains was assessed using a Friedman test, followed by Wilcoxon signed-rank tests for post hoc comparisons. Additionally, a Pearson correlation analysis examined the relationship between learners' self-reported improvement and overall satisfaction with the platform.

Together, these statistical procedures offered a multidimensional understanding of learners' interactions with the tool, the challenges they faced, and the influence of background factors on both user experience and learning performance. Moreover, this analytical approach aligns with established methodologies in previous research on CALL, CAPT, and serious games (e.g., Chen et al. [16]; Girard et al. [32]; Subhash & Cudney [33]).

3. Results

3.1. Pre-/Post-Tests

To address RQ1, Figure 16 presents the quantitative results of the pre- and post-tests, illustrating both average learner performance and score variability across the cohort.

Figure 16 displays the mean percentage of correct responses for each phonetic skill—perception, production, and transcription—before and after the intervention. All comparisons yielded statistically significant improvements ($p < 0.001$), which were further corroborated by Cohen's d effect size calculations. These indicated large effects across all three domains, with particularly notable gains in perception [6,24,25,45].

To assess the consistency and dispersion of student performance, standard deviation values were calculated for each skill category and test phase. Figure 16 shows the standard error bar for each category. The results revealed a marked reduction in variability across the board:

- In Perception, the standard deviation decreased from 0.0284 to 0.0070.
- In Production, it dropped from 0.0669 to 0.0197.
- In Transcription, it declined from 0.0444 to 0.0111.

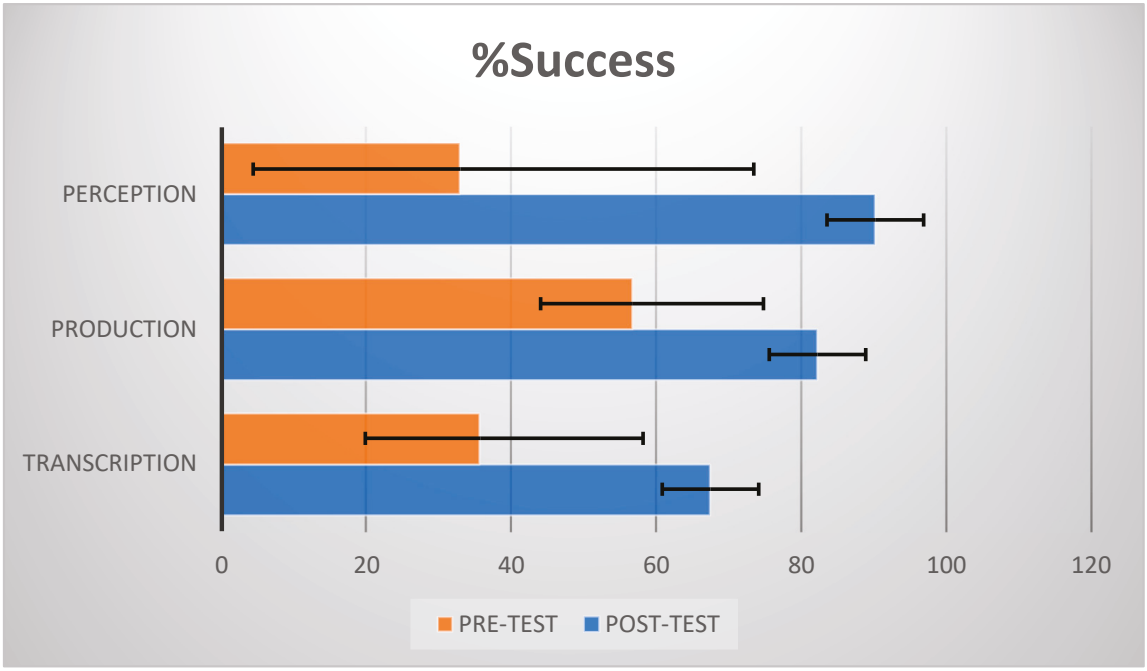


Figure 16. Mean percentage of correct responses in pre-test and post-test with standard error bars.

The observed reduction in score variability, alongside improved mean performance, indicates that learners' outcomes became more homogeneous after using *e-SoundWay*. Beyond facilitating overall progress, the tool appears to promote greater equity and consistency in learning, helping to mitigate disparities in individual achievement—a pedagogically significant benefit for classroom implementation. This pattern aligns with prior research on the convergence effects of serious games in language learning contexts [32,33]. In particular, the joint presence of substantial performance gains and reduced variability mirrors findings from earlier studies on pronunciation training tools, which emphasize the benefits of scaffolded and adaptive learning environments for enhancing both individual and group outcomes (Rogerson-Revell [24]; Neri et al. [25]). Taken as a whole, the current findings reinforce existing literature and provide further empirical support for the pedagogical value of game-based CAPT systems in EFL instruction.

3.2. User Experience Questionnaire

To address RQ2, on the other hand, it should first be noted that the User Experience Questionnaire was completed by 91% of the participants, the vast majority (also 91%) of whom were undergraduate students. Only 9% indicated other occupations, suggesting a relatively homogenous academic cohort. Daily internet usage was high: 64% of students reported being online for between 2 and 4 h daily, 18% between 1 and 2 h, and 18% more than 4 h. Despite this, their engagement with online games was modest, with 64% spending less than 30 min daily, 27% between 30 and 60 min, and only 9% exceeding two hours. While only 21% had prior experience with serious games, this background did not appear to significantly impact their perceived usability or engagement with e-SoundWay. Table 4 summarizes our main results.

Table 4. Summary of user experience questionnaire results.

Question	Summary/Distribution	Interpretive Notes
Q1: Participant type	91% undergraduate	Homogeneous academic cohort
Q2: Daily internet use	64% (2–4 h), 18% (1–2 h), 18% (>4 h)	High internet usage
Q3: Daily gaming time	64% (<30 min), 27% (30–60 min), 9% (>2 h)	Low gaming engagement
Q4: Prior experience with serious games	21% Yes, 79% No	No significant impact on Q5
Q5: Instruction clarity	73% Clear, 18% Not always clear	Generally positive
Q6: Prior experience vs Q5 (Chi-square)	$\chi^2(2, N = 91) = 1.84$, $p = 0.398$	No significant association
Q7: Daily gaming vs engagement (ANOVA)	$F(2, 88) = 4.67$, $p = 0.012$ (Tukey: 30–60 min > rare)	Moderate gamers more engaged
Q8: Perceived game-likeness	55% Game-like, 27% Challenging, 9% Exercise-like	Subjective perceptions varied
Q9: Most difficult phonetic task	64% Transcription, 18% Production, 9% Perception	Friedman $\chi^2(2) = 36.5$, $p < 0.001$; transcription hardest
Q10: Perceived improvement	Strongly Agree (39%), Agree (42%), Neutral (12%), Disagree (6%), No answer (1%)	Positive overall; high agreement
Q11: Overall satisfaction	Very satisfied (36%), Satisfied (45%), Neutral (13%), Dissatisfied (4%), No answer (2%)	Correlated with Q10; generally high satisfaction
Q12: Willingness to recommend	Definitely (49%), Probably (36%), Not sure (10%), Probably not (4%), No answer (1%)	The majority would recommend the tool
Q10 vs Q11 (Pearson correlation)	$r = 0.52$, $p < 0.001$	Moderate, significant correlation

As shown in Table 4, a chi-square test of independence was conducted to examine whether prior experience with serious games was associated with learners' perceptions of usability (Question 5: "Are the instructions to play clear enough?"). The analysis revealed no statistically significant association, $\chi^2(2, N = 91) = 1.84$, $p = 0.398$. This finding supports previous research by Chen et al. [16], which indicated that learners with limited exposure to serious games can still navigate such environments effectively when the interface design is intuitive.

Additionally, one-way ANOVA was conducted to assess whether daily gaming time predicted perceived challenge and enjoyment (Question 7). Results indicated a statistically significant effect of gaming frequency, $F(2, 88) = 4.67$, $p = 0.012$. Post hoc Tukey tests showed that participants who engaged in online gaming for 30–60 min per day rated e-SoundWay as significantly more engaging than those who rarely played. This suggests that moderate gaming experience may enhance appreciation for game-based instructional design, corroborating findings by Davis et al. [34] on learner profiles and engagement in gamified learning contexts.

Regarding the perceived "game-like" nature of e-SoundWay (Question 8), responses were varied. While 55% of participants described the platform as game-like, others noted aspects of monotony or perceived it as more exercise-oriented. These diverse evaluations align with Rogerson-Revell [24], who emphasized that perceptions of game-likeness often depend on factors such as interactivity, novelty, and feedback mechanisms.

In terms of perceived difficulty across the three phonetic categories, most learners identified Transcription as the most challenging. A Friedman test revealed a statistically

significant difference in perceived difficulty, $\chi^2(2) = 36.5$, $p < 0.001$. Subsequent pairwise comparisons using Wilcoxon signed-rank tests (with Bonferroni correction) confirmed that Transcription was perceived as significantly more difficult than both Production ($p < 0.01$) and Perception ($p < 0.001$). These findings are consistent with prior research by Derwing and Munro [6] and the authors in [20], which highlighted the cognitive and metalinguistic demands of transcription tasks.

Finally, learners' responses regarding perceived learning benefits (Q10), overall satisfaction (Q11), and willingness to recommend the platform (Q12) revealed broadly positive evaluations. A Pearson correlation analysis indicated a moderate, statistically significant relationship between perceived improvement and overall satisfaction, r = 0.52, $p < 0.001$. This suggests that students who felt they had improved their phonetic skills were also more satisfied with the platform, a pattern similarly observed in Girard et al. [32], among others.

4. Discussion

The results derived from the pre- and post-tests (RQ1), supported by the user experience questionnaire (RQ2), offer compelling evidence of the platform's pedagogical value and learner acceptability. The quantitative data reveal statistically significant improvements in all three phonetic domains, i.e., perception, production, and transcription, with the former showing the most marked increase. These results confirm that scaffolded, interactive environments, especially those incorporating immediate feedback, can have a substantial impact on learners' auditory discrimination skills [25,45]. Gains in production and transcription similarly indicate that the multimodal structure of *e-SoundWay* effectively supports productive phonological learning by integrating visual, acoustic, and kinesthetic cues [6,24].

In addition to performance improvements, a significant reduction in standard deviation values across all categories in the post-test indicates a narrowing of learner variability. This result suggests not only that the tool was effective but also that it supported more equitable learning outcomes, a finding consistent with prior research on serious games and performance convergence [32,33].

The results concerning learner perception data provide further insight into the usability and motivational dimensions of the tool. Most participants found the platform accessible and the instructions clear, with many highlighting the value of the immediate corrective feedback provided throughout the experience. These features have been widely recognized as crucial to user engagement and sustained attention in digital learning environments [16,24,33]. While some participants described the experience as somewhat monotonous, others specifically praised the game's integration of accent variation, which helped them understand how phonetic differences influence meaning. This aligns with broader findings that exposure to phonetic diversity enhances learners' perceptual tolerance and intelligibility in real-world communication [44,45].

The platform's strong reception among users is also reflected in quantitative indicators from the user experience questionnaire. Notably, 64% of the participants reported being "satisfied" or "very satisfied" with the improvement in their level of phonetic competence after using *e-SoundWay*. Furthermore, 91% of respondents indicated they would recommend the tool to others for learning English pronunciation. These findings further validate the platform's effectiveness and perceived relevance in authentic learning contexts. They also reinforce the conclusion that serious games, when appropriately designed, can significantly enhance learner motivation and engagement, factors critical to successful language acquisition [32,33].

One of the strengths of this initiative lies in its cross-disciplinary design, which brought together technologists, linguists, educators, artists, and native speakers. This collaborative approach ensured both academic rigor and pedagogical relevance, resulting in a tool that is linguistically accurate, culturally meaningful, and technologically accessible. The inclusion of accented characters guiding learners along the Camino de Santiago not only enriched the learning narrative but also promoted phonemic awareness and listening flexibility, both of which are essential in today's global English environments [41,44,47,48].

Importantly, the scope of *e-SoundWay* extends beyond the traditional EFL learner. Its design and functionality make it a valuable resource for educators, speech therapists, and autonomous learners seeking to improve pronunciation. The combination of game-based challenges, cultural storytelling, and multimodal feedback reflects current trends in phonetic pedagogy and responds to calls for accessible, scalable, and contextually sensitive solutions in pronunciation instruction [7,8,29].

5. Conclusions

The findings of this study confirm the potential of *e-SoundWay* as an effective, engaging, and equitable tool for phonetic instruction in EFL contexts. Learners demonstrated statistically significant gains across all three core phonetic competencies, i.e., perception, production, and transcription, and performance disparities were notably reduced after the intervention. These improvements, combined with high levels of user satisfaction, underscore the tool's instructional efficacy. Specifically, 64% of learners reported being satisfied or very satisfied with their progress in English phonetics, and 91% indicated they would recommend *e-SoundWay* to peers, a strong endorsement of the platform's educational impact.

Beyond its effectiveness, the platform represents a notable innovation in the field of digital language learning. It brings together over 600 interactive multimedia minigames, real-time feedback mechanisms, and an immersive narrative experience grounded in the culturally meaningful route of the *Camino de Santiago*. This fusion of pedagogy, culture, and game mechanics introduces a novel model for pronunciation training that goes well beyond traditional drill-based approaches. By integrating accented characters and situational diversity, the platform promotes phonological flexibility and real-world listening comprehension, preparing learners to navigate global English interactions.

Importantly, *e-SoundWay* reimagines how phonetics can be taught, not as an isolated theory, but as an interactive, exploratory, and personalized learning journey. Its use of multimodal input, including visual, auditory, and kinesthetic elements, supports differentiated learning and encourages learner autonomy. The cross-disciplinary development model behind the platform, uniting technologists, linguists, educators, and native speakers, ensures that it is both pedagogically sound and adaptable across instructional settings.

By positioning phonetic training within a serious game framework, the tool responds to increasing calls for more motivational, learner-centered, and scalable educational tools. It challenges conventional limitations in pronunciation pedagogy by offering a flexible, data-informed, and culturally enriched alternative that aligns with modern language learners' needs.

Overall, *e-SoundWay* changes the educational landscape by demonstrating how multimodal technologies can deliver phonetic instruction that is both academically rigorous and emotionally engaging. Its novel approach holds promise not only for language students and educators but also for speech therapists and lifelong learners. Future work should explore its longitudinal impact, classroom integration, and adaptability to other L1 back-

grounds, contributing to the continued evolution of pronunciation teaching in multilingual and multicultural contexts.

Author Contributions: Conceptualization, M.Á.G.-G.; methodology, M.Á.G.-G.; software, A.L.-F. and J.C.L.-A.; validation, M.Á.G.-G. and A.L.-F.; formal analysis, M.Á.G.-G. and A.L.-F.; investigation, M.Á.G.-G., A.L.-F. and J.C.L.-A.; resources, M.Á.G.-G.; writing—original draft preparation A.L.-F.; writing—review and editing, M.Á.G.-G., A.L.-F. and J.C.L.-A.; project administration, A.L.-F. and M.Á.G.-G. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported through a project funded by the Spanish Ministry of Science and Innovation (PID2019-105678RB-C21/C22), a grant corresponding to the 2021 call for "Strategic Projects Oriented towards Ecological Transition and Digital Transition" of the State Programme to Promote Scientific-Technical Research and its Transfer, of the State Plan for Scientific, Technical and Innovation Research 2021–2023, within the framework of the Recovery, Transformation and Resilience Plan. (TED 2021-130283B-C21/C22) and a grant the groups of Competitive Reference of the Xunta de Galicia (ED431C2023/15).

Institutional Review Board Statement: The study approved by the COMITÉ DE ÉTICA NA INVESTIGACIÓN DA USC (protocol code USC 96/2024, approval date 3rd March 2023).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

Acknowledgments: The authors would like to thank Teresa Sánchez Roura for allowing the students of the second-year English Phonetics and Phonology course of the English Language and Literature degree at the Universidade de Santiago de Compostela to work with this educational resource and to carry out the surveys necessary to obtain the results discussed in the previous section.

Conflicts of Interest: The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

1. Language Policy Programme Education Policy Division, Education Department, Council Europe. *Common European Framework of Reference for Languages: Learning, Teaching, Assessment—Companion*; Editorial Council of Europe Publishing: Strasbourg, France, 2020; 278p.
2. Camacho Marti, M.; Estévez-González, V.; Esteve Mon, F.M.; Gisbert Cervera, M.M.; Lázaro Cantabrana, J.L. Capítulo XIV. Indicadores de Calidad para el uso de las TIC en los Centros Educativos. Compartir Aprendizaje. Editorial Universitat Rovira i Virgili, Tarragona, Sapin. 2014. Available online: https://hdl.handle.net/20.500.11797/imarina9149374 (accessed on 3 June 2025).
3. Salinas, J. Innovación Educativa y Uso de las TIC. Editorial Universidad Internacional de Andalucía, Andalucía, Spain, 2008-09, 15–30. Available online: http://hdl.handle.net/10334/136 (accessed on 12 May 2025).
4. Clark, R.C.; Mayer, R.E. *E-Learning and the Science of Instruction: Proven Guidelines for Consumers and Designers of Multimedia Learning*, 3rd ed.; Pfeiffer: Zurich, Switzerland, 2011; 528p.
5. Jordan, A.; Carlile, O.; Stack, A. *Approaches to Learning: A Guide for Teachers*; McGraw Hill Education: London, UK, 2008; 278p.
6. Derwing, T.M.; Munro, M.J. Pronunciation Fundamentals: Evidence-based perspectives for L2 teaching and research. In *Language Learning & language Teaching*; John Benjamin Publishing Company: Amsterdam, The Netherlands, 2015; 208p.
7. Levy, M. Culture, culture learning and new technologies: Towards a pedagogical framework. *Lang. Learn. Lang. Teach.* **2007**, *11*, 104–127.
8. Hou, Z.; Avadoust, V. A review of the methodological quality of quantitative mobile-assisted language learning research. *System* **2021**, *100*, 102568. [CrossRef]
9. Lago Ferreiro, A.; Gómez González, M.Á.; Fragueiro Agrelo, Á.; Llamas Nistal, M. Mulit-platform application for learning English phonetics: Serious Game. In Proceedings of the XV Technologies Applied to Electronics Teaching Conference, Teruel, Spain, 29 June–1 July 2022. [CrossRef]

10. O'Malley, C.; Vavoula, G.N.; Glew, J.P.; Taylor, J.; Sharples, M.; Lefrere, P.; Lonsdale, P.; Naismith, L.; Waycott, J. Guidelines for Learning/Teaching/Tutoring in a Mobile Environment. Mobilearn Project Deliverable. 2005. Available online: https://www.researchgate.net/publication/280851673_Guidelines_for_learningteachingtutoring_in_a_mobile_environment (accessed on 1 December 2024).
11. Hwang, W.Y.; Chen, H.S.L.; Shadiev, R.; Huang, R.Y.-M.; ChenImproving, C.-Y. English as a foreign language writing in elementary schools using mobile devices in familiar situational context. *Comput. Assist. Lang. Learn.* **2014**, *27*, 359–378. [CrossRef]
12. Hwang, W.Y.; Chen, H.S. Users' familiar situational contexts facilitate the practice of EFL in elementary schools with mobile devices. *Comput. Assist. Lang. Learn.* **2013**, *26*, 101–125. [CrossRef]
13. Wang, Y.H. Integrating self-paced mobile learning into language instruction: Impact on reading comprehension and learner satisfaction. *Interact. Learn. Environ.* **2017**, *25*, 397–411. [CrossRef]
14. Wrigglesworth, J. Using smartphones to extend interaction beyond the EFL classroom. *Comput. Assist. Lang. Learn.* **2019**, *33*, 413–434. [CrossRef]
15. Burston, J. Twenty years of MALL project implementation: A meta-analysis of learning outcomes. *ReCALL* **2015**, *27*, 4–20. [CrossRef]
16. Chen, C.M.; Chen, L.C.; Yang, S.M. An English vocabulary learning app with self-regulated learning mechanism to improve learning performance and motivation. *Comput. Assist. Lang. Learn.* **2019**, *32*, 237–260. [CrossRef]
17. Dehghan, F.; Rezvani, R.; Faceli, S. Social networks and their effectiveness in learning foreign language vocabulary: A comparative study using WhatsApp. *CALL-EJ* **2017**, *18*, 1–13.
18. Elekaei, A.; Tabrizi, H.H.; Chalak, A. Investigating the Effects of EFL Learners'vocabulary gain and retention levels on their choice of memory and compensation strategies in an e-learning project. *CALL-EJ* **2019**, *20*, 1–18.
19. Gómez González, M.Á.; Sánchez Roura, T. *English Pronunciation for Speakers of Spanish: From Theory to Practice*; Editorial De Gruyter Mouton: Berlin, Germany, 2016.
20. Gómez González, M.Á.; Lago Ferreiro, A. Computer-Assisted Pronunciation Training (CAPT): An empirical evaluation of EPSS Multimedia Lab. *Lang. Learn. Lang. Teach.* **2024**, *28*, 1–44.
21. Gómez González, M.Á.; Lago Ferreiro, A. Web-assisted instruction for teaching and learning EFL phonetics to Spanish learners: Effectiveness, perceptions and challenges. *Comput. Educ. Open* **2024**, *7*, 100214. [CrossRef]
22. Hardisty, D.; Windeatt, S. *Computer-Assisted Language Learning*; Oxford University Press: Oxford, UK, 1989; 165p.
23. Chansarian-Dehkordi, F.; Ameri-Golestan, A. Effects of mobile learning on acquisition and retention of vocabulary among Persian-speaking EFL learners. *CALL-EJ* **2016**, *17*, 43–56.
24. Rogerson-Revell, P.M. Computer-Assisted Pronunciation Training (CAPT): Current issues and future directions. *RELC J.* **2021**, *52*, 189–205. [CrossRef]
25. Colferai, E.; Gregory, S. Minimizing attrition in online degree courses. *J. Educ. Online* **2015**, *12*, 62–90. [CrossRef]
26. Stracke, E. A road to understanding: A qualitative study into why learners drop out of a blended language learning (BLL) environment. *ReCALL* **2007**, *19*, 57–78. [CrossRef]
27. Vaz de Carvalho, C.; Coelho, A. Game-Based Learning, Gamification in Education and Serious Games. *Computers* **2022**, *11*, 36. [CrossRef]
28. Ko, M.H. Students' reactions to using smartphones and social media for vocabulary feedback. *Comput. Assist. Lang. Learn.* **2019**, *32*, 920–944. [CrossRef]
29. Fouz-González, J. Pronunciation instruction through twitter: The case of commonly mispronounced words. *Comput. Assist. Lang. Learn.* **2017**, *30*, 631–663. [CrossRef]
30. Sun, Z.; Lin, C.-H.; You, J.; Shen, H.J.; Qi, S.; Luo, L. Improving the English-speaking skills of young learners through mobile social networking. *Comput. Assist. Lang. Learn.* **2017**, *30*, 304–324. [CrossRef]
31. Maskeliünas, R.; Kulikajevas, A.; Blažauskas, T.; Damaševičius, R.; Swacha, J. An Interactive Serious Mobile Game for Supporting the Learning of Programming in JavaScript in the Context of Eco-Friendly City Management. *Computers* **2020**, *9*, 102. [CrossRef]
32. Girard, C.; Ecalle, J.; Magnan, A.A. Serious games as new educational tools: How effective are they? A meta-analysis of recent studie. *J. Comput. Assist. Learn.* **2013**, *29*, 207–219. [CrossRef]
33. Subhash, S.; Cudney, E.A. Gamified learning in higher education: A systematic review of the literature. *Comput. Hum. Behav.* **2018**, *87*, 192–206. [CrossRef]
34. Davis, K.; Sridharan, H.; Koepke, L.; Singh, S.; Boiko, R. Learning and engagement in a gamified course: Investigating the effects of student characteristics. *J. Comput. Assist. Learn.* **2018**, *34*, 492–503. [CrossRef]
35. Da Silva, J.P.; Silveira, I.F. A Systematic Review on Open Educational Games for Programming Learning and Teaching. *Int. J. Emerg. Technol. Learn.* **2020**, *15*, 156–172. [CrossRef]

36. Unity User Manual 2023.1. Available online: https://docs.unity3d.com/2023.1/Documentation/Manual/UnityManual.html (accessed on 5 January 2024).
37. Recognissimo Documentation. Available online: https://bluezzzy.github.io/recognissimo-docs/ (accessed on 5 January 2024).
38. Allen, M.L.; Hartley, C.; Cain, K. Do iPads promote symbolic understanding and word learning in children with autism? *Front. Psychol.* **2015**, *6*, 138. [CrossRef] [PubMed]
39. Lems, K. New Ideas for Teaching English Using Songs and Music. *Engl. Teach. Forum* **2018**, *56*, 14–21. Available online: https://americanenglish.state.gov/files/ae/resource_files/etf_56_1_pg14-21_0.pdf (accessed on 3 June 2025).
40. Peterson, M.; Jabbari, N. *Digital Games in Language Learning: Case Studies and Applications*; Routledge Taylor & Francis Group: Oxfordshire, UK, 2023; 202p.
41. Strange, W.; Shafer, V.L. Speech perception in second language learners: The re-education of selective perception. In *Phonology and Second Language Acquisition*; John Benjamins Publishing Company: Amsterdam, The Netherlands, 2008; pp. 153–191.
42. Bradlow, A.R.; Pisoni, D.B.; Akahane-Yamada, R.; Tohkura, Y. Training Japanese listeners to identify English /r/ and /l/: Long-term retention of learning in perception and production. *Percept. Psychophys.* **1997**, *59*, 87–100. [CrossRef]
43. Lengeris, A.; Hazan, V. The effect of native vowel processing ability and frequency discrimination acuity on the phonetic training of English vowels for native Greek listeners. *J. Acoust. Soc. Am.* **2010**, *128*, 3757–3768. [CrossRef]
44. Jenkins, J.; Setter, J. Teaching English pronunciation: A state of the art review. *Lang. Teach.* **2005**, *38*, 1–17. [CrossRef]
45. Munro, M.J.; Derwing, T.M. The functional load principle in ESL pronunciation instruction: An exploratory study. *System* **2006**, *34*, 520–531. [CrossRef]
46. Wang, Y.; Munro, M.J. Computer-based training for learning English vowel contrasts. *System* **2004**, *32*, 539–552. [CrossRef]
47. Lively, S.E.; Logan, J.S.; Pisoni, D.B. Training Japanese listeners to identify English /r/ and /l/: II. The role of phonetic environment and talker variability in learning. *J. Acoust. Soc. Am.* **1993**, *94*, 1242–1255. [CrossRef]
48. Pardo, J.S. On phonetic convergence during conversational interaction. *J. Acoust. Soc. Am.* **2006**, *119*, 2382–2393. [CrossRef] [PubMed]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

Multimodal Technologies and Interaction

MDPI

Article

Developing Comprehensive e-Game Design Guidelines to Support Children with Language Delay: A Step-by-Step Approach with Initial Validation

Noha Badkook *, Doaa Sinnari and Abeer Almakky

Information Technology Department, Faculty of Computing and Information Technology, King AbdulAziz University, Jeddah 21589, Saudi Arabia; dsinnari@kau.edu.sa (D.S.); aalmakky@kau.edu.sa (A.A.)
* Correspondence: nbadkook0008@stu.kau.edu.sa

Abstract: e-Games have become increasingly important in supporting the development of children with language delays. However, most existing educational games were not designed using usability guidelines tailored to the specific needs of this group. While various general and game-specific guidelines exist, they often have limitations. Some are too broad, others only address limited features of e-Games, and many fail to consider needs relevant to children with speech and language challenges. Therefore, this paper introduced a new collection of usability guidelines, called eGLD (e-Game for Language Delay), specifically designed for evaluating and improving educational games for children with language delays. The guidelines were created based on Quinones et al.'s methodology, which involves seven stages from the exploratory phase to the refining phase. eGLD consists of 19 guidelines and 131 checklist items that are user-friendly and applicable, addressing diverse features of e-Games for treating language delay in children. To conduct the first validation of eGLD, an experiment was carried out on two popular e-Games, "MITA" and "Speech Blubs", by comparing the usability issues identified using eGLD with those identified by Nielsen and GUESS (Game User Experience Satisfaction Scale) guidelines. The experiment revealed that eGLD detected a greater number of usability issues, including critical ones, demonstrating its potential effectiveness in assessing and enhancing the usability of e-Games for children with language delay. Based on this validation, the guidelines were refined, and a second round of validation is planned to further ensure their reliability and applicability.

Keywords: usability; usability guidelines; e-Game; usability guidelines for children with language delay; eGLD

1. Introduction

The ability of a child to speak is one of the most crucial stages of language development in early childhood. Language development is considered a very important aspect of a child's life and one of the best indicators of educational achievement [1]. A "delay in language" describes the proper sequence of speech and language development that occurs at a slower rate than expected [2]. Children with language delays experience delays in the speaking process when compared to other children in their age; children who experience language delays will exhibit difficulties with letter-arrangement completeness, articulation, sound, and fluency [3]. There are factors that cause language delay in children, such as psychological factors, which include meeting the child's needs without verbal interactions or the child living in a quiet environment

without much communication with others. Other causes of language delay could be related to intrinsic factors such as autism, hearing loss, or mental disability [4]. Language delays among children are a serious problem that requires immediate attention. It is one of the most common causes of developmental disorders in children [5]. A delay in treating this problem may cause the child to become frustrated, withdraw from social relationships, and isolate. It may also have a negative impact on the child's academic achievements and quality of life [4]. When a child reaches the age of one year and cannot pronounce their first words, or their speech is unclear, or they lag behind their peers, fears arise among parents and those around him, who begin searching for the appropriate evaluations and treatments to address the problem [6]. Traditional treatment methods that are followed in therapeutic sessions to improve communication skills in children include group language intervention, individual speech and language therapy, language intervention implemented by parents, and main intervention techniques such as observation, waiting, listening, paraphrasing, focused stimulation, and visual input use [7]. However, the numerous treatment sessions and the extended duration of therapy can make the child feel bored and as if they are not having fun. Therefore, the use of modern technology to treat linguistic delay in children leads to an ideal and satisfactory result [8]. Nowadays, technical solutions and modern techniques are used in the form of e-Games or applications created specifically to treat language delay in children [4]. The therapeutic approach through e-Games has garnered great interest from healthcare practitioners and has proven its effectiveness when integrating e-Games and therapeutic goals [8], which has a positive impact on learning through motivation and engagement development [9]. Therefore, researchers confirm that combining therapeutic content with e-Games leads to the creation of engaging and stimulating therapeutic environments, and that e-Games have great potential in stimulating the cognitive capacities of various target groups. e-Games also find excellent acceptance among people with special needs because they feel safe and comfortable exploring the virtual world and receiving prompt feedback, free from concerns about social exclusion [10].

Despite the rise in recent research on the utilization of therapeutic e-Games to stimulate language abilities, there is a gap in modern research in the field of e-Games. This gap is related to the fact that existing e-Games used to treat linguistically delayed children are not based on any treatment or design guidelines [10], which prevents these e-Games from achieving the intended therapeutic outcomes effectively.

We hypothesize that the proposed eGLD guidelines will detect a higher number of and more critical usability issues than existing guidelines such as Nielsen and GUESS, particularly when applied to therapeutic e-Games for children with language delay. This study aims to close this gap by outlining a comprehensive set of guidelines for e-Games, which will aid future designers and developers in creating e-Games specifically for children with language delays, titled eGLD, and will also support evaluating existing e-Games; eGLD guidelines meet the following criteria:

1. Simple to follow and understand for designers, developers, and evaluators.
2. Suitable for all users, especially children with linguistic delays.
3. Specialized in designing and evaluating e-Games that support language development and incorporate diverse interactive game and therapy features to enhance learning and engagement.

This study has two primary objectives. The initial objective is to create eGLD using Quinones et al.'s methodology [11]. The second objective is to assess the effectiveness of eGLD as part of the first stage of validation, according to the Quinones methodology, by comparing it with the well-known Nielsen and GUESS guidelines. This comparison was carried out in Saudi Arabia using two popular e-Games designed for language development in children: "MITA" and "Speech Blubs".

The rest of the paper is organized as follows: Section 2 provides background information and a literature review on language delay, usability guidelines, and the methodology for developing new heuristics. Section 3 describes the development of eGLD. Section 4 presents the results of developing and validating the proposed guidelines. Section 5 discusses the key findings, while Section 6 concludes the paper and outlines future work.

2. Background Information and Related Work

Children often begin to pronounce their first words at the age of 12 months and then start combining words to form sentences before the age of two years [12]. These periods are considered sensitive for language development in children, and those who fail to achieve linguistic development in their early years are vulnerable to problems in receptive and expressive language during school age and beyond [12]. Children who have a language delay suffer from a delay in the speaking process compared to healthy children of the same age [3]. The affected children will show limitations in voice, speech, letter order, and fluency. Language delay can also be described by the children's inability to express or speak clearly, limited vocabulary, difficulty engaging in conversations with parents, and difficulties in understanding spoken words [3].

Two basic factors affect Speech and language abilities: intrinsic factors and psychosocial factors. Intrinsic factors include innate conditions such as the physiology of the organs involved in language and speech skills, including cognitive impairment, hearing loss, autism, selective mutation, cerebral palsy, and abnormalities in the speech organs. Psychosocial factors can take the form of stimuli surrounding the child, such as a quiet environment, twins, bilingualism, incorrect teaching methods, and television viewing patterns [5].

2.1. Usability and Guidelines

Usability is considered one of the important areas in human–computer interaction (HCI), and it is an essential feature because it is necessary and important for the system to be well designed and highly usable [8,13]. In 2016, the ISO updated the standard 9241-11, which focuses on ergonomics in human–system interactions. This standard defines usability as how well a system, product, or service helps specific users achieve their goals effectively, efficiently, and satisfactorily within a given context [14]. This standard was reviewed and confirmed in 2023 [15].

Usability can be improved and tested by following specific guidelines. These guidelines are helpful because they can be used early in the design process to evaluate and discover problems before the app is released [16]. Nielsen's principles are some of the first and most well-known usability guidelines. They were created by Nielsen and Molich in 1990, and these guidelines are widely applicable to most user interfaces [17]. However, traditional usability rules might not assess unique features for specific domains. Because of this, some experts have created new usability rules by updating Nielsen's guidelines or adding new ones to better fit specific needs [11]. Moreover, M. Muhanna et al. [18] proposed that Nielsen's criteria are insufficient for evaluating games—especially Arabic mobile games—because they are general and do not cover all aspects of the game. Therefore, we need specific criteria for e-Games to ensure that important aspects are not overlooked when evaluating game usability and enjoyment.

Usability guidelines often come with a checklist that helps in assessing applications tailored to specific domains [11]. The checklist items can add value to categorization and evaluation by identifying unique game characteristics and guiding the design process. They provide a structured approach for developers to ensure all essential features are met,

define requirements, and support more effective communication and discussion among game developers about game design. This shared language fosters collaboration across teams and across the various disciplines involved in game development [19].

Several studies have used different guidelines to design and evaluate e-Games. Nielsen's principles [20] comprise ten usability guidelines, which are:

1. Visibility of system status.
2. Similarity between the system and the real world.
3. User control and freedom.
4. Consistency and standards.
5. Error prevention.
6. Recognition rather than recall.
7. Flexibility and efficiency of use.
8. Aesthetic and minimalist design.
9. Help users recognize and recover from errors.
10. Help and documentation.

J. Phan et al. [21] developed and validated the GUESS, with 9 guidelines and 55 checklist items for gauging video-game satisfaction. Overall, the GUESS was shown to have satisfactory convergent and discriminant validity, as well as good content validity and internal consistency. Additionally, the GUESS was created and verified using evaluations of more than 450 distinct video-game titles from a variety of well-known genres (such as action adventure and role-playing). The GUESS comprises nine usability guidelines, which are:

1. Usability/playability.
2. Narratives.
3. Play engrossment.
4. Enjoyment.
5. Creative freedom.
6. Audio aesthetics.
7. Personal gratification.
8. Social connectivity.
9. Visual aesthetics.

The results showed that the GUESS can be applied to video-game players with different levels of experience (e.g., novice, professional, expert) who play different kinds of games. To generalize the GUESS to further gaming genres, more investigation is also required. In particular, the majority of the games assessed in studies were commercial games created just for entertainment purposes. Furthermore, most of these games were enjoyed by players rather than disliked [21,22]. This means that GUESS might not cover all the important usability points when it comes to serious or therapeutic e-Games. These types of e-Games have different goals and need to be looked at in a different way. So, using GUESS alone might not be enough for e-Games made for learning or therapy.

2.2. Developing New Usability Guidelines

Guidelines are the set of important concepts and rules for user interface (UI) design. These guidelines are based on the human psyche [10]. To create a new set of usability guidelines, specific methodologies must be followed that are systematic and formal such as Quinones [23]. Quinones' methodology is divided into seven primary phases, which include:

1. *Exploratory stage:* gathering information from previous research about the application domain for which a new set of heuristics will be created, the application's features and attributes.
2. *Descriptive Stage:* highlighting the most important information collected from the previous stage to formulate the main concepts of the research.
3. *Correlational stage:* identify specific features of the application that require evaluation using usability heuristics.
4. *Selection stage:* determine the process used for each guideline. Using the existing heuristics collected, you have the option to retain, modify, or remove the heuristics, or develop new guidelines.
5. *Specification stage:* formally define the new set of proposed guidelines using a standard template.
6. *Validation stage:* validation of the guideline set through multiple experiments.
7. *Refinement stage:* after obtaining feedback from the previous phase, the new set of guidelines is improved.

The benefit of using Quinones et al.'s method is that it gives a clear and organized way to explain each guideline. Quinones et al.'s method includes details such as the guideline's ID, priority, name, definition, application features, where it can be used, examples, benefits, checklists, and usability features. Additionally, you can repeat and improve different parts as many times as needed to make the guidelines better or to improve how experiments are carried out during testing [23].

Validating Usability Guidelines

In Quinones et al.'s [23] methodology, there are three ways to verify the methodology.

1. *Guideline evaluation:* to check the proposed guidelines against traditional or specialized guidelines in terms of the number of usability issues detected (as recommended by Quinones).
2. *Expert judgment:* use a questionnaire that assesses the perceptions and opinions of evaluators and experts of the proposed usability guidelines (to obtain additional feedback).
3. *User testing:* to obtain user opinions about a set of usability guidelines (to receive additional feedback).

This study applied the first type of validation by comparing eGLD with existing guidelines to assess its effectiveness. The findings are presented below. We plan to use the remaining two methods, expert judgment and user testing, in future work to further enhance the guidelines.

3. Materials and Methods

This section outlines the process for developing and creating the eGLD guidelines, as well as the procedures for validating the results from the first stage.

3.1. Developing eGLD

The literature review identified several usability guidelines that are used to assess e-Games and applications. Some of these guidelines, such as Nielsen's usability heuristics [20], are universally applicable and used in a variety of fields, including e-Games. Others, like GUESS [21], are designed especially to evaluate user experience and satisfaction in video games, emphasizing important elements like fun, involvement, and creative freedom. As discussed in Section 2, each of these sources has limits even though they all offer insightful information about usability and user experience.

Inspired by these earlier researches, we created a comprehensive set of usability guidelines known as eGLD that are specifically designed to evaluate and guide the development of e-Games for children with language delays. Nielsen's heuristics, GUESS, and the results of pertinent e-Game experiments were among the well-established studies from which the eGLD guidelines and checklist items were derived. In order to guarantee a structured and systematic development process, we used the formal approach suggested by Quiñones et al. [23], which was previously discussed in Section 2.

3.1.1. Stage 1: Exploratory Stage

In this stage, a systematic literature review was carried out to obtain relevant information pertinent to the design and evaluation of e-Games for children with language delays. The review aimed to identify critical gaming features, usability attributes, and design guidelines that promote effective learning for children with language delays. To perform the review, we searched for the following sets of keywords:

1. For features: ("e-Games" OR "mobile game") AND "features";
2. For attributes: ("e-Games" OR "mobile game") AND "usability attributes";
3. Guidelines and recommendations: ("e-Games" OR "mobile game") AND ("usability guidelines" OR "usability principles").

We looked for relevant material published between 2017 and 2024 in several academic databases, including IEEE, Springer Link, ScienceDirect, and Google Scholar.

The inclusion criteria for choosing studies are as follows: (1) focus on e-Games; (2) detailed descriptions of relevant characteristics, attributes, guidelines, or suggestions; (4) publicly available full-text studies; and (5) English-language articles. The exclusion criteria were studies (1) without complete text, (2) articles unrelated to e-Games, (3) papers lacking an extensive explanation of essential design features, and (4) broad survey papers with no specific conclusions.

As shown in Figure 1, a total of 42 records were identified through database searching. After title and abstract screening, 31 records were retained, while 9 were excluded based on the predefined inclusion and exclusion criteria. Then, 18 full-text articles were assessed for eligibility, and 8 were excluded according to the same criteria. In the end, 10 studies were included in the final analysis: 5 studies focused on features, 3 on usability attributes, and 2 on design guidelines.

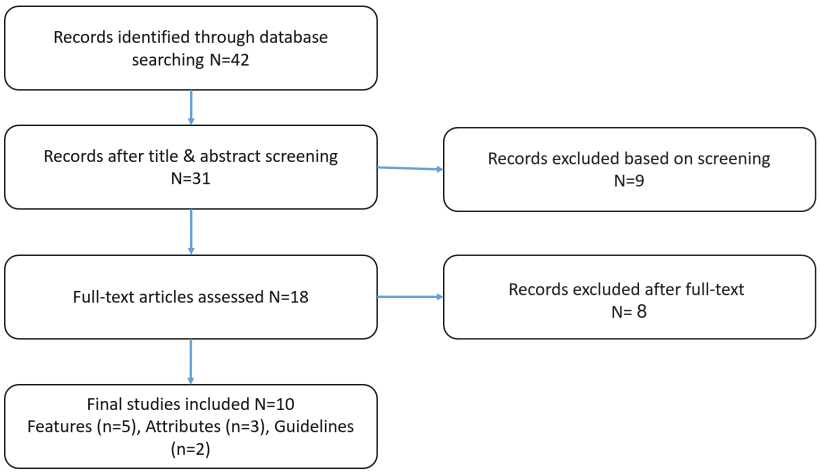


Figure 1. A flowchart of the systematic literature review selection for eGLD.

The information gathered from this stage is displayed below.

- **Features:** Features refer to the unique functions or capabilities of an application that distinguish it from other types of apps. These features are particularly relevant for design e-Games for children with language delay or disabilities [10,24–27], and will be examined in the next stage.
- **Usability attributes:** Attributes are the criteria used to evaluate the quality of mobile applications during usability testing, using specific usability metrics. Which include:
  - Nielsen's attributes: learnability, efficiency, memorability, errors and satisfaction [28].
  - ISO attributes: effectiveness, efficiency, and satisfaction) [29].
  - PACMAD (People At the Centre of Mobile Application Development) attributes: combine elements from both ISO and Nielsen usability models, while also adding the cognitive load aspect. These attributes include effectiveness, efficiency, learnability, memorability, errors, satisfaction, and cognitive load [30].
- **Usability guidelines:** These serve as the basic principles, and include:
  - Nielsen's usability heuristics [20].
  - GUESS guidelines for the evaluation of e-Games [21].

3.1.2. Stage 2: Descriptive Stage

- **For features:** We collected all the features from [10,24–27]. We added four language delay-specific features, combining them with the existing ones to create a comprehensive list of 28 integrated features, as shown below:
  1. Identification with the game [10].
  2. Interface design [10].
  3. Layout [10].
  4. Demonstrations [10].
  5. Reward/encouragement [10,25–27].
  6. Performance feedback and guidance [10,24–27].
  7. Personalization [10].
  8. Adaptive games and challenges [10,26,27].
  9. Social interaction [10].
  10. Mobility [10].
  11. Time management/restriction [10].
  12. Repetition and rehearsal of skills [10].
  13. Motivation and engagement [10,24].
  14. Motor skill [10].
  15. Cognitive development [10].
  16. Speech recognition [24].
  17. Emotional attachment [24].
  18. Animation and engaging graphics [25].
  19. Background music and sound effects [25].
  20. Points [26,27].
  21. Story or theme [26].
  22. Show progress [26].
  23. Provides badges for achievements [26].
  24. Provides clear goals [26].
  25. Gradual language complexity.
  26. Review of previous words.
  27. Play-based therapy.

28. Syntax and sentence formation.

- **Usability attributes:** We focused exclusively on attributes defined by PACMAD attributes [30]. Below are the definitions of each attribute [30].
  - Effectiveness: the capacity to accomplish the intended results.
  - Efficiency: minimizing wasted time, effort, money, and resources.
  - Satisfaction: meeting the user's expectations and making them feel fulfilled.
  - Learnability: how quickly and easily users can figure out how to use the app.
  - Memorability: how well users can remember how to use the app after not using it for a while.
  - Errors: reducing the user's error rate when using the program.
  - Cognitive load: the capacity to operate a mobile application while carrying out daily tasks.
- **For usability guidelines:** We divided the relevant research into two categories: main guidelines, which will serve as the basic principles for the next stages, and checklist items, which will customize the main guidelines more particularly to e-Games for children with language delay.
  - **Main guidelines:** we selected all Nielsen's usability heuristics [20], and all GUESS guidelines [22].
  - **Checklist items:** we selected all the checklists for designing e-Games for children with language delays [10].

3.1.3. Stage 3: Correlational Stage

Table 1 presents how the identified features, attributes, and key usability guidelines were aligned during this stage. This step aims to define the unique features clearly and attributes of eGLD that should be evaluated using the main set of guidelines [23]. The matching process is feature-driven, meaning we first matched attributes that can measure each feature, then chose the relevant main guidelines (from Nielsen and GUESS) labeled N and G, that relate to those attributes. The codes N1–N10 refer to Nielsen's ten usability heuristics, and G1–G9 correspond to the GUESS game usability guidelines. Their definitions are as follows:

- **Nielsen:** N1—Visibility of system status; N2—Similarity between the system and the real world; N3—User control and freedom; N4—Consistency and standards; N5—Error prevention; N6—Recognition rather than recall; N7—Flexibility and efficiency of use; N8—Aesthetic and minimalist design; N9—Help users recognize and recover from errors; N10—Help and documentation.
- **GUESS:** G1—Usability/Reproduction; G2—Narratives; G3—Playing involvement; G4—Pleasure; G5—Creative freedom; G6—Audio aesthetics; G7—Personal gratification; G8—Social connectivity; G9—Visual aesthetics.

As an example, we examined the "Adaptive games and challenges" feature. The most suitable attributes and guidelines that it may be compared to:

- Attributes:
  - Cognitive load: to determine if the game adapts its level of difficulty to the player's skill level, making sure it is neither too easy nor too challenging.
  - Satisfaction: to gauge how gratifying and pleasurable the gaming is when the difficulties are suitably distributed.
- Guidelines:

- Similarity between the system and the real world (N2) in Nielsen: guarantees that the game's complexity and substance are in line with user comprehension and real-world expectations.
- Help users recognize recover from errors (N9): when players are having trouble, the game may gently adapt or provide guidance, turning challenging situations into teaching opportunities rather than frustrating ones.
- Play engrossment (G7), creative freedom (G5), and playing involvement (G3) in GUESS: adaptive gameplay keeps players deeply engaged throughout the game, boosts user motivation, and enables them to make meaningful decisions.

Table 1. Mapping between features, attributes, and usability guidelines.

No.	Feature	PACMAD Attributes	Nielsen	GUESS
F1	Identification with the game	Satisfaction	N2	G2, G7
F2	Interface design	Effectiveness, Efficiency, Learnability	N4	G1, G9
F3	Layout	Efficiency, Satisfaction, Memorability	N3, N8	G9
F4	Demonstrations	Learnability, Cognitive Load	N1, N10	G1
F5	Reward/encouragement	Satisfaction	–	G7, G4
F6	Performance feedback and guidance	Effectiveness, Errors	–	–
F7	Personalization	Effectiveness, Satisfaction	N7	G3, G7
F8	Adaptive games and challenges	Cognitive Load, Satisfaction	N2, N9	G3, G5, G7
F9	Social interaction	Satisfaction	–	G8
F10	Mobility	Efficiency	N7	–
F11	Time management/restriction	Efficiency, Cognitive Load	N3	–
F12	Repetition and rehearsal of skills	Memorability, Learnability	–	–
F13	Motivation and engagement	Satisfaction	N6	G3, G4
F14	Motor skill	Effectiveness	N3	G1
F15	Cognitive development	Effectiveness, Cognitive Load	–	G3, G7
F16	Speech recognition	Effectiveness, Efficiency	N1, N2	G1
F17	Emotional attachment	Satisfaction	–	G7
F18	Animation and engaging graphics	Satisfaction, Memorability	N8	G9
F19	Background music and sound effects	Satisfaction, Cognitive Load	–	G6
F20	Points	Satisfaction	–	G4, G7
F21	Story or theme	Satisfaction, Learnability	–	G2, G5
F22	Show progress	Effectiveness, Satisfaction	N1	G1
F23	Provides badges for achievements	Satisfaction	–	G4, G7
F24	Provides clear goals	Effectiveness	N2, N5	–
F25	Gradual language complexity	Effectiveness, Cognitive Load	–	–
F26	Review of previous words	Memorability	–	–
F27	Play-based therapy	Satisfaction	–	–
F28	Syntax and sentence formation	Effectiveness	–	–

3.1.4. Stage 4: Selection Stage

In this step, we evaluated and categorized every usability guideline found in the previous stage according to how well it addressed the unique features of e-Games designed for children with language delay. Four categories were used for classification [23]:

- Keep: since the guideline was already in accordance with the desired qualities, it was kept exactly as is.
- Adapt: to properly represent the context of e-Game usability, the checklist elements in the guideline were expanded or improved.
- Eliminate: because the guideline was redundant or irrelevant to the target domain, it was eliminated.
- Create: to assess a feature not addressed by the previous rules, a new guideline was developed.

We kept all of Nielsen's guidelines, but we made some changes to the checklists to make them better fit e-Games, especially games that focus on learning and interaction. From the GUESS guidelines, we only kept the ones that made sense for our project. For example, we kept things like *play involvement*, *pleasure*, *creative freedom*, and *internal gratification* because they help make games more fun and engaging. But we removed others like *narrative* and *visual aesthetics* because they were either too similar to what we already had in Nielsen's guidelines.

While working on this, we noticed that some important features like *gradual language complexity* and *interactive feedback* were not really covered by either Nielsen or GUESS. So, we created five new guidelines to focus on these important areas that help kids learn through games.

Each guideline (whether old or new) was given a priority: (1) useful, (2) important, or (3) critical—depending on how much it helps with the learning goals of the e-Game. The final results of this stage in Tables 2–4.

Table 2. Choosing guidelines from Nielsen's.

ID	Guideline Name	Action	Covered Features	Applicability
N1	Visibility of system status	Adapt *	F4, F16, F22	(3) Critical
N2	Similarity between the system and the real world	Adapt *	F1, F8, F16, F24	(2) Important
N3	User control and freedom	Adapt *	F3, F11, F14	(2) Important
N4	Consistency and standards	Adapt *	F2	(1) Useful
N5	Error prevention	Adapt *	F24	(3) Critical
N6	Recognition rather than recall	Adapt *	F13	(2) Important
N7	Flexibility and efficiency of use	Adapt *	F7, F10	(2) Important
N8	Aesthetic and minimalist design	Adapt *	F3, F18	(1) Useful
N9	Help users recognize and recover from errors	Adapt *	F8	(1) Useful
N10	Help and documentation	Adapt *	F8	(2) Important

* Add more checklist items about e-Games from other research.

Table 3. Choosing guidelines from GUESS.

ID	Guideline Name	Action	Covered Features	Priority
G1	Usability/ergonomics	Eliminate (moved to N1)	F1, F2, F4, F16, F22	–
G2	Narrative	Eliminate (moved to N2)	F11, F21	–
G3	Playing involvement	Adapt *	F8, F9, F13, F15	(3) Critical
G4	Pleasure	Adapt *	F5, F8, F13, F20	(2) Important
G5	Creative freedom	Adapt *	F8, F21	(2) Important
G6	Audio aesthetics	Adapt *	F19	(2) Important
G7	Internal gratification	Adapt *	F1, F9, F13, F15, F20	(2) Important
G8	Social connectivity	Adapt *	F9	(2) Important
G9	Visual aesthetics	Eliminate (moved to N8)	F2, F3, F18	–

* Add more checklist items about e-Games from other research.

Table 4. Newly created guidelines derived from literature review and key features of effective e-Games children with language delay.

ID	Guideline Name	Action	Covered Features	Priority
NW1	Gradual language complexity	Create	F25	(3) Critical
NW2	Interactive feedback	Create	F6, F26	(3) Critical
NW3	Play-based engagement	Create	F27	(2) Important
NW4	Words review and sentence formation	Create	F26	(2) Important
NW5	Repetition and rehearsal of skills	Create	F12	(3) Critical

Note: NW1–NW5 represent new guidelines proposed in this study.

3.1.5. Stage 5: Specification Stage

In this stage, a new set of eGLD guidelines was formally defined, comprising 19 usability guidelines with 130 associated checklists. These guidelines were developed through the integration of Nielsen's and GUESS guidelines, along with an additional newly created guidelines. Furthermore, the eGLD guidelines were adapted and extended by incorporating additional checklists from one study [10] to enhance their relevance and applicability. These checklists were added to better align the main guidelines with the specific features of e-Games designed to support language development in children with language delays. Some of these features are created with therapeutic goals to help improve children's language skills through interactive and engaging activities. Below are the 19 guidelines of eGLD, including their respective IDs and names (IDs start with eGLD, representing e-Game for Language Delay).

- **eGLD1:** Visibility of system status.
- **eGLD2:** Similarity between the system and the real world.
- **eGLD3:** User control and freedom.
- **eGLD4:** Consistent and minimalist design.
- **eGLD5:** Minimize and manage errors.
- **eGLD6:** Recognition rather than recall.
- **eGLD7:** Flexibility and efficiency of use.
- **eGLD8:** Help and documentation.
- **eGLD9:** Playing involvement.
- **eGLD10:** Pleasure.
- **eGLD11:** Creative freedom
- **eGLD12:** Audio aesthetics.
- **eGLD13:** Personal gratification.
- **eGLD14:** Social connectivity.
- **eGLD15:** Gradual language complexity.
- **eGLD16:** Interactive feedback.
- **eGLD17:** Play-based engagement.
- **eGLD18:** Words review and sentence formation.
- **eGLD19:** Repetition and rehearsal of skills.

3.2. Validating eGLD

Stage 6: Validation Stage

The validation stage follows the three steps proposed by Quinones et al.'s methodology [23]: (1) guideline evaluation, (2) expert judgment, and (3) user testing. In this study, we completed only the first step, which involved comparing the new eGLD guidelines with Nielsen and GUESS to assess how many usability issues were identified and how severe they were. The expert judgment and user testing steps will be conducted in future work. Below are the details of the first step.

- **Experiment design**
  - **Participants:** The evaluation was carried out by two of the researchers who worked on this study. The first one, Mrs. Noha Badkook, is a master's student at the Faculty of Computing and Information Technology (FCIT) at King Abdulaziz University. She has a moderate level of experience with usability testing. The second evaluator is Dr. Duaa Sinnari, an assistant professor in FCIT at King Abdulaziz University, and she has high experience in usability testing and human–computer interaction (HCI).

- **Guidelines:** Three different sets of guidelines were tested. The first one, eGLD, is a detailed set of 19 usability guidelines that includes 130 checklist items specifically created for apps designed to help with language delay. The second set, Nielsen's guidelines, is a well-known group of 10 general usability guidelines aimed at improving how apps are designed and used. Lastly, GUESS includes 9 guidelines that focus on evaluating games, especially how engaging and visually appealing they are.
- **Applications:** We chose two of the most popular games for treating language delay in children, which are "Speech Blubs" and "MITA", to conduct the experiment. The English language was the default interface and the mobile phone used was running on the iOS system. Reasons for choosing these two apps: Both apps are easy to use, popular, and specifically designed for children's language development. Figure 2 shows the home screen of the "Speech Blubs" game, and Figure 3 shows the home screen of the "MITA" game.

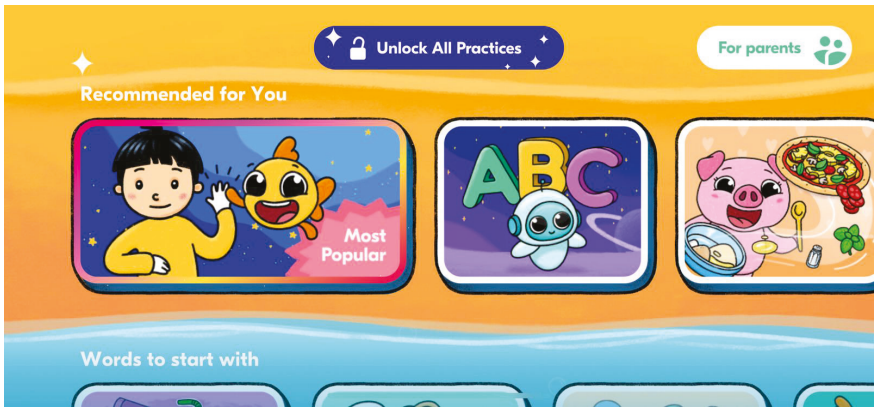


Figure 2. "Speech Blubs" home screen.



Figure 3. "MITA" home screen.

- **Experiment Procedure:** Each evaluator started by testing the "Speech Blubs" app using the eGLD guidelines and recorded the usability issues they found. This was followed by evaluations using Nielsen's and GUESS guidelines separately for the same app. Once the "Speech Blubs" app was evaluated, the process was repeated for the "MITA" app, following the same steps. This resulted in six evaluations in total: three for each app using the three sets of guidelines, as shown below:
  1. List of discovered usability issues on **"Speech Blubs"** using **eGLD**.
  2. List of discovered usability issues on **"Speech Blubs"** using **Nielsen**.
  3. List of discovered usability issues on **"Speech Blubs"** using **GUESS**.

4. List of discovered usability issues on **"MITA"** using **eGLD**.
5. List of discovered usability issues on **"MITA"** using **Nielsen**.
6. List of discovered usability issues on **"MITA"** using **GUESS**.

After completing these evaluations, the evaluators combined their findings to create unified lists of usability issues for each app and each set of guidelines. To prioritize which issues needed attention, the evaluators used the Severity Rating Scale (SRS), which ranks problems on a scale from 0 (not a problem) to 4 (critical issue) [31] as shown in Table 5.

Table 5. Severity Rating Scale (SRS).

Rating	Description
0	No problem at all.
1	Cosmetic issue; only fix if there is extra time.
2	Minor usability problem; low priority for fixing.
3	Major usability issue; should be fixed as a high priority.
4	Critical usability problem; must be fixed immediately before release.

4. Results

This section outlines the results obtained from the development and validation of the eGLD guidelines.

4.1. Results of eGLD Development

The Quinones et al. [23] template offers various elements. It is the discretion of the researcher to decide whether to use the entire template or just specific parts [32]. In our study, we chose to include the ID, priority, name, definition, related features, and checklist items for each guideline, as shown in Table 6. A detailed breakdown of each guideline in eGLD is provided in Appendix A.

Table 6. The descriptive elements for the eGLD guidelines.

Element	Description
ID	Starts with eGLD (e-Game for Language Delay).
Priority	(3) Critical, (2) important, or (1) useful.
Name	The name of the guideline, derived from Nielsen [20] and GUESS [22].
Definition	Describes the guideline and clarifies its intended goal.
Features	Lists the features covered by the guideline.
Checklists	Checklist items are taken from [10] to add more details, making the guidelines relevant for e-Game applications for children with language delay.

4.2. Results of the eGLD Validation

From the usability results collected during the sixth stage (validation), as outlined in Section 3, the main findings are summarized below:

4.2.1. Usability Issues Among the Three Guidelines

As shown in Table 7, eGLD discovered more usability issues ("MITA": 96, "Speech Blubs": 62) than Nielsen ("MITA": 16, "Speech Blubs": 13) and GUESS ("MITA": 39, "Speech Blubs": 13) across both games.

Table 7. Numbers of usability issues in both games based on the three guidelines.

Guideline	MITA	Speech Blubs	Total
Nielsen	16	13	29
GUESS	39	13	52
eGLD	96	62	158

4.2.2. Severe Usability Issues Among the Three Guidelines

As shown in Table 8 and Figure 4, eGLD discovered more catastrophic issues ("MITA": 6, Speech Blubs: 5) than Nielsen ("MITA": 2, "Speech Blubs": 0) and GUESS ("MITA": 0, "Speech Blubs": 0) across both games.

Table 8. Severity of usability issues in "MITA" and "Speech Blubs" games.

Rating	MITA (eGLD)	MITA (GUESS)	MITA (Nielsen)	Speech Blubs (eGLD)	Speech Blubs (GUESS)	Speech Blubs (Nielsen)
No Problem	7	0	0	6	0	0
Cosmetic	4	10	5	5	3	1
Minor	31	20	4	17	5	8
Major	48	9	5	29	5	4
Catastrophic	6	0	2	5	0	0
Total	96	39	16	62	13	13

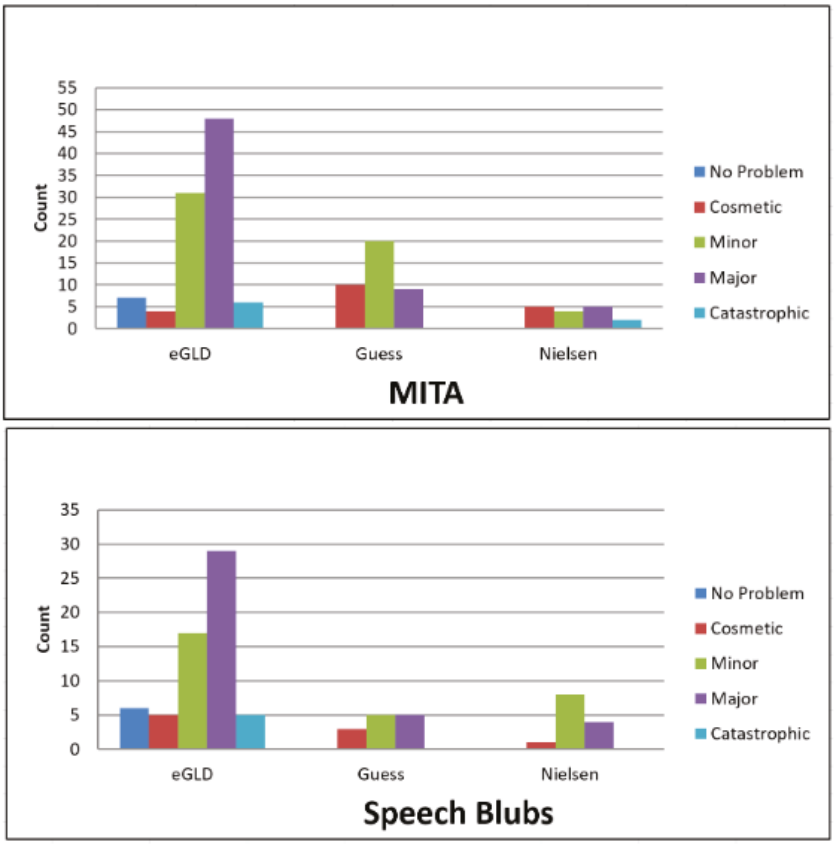


Figure 4. Severity of issues in the "MITA" and "Speech Blubs" games among the three guidelines.

The comparative results presented in Tables 7 and 8 clearly demonstrate the superior performance of the proposed eGLD guidelines over Nielsen and GUESS. Specifically, eGLD

identified a total of 158 usability issues across both games, compared to only 29 issues using Nielsen and 52 using GUESS. Moreover, eGLD detected 11 critical (catastrophic) issues, while Nielsen identified only two and GUESS none. These metrics quantitatively support the hypothesis that eGLD is more effective in capturing both the quantity and severity of usability problems, particularly those relevant to therapeutic e-Games for children with language delay.

4.2.3. Usability Issues in the eGLD Guidelines

The results in Figure 5 highlight key usability issues in both games, particularly in areas like creative freedom, repetition and rehearsal of skills, and minimizing and managing errors within the "MITA" game. Users may have faced challenges in expressing creativity, reinforcing learned skills, or recovering from errors due to insufficient error management. Similarly, the "Speech Blubs" game revealed notable issues in gradual language complexity, play-based engagement, and social connectivity, suggesting that the applications may not have effectively supported language progression, interactive play, or social interaction. These findings emphasize the need for applications to balance structured learning with creativity, engagement, and social interaction, ensuring that users can progress smoothly while staying motivated.

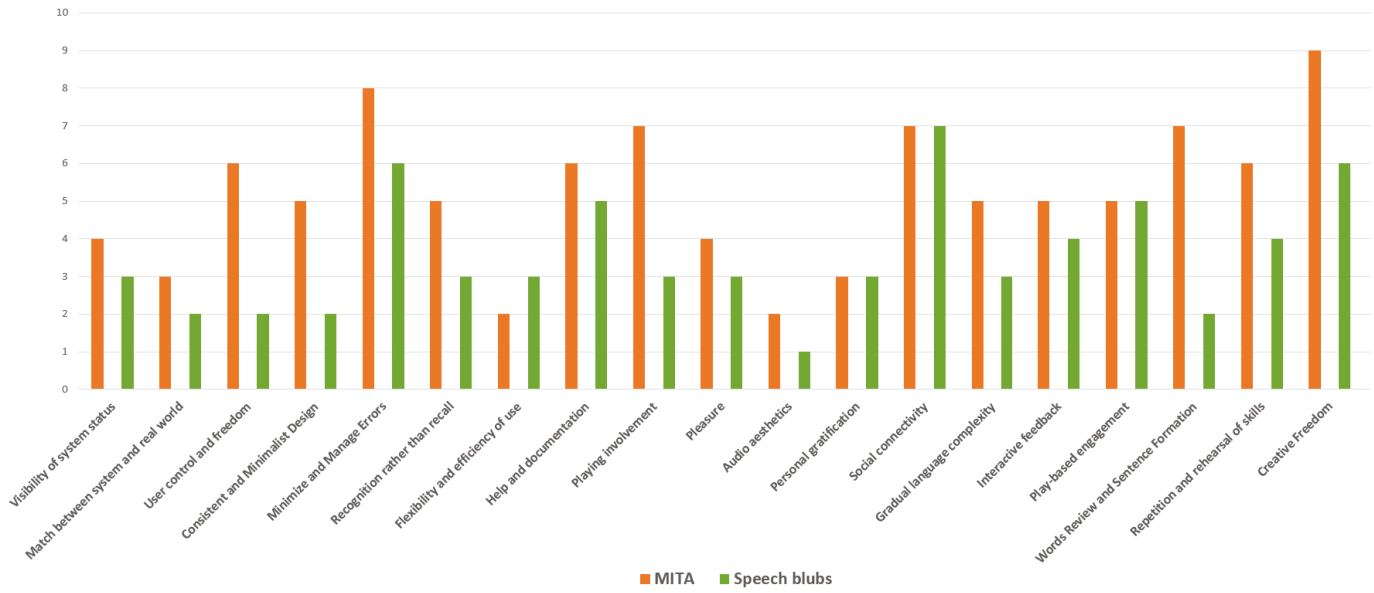


Figure 5. Severity of issues in the "MITA" and "Speech Blubs" games among the three guidelines.

eGLD was able to identify a greater number of usability issues, including more severe issues, across both apps. This is because eGLD includes a more comprehensive set of guidelines and checklists specifically designed for educational gaming apps, making it more effective at assessing usability challenges in this context. Based on Quinones methodology [11], we conclude that eGLD demonstrated higher effectiveness than Nielsen and GUESS.

5. Discussion

This section discussed the findings from the development and validation of the eGLD guidelines.

5.1. eGLD Development Discussion

When we compared the eGLD guidelines with those from Nielsen and GUESS, it was clear that eGLD provided a broader set of 19 guidelines, compared to 10 in Nielsen and 9 in GUESS. This is because eGLD was developed using Quinones et al.'s methodology [23], which involved adapting and merging guidelines from both Nielsen and GUESS, and introducing new ones to cover features not addressed by the original sets. Table 9 shows how each eGLD guideline corresponds to those from Nielsen and GUESS.

Table 9. Connection between eGLD guidelines, Nielsen, and GUESS.

ID	eGLD Guideline	Nielsen's Guideline	GUESS Guideline
eGLD1	Visibility of system status	Visibility of system status	Usability/Reproduction
eGLD2	Similarity between the system and the real world	Similarity between the system and the real world	Narratives
eGLD3	User control and freedom	User control and freedom	–
eGLD4	Consistent and minimalist design	Consistency and standards; aesthetic and minimalist design	Visual Aesthetics
eGLD5	Minimize and manage errors	Error prevention; help users recognize and recover from errors	–
eGLD6	Recognition rather than recall	Recognition rather than recall	–
eGLD7	Flexibility and efficiency of use	Flexibility and efficiency of use	–
eGLD8	Help and documentation	Help and documentation	–
eGLD9	Playing involvement	–	Playing involvement
eGLD10	Pleasure	–	Pleasure
eGLD11	Creative freedom	–	Creative freedom
eGLD12	Audio aesthetics	–	Audio aesthetics
eGLD13	Personal gratification	–	Personal gratification
eGLD14	Social connectivity	–	Social connectivity
eGLD15	Gradual language complexity	–	–
eGLD16	Interactive feedback	–	–
eGLD17	Play-based engagement	–	–
eGLD18	Words review and sentence formation	–	–
eGLD19	Repetition and rehearsal of skills	–	–

The eGLD guidelines include 131 checklist items, compared to 55 items in GUESS, while Nielsen's guidelines do not include checklist items at all. This is because Nielsen's guidelines are general, and only a few of GUESS's checklist items focus on the general interaction and engagement aspects of e-Games, rather than addressing the specific needs of children with language delays. To fill this gap, eGLD introduced detailed and specialized checklist items designed to evaluate usability and engagement in e-Games created for children with language delays. These items focus on key features that meet the specific needs of this target group and support their language development.

This underlines the major contribution of eGLD, which provides a greater number of guidelines and checklist items than Nielsen and GUESS, as seen in Table 10.

Table 10. Comparison of guidelines and checklist items across eGLD, GUESS, and nielsen.

	eGLD	**GUESS**	**Nielsen**
Number of guidelines	19	9	10
Number of checklist items	131	55	0

5.2. eGLD Validation Discussion

The results of the first experiment to validate eGLD showed that it detected more usability issues than both Nielsen and GUESS across the two games tested. This can be attributed to eGLD's broader and more customized set of guidelines and checklist items. Interestingly, "MITA" revealed more usability problems than "Speech Blubs" under all three guideline sets, demonstrating that eGLD, like Nielsen and GUESS, can effectively help evaluate which game provides a better user experience. Furthermore, in terms of severity, eGLD also uncovered more critical issues, which is reasonable given that its checklist items are specifically crafted for e-Games designed for children with language delays. Looking deeper into the types of problems, "MITA" had the most difficulties with guidelines such as creative freedom, repetition and rehearsal of skills, and error management. These suggest that players may have struggled to express themselves creatively, reinforce learned content, or recover from mistakes due to limited support within the game. On the other hand, "Speech Blubs" showed key issues in areas like gradual language complexity, play-based engagement, and social connectivity, indicating that the game may not have done enough to support smooth language development, interactive learning, or meaningful social interactions. Collectively, these findings emphasize the need to create e-Games that strike a balance between controlled learning and engaging, creative experiences, as well as encouraging social engagement to maintain children's engagement and support their developmental progress.

Overall, the validation results highlight the strength of the eGLD framework in detecting a broader range and higher severity of usability issues compared to Nielsen and GUESS. This suggests that eGLD is more tailored to the context of therapeutic e-Games and better aligned with the needs of children with language delay. The presence of more critical and major issues identified by eGLD indicates that general-purpose guidelines may overlook domain-specific challenges, such as cognitive load management and play-based engagement. These insights reinforce the necessity of having a specialized set of heuristics like eGLD, and emphasize that usability evaluation in sensitive therapeutic contexts requires domain-specific depth that generic guidelines may lack.

Based on the results obtained in this study, the initial hypothesis was supported. The eGLD framework successfully identified a greater number and higher severity of usability issues compared to the Nielsen and GUESS guidelines, especially those relevant to therapeutic e-Games. This confirms that eGLD is more effective in evaluating domain-specific usability aspects for children with language delays.

One of the main challenges encountered was the delay in completing phases 2 and 3 of the validation stage. This delay occurred due to pending feedback from speech-language specialists, app developers, and HCI experts who were invited to review the proposed guidelines. Their insights are crucial to ensure the guidelines are practical, relevant, and effective for children with language delays. Pending their input, these phases have been temporarily postponed and will resume in the near future.

6. Conclusions

e-Games have increasingly become a crucial component of learning and therapy, particularly for children with language delay. It is really important that these games are easy to use and helpful. If they are confusing or hard to play, they may fail to achieve their intended therapeutic outcomes.

Therefore, this study developed eGLD, a specialized set of usability guidelines designed to support the design and evaluation of e-Games for children with language delay. It includes 19 guidelines and 131 checklist items, based on established guidelines such as Nielsen and GUESS, but further refined and tailored to meet the needs of this target group.

To verify the validity of the eGLD guidelines, we tested them on two popular games: "MITA" and "Speech Blubs". We also used Nielsen and GUESS guidelines for comparison. The results showed that eGLD uncovered more usability issues—many of them critical—than the other guidelines. This highlights eGLD's strength in identifying areas that could directly impact user experience, engagement, and learning outcomes for children with language delays.

Overall, eGLD helps developers, designers, and evaluators to build better, more engaging, and more effective learning e-Games for children who need additional language support. However, it is important to acknowledge some limitations of the current eGLD guidelines. The checklist items, while comprehensive, have not yet undergone large-scale testing with diverse user groups. Additionally, further iterations and broader evaluations are required to enhance the generalizability of the guidelines. In addition, one practical limitation of the current version of eGLD is its length; the large number of checklist items may make the evaluation process time-consuming and challenging to implement consistently. However, this can be addressed through iterative refinement and expert feedback, allowing future versions to be more concise and streamlined.

For future work, we will focus on further validation and enhancement of eGLD by:

- Offering design recommendations based on the usability issues found in "MITA" and "Speech Blubs" and gathering feedback from development teams.
- Involving UX experts, speech-language therapists, and HCI professionals to review the guidelines and contribute expert insights during the refining stage.
- Enhancing a selected game using the refined guidelines and conducting user testing to evaluate the impact of these improvements in real-world use.

Author Contributions: Conceptualization, N.B., D.S. and A.A.; methodology, N.B.; validation, N.B. and D.S.; formal analysis, N.B.; resources, N.B.; writing—original draft preparation, N.B.; writing—review and editing, N.B., D.S. and A.A.; supervision, D.S. and A.A.; project administration, D.S. and A.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: This study did not involve direct interaction with children or the collection of personal data from minors. Instead, the evaluation was conducted by adult expert reviewers (e.g., usability specialists and HCI experts), who were informed of the study's objectives and voluntarily participated in the evaluation. Therefore, based on the guidelines of our institution (King Abdulaziz University), formal ethical approval was not required.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The raw data supporting the conclusions of this article will be made available by the authors on request.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

e-Game	Electronic game
eGLD	e-Game for Language Delay
HCI	Human–computer interaction

Appendix A

Appendix A.1

This appendix shows 19 usability guidelines that were created to help design and evaluate e-Games made for children with language delay (eGLD). Each guideline is explained using a special format from Quinones et al. [23]. For every guideline, we include its ID, name, what it means, which game features it covers, and a checklist to help people evaluate how well the guideline is being used.

Table A1. Descriptive details for eGLD1: Visibility of System Status.

ID	eGLD1
Priority	(3) Critical
Name	Visibility of System Status
Definition	Ensure that players can always see their progress and the game status. Show feedback on what is happening in the game in real time.
Features	Progress indicators, feedback systems, motivational messages
Checklists	1. Use language that is straightforward, clear, and easy for children to understand. 2. Make sure the game gives immediate and helpful feedback during play. 3. Include a scoring system to help monitor the child's progress. 4. The game should incorporate performance feedback to show children their progress and achievements. 5. Use sound-based feedback, especially spoken language, to support children with communication challenges by focusing on vocabulary and understanding. 6. The interface should offer extra support like hints, scaffolding, or hidden cues to help children complete tasks that are just beyond their reach. 7. Use clear instructions, simple navigation, and smooth steps so children can easily understand and complete tasks without feeling overwhelmed.

Note: Checklist items from 1 to 6 in the following table are derived from [10].

Table A2. Descriptive details for eGLD2: Similarity Between the System and the Real World.

ID	eGLD2
Priority	(2) Important
Name	Similarity Between the System and the Real World
Definition	Make the game relatable by incorporating real-world tasks, language, and visuals that players recognize.
Features	Real-life examples, relatable scenarios
Checklists	1. The game should use the child's native language. 2. It must avoid any aggressive content or violent themes. 3. The game should be designed according to the cultural environment of the player. 4. The game needs to be very straightforward, simple, and easy to learn to avoid confusion. 5. Include playful and therapy-related activities that spark imagination and keep children engaged.

Note: All checklist items in the following table are derived from [10].

Table A3. Descriptive details for eGLD3: User Control and Freedom.

ID	eGLD3
Name	User Control and Freedom
Priority	(2) Important
Definition	Allow players to explore the game freely, retry tasks, and fix mistakes without penalty.
Features	Back/Retry buttons, skip features
Checklists	1. Use clear, large-sized fonts that are easy to read. 2. Give easy-to-follow directions during the game to help guide children through each step. 3. Ensure that instructions are presented in a format appropriate for the child's age. 4. The text should be easy to read and clearly distinguishable from the background. 5. Minimize the amount of text and avoid displaying dense blocks of writing on the screen. 6. Use verbal messages to capture the child's attention and encourage their focus and engagement. 7. Make sure instructions and buttons are easy to see and consistently placed in the same location throughout the game. 8. Organize borders and spacing to make icons and information easier for children to recognize and locate. 9. Include separate volume or mute controls for voice and background music to suit individual preferences.

Note: All checklist items in the following table are derived from [10].

Table A4. Descriptive details for eGLD4: Consistency and Standards.

ID	eGLD4
Name	Consistency and Standards
Priority	(2) Important
Definition	Keep the design, instructions, and interface consistent across the game to make it easier for players to understand.
Features	Unified interface elements, consistent visuals
Checklists	1. Create interfaces that motivate users to explore and interact with different elements. 2. Create clean and straightforward layouts to help users stay focused. 3. Use symbols and icons that people already recognize and understand in the usual way. 4. A clear sign or indicator should be provided to show players how to exit the game. 5. Make sure that every feature is properly implemented and performs as expected. 6. Add pictures that are labeled with specific names to help users remember the information better. 7. Create interfaces that help users recognize options easily instead of having to remember them, and make sure the visuals are clear and engaging. 8. Keep the game window fixed to prevent confusion or frustration for children during the game. 9. Use consistent colors for objects or texts that share the same meaning. 10. Keep options menus simple and free of clutter. 11. Make sure that icons and graphics are visually meaningful and easily understandable for children. 12. Design interfaces to be highly visual, minimizing the use of text wherever possible. 13. Use minimal interface items to avoid overwhelming the user. 14. Ensure color harmony with enough contrast between foreground and background. 15. Keep menus and buttons in consistent positions. 16. Create clean layouts free from unnecessary clutter. 17. Use visual highlights to emphasize important elements. 18. Include a dark mode option to reduce eye strain. 19. Buttons and elements should align with natural hand/finger posture.

Note: Checklist items from 1 to 13 in the following table are derived from [10].

Table A5. Descriptive details for eGLD5: Minimize and Manage Errors.

ID	eGLD5
Name	Minimize and Manage Errors
Priority	(3) Critical
Definition	Minimizing mistakes made by children while interacting with the game. This is achieved by providing clear instructions, validating their inputs, and guiding them through the correct steps to achieve success.
Features	Warnings and alerts for incorrect actions, immediate feedback to correct mistakes, guided steps for error recovery, interactive prompts to clarify actions.
Checklists	1. Design interfaces that help guide users and limit their choices to appropriate actions [10]. 2. Provide visual or audio cues to indicate incorrect actions and offer helpful hints or guidance to correct them. 3. Add visual or audio cues to indicate incorrect actions and suggest alternatives. 4. Provide undo or redo options to allow children to fix mistakes without penalty. 5. Create a "hint" button to assist children when they face difficulty. 6. Include tutorials or practice rounds to help children become familiar with game mechanics before progressing. 7. Provide clear messages to help users identify and correct mistakes. 8. Avoid ambiguous error messages. 9. Provide detailed explanations for critical errors to enhance learning opportunities.

Table A6. Descriptive details for eGLD6: Recognition Rather than Recall.

ID	eGLD6
Name	Recognition Rather than Recall
Priority	(2) Important
Definition	Reduce the cognitive load on children by ensuring that they can recognize options or choices rather than relying on memory.
Features	Dropdown menus, visual prompts, consistent and familiar icons across the game
Checklists	1. Use familiar icons and design elements consistently. 2. Provide clear labels or tooltips to describe functions of buttons and icons. 3. Use visual storytelling or animations to guide children through tasks without requiring verbal instructions. 4. Include color-coded options for better differentiation and recognition. 5. Highlight the most frequently used options to make them easy to locate. 6. Allow children to see a preview of their choices (e.g., sample animations or sounds) before making a decision.

Table A7. Descriptive details for eGLD7: Flexibility and Efficiency of Use.

ID	eGLD7
Name	Flexibility and Efficiency of Use
Priority	(2) Important
Definition	Adapting the game to the player's speed, preferences, and abilities. It ensures that the game remains accessible to diverse players while maintaining a smooth and efficient experience.
Features	Adjustable difficulty settings, progress-tracking tools
Checklists	1. The game should provide different levels of difficulty. 2. Games should be adaptive to match the skills and abilities of children. 3. Games should maintain a balance between success and challenge. 4. Enable parents or supervisors to modify game settings to fit the child's developmental needs. 5. Allow for skip or repeat functions, enabling players to revisit tasks or move past challenges they find too easy.

Note: Checklist items from 1 to 3 in the following table are derived from [10].

Table A8. Descriptive details for eGLD8: Help and Documentation.

ID	eGLD8
Name	Help and Documentation
Priority	(2) Important
Definition	Provide helpful instructions and resources to assist players when they need guidance.
Features	Help menus, tutorials
Checklists	1. Include a help menu or tutorial for first-time users. 2. Provide visual cues for troubleshooting common problems. 3. Include a help button on all screens. 4. Provide an FAQ section to address common user challenges. 5. Integrate a search function in the help section to quickly locate relevant topics. 6. The game should have easy-to-follow demos that you can watch anytime [10]. 7. There should be a quick tutorial that shows you how to play the game step by step [10].

Table A9. Descriptive details for eGLD9: Playing Involvement.

ID	eGLD9
Name	Playing Involvement
Priority	(2) Important
Definition	Engage players by including interactive and challenging tasks to maintain their interest.
Features	Interactive elements, increasing difficulty levels
Checklists	1. Define a clear and attainable goal to focus the player's attention and maintain their interest. 2. Add animated characters on the screen to guide players, keep them entertained, and encourage them during tasks. 3. The game should engage players so they complete tasks because they are interested, not just for rewards. 4. Minimize delays to keep the children engaged with the interaction, tasks, and learning experience. 5. Offer diverse and enjoyable activities to keep children motivated and help them learn new skills. 6. Provide tasks that estimate and adapt to the child's cognitive abilities to ensure progress. 7. Offer guidance through decision-making processes in a step-by-step manner. 8. Encourage children to identify objects, actions, and events in their environment. 9. Use storytelling elements or themes to make gameplay more immersive and relatable. 10. Incorporate drag-and-drop tasks and other interactive game elements to make learning engaging. 11. Introduce levels with increasing complexity for sustained interest and long-term learning.

Note: Checklist items from 1 to 8 in the following table are derived from [10].

Table A10. Descriptive details for eGLD10: Pleasure.

ID	eGLD10
Name	Pleasure
Priority	(3) Critical
Definition	Make the game enjoyable with rewards and activities that keep players happy and motivated.
Features	Rewards, animations, creative tasks
Checklists	1. Rewards should be offered consistently across all levels of the game [10]. 2. Provide immediate feedback and rewards for correct actions. 3. Use meaningful and engaging rewards like badges and trophies that motivate players. 4. Add animations and creative tasks to maintain interest. 5. Incorporate customizable avatars or themes to allow players to personalize their gaming experience.

Table A11. Descriptive details for eGLD11: Creative Freedom.

ID	eGLD11
Name	Creative Freedom
Priority	(2) Important
Definition	Provide players with opportunities to express creativity and make unique choices to shape their gameplay experience.
Features	Customizable gameplay elements, open-ended decision-making, and creative problem-solving tools.
Checklists	1. Allow players to customize and design their characters, including choosing colors, outfits, and accessories. 2. Create opportunities for players to solve open-ended challenges and express creativity, like building structures, creating patterns, or writing and narrating their own stories with customizable options. 3. Offer interactive tools like drawing boards or musical instruments to encourage experimentation with creative features. 4. Allow players to make gameplay decisions that lead to different outcomes, ensuring their choices feel meaningful and unique. 5. Add replay options for tasks or levels, allowing players to approach challenges with different strategies or ideas.

Table A12. Descriptive details for eGLD12: Audio Aesthetics.

ID	eGLD12
Name	Audio Aesthetics
Priority	(2) Important
Definition	Use sound and music to enhance the gaming experience without distracting players.
Features	Background music, sound effects
Checklists	1. Sounds should be clear and easy for players to recognize and understand [10]. 2. Ensure that the game's audio aligns with the game's theme. 3. Avoid loud or distracting sounds. 4. Allow players to select their preferred background music or sound themes for enhanced immersion. 5. The game should combine both sound and visuals to support an effective learning experience [10].

Table A13. Descriptive details for eGLD13: Personal Gratification.

ID	eGLD13
Name	Personal Gratification
Priority	(2) Important
Definition	Reward players for their achievements to make them feel accomplished and motivated.
Features	Virtual trophies, leaderboards, progress tracking
Checklists	1. Showing the child's information on the screen can help capture their interest [10]. 2. Design characters that look like the child to build a sense of connection and encourage achievable actions [10]. 3. Track user progress visibly using leaderboards and trophies. 4. Include a reward system that allows users to redeem achievements for in-game items or benefits.

Table A14. Descriptive details for eGLD14: Social Connectivity.

ID	eGLD14
Name	Social Connectivity
Priority	(2) Important
Definition	The ability of the game to encourage meaningful social interactions among children, their peers, and family members. It promotes collaboration, communication, and relationship-building through engaging activities that require teamwork or shared experiences.
Features	Two-way interactive communication tools, multiplayer or cooperative modes, shared progress and achievements
Checklists	1. Integrate features that allow adults to guide, manage, and supervise children's gameplay, and provide feedback through in-app tools. 2. The game should allow for two-way interaction that helps children improve their speech and language skills [10]. 3. Promote interaction and collaboration by encouraging children to complete tasks with their friends and family [10]. 4. Provide shared progress or reward systems (e.g., family leaderboards or group achievements) to motivate collective participation. 5. Encourage role-playing scenarios that involve family or friends.

Table A15. Descriptive details for eGLD15: Gradual Language Complexity.

ID	eGLD15
Name	Gradual Language Complexity
Priority	(2) Important
Definition	Start with simple words and phrases, gradually introducing more complex language as the player progresses.
Features	Increasing difficulty in language tasks
Checklists	1. Start with basic vocabulary and gradually increase linguistic difficulty. 2. Align word complexity with user progress. 3. Provide hints or examples for difficult vocabulary tasks to ensure learning progression. 4. Provide clear and slow pronunciations of words to ensure clarity. 5. Allow multiple attempts to encourage repetition and learning. 6. Automatically adjust the game's difficulty based on the child's progress and performance.

Table A16. Descriptive details for eGLD16: Interactive Feedback.

ID	eGLD16
Name	Interactive Feedback
Priority	(3) Critical
Definition	Provide real-time feedback to guide players, correct mistakes, and encourage learning.
Features	Hints, feedback messages, animations
Checklists	1. Provide real-time feedback and notifications to highlight mistakes, including comparing the child's pronunciation with the correct version. 2. Offer tips and visual animations for corrective actions. 3. Use adaptive feedback that changes based on the number of user mistakes or confusion in completing tasks. 4. Incorporate a recording feature to capture the child's pronunciation for evaluation.

Table A17. Descriptive details for eGLD17: Play-Based Engagement.

ID	eGLD17
Name	Play-Based Therapy
Priority	(3) Critical
Definition	Incorporate fun elements like games and challenges to make learning enjoyable.
Features	Mini-games, animated characters
Checklists	1. Repeat the word to the child at least three times to reinforce recognition. 2. Use the new word in various sentences to help the child understand how and where it can be used. 3. Add a word-matching game after the child has learned a set of words to ensure they have mastered those words. 4. Include sentence-forming exercises to teach the child how and where to use the words.

Table A18. Descriptive details for eGLD18: Review of Previous Words.

ID	eGLD18
Name	Words Review and Sentence Formation
Priority	(3) Critical
Definition	Provide players with opportunities to review and practice vocabulary they have learned previously, and help players move from learning individual words to forming complete sentences with correct syntax.
Features	Flashcards, practice games, reminders, sentence-building tasks, visual aids
Checklists	1. Offer a section to revisit learned vocabulary using flashcards and reminders. 2. Encourage regular reviews with practice games. 3. Highlight words or lessons that the player struggles with for targeted practice sessions. 4. Show a new story after the child learns new words, incorporating those words into the story with engaging visuals, animations, and interactive elements to encourage practice and maintain interest. 5. Incorporate drag-and-drop tasks to build sentences. 6. Use visual aids to demonstrate syntax. 7. Provide exercises for sentence construction.

Table A19. Descriptive details for eGLD19: Repetition and Rehearsal of Skills.

ID	eGLD19
Name	Repetition and Rehearsal of Skills
Priority	(3) Critical
Definition	Allow repeated practice of key skills to help players reinforce their learning and improve memory.
Features	Repeated tasks, varied practice methods
Checklists	1. Allow repeated task attempts for skill reinforcement. 2. Use varied repetition formats like quizzes. 3. Provide visual aids (e.g., waveforms or phonetic representations) to guide pronunciation improvement. 4. Incorporate engaging and repetitive features like animations, sound effects, or interactive tasks to capture and maintain children's interest, encouraging them to stay involved and actively participate in the game [10]. 5. If players cannot overcome certain challenges, the game should provide easier opportunities for them to rehearse and apply new skills.

References

1. Manipuspika, Y.S.; Sudarwati, E. Phonological development of children with speech delay. *RETORIKA J. Ilmu Bhs.* **2019**, *5*, 12–22. [CrossRef]
2. Feltner, C.; Wallace, I.F.; Nowell, S.W.; Orr, C.J.; Raffa, B.; Middleton, J.C.; Vaughan, J.; Baker, C.; Chou, R.; Kahwati, L. Screening for speech and language delay and disorders in children 5 years or younger: Evidence report and systematic review for the US Preventive Services Task Force. *JAMA* **2024**, *331*, 335–351. [CrossRef] [PubMed]
3. Mirnawati, M.; Ilahiah, R. Development of Toupic (Touch Picture) Applications to Improve Expressive Language in Speech-Delayed Children. *AL-ISHLAH J. Pendidik.* **2023**, *15*, 5067–5077. [CrossRef]
4. Warmbier, W.A.; Podgórska-Bednarz, J.; Gagat-Matuła, A.; Perenc, L. The Use of the "Talk To Me" Application in the Therapy of Speech Development Delays. Available online: https://repozytorium.ur.edu.pl/items/dec0831d-6893-40aa-af17-be743963a53f (accessed on 26 April 2024).
5. Akbar, M.A.; Ismail, H. Language acquisition in child who speech delay. *J. Xi'an Univ. Archit. Technol.* **2021**, *13*, 392–402.
6. Diamante, R.E.J. Interactive Learning Tool Using a Neural Network Algorithm for Speech-Delayed Child. *Am. J. Agric. Sci. Eng. Technol.* **2022**, *6*, 124–130. [CrossRef]
7. Vermeij, B.A.; Wiefferink, C.H.; Knoors, H.; Scholte, R.H. Effects in language development of young children with language delay during early intervention. *J. Commun. Disord.* **2023**, *103*, 106326. [CrossRef]
8. Zaki, N.A.A.; Wook, T.S.M.T.; Ahmad, K. A usability testing of ASAH-/for children with speech and language delay. In Proceedings of the 2017 6th International Conference on Electrical Engineering and Informatics (ICEEI), Langkawi, Malaysia, 25–27 November 2017; IEEE: Piscataway, NJ, USA, 2017; pp. 1–6.
9. Gavriushenko, M.; Porokuokka, I.; Khriyenko, O.; Kärkkäinen, T. On developing adaptive vocabulary learning game for children with an early language delay. *GSTF J. Comput.* **2018**, *6*, 1.
10. Zaki, N.A.A.; Wook, T.S.M.T.; Ahmad, K. Therapeutic serious game design guidelines for stimulating cognitive abilities of children with speech and language delay. *J. Inf. Commun. Technol.* **2017**, *16*, 284–312.
11. Otey, D.Q. A methodology to develop usability/user experience heuristics. In Proceedings of the XVIII International Conference on Human Computer Interaction, Cancun, Mexico, 25–27 September 2017; pp. 1–2.
12. Preston, J.L.; Frost, S.J.; Mencl, W.E.; Fulbright, R.K.; Landi, N.; Grigorenko, E.; Jacobsen, L.; Pugh, K.R. Early and late talkers: School-age language, literacy and neurolinguistic differences. *Brain* **2010**, *133*, 2185–2195. [CrossRef]
13. Nasr, E.; Alsaggaf, W.; Sinnari, D. Developing Usability Guidelines for mHealth Applications (UGmHA). *Multimodal Technol. Interact.* **2023**, *7*, 26. [CrossRef]
14. Dietlein, C.S.; Bock, O.L. Development of a usability scale based on the three ISO 9241-11 categories "effectiveness,""efficacy" and "satisfaction": A technical note. *Accredit. Qual. Assur.* **2019**, *24*, 181–189. [CrossRef]
15. *ISO 9241-11:2018*; Ergonomics of Human-System Interaction—Part 11: Usability: Definitions and Concepts. International Organization for Standardization: Geneva, Switzerland, 2018; Reviewed and Confirmed in 2023. Available online: https://www.iso.org/standard/63500.html (accessed on 21 October 2014).
16. Bevana, N.; Kirakowskib, J.; Maissela, J. What is usability. In Proceedings of the 4th International Conference on HCI, 1991; Citeseer: University Park, PA, USA, 1991; pp. 1–6.
17. Quiñones, D.; Rusu, C. How to develop usability heuristics: A systematic literature review. *Comput. Stand. Interfaces* **2017**, *53*, 89–122. [CrossRef]
18. Muhanna, M.; Masoud, A.; Qusef, A. Usability heuristics for evaluating Arabic mobile games. *Int. J. Comput. Games Technol.* **2022**, *2022*, 5641486. [CrossRef]
19. Valente, L.; Feijó, B.; Leite, J.C. Features and checklists to assist in pervasive mobile game development. In Proceedings of the SBGames 2013, São Paulo, Brazil, 16–18 October 2013; pp. 90–99.
20. Nielsen, J. Ten Usability Heuristics. 2005. Available online: http://www.nngroup.com/articles/ten-usability-heuristics (accessed on 14 September 2024).
21. Phan, M.H.; Keebler, J.R.; Chaparro, B.S. The development and validation of the game user experience satisfaction scale (GUESS). *Hum. Factors* **2016**, *58*, 1217–1247. [CrossRef] [PubMed]
22. Vieira, E.A.O.; SILVEIRA, A.C.d.; Martins, R.X. Heuristic evaluation on usability of educational games: A systematic review. *Inform. Educ.* **2019**, *18*, 427–442. [CrossRef]
23. Quiñones, D.; Rusu, C.; Rusu, V. A methodology to develop usability/user experience heuristics. *Comput. Stand. Interfaces* **2018**, *59*, 109–129. [CrossRef]
24. Saeedi, S.; Bouraghi, H.; Seifpanahi, M.S.; Ghazisaeedi, M. Application of digital games for speech therapy in children: A systematic review of features and challenges. *J. Healthc. Eng.* **2022**, *2022*, 4814945. [CrossRef]

25. Tommy, C.A.; Minoi, J.L.; Sian, C.S. Pre-School Children with Speech Delay: Case Control Study Based on Guidelines for Game Designs. In Proceedings of the European Conference on Games Based Learning, Sophia Antipolis, France, 4–5 October 2018; Academic Conferences International Limited: Reading, UK, 2018; p. 713-XXI.
26. Mahmoudi, E.; Yoo, P.Y.; Chandra, A.; Cardoso, R.; Dos Santos, C.D.; Majnemer, A.; Shikako, K. Gamification in Mobile Apps for Children with Disabilities: Scoping Review. *JMIR Serious Games* **2024**, *12*, e49029. [CrossRef]
27. Hammady, R.; Arnab, S. Serious gaming for behaviour change: A systematic review. *Information* **2022**, *13*, 142. [CrossRef]
28. Nielsen, J. Usability 101: Introduction to Usability (2012). 2012; Volume 9, p. 35. Available online: http://www.nngroup.com/articles/usability-101-introduction-to-usability/ (accessed on 14 September 2024).
29. Iso, W. 9241-11. Ergonomic requirements for office work with visual display terminals (VDTs). *Int. Organ. Stand.* **1998**, *45*, 22.
30. Harrison, R.; Flood, D.; Duce, D. Usability of mobile applications: Literature review and rationale for a new usability model. *J. Interact. Sci.* **2013**, *1*, 1–16. [CrossRef]
31. Group, N.N. How to Rate the Severity of Usability Problems. Available online: https://www.nngroup.com/articles/how-to-rate-the-severity-of-usability-problems/ (accessed on 15 April 2025).
32. Quiñones, D.; Rusu, C. Applying a methodology to develop user eXperience heuristics. *Comput. Stand. Interfaces* **2019**, *66*, 103345. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

Multimodal Technologies and Interaction

MDPI

Article

Design and Evaluation of a Serious Game Prototype to Stimulate Pre-Reading Fluency Processes in Paediatric Hospital Classrooms

Juan Pedro Tacoronte-Sosa [1,2] and María Ángeles Peña-Hita [3,*]

[1] Department of Education, International Iberoamerican University, Campeche 24560, Mexico; jpts0002@red.ujaen.es or juanpedro.tacoronte.edu@juntadeandalucia.es
[2] Territorial Office for Education, Vocational Training, Universities, Research and Innovation, Regional Government of Andalusia, 18071 Granada, Spain
[3] Department of Pedagogy, University of Jaen, 23071 Jaen, Spain
* Correspondence: mapena@ujaen.es

Abstract: Didactic digital tools can commence, enhance, and strengthen reading fluency in children undergoing long-term hospitalization due to oncology conditions. However, resources specifically designed to support rapid naming and decoding in Spanish remain scarce. This study presents the design, development, and evaluation of a game prototype aimed at addressing this gap among Spanish-speaking preschoolers in hospital settings. Developed using Unity through a design-based research methodology, the game comprises three narratively linked levels targeting rapid naming, decoding, and fluency. A sequential exploratory mixed-methods design (QUAL-quan) guided the evaluation. Qualitative data were obtained from a focus group of hospital teachers (N = 6) and interviews with experts (N = 30) in relevant fields. Quantitative validation involved 274 experts assessing the game's contextual, pedagogical, and technical quality. The prototype was also piloted with four end-users using standardised tests for rapid naming, decoding, and fluency in Spanish. Results indicated strong expert consensus regarding the game's educational value, contextual fit, and usability. Preliminary findings suggest potential for fostering and supplementing early literacy skills in hospitalised children. Further research with larger clinical samples is recommended to validate these outcomes.

Keywords: video game; educational video game; serious game; Spanish literacy; paediatric oncology classrooms; hospitalisation

1. Introduction

In the complex landscape of paediatric healthcare, educational interventions face distinctive challenges. This is particularly prominent within oncology wards, where maintaining children's engagement and learning continuity is critical yet frequently disrupted. Game-based learning and gamification have emerged as powerful tools for fostering motivation and engagement, with a growing body of research demonstrating their efficacy across various literacy-related processes, educational stages, and inclusive learning environments [1–4]. One of the most notable examples of this are serious games. These digital games, distinct from gamification and general game-based learning approaches, are designed with educational, therapeutic, or training purposes in mind. Their most prominent characteristic is, integrating learning objectives within interactive and engaging game mechanics to promote active participation and skill development in specific

contexts. Despite their potential benefits, their application in non-traditional settings, such as hospital classrooms, remains notably underexplored, especially in the context of paediatric oncology.

Recent studies have shown that neurological disruptions—such as those caused by developmental disorders, brain injuries, or central nervous system (CNS) conditions—can affect brain regions that are essential for developing reading fluency [5]. These disruptions are particularly prominent in children diagnosed with CNS-related conditions, where any changes to the tissue, including tumours, often impair neural areas responsible for critical reading components such as speed, accuracy, and prosody [6–8]. Consequently, this overlap between affected brain regions and those involved in reading fluency may result in long-term neurocognitive consequences that hinder both short- and long-term literacy outcomes, as seen in Figure 1.

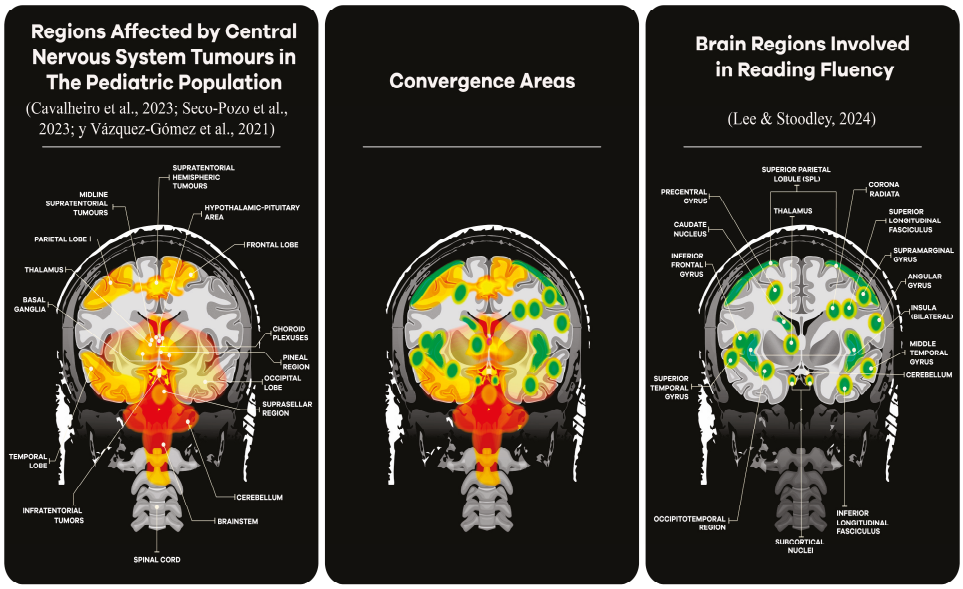


Figure 1. Neuroanatomical convergence between regions affected by paediatric brain tumours [6–8] and areas involved in reading fluency [5].

These impairments pose a dual challenge—both clinical and educational—by disrupting executive functions that are essential for literacy development, such as processing speed, working memory, and sustained attention. The children's learning trajectories are only further disrupted with additional barriers such as intensive medical treatments, hospital-induced stress, and the emotional toll of their illness. In this context, hospital-based classrooms must provide instruction that is not only flexible and individualised but also emotionally responsive to the needs of this vulnerable paediatric population.

Consequently, the intersection of paediatric oncology, serious games, and reading fluency represents a vital yet understudied frontier in educational technology. This notion is supported by a review of existing literature, which highlights a significant gap in research specifically examining the use of video games to improve reading fluency in hospital classroom settings—particularly among children receiving treatment for CNS conditions [9].

2. Background and Related Works

2.1. Pedagogy and Characteristics of Paediatric Oncology

Pedagogical practice in paediatric oncology is grounded in addressing the unique educational and emotional needs of children with cancer within hospital classrooms. This requires a highly adaptable approach that accommodates the varying stages of treatment, fluctuating physical health, and evolving emotional well-being of each child. Given the prolonged and often intensive nature of cancer therapies—frequently extending over months or years and involving intricate medical protocols—educational interventions must be flexible, responsive, and deeply personalised to support meaningful learning throughout the treatment journey.

Empathy plays a central role in this pedagogical model. The psychological strain of illness, combined with the disruption of normal routines, necessitates a learning environment that is not only academically supportive but also emotionally nurturing. Therefore, cultivating a warm, welcoming atmosphere is essential to promote both cognitive development and overall well-being, benefiting children and their families alike.

Within this framework, the 'hospital classroom' becomes a critical educational space [10]. These classrooms are designed to allow hospitalised or outpatient children to continue their academic journey and participate in recreational activities during treatment. Their key goal is to prevent educational marginalisation while facilitating smooth reintegration into mainstream schooling following recovery.

Paediatric oncology areas within these classrooms are typically located near or within the oncology unit and are designed to meet the specific needs of young patients. Attention to lighting, colour schemes, and child-friendly furnishings contributes to a calm, engaging, and emotionally supportive environment. Moreover, such areas are also often equipped with specialised educational technologies, such as computers, tablets, and interactive tools, alongside resources for emotional and psychological support. Collectively, these features work to sustain academic progress and promote positive learning experiences during a profoundly challenging period in a child's life.

2.2. Enhancing Reading Fluency Through Video Games

Reading fluency is a foundational component of early literacy. Among its key predictors is Rapid Automatized Naming (RAN)—the ability to name familiar items quickly and accurately—which plays a crucial role in the development of fluent reading and overall literacy success [11]. As specialised reading circuits in the brain become automated, cognitive resources are freed for higher-level processes such as reading comprehension [12,13]. Building on this, numerous studies emphasise that reading fluency not only facilitates accurate word recognition but also serves as a strong predictor of future reading achievement [14–24]. In particular, RAN has been identified as one of the most reliable indicators of subsequent reading fluency development [25–28].

Digital games have gained recognition as promising tools for supporting literacy by offering highly interactive, engaging environments that can motivate and challenge learners. However, despite their growing application in education, only a few are specifically designed to enhance reading acquisition. In particular, a review of 18 studies found only one that directly addressed reading acquisition through gameplay, underscoring the lack of attention this domain has received in serious game research [29,30].

One notable example of this is *The Adventures of Amaru*, which integrates word decoding and vocabulary development through narrative gameplay, character engagement, and a balanced reward system following the completion of mini-games [31]. Additional examples reflect a growing awareness of games' educational potential, although the field

remains relatively unexplored. For instance, *Tradislexia* aims to enhance reading speed and accuracy in children with dyslexia [32], while a separate preschool-focused game has shown greater improvements in early reading outcomes compared to traditional textbook-based learning [30]. Moreover, and in conflict-affected regions, significantly enhanced decoding and fluency in Arabic among Syrian refugee children aged 5 to 10 has been noted following engagement with *Antura and the Letters* [33].

In the Netherlands, improved reading fluency and motivation in Dutch-speaking primary students was noted following student's interaction with *Letterprins* and *Reading Race* [34,35]. Likewise, the employment of *ELAN* in France demonstrated measurable gains in phonological awareness and grapheme–phoneme correspondence [36]. Positive outcomes were also reported following the use of other notable tools such as a therapeutic video game designed for children with special learning difficulties [37]. Similar results were reported following the use of the *ABC Adventure* virtual reality game, developed for early childhood reading in Early Childhood Education, leading to significant improvements in reading skills [38]. A significant improvement in core literacy skills was also observed after just three months of using *Escribo Play*, a Portuguese app featuring 20 mini-games, including a 68% increase in reading proficiency and a 48% increase in writing skills [39].

Comparable initiatives have also emerged in other regions—for example, in Mexico, a game was developed specifically to enhance reading comprehension among third-grade students [40]. Additionally, *Saving the Word* was designed to enhance reading abilities in primary school learners [41], and *Isla Secreta* (Secret Island) was created to support reading and writing development in first and second grade students [42]. In Spain, *Galexia* targeted fluency in children aged eight and older with dyslexia [43], *PetitUBinding* intervention program reported fluency gains in early readers [44], and *Leobien* supported first-grade reading comprehension, prioritising progress over initial outcomes [45].

Despite some advancements in the field, the current literature reveals a notable scarcity of serious games specifically designed to support reading enhancement, particularly in initiating and developing reading fluency. Within the context of Early Childhood Education (ECE), the focus of the present study, only a handful of recent developments have targeted early reading acquisition [30,38,39]. Among these, just one game has been explicitly designed to improve reading fluency, and it is intended for children aged eight and above [43]. Moreover, it is also worth noting that only two games, one digital and one virtual reality-based, have been developed with the specific aim of fostering reading in five-year-olds [30,38]. In particular, *Escribo Play*, while effective in supporting general reading acquisition through a series of mini-games, was designed for Portuguese-speaking learners in mainstream classroom settings and does not specifically target fluency [39]. As for Spanish-language tools, *Galexia* is the only game focused exclusively on reading fluency, but it is intended for children from eight years old onwards [43].

In addition to the scarcity of serious games specifically designed to target distinct subsets of literacy skills, a recurring critique in the field highlights the simplistic design of many educational games, especially 2D point-and-click formats, which often depend on basic response mechanics and lack opportunities for meaningful exploration or cohesive storytelling. These games, often described by the metaphor "broccoli with chocolate," fail to integrate gameplay with educational content in a way that fosters authentic engagement [46]. Notably, when game mechanics fail to align with learning objectives, both the instructional effectiveness and motivational appeal of these games are significantly reduced. As a result, there is a noticeable lack of serious games targeting reading fluency in ECE. This gap not only highlights the need for more specialized game development but also reveals a

promising and underexplored domain—one with significant potential for academic inquiry and innovation.

2.3. Serious Games in Paediatric Healthcare Education

The integration of ICT resources and video games in paediatric healthcare education, particularly within oncology-focused hospital classrooms, remains notably underresearched. Despite growing interest in game-based learning, few studies have examined the potential of digital games to support early reading fluency in clinical educational settings [5]. This gap is especially pronounced in relation to young children with oncological conditions, where tailored game-based interventions for initiating reading instruction are virtually absent.

Nevertheless, some research has explored the broader educational and emotional potential of video games in paediatric hospital contexts [47–49]. This includes projects such as SAVEH, VIDEM, and SALUD-in, which aim to develop digital games that promote social interaction, collaboration, and cognitive stimulation among hospitalised children [48]. Specifically, examples like *Mundo Isla* and *Tango.H* exemplify this approach, focusing on activities such as cooperative learning, physical engagement, and cognitive exercises through interactive platforms [47].

The emotional and psychological benefits of immersive technologies have also been highlighted in studies on virtual reality, which has been shown to reduce distress and anxiety among hospitalised children and adolescents [50]. Similarly, platforms like Minecraft have been employed to facilitate role-playing and creativity, enhancing emotional well-being through playful, open-ended learning experiences [51]. Another emerging strategy involves involving children in co-creating their own video games, a practice shown to increase motivation, autonomy, and satisfaction [52].

While the use of video games as educational tools in hospital classrooms is not extensively explored, existing research provides a beneficial foundation for developing investigations that combine the didactic potential of video games with their educational value to develop reading skills [53], particularly for initiating reading and improving fluency and in such unique learning environments.

2.4. Serious Games Prototypes Evaluation

Designing a serious game is a multifaceted process that requires a careful integration of pedagogical goals, technological capabilities, and game design principles. Throughout this process, a clear educational objective is essential to effectively guide the design strategy from the onset [54–56]. Moreover, despite the rising prominence of serious games in education, many still lack a solid theoretical foundation—an oversight made more apparent by literature reviews that point out the importance of engaging narratives, intuitive controls, and purposeful gameplay in sustaining user engagement and reinforcing instructional content [57]. Yet, numerous games incorporate overly complex exploratory elements that risk diluting learning outcomes [58]. In contrast, clearly defined objectives and accessible interfaces have been shown to enhance user attention and educational focus [59].

As a key driver of learner-centred design, prototyping is one of the most crucial elements in game development. It enables the creation of preliminary versions that allow developers to refine gameplay and instructional features through demonstration and usability testing—ultimately enhancing the game's value for the end user through iterative improvement [60,61]. Similarly, prototyping serves several purposes in educational game design, from it facilitates early evaluation of core game elements, provides insights into

user interaction, and supports the assessment of key aspects such as usability, engagement, and content alignment.

Prototypes range from low-fidelity models (e.g., paper sketches) used for conceptual validation to high-fidelity versions that closely resemble the final product. While low-fidelity prototypes allow for rapid iteration, high-fidelity prototypes offer detailed, realistic gameplay experiences that enable nuanced evaluations of functionality, aesthetics, and pedagogical integration. However, developing such advanced prototypes is time- and resource-intensive, a challenge reflected in the present study, which employed a fully functional, forward-designed high-fidelity prototype. Prior research using tools like the Fun Toolkit suggests that prototype fidelity strongly influences user perception, particularly among school-aged users [62].

Numerous studies and review articles have proposed diverse evaluation heuristics and criteria for assessing educational games [63–84]. Many studies [85–95] and review studies [57,96–104] have reported the use of a wide diversity of evaluation aspects focusing on educational games. Despite this, there is no universally accepted framework or scale for evaluating serious game design contexts [78,99,101,103,104], particularly within hospital classrooms. This lack of standardisation reflects the complexity and diversity of serious game evaluation. Moreover, general usability criteria are often insufficient in educational settings, prompting researchers to advocate for pedagogical usability considerations that address learning goals and learner characteristics [105,106].

Some studies have even proposed tailored design heuristics for digital games intended for children with language delays, further illustrating the need for population-specific evaluation approaches [107]. Yet, most assessments of educational software continue to rely either on conventional engineering metrics or on purely pedagogical indicators, an analytical limitation noted across multiple comparative studies [108,109].

Given the absence of validated scales for assessing serious games or high-fidelity prototypes in hospital classroom environments, we developed a custom evaluation tool for this study: the Questionnaire for the Quality Evaluation of the Serious Game Prototype Design within Hospital Schools. This instrument was derived and adapted from existing validated frameworks [76,78,80,81,86–88,96,98,105,110–112], which were selected and mapped based on their relevance to serious game design evaluation and research information in the qualitative phase, all in line with the principles of exploratory sequential mixed methods designs. Therefore, our approach aligns with a common practice in the field, constructing ad hoc tools informed by established heuristics and data, particularly in contexts where no validated instruments exist for the target environment.

This study presents a novel contribution by designing and evaluating a serious game prototype aimed at supporting rapid naming, decoding, and early reading fluency in paediatric oncology hospital classrooms. The primary research objective was to capture and assess expert perceptions of the prototype's design quality, with particular emphasis on its suitability for preschool-aged children undergoing cancer treatment. To address this aim, the following sections outline the methodological framework and instructional design (Section 3), present and integrate findings from both qualitative and quantitative phases (Section 4), and discuss the results, limitations, and broader implications (Section 5).

3. Materials and Methods

3.1. Instructional Serious Game Design

In accordance with the pedagogical framework outlined for game-based learning, the game design integrates playful and academic activities essential for fostering interactive and meaningful learning in formal education settings [113]. This approach aims

to immerse students in an engaging educational experience that merges enjoyment with essential literacy outcomes. Central to the educational framework is the development of phonological awareness, which numerous studies identify as foundational to literacy acquisition [114–123]. This skill is supported through the use of technology and game-based methods specifically tailored for younger learners [124–127].

Additionally, the game's design is informed by vocabulary research, particularly the strategy of repeated exposure to target words. This approach—integrated into gameplay mechanics—encourages multiple interactions with selected vocabulary items. Such repetition has been shown to enhance reading acquisition by fostering deeper language engagement and improving comprehension [128–133]. Moreover, repeated reading practices, implemented within immersive digital environments, also play a crucial role in enhancing reading fluency and accuracy [16,134–142]. These practices are only further enriched by incorporating technological learning environments in early childhood education and the use of serious games [143–146].

Building on this foundation, the vocabulary enrichment strategies employed in the game align closely with contemporary recommendations in early childhood education. These guidelines advocate for focused instruction on 3–5 high-utility words per week, coupled with explicit teaching of word meanings. Moreover, the game's activities are closely designed to promote contextualised word usage, a practice shown to enhance both vocabulary retention and the ability to apply new words effectively [147]. On top of this, the game's pedagogical structure also includes phonological and syllabic awareness. More specifically, it follows an integrated method that includes multiple phonological and syllabic components [148,149], complemented by a developmental sequence that progresses from larger to smaller phonological units in line with the natural evolution of phonological skills [150].

The game's instructional model is grounded in an eclectic approach, drawing on research that supports the effectiveness of holistic and mixed methodologies in enhancing reading outcomes [42,151–154]. In this context, *eclectic* refers to the intentional integration of techniques from various instructional models—including rapid naming, syllabic decoding, and visual word recognition—tailored to address the neurocognitive needs of the target paediatric population.

Central to the game's pedagogical delivery are guided instruction and explicit teaching, both of which are rooted in well-established educational practices. These methods emphasize breaking down tasks into manageable steps, modelling desired skills, providing scaffolded support, delivering timely and constructive feedback, and ensuring repeated opportunities for practice [155–160]. Among these, explicit teaching has shown particular efficacy in literacy interventions and remains especially critical in early reading development [18,161].

Aside from explicit teaching, the game design also incorporates principles of active recall and cognitive load management. Drawing from retrieval-based learning, and the concept of desirable difficulties, the game supports memory retention and transfer [162–164]. It also applies cognitive load theory [165], by ensuring clear user guidance and minimising extraneous load—practices supported by recent research on serious games in literacy and educational training [166,167].

Multimedia learning principles are embedded throughout the game to enhance comprehension and retention [168]. The design also integrates spaced practice, as recommended by cognitive psychology literature, to promote long-term learning [169–171]. Finally, scaffolding strategies, grounded in research on cognitive development and educational

games, facilitate knowledge transfer across contexts and support skill consolidation and retention [172–176].

3.2. Game Structure: Characterisation of the Video Game Yuki's Adventures: Hidden Words

Name: *Yuki's Adventures: Hidden Words*.

Knowledge Area: Reading literacy: Decoding and fluency.

Topic: Enhancing decoding and initial reading fluency in Spanish through interactive gaming among paediatric oncology patients.

Taxonomy: Prototype of a serious/educational game. Edu Game or Edutainment [177]. It is a linear game, focused on basic learning and the development of early skills through mini-games and elements of action, adventure, card challenges, strategy challenges, puzzle games, sports games, and simulation games [178–180]. Moreover, it also features balanced exploration, helping prevent overstimulation from visually striking components that could distract from the learning objectives [58].

Methodology used for the design: Integrative methodological model [181].

Purpose: To facilitate the development of foundational reading skills in young oncology patients, leveraging the motivational aspects through a serious game adapted specifically to their needs.

Target audience: Preschool children aged 5–6 from oncology areas with CNS conditions.

Objectives: Initiate and improve word recognition, rapid naming, decoding, early fluency and prosody as a predictor of reading success when starting Primary Education.

Environment: Interactive adventures in virtual scenarios designed for children's engagement.

Game rules: The player begins in a rural village, taking on the role of a character who has forgotten all the knowledge related to reading. The journey unfolds as a solo adventure starting in World 1. Upon awakening in a village house, the character interacts with surrounding objects and environmental elements, each representing high-frequency words. Recognising these elements enables the character to "remember" them, earning items that can be used in subsequent settings or sold in an in-game shop.

Following this initial stage, the character enters a nearby dungeon, where they must complete a series of reading-based challenges (Figure 2). Successfully navigating the dungeon opens a door leading to a forest. Before progressing further, the character encounters the game's antagonist, Zote, a cosy ghost-like figure who steals reading memories from others.



Figure 2. Example of a dungeon scene inside a magical scale-model house in World 1.

The forest, known as Bosque Luminaria (Figure 3), represents the second stage of the learning journey. Here, the character interacts with forest elements to develop skills such as recognising capital letters, decoding syllables, and reading full words. This world includes a variety of challenges: engaging in dialogue with two characters, solving a witch's riddle, locating a hidden treasure chest, fishing, using virtual lenses to uncover concealed elements, and navigating a labyrinth. Upon completing these tasks, the character reaches a hut containing a computer that presents the final challenge of this world. Completing it unlocks the door to World 3.



Figure 3. World 2: Forest setting for capital-letter recognition, as well as word and syllables decoding.

World 3 (Figure 4) is a vibrant, hyper-casual city where the character attends school each morning and receives guidance from a virtual teacher. In the afternoon, the player visits the club library, where instruction becomes more explicit, covering letters, syllables, corresponding vocabulary, and complete sentences. The character also explores various city environments, including a sports centre, a beach, and a theme park, completing context-based tasks tied to literacy goals.



Figure 4. World 3: hyper-casual city unlocked after completing the multimodal quest.

All challenges are designed to foster engagement and immersion, promoting word recognition and reading within a meaningful and motivating narrative. Mini-games are embedded throughout the gameplay and target various literacy levels, including whole-word, syllable, letter, and phoneme-based tasks.

Following core instructional segments, the game offers a series of playful wrap-up activities (Figure 5). While these are primarily recreational rather than pedagogical, they are designed to provide children with moments of enjoyment and emotional relief. These light-hearted segments serve as positive emotional pauses, a particularly important consideration in hospital settings, where supporting emotional well-being is integral to the learning experience.



Figure 5. Example of recreational activity fostering emotional relief.

Challenges: The game's levels are designed with increasing complexity to promote progressive learning. Core instructional principles such as scaffolding, spaced repetition, and active recall are embedded throughout gameplay. Children receive real-time feedback on their advancement via a dynamic progress map (Figure 6), which visualises their journey and reinforces a sense of achievement.



Figure 6. Map illustrating the child's progress.

Gamification mechanics are incorporated to support both intrinsic and extrinsic motivation. Upon completing each world, players earn a corresponding digital sticker, which

is stored in a rewards album (Figure 7), and receive a tangible prize from a teacher or caregiver. This dual-reward system is intended to enhance engagement through visible and meaningful accomplishments—particularly in clinical contexts where sustaining attention and motivation is vital. While the motivational impact of specific reward types was not empirically tested in this pilot, informal teacher feedback indicated positive reception; future research will explore age-related responses and the effectiveness of adaptive reward mechanisms.



Figure 7. Reward album with digital stickers and tangible prizes after completing each world.

To support individualised instruction, the prototype features a learning evidence log that registers correct and incorrect responses for each task completed by the child (Figure 8). This interface, currently implemented at an initial level for demonstration purposes, presents structured feedback by game world, task completion status, and item-level performance, allowing hospital-based teachers to identify specific areas of difficulty and adjust instruction accordingly. While the current prototype offers a simplified display layer, the final system architecture is being designed to support advanced learning analytics through the integration of usage records encoded as the current .json data. This will be made operational via a middleware layer deployed on Heroku (https://www.heroku.com/) and linked to a MongoDB database (https://www.mongodb.com/) for the persistent storage of interaction data. This infrastructure will enable the systematic collection of fine-grained information on session duration, user behaviour, and response latency to reading-related stimuli, thereby establishing a robust foundation for the temporal analysis of emergent fluency indicators within gameplay. Visual and colour-coded elements enhance interpretability, facilitating its use in clinical-educational settings. This feature supports targeted reinforcement and fosters vocabulary consolidation tailored to each learner's progress and instructional needs.



Figure 8. Learning evidence log with correct and incorrect responses.

Furthermore, the system incorporates a spaced repetition mechanism in the menu, as illustrated in Figure 9, that retrieves previously incorrect responses and reintroduces them later through flashcard-based activities, reinforcing retention and supporting long-term acquisition.



Figure 9. Flashcard system with spaced repetition for reviewing incorrect words.

In addition, mini-games can be revisited independently through a dedicated selection menu (Figure 10), which reinforces retrieval-based learning and spaced practice. This feature allows players to access specific reading challenges at any time, supporting memory consolidation and transfer through repeated, distributed exposure.



Figure 10. Mini-game selection menu designed to support retrieval-based learning and spaced practice.

Interface Design and Accessibility Features: The game interface is designed to be child-friendly and intuitive, featuring visually appealing elements in a hyper-casual, cartoon-like style to sustain engagement. The gameplay is based on simple touchscreen interactions, allowing players to perform core actions such as moving in four directions, selecting objects, or dragging items during in-game tasks. Given the clinical context, accessibility was a key principle in the design process. Namely, the prototype was specifically adapted for five-year-old children with CNS tumours undergoing oncological treatment, whose neurocognitive profile may involve reduced working memory, slower processing speed, and fluctuating attention spans. To accommodate these characteristics, the game employs minimal cognitive load interfaces, large visual targets, and one-touch navigation compatible with desktop simulating touch-based interaction through mouse input in anticipation of its future deployment on tablets (see Figure 11).

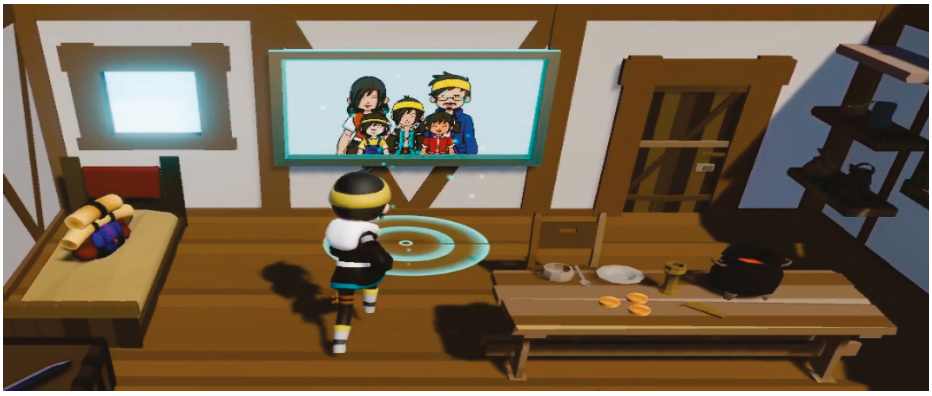


Figure 11. Touch-based interface for easy navigation.

Visual design features were intentionally developed to create cosy, high-contrast environments in warm tones, combined with smooth animations and child-friendly visuals. These elements were designed to minimise visual fatigue and foster a sense of emotional comfort and security for young users in hospital settings, particularly boasting bright, affectively supportive colour palettes as depicted in Figure 12.



Figure 12. Warm-toned interface and high-contrast visual design aimed at reducing visual fatigue and promoting emotional comfort.

The game's usability is further promoted by the addition of supportive elements such as manual navigation, constant in-game narrator guidance, and in-game elements helpers (Figure 13). While advanced personalisation settings were not implemented at this prototype stage, final development aims to include additional accessibility tools, such as scalable text, colour adjustment, or audio support.

The settings menu was designed to allow users to adjust key technical parameters relevant to gameplay performance and user comfort. These include screen resolution, windowed or full-screen display, overall game quality according to hardware capacity (managed via level-of-detail optimisation), volume control, and interaction mode selection, enabling either keyboard-based navigation or point-and-click control with a mouse or touchscreen. These options were included to accommodate diverse testing environments and devices used by participants during evaluation. As shown in Figure 14, the current interface already incorporates basic accessibility-focused adjustments aimed at supporting varied user needs. Although these features were positively received by expert reviewers, the final version is intended to include additional features including scalable text, manual brightness adjustment, alternative contrast modes, and expanded visual customisation tools.



Figure 13. Supportive accessibility elements: navigation bar, narrator guidance, and in-game helper icon cosy star.



Figure 14. Accessibility-focused technical settings menu for customising resolution, quality, controls, and display configuration.

3.3. Learning and Game Mechanics Alignment

A key strength of the developed prototype lies in its deliberate harmonisation of educational objectives with game mechanics—an essential principle in serious game design. Organized into three uniquely themed worlds, the game's prototype is carefully structured to support literacy development, guiding players from foundational word recognition toward early stages of reading fluency [53]. Moreover, each game element is intentionally designed to be both engaging and pedagogically meaningful, ensuring entertainment and instruction are seamlessly integrated throughout the gameplay experience.

Each in-game activity reflects the eclectic instructional method, with gameplay mechanics directly mapped to specific pedagogical strategies. These activities are not only interactive but are also grounded in well-defined educational purposes, an essential characteristic of serious game frameworks. This alignment ensures that learning experiences are meaningfully embedded within the gameplay, fostering a coherent environment in which educational theory and interactive design are inextricably linked. Furthermore, the storyline and learning content were purposefully integrated across the game's three narrative environments, each corresponding to specific stages in the reading acquisition process. This coupling was grounded in pedagogical models such as Bloom's taxonomy and the Learning Mechanics-Game Mechanics (LM-GM) framework [54], and further validated by expert feedback, which highlighted the narrative's coherence and its capacity to support progression through literacy tasks. Detailed mapping of game levels, learning objectives, and didactic strategies is provided in Sections A–C.

3.4. Evaluation and Validation Methods

This study employed an exploratory mixed-methods approach, combining qualitative and quantitative data to evaluate the serious game prototype. Central to the evaluation was the expert judgment technique, a widely recognized method in educational research for assessing technological resources, including serious games [182,183]. By leveraging expert insights, the study ensured and rigorous and informed appraisal of the prototype's educational and design quality. A critical aspect of this method involves the selection and qualification of experts, addressed here through the Expert Competence Coefficient (CCE). According to the classification criteria:

- A coefficient between 0.8 and 1.0 indicates high competence.

- A coefficient between 0.5 and 0.8 indicates a medium level of competence.
- A coefficient below 0.5 indicates a low level of competence.

To ensure methodological rigour in the expert validation process, the Expert Competence Coefficient (CCE) was applied, calculated using the formula $K = \frac{1}{2}$ ($Kc + Ka$). This coefficient is widely used in educational and technological evaluation studies to ensure the qualification of participants in expert judgement processes. The first component, *Kc* (knowledge coefficient), was derived from the expert's self-assessed level of expertise in relation to the topic on a scale from 0 (no knowledge) to 10 (full mastery).

The second component, *Ka* (argumentation coefficient), was calculated based on the expert's assessment of the influence of different sources on their knowledge and judgement. Each source was assigned a weight depending on the degree of influence indicated: theoretical analysis carried out by the expert (0.10 low, 0.20 medium, 0.30 high), practical professional experience (0.20 low, 0.40 medium, 0.50 high), and other sources, including national and international literature, awareness of developments abroad, and personal intuition, which each contributed a fixed value of 0.05. The *Ka* was obtained by summing the weights corresponding to the expert's chosen levels across all sources.

Therefore, the final CCE value results from averaging *Kc* and *Ka*, resulting in a coefficient ranging from 0 to 1. In this study, only experts with a CCE equal to or higher than 0.8, indicating a high level of competence, were considered for inclusion in the validation process.

For the qualitative evaluation (QUAL), a focus group session was conducted, supplemented by individual interviews with experts who met the CCE threshold. These sessions aimed to capture in-depth insights into the design quality and educational alignment of the prototype. The focus group was composed of six teachers from paediatric oncology hospital classrooms in Andalusia, the Canary Islands, and the Region of Murcia. Participants were selected based on their extensive experience with children aged 3 to 13 diagnosed with CNS tumours. The group met online for 52 min and provided feedback on the prototype's pedagogical basis, visual and auditory design, clarity and accessibility of content, alignment of objectives and challenges, and overall usability in hospital settings.

For the quantitative evaluation (QUAN), an ad hoc questionnaire was developed and subjected to a pilot study, which confirmed its strong content validity, clear factorial structure, and high internal consistency. To validate the instrument, a group of ten experts specializing in Educational Technology and Game Design assessed the relevance and clarity of each item using the expert judgment technique. The result, measured through Aiken's V coefficient, showed that all items exceeded the threshold value of 0.8, demonstrating a high degree of clarity and pertinence.

To further examine the questionnaire's structure, an Exploratory Factor Analysis (EFA) was conducted, revealing a robust three-factor structure that accounted for a significant proportion of the variance. The Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy was 0.822, and Bartlett's test of sphericity was significant ($p < 0.001$), supporting the suitability of factor analysis. Reliability was subsequently evaluated using both Cronbach's Alpha and McDonald's Omega, with each factor demonstrating strong internal consistency: Factor I ($\alpha = 0.926$, $\omega = 0.928$), Factor II ($\alpha = 0.851$, $\omega = 0.861$), and Factor III ($\alpha = 0.815$, $\omega = 0.814$).

Convergent validity was demonstrated by Composite Reliability, with values exceeding 0.70, and Average Variance Extracted (AVE) values above 0.40 [184]. Moreover, discriminant validity, assessed using the Fornell-Larcker criterion, showed that the square root of the AVE for each factor was greater than its correlations with other factors: Factor I

(0.744), Factor II (0.753), and Factor III (0.733). These results confirmed the distinctiveness and structural validity of the instrument.

The final questionnaire included 19 items distributed across three constructs: (1) quality of design in the adaptation of content to context, (2) pedagogical quality of design, and (3) technical quality of design. The instrument was administered online via Google Forms and utilised a five-point Likert scale with the following options:

1. MN = Inadequate/Strongly Disagree;
2. N = Deficient/Disagree;
3. R = Regular;
4. P = Good/Agree;
5. MP = Very good/Strongly agree.

To preliminarily explore the potential impact and feasibility of the game prototype, a flexible twelve-week intervention (February–April 2025) was conducted. This pilot implementation was carried out in a hospital classroom with four children undergoing oncology treatment. The prototype was integrated into the teacher's regular pedagogical planning. Sessions were brief, lasting between 5 and 10 min, and were distributed across up to three moments during the day. The frequency and timing were adapted to each child's condition and the rhythm of the hospital classroom, aiming to explore the potential and contextual relevance of the game in a real paediatric oncology setting. Participants used the game following the provided guidelines, with scheduling adjusted to classroom routines. The initiative originated from a teacher in charge of an oncology hospital classroom. Participation was voluntary and included informed consent from both families (through a collaborating association) and the teacher. The intervention was assessed using the Rapid Naming Test (TDR—*Test de Denominación Rápida* in Spanish), the Basic Instrumental Aspects in Language and Mathematics test (PAIB-1—*Prueba de Aspectos Instrumentales Básicos en Lenguaje y Matemáticas*), and the validated Scale of Reading Fluency in Spanish (SRFS).

4. Results

The serious game prototype evaluated in this study was developed using Unity (version 2022.3.3f1) and incorporates key learning mechanisms to support literacy development. Scaffolding is achieved through the progressive sequencing of tasks across the three themed worlds, while spaced repetition and active recall are embedded through the structured revisits of previously learned graphemes, syllables, and words—both within the Reading Club and through retrieval-based mini-games. The game was designed for low- to mid-range devices to ensure accessibility and feasibility across evaluation contexts. Minimum hardware specifications included the following: Windows 10 operating system, Intel Core i3 processor (or equivalent), 4 GB RAM, integrated GPU (e.g., Intel HD Graphics 4000), and 1 GB of available storage.

For testing purposes, all expert reviewers accessed the downloadable build on desktop computers, while the hospital-based pilot was implemented on a 2-in-1 convertible device (tablet-laptop hybrid) operating in tablet mode to simulate the touchscreen tablets intended for clinical use in hospital classrooms.

Following the technical deployment of the prototype, qualitative analyses were conducted using Atlas.ti software (Atlas.ti v.24, ATLAS.ti Scientific Software Development GmbH, Berlin, Germany) was used for all qualitative data analyses. Qualitative findings indicated that hospital-based teachers anticipated strong emotional engagement and perceived the serious game as a potential motivator for children to explore and participate actively. According to both the focus group with six hospital-based teachers and the individual interviews conducted with 30 design experts, the game was considered to be

well-aligned with the intended learning objectives and appropriate for the hospital context. The expert sample included five specialists from each of six key fields involved in serious games design of this nature: educational game design; educational technology; psycholinguistics and reading processes applied to digital environments; computer engineering, programming and development; hospital education and health-related technologies; and accessibility in reading-focused games.

In the focus group discussion (N = 6), participants were invited to share their experiences with the game and assess its design, narrative, pedagogy, and challenges in relation to the intended educational outcomes. There was broad consensus that the game was well-suited to the needs of children undergoing cancer treatment. Specifically, the participants agreed that the objectives, activities and tasks, theoretical and didactic underpinnings, language, content, and presentation were all appropriate for the paediatric oncology context. From the perspective of hospital-based teachers, the alignment of the game with the clinical and emotional trajectories of paediatric cancer patients was regarded as one of its key strengths, contributing not only to literacy development but also to children's social, emotional, and curricular growth.

Participants also highlighted the game's visual and musical aesthetics as particularly effective in capturing children's attention. The technological nature of the tool was seen as an advantage over traditional paper-and-pencil methods, increasing learner engagement and making the learning experience more enjoyable. Additional strengths cited included the game's intuitiveness and ease of use, its accessibility and inclusivity, and the originality and challenge level of the embedded activities.

Nonetheless, several design and technical challenges emerged during the evaluation. Teachers reported difficulties in exiting the game and returning to the start menu and noted that the time required to access interactive segments or mini-games could be improved. These points of feedback were addressed in the subsequent iteration of the game prototype.

In addition to the focus group with hospital teachers, individual interviews were conducted with the 30 design experts previously described, representing complementary domains involved in serious game design development, with gender representation being relatively balanced. These experts affirmed the strength of the theoretical foundation underlying the prototype, its contextual suitability for paediatric hospital settings, and its potential to foster intrinsic motivation among young patients. They also praised the integration of playful and educational elements, the clarity and simplicity of the narrative, and language deemed developmentally appropriate for the target audience and the overall usability of the interactive environment. The visual and auditory quality of the game, along with its feasibility for implementation in real hospitalisation conditions, were also highlighted as key advantages.

Areas identified for further development included the need to enhance technical performance, diversify the range of tasks, and provide clearer, more accessible instructions to support user autonomy. These suggestions were implemented in the final version of the video game to improve functionality, educational effectiveness, and alignment with the hospital learning environment.

Concerning the QUAN investigations, 274 experts (N = 274) participated in the evaluation (Table 1), with over 50% of participants being male and middle-aged. Participants received access to the serious game prototype through a downloadable build provided via Google Drive, which they explored prior to completing the online questionnaire. All statistical analyses were conducted using SPSS software (SPSS v.27, IBM Corporation, Armonk, NY, USA).

Table 1. Descriptive statistics for demographic data (N = 274).

Variable	Category	*n*	%
Gender	Female	92	33.6
	Male	176	64.2
	Non-binary	2	0.7
	Gender-fluid	1	0.4
	Prefer not to say	3	1.1
Age group (years)	<30	36	13.1
	30–39	84	30.7
	40–49	110	40.1
	50–59	37	13.5
	≥60	7	2.6
Country of residence	Spain	204	74.5
	Argentina	22	8.0
	Colombia	15	5.5
	Mexico	15	5.5
	Venezuela	5	1.8
	Ecuador	5	1.8
	Chile	2	0.7
	Peru	2	0.7
	Cuba	1	0.4
	Guatemala	1	0.4
	Dominican Republic	1	0.4
	Uruguay	1	0.4
Primary workplace	University	144	52.6
	Video-game production company	112	40.9
	Non-university	11	4.0
	Training company	7	2.6
Primary professional activity	Professional video-game designer	126	46.0
	Lecturer	110	40.1
	Lecturer/Researcher	38	13.9
Expert competence coefficient (CCE)	Mean (SD)	0.871 (0.058)	–

This study also explored participants' views of serious game design (Table 2), with the experts' responses in the distributed instrument being overly positive and consistent (Cronbach's alpha = 0.86) across all the examined constructs.

Table 2. Descriptive statistics for the serious game prototype (N = 274).

Dimensions	M	SD
D1. Quality of design in the adaptation of content to context	4.28	0.396
D2. Pedagogical quality of design	4.16	0.444
D3. Technical quality of design	4.27	0.433

Furthermore, item-level scores were analysed within each of the three evaluated dimensions. Table 3 presents the mean values and standard deviations for the dimension titled "Quality of Design in the Adaptation of Content to Context."

Table 3. Mean evaluation and standard deviation carried out by the experts across items of the dimension: "Quality of design in the adaptation of content to context".

Quality of Design in the Adaptation of Content to Context	M	SD
The prototype's content is based on a solid and coherent theoretical framework.	4.44	0.604
The language used in the prototype is tailored to the target student audience.	4.34	0.649
The content delivery facilitates the intended learning outcomes.	4.15	0.673
The prototype responds accurately and as expected to user interactions.	4.16	0.729
The auditory and visual representations employed are aesthetically pleasing.	4.23	0.646
The prototype's design offers a reasonable solution to the specific needs of the students.	4.34	0.683
The prototype design is adapted to the specific location and time of use.	4.34	0.690
All necessary factors for ensuring successful game implementation in hospital classrooms have been considered in the design.	4.21	0.651
The prototype is designed for use in the intended space, considering the learning times of the target students and the resources available in hospital classrooms.	4.34	0.655

The mean rating for all items within this construct was M = 4.28 (SD = 0.396), with individual item scores ranging from 4.15 to 4.44. These results reflect a high overall assessment of the prototype's quality in adapting content to the specific educational context. Furthermore, the experts indicated that the prototype is grounded in a robust theoretical framework and makes effective use of language and methodology that are well-aligned with the needs of the target learner population and the characteristics of the hospital-based learning environment. Additionally, the game demonstrated adequate responsiveness and featured well-balanced auditory and visual aesthetics. Its development also accounted for key factors necessary for successful implementation within the intended educational context.

The highest-rated item was the theoretical soundness and internal coherence of the prototype (Item 1; M = 4.44, SD = 0.604), followed by its contextual and temporal adaptability (Item 7; M = 4.35, SD = 0.690). The lowest score within this dimension was assigned to the prototype's capacity to facilitate the intended learning outcomes (Item 3; M = 4.15, SD = 0.673). Nevertheless, this score still reflects a clearly positive expert perception of the didactic methodology employed.

Overall, the aforementioned results suggest a high level of agreement among respondents, as demonstrated by the relatively low standard deviations. The strongest consensus was observed for the item on theoretical coherence (Item 1), while slightly more variation in responses was found for the item addressing the prototype's responsiveness to player actions (Item 4; M = 4.16, SD = 0.729).

Table 4 presents the mean scores and standard deviations for the dimension "Pedagogical Quality of Design."

Table 4. Expert assessment of the items of the dimension "Pedagogical quality of design".

Pedagogical Quality of Design	M	SD
The content is narratively diverse enough to stimulate learning.	4.09	0.605
The prototype has specific, didactic, and relevant objectives for educational use.	4.39	0.667
Relevant and sufficient activities and challenges are presented.	4.13	0.650
The prototype elicits or helps elicit students' interest.	4.09	0.611
The prototype enables user assistance. It provides training levels that facilitate learning the game mechanics.	4.10	0.597

The mean rating for all items within the Pedagogical Quality of Design construct was M = 4.16 (SD = 0.444), with individual item scores ranging from 4.09 to 4.39. These results reflect a high perceived pedagogical quality in the design of the serious game prototype. According to expert evaluations, the learning objectives were considered appropriate, the instructional activities relevant and sufficient, and the user support mechanisms effective

in facilitating learning. Additionally, the content was viewed as varied and engaging, with the potential to stimulate learners' interest.

The highest-rated elements within this dimension were those related to the game's educational objectives and instructional activities. Experts regarded the objectives as specific, pedagogically grounded, and relevant for the target audience (Item 11; M = 4.39, SD = 0.667). Similarly, the activities were rated as both pertinent and sufficient for achieving intended learning outcomes (Item 12; M = 4.13, SD = 0.650). Conversely, the lowest-rated aspects were the variety of content (Item 10; M = 4.09, SD = 0.605) and the prototype's capacity to generate learner interest (Item 13; M = 4.09, SD = 0.611). Despite these lower ratings, the scores remain within a high and favourable range.

Low variability in responses across most items suggests a strong consensus among experts. Notably, there was high agreement regarding the level of training and support provided by the game (Item 14; M = 4.10, SD = 0.597). The greatest variation in expert opinion occurred with respect to the adequacy and relevance of the educational objectives (Item 11; SD = 0.667), though the mean rating remained the highest within the construct.

Table 5 reports the mean scores and standard deviations for the dimension "Technical Quality of Design."

The technical quality of the prototype was likewise positively assessed, with an overall mean of M = 4.27 (SD = 0.433). Individual item means ranged from 4.13 to 4.41, indicating consistently strong expert approval. Experts affirmed that the prototype demonstrated adequate technical performance, responsive interaction, intuitive controls, and satisfying feedback mechanisms. They also emphasised the importance of the prototype's evaluability, noting that the system architecture and content components lend themselves well to assessment and future improvements. In this context, *evaluability* refers to the extent to which the prototype's structure, functionalities, and pedagogical elements are clearly defined, observable, and measurable, enabling systematic assessment and iterative refinement. Furthermore, the clarity of the information provided about the design was seen as a strength, supporting the potential for iterative refinements based on user feedback and contextual requirements.

The highest technical ratings were associated with the evaluability of the prototype's content and components (Item 18; M = 4.41, SD = 0.716), followed by its accessibility, stability, and overall system performance (Item 15; M = 4.31, SD = 0.631). The lowest-rated technical aspect, though still favourably evaluated, was the ease of interaction with the prototype to receive pleasant feedback (Item 17; M = 4.13, SD = 0.565).

Table 5. Expert assessment of the items of the dimension "Technical quality of design".

Technical Quality of Design	M	SD
The prototype operates properly in terms of accessibility, stability, and load speed (technical performance), without significant delays in user response.	4.31	0.631
The controlled units (elements and interface) react to user actions.	4.28	0.662
The prototype allows the user to interact easily and adaptively, providing a pleasant feedback experience.	4.13	0.565
The prototype's content and its components can be evaluated.	4.41	0.716
During game development, information on its design can be obtained for improvement.	4.20	0.736

Although responses across this construct remained relatively consistent, variability was slightly higher than in the previous two dimensions. Greater dispersion was observed, particularly in items related to the availability of information for prototype improvement (Item 19; M = 4.20, SD = 0.736) and the evaluability of content and components (Item 18; M = 4.41, SD = 0.716).

Taken together, these findings indicate that the prototype's content is appropriate and relevant for the hospital classroom context. The educational approach is well-structured, coherent, and tailored to the developmental and emotional needs of young learners in clinical settings. Furthermore, the technical design demonstrates a high level of robustness, contributing to a smooth and accessible user experience. Collectively, these results validate the prototype's promise as an effective educational resource in hospital-based learning environments.

In addition to analysing expert responses for each construct, correlations among the three dimensions of the questionnaire were also examined using Spearman's rank-order correlation coefficient. The analysis revealed statistically significant positive correlations among all constructs ($p < 0.001$). The strongest association was observed between content-context adaptation and technical quality of design (r = 0.63), followed by the correlation between content-context adaptation and pedagogical quality of design (r = 0.52). The correlation between technical and pedagogical quality was also significant, though more moderate (r = 0.45). These relationships are visually illustrated in Figure 15.

These correlation patterns suggest that expert ratings of contextual adaptation were strongly linked to perceptions of both pedagogical and technical quality. Additionally, positive evaluations of pedagogical design were often accompanied by favourable assessments of technical performance. The observed differences in correlation strength highlight the centrality of contextual alignment in expert evaluations, indicating that well-contextualised design contributes meaningfully to both pedagogical coherence and technical functionality.

The experts who participated in the quantitative phase of the study represented two primary professional backgrounds: individuals from the educational field (e.g., professors and researchers) and professionals from the video game design sector. Given the differing domains of expertise, it was hypothesised that their evaluations of the prototype might also diverge, particularly with regard to aspects most aligned with their respective areas of specialisation. For example, it was anticipated that designers might be more critical when evaluating technical features, while educators might apply more stringent criteria to the assessment of pedagogical elements.

To examine potential differences in evaluation between the two groups, the Mann–Whitney U test was conducted, appropriate for ordinal data and non-parametric distributions. Bonferroni correction was applied to adjust for multiple comparisons and reduce the risk of Type I error.

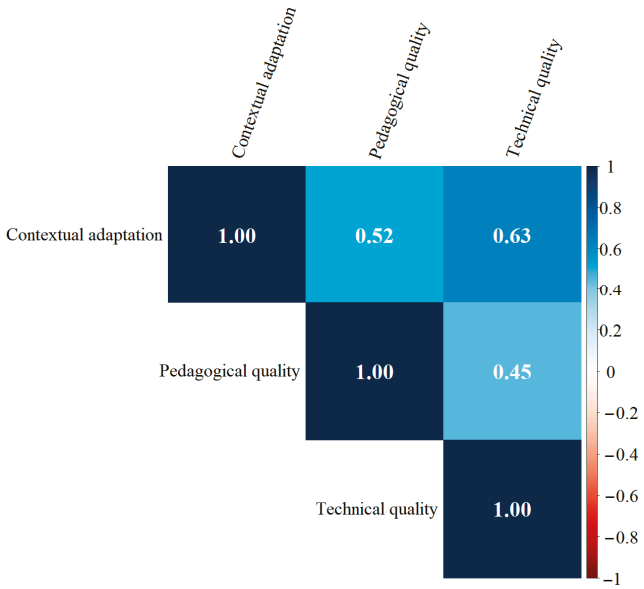


Figure 15. Correlation matrix among the three questionnaire constructs.

The results revealed no statistically significant differences in the item-level ratings between the two groups ($p > 0.05$ across all comparisons). As shown in Table 6, both educational experts and video game designers provided consistently high ratings across all dimensions of the questionnaire. This convergence in responses despite professional background differences further reinforces the perceived overall quality and coherence of the serious game prototype.

Table 6. Mean and standard deviation (in parentheses) per item and group analysed.

Construct	Item	Activity	
		Professional Game Design (*n* = 126)	**Teaching/Research (*n* = 148)**
Quality of design in the adaptation of content to context	1	4.524 *(0.547)*	4.372 *(0.642)*
	2	4.397 *(0.608)*	4.284 *(0.681)*
	3	4.246 *(0.666)*	4.074 *(0.671)*
	4	4.095 *(0.774)*	4.216 *(0.686)*
	5	4.151 *(0.633)*	4.291 *(0.652)*
	6	4.460 *(0.615)*	4.236 *(0.722)*
	7	4.444 *(0.627)*	4.264 *(0.732)*
	8	4.310 *(0.600)*	4.128 *(0.683)*
	9	4.444 *(0.573)*	4.243 *(0.706)*
Pedagogical quality of design	10	4.159 *(0.557)*	4.027 *(0.638)*
	11	4.429 *(0.638)*	4.365 *(0.692)*
	12	4.190 *(0.678)*	4.081 *(0.623)*
	13	4.119 *(0.560)*	4.061 *(0.652)*
	14	4.167 *(0.603)*	4.047 *(0.587)*
Technical quality of design	15	4.357 *(0.613)*	4.270 *(0.645)*
	16	4.325 *(0.617)*	4.250 *(0.699)*
	17	4.198 *(0.537)*	4.074 *(0.584)*
	18	4.429 *(0.709)*	4.385 *(0.724)*
	19	4.294 *(0.658)*	4.122 *(0.790)*

Taken together, the results of the quantitative evaluation indicate a highly favourable appraisal of the prototype across its three core dimensions: content-context adaptation, pedagogical quality, and technical quality. Importantly, these positive assessments were consistent regardless of the respondent's area of expertise. That is, professors and researchers were not disproportionately critical of pedagogical aspects, nor were game designers more exacting in their assessment of technical components.

After integrating and triangulating these findings with the qualitative data, it can be concluded that the educational game prototype, *Yuki's Adventure: Hidden Words*, demonstrates strong contextual alignment, pedagogical soundness, and technical robustness. These characteristics collectively support its potential application within paediatric oncology classroom settings. Specifically, the game offers promising benefits for fostering rapid naming, initial decoding, and automatization in the recognition of syllables, words, and basic sentences, core components in early literacy development for hospitalised learners.

The mixed-methods integration was a central phase in the evaluation of the prototype, consistent with the exploratory sequential design adopted for this study. The integration process was not limited to surface comparison but involved a systematic cross-analysis to identify convergences, logical divergences due to iterative development, and expansions that enriched the interpretation of results. To facilitate this process, a joint display strategy was employed, aligning analytical dimensions across both strands using MAXQDA

24 software (MAXQDA 2024, version 24.4.1, VERBI Software, Berlin, Germany). Qualitative findings from expert interviews and the focus group guided iterative refinements of the prototype, and these changes were subsequently validated through quantitative ratings. This developmental logic resulted in patterns of partial convergence, logical divergences by development, and confirmatory expansion across constructs.

Three joint displays were constructed, as can be seen in Table 7, corresponding to the study's core constructs: contextual adaptation, pedagogical quality, and technical quality. These displays demonstrated high coherence across data strands. Experts' initial concerns during the qualitative phase, such as clarity of instruction, progression, or visual consistency, were addressed in the redesign, and high quantitative ratings (all means above 4 on a 5-point Likert scale) confirmed the success of these adjustments.

Table 7. Summary of mixed-methods integration types by construct.

Construct	Integration Outcome	Nature of Integration
Contextual Adaptation	Converging and expanding	Qualitative insights guided design refinement (e.g., language, flexibility, aesthetics). Likert ratings confirmed adequacy. Iterative adjustments resolved early concerns.
Pedagogical Quality	Logical divergence + confirmation	Experts identified areas for improvement (challenge variety, progression); iterative adjustments led to converging positive ratings.
Technical Quality	Logical divergence + confirmation and expansion	Experts initially identified technical issues (bugs, navigation). These informed refinements were later validated by high usability and performance ratings. Integration revealed expansion and convergence.

The integration process yielded the following outcomes:

- **Convergence:** Both data strands confirmed the strengths of the prototype in terms of contextual relevance, ease of use, and motivational appeal. For example, the prototype's alignment with paediatric oncology conditions (e.g., short sessions, reduced load) was consistently praised across methods.
- **Expansion:** Qualitative narratives provided depth to quantitative trends. While visual and auditory aesthetics received high scores, experts elaborated on specific tensions (e.g., background music monotony, visual overstimulation) that informed future improvements.
- **Divergence by development:** In areas such as feedback mechanisms or instructional clarity, qualitative critiques prompted design revisions. These refinements were later validated by positive quantitative ratings, illustrating a cycle of development-confirmation that aligns with Design-Based Research principles.
- **Cross-construct coherence:** The analysis revealed interdependencies among the three core constructs. For instance, high technical quality was necessary to deliver pedagogical content effectively, and the contextual adaptation of the interface supported both usability and didactic effectiveness. Adaptation to context emerged as the central axis reinforcing both pedagogical and technical quality.
- **Meta-inferences:** The integrated results validated the relevance of Design-Based Research as a methodological framework for creating digital learning tools tailored to clinical-educational environments. The iterative integration of expert input contributed to a high-fidelity prototype evaluated as pedagogically sound, technically functional, and contextually appropriate for hospital classrooms.

Additionally, the analysis revealed strong interdependence among the constructs. For instance:

- The technical performance of the prototype was directly linked to its contextual feasibility in hospital settings. Stability and ease of interaction were essential to avoid frustration in vulnerable learners.
- The pedagogical strength of the game depended on its contextual adaptation. Tailored pacing, language, and representation of content were crucial to addressing the cognitive and emotional needs of children undergoing cancer treatment.
- Technical and pedagogical dimensions were mutually reinforcing. Experts noted that intuitive interaction design supported educational understanding, while visual and auditory coherence facilitated cognitive engagement.

Finally, this phase enabled the formulation of further meta-inferences about the success of the prototype:

- The presence of multiple logical divergences by development confirmed that initial qualitative criticisms triggered meaningful improvements, validated later through high Likert ratings (all means > 4).
- The most consistent point of integration was the adaptation to the paediatric oncology context, which acted as a central pillar in enhancing both pedagogical relevance and technical feasibility.
- The concept of confirmation with expansion emerged repeatedly: quantitative validation was complemented by rich qualitative insight that explained why the design was perceived as effective.

This integrated analysis highlighted how iterative refinement, grounded in expert feedback, resulted in a final prototype that is pedagogically, contextually, and technically sound. The mixed-methods process validated the logic of the design-based research (DBR) framework, where expert-informed adjustments during the qualitative phase were corroborated by positive quantitative evaluation scores. Additionally, the analysis also confirmed the multidimensional strength of the final prototype, which validated the prototype's pedagogical, contextual, and technical adequacy for use in hospital-based pre-fluency interventions targeting young children with central nervous system tumours.

In light of the highly positive expert evaluations and the lack of statistically significant differences across professional backgrounds, an exploratory assessment of the prototype's potential impact and feasibility in real-world clinical-educational settings was considered appropriate. This initiative was encouraged by a teacher working in an oncology hospital classroom in the Canary Islands, who recognised the educational promise of the prototype and advocated for its integration into daily instructional practice. Notably, this application took place despite the prototype still being under development and not yet finalised.

Given the early educational stage of the participants (ages 5–6, enrolled in the final year of Infant Education) and the limited availability of validated literacy assessment tools designed for this age group in the Spanish context, careful methodological consideration was required for instrument selection. In Spain, the official Infant Education curriculum promotes a global, exploratory introduction to literacy, with formal reading instruction typically beginning in Primary Education. As such, administering comprehensive standardized reading tests at this stage would be both pedagogically inappropriate and psychometrically unreliable.

Moreover, the exploratory nature of this pilot, implemented as a single-group post-test design, was not intended to establish causal relationships or generate generalizable findings. Instead, it aimed to offer preliminary insights into the prototype's educational

potential. To achieve this, three validated instruments were selected for their alignment with foundational literacy skills: the Rapid Naming Test (TDR), the reading subtest of the Spanish Test of Basic Instrumental Aspects in Language and Mathematics (PAIB-1), and section 3A of the Spanish Reading and Writing Test (LEE), focusing on basic sentence-level reading. These tools specifically targeted rapid naming, initial decoding, and early indicators of fluency, thus aligning with both developmental appropriateness and curricular expectations.

In addition, fluency performance was assessed using the Scale of Reading Fluency in Spanish (SRFS; Escala de Fluidez Lectora en Español, EFLE) [185]. The TDR measured rapid automatized naming, widely acknowledged as a strong predictor of reading fluency. The PAIB-1 captured syllable and word-level reading skills in Spanish-speaking preschoolers, while the selected LEE items assessed sentence-level decoding. The SRFS was applied across both PAIB-1 and LEE tasks, offering a multicomponential fluency evaluation encompassing reading speed, accuracy, prosody (including volume, intonation, pauses, and phrasing), and an additional component (reading quality) to provide a holistic perspective. Each component was rated using a four-point scale (1 = lowest, 4 = highest), with standardised descriptors outlined in Appendix 1 of the EFLE scoring scale [185].

Four children (1 male, 3 female) participated in this pilot implementation, selected according to strict inclusion criteria: age (5–6 years), enrolment in preschool, no prior formal literacy instruction, ongoing treatment for oncology conditions, and diagnosis involving the central nervous system.

As presented in Table 8, results from the Rapid Naming Test indicated that two participants (P1 and P4) exhibited high naming speed ($\geq$80th percentile), while the remaining two (P2 and P3) demonstrated moderate performance, falling between the 40th and 50th percentiles. The administered subtests included object and colour naming, standardised for this age group, and letter and number naming, which were used qualitatively and interpreted using first-grade reference norms due to the participants' pre-literacy stage.

Table 8. Rapid Naming Test (TDR in Spanish) Results.

Participant	Objects	Colours	Numbers	Letters	Level (Estimated Percentile)
P1	49″	44″	39″	41″	High ($\geq$p90)
P2	65″	69″	42″	48″	Medium (~p50)
P3	67″	72″	44″	47″	Medium (~p40)
P4	51″	46″	42″	43″	High (p80–90)

In the PAIB-1, designed to assess isolated word reading, three out of four participants performed within the medium range on all indicators (direct score, centile, standardised score, and T score). One participant (P3) fell into the low range, indicating difficulties in early reading processes (Table 9).

Table 9. PAIB-1 Reading Scores.

Participant	Direct Score	Percentile	Standardised (Score)	T Score	Interpretation
P1	23	35 (medium)	7 (medium-high)	44 (medium)	Medium
P2	21	25 (medium)	6 (medium-high)	42 (medium)	Medium
P3	16	15 (low)	4 (medium-low)	38 (low)	Low
P4	22	30 (medium)	6 (medium-high)	43 (medium)	Medium

For the assessment of sentence-level reading fluency, two sentences from Section 3A of the LEE (Test de Lectura y Escritura en Español) were selected. Although the LEE is standardised for students in Grades 1 to 4 of Primary Education, the chosen items were specifically selected for their brevity and syntactic simplicity, rendering them suitable for the developmental level of the participants, who were five-year-old children undergoing oncology treatment who had not yet commenced formal literacy instruction.

Given these developmental considerations, the administration of full subtests or longer passages would have been inappropriate, both pedagogically and methodologically. Moreover, since the game prototype was still in its early stages and not designed to produce significant gains in complex reading comprehension, these simplified sentences served as controlled indicators of emergent sentence-level fluency. To complement this, word-level reading was independently assessed using the PAIB-1 test, which includes isolated lexical items appropriate for learners at the pre-reading stage.

The adaptation of the LEE—limiting it to two brief sentences—was both ethically justified and methodologically necessary to ensure developmental validity, reduce cognitive burden, and maintain alignment with the study's exploratory aims.

Based on the application of the SRFS-EFLE (Escala de Fluidez Lectora en Español) across performance on the PAIB-1 and LEE assessments, 75% of participants (3 out of 4) achieved an overall fluency score equal to or above the group mean (M = 3.0). These students (Students 1, 2, and 4) obtained ratings of 3 or higher in at least three of the four assessed components: reading speed, accuracy, prosody, and reading quality. In contrast, Student 3 showed below-average performance, particularly in speed, prosody, and reading quality, which is consistent with their lower scores on the PAIB-1 subtest, as detailed in Tables 10 and 11.

Table 10. Reading Fluency Components—SRFS-EFLE.

Component	P1	P2	P3	P4
1. Speed	4	3	2	3
2. Accuracy	3	3	3	3
3. Prosody	3	3	2	3
4. Reading Quality	4	3	2	3

Table 11. Descriptive Statistics—SRFS-EFLE (N = 4).

Component	Mean (M)	Median (Md)	Mode (Mo)	SD
1. Speed	3.00	3.00	3.00	0.82
2. Accuracy	3.00	3.00	3.00	0.00
3. Prosody	2.75	3.00	3.00	0.50
4. Reading Quality	3.00	3.00	3.00	0.82

Given the limited sample size (N = 4), which reflects the exploratory nature of the pilot within a hospital classroom setting, the analysis was restricted to descriptive statistics, without inferential testing. The results suggest a generally favourable trend in early reading fluency development following the intervention. Specifically, three out of four participants achieved scores at or above the group mean in overall initial fluency, as measured by the SRFS-EFLE scale.

The strongest performance was recorded in the accuracy component, where all participants received an identical score (M = 3.0, SD = 0.00), indicating a high and consistent ability to decode words or sentences correctly. In contrast, greater variability was observed in the components of reading speed and reading quality (SD = 0.82). While not designed to address effectiveness, this pilot study aimed to explore feasibility and inform final development; subsequent research should, therefore, incorporate larger samples and control group designs where feasible.

5. Discussion, Conclusions and Future Research

This study presented the instructional design, development, and evaluation of a serious game prototype aimed at fostering rapid naming, decoding, and reading automatization in preschool-aged children undergoing long-term hospitalisation in oncology settings.

The game was grounded in an extensive three-year process of literature review and iterative development, informed by evidence-based pedagogical practices, principles of game design, and the experiential insights of hospital educators, alongside experts in Educational game design; educational technology; computer engineering, programming and development; psycholinguistics and reading processes applied to video games; hospital education and health with technology; and accessibility in reading games.

Drawing on evidence-based practices from paediatric oncology classrooms, the resulting prototype was designed to merge playful engagement with targeted literacy outcomes. Its evaluation revealed that the game aligns well with the educational and clinical needs of its intended context. This alignment was further validated through highly favourable feedback from a diverse group of experts—including preschool hospital teachers and game design experts and professionals—who praised its pedagogical coherence, usability, and contextual relevance. The convergence of insights from these varied perspectives lends strong credibility to the qualitative assessment and reinforces the prototype's potential as a valuable supplementary tool for supporting instructional goals in paediatric hospital classrooms.

Quantitative results mirrored this positive perception: across all three evaluated dimensions (contextual adaptation, pedagogical quality, and technical quality), mean scores ranged from 4.16 to 4.28 out of 5, indicating uniformly high levels of approval among 274 expert respondents. These data corroborate the prototype's potential to bridge gaps in stimulating pre-reading fluency processes under clinical and developmental constraints.

Although the post-test pilot sample was limited to four children—reflecting the exploratory nature of the study—the primary aim was not to evaluate effectiveness, but rather to assess the feasibility of implementing the prototype within real hospital classroom settings and to inform its future development. Preliminary findings suggest that the game may support early gains in reading fluency, particularly in naming accuracy and basic decoding skills. These encouraging results highlight the need for further research with larger and more diverse clinical samples to evaluate the game's generalizability and its potential for measurable educational impact.

Importantly, the small pilot cohort was not pre-planned but emerged from a contextual opportunity, encouraged by a teacher who saw value in the game's educational potential even at a pre-final stage. Ethical considerations, especially the participants' vulnerability due to age, health condition, and treatment status, necessitated a cautious, non-intrusive exploratory approach, consistent with the British Educational Research Association (BERA) Risk–Benefit Guidelines. As such, the study was exploratory and descriptive in nature, with a focus on validating design quality rather than establishing statistical significance in learning outcomes. Although children's subjective experiences and motivational indicators were not systematically assessed using formal instruments, informal teacher feedback indicated a generally positive reception, with boys showing particular enthusiasm. This anecdotal input, though not empirically analysed, offers initial insight into engagement and will inform future research.

Furthermore, this approach facilitated the extraction of evidence-informed design principles while simultaneously identifying key areas for improvement—such as accessibility, interface usability, narrative coherence, and motivational appeal—all of which have been implemented in the more recent version of the prototype. Building on these insights, this research calls for a re-examination of traditional educational design paradigms, advocating

for pedagogical innovation through interactive and inclusive technologies. In doing so, it also surfaces important epistemological, methodological, and ethical considerations that future studies must address to establish a rigorous framework for developing serious games in sensitive clinical contexts.

In addition to its methodological contribution, this study presents a serious game whose design embodies a set of distinctive features that clearly differentiate it from existing literacy-oriented educational games. First, unlike general-purpose applications, the prototype was conceived specifically for preschool children undergoing oncological treatment for CNS tumours, a population that frequently exhibits neurocognitive vulnerabilities in rapid naming, processing speed, and fluency. The game is not intended as a vehicle for formal reading instruction; rather, it serves as a pre-reading cognitive-linguistic training tool, focusing on foundational processes such as syllable segmentation, rapid naming, and decoding automatization, which are critical precursors to fluent reading. Importantly, these tasks are not presented in isolation, but are embedded in a high-fidelity, story-driven role-playing environment that substitutes traditional drill-based approaches with quests, virtual classrooms and interactive scenarios. This integration of pedagogy and gameplay is supported by an eclectic instructional model that aligns each learning objective with specific mini-games and narrative mechanics (see Sections A–C). Furthermore, the aesthetic design—including neutral characters, adaptive pacing, and emotionally safe audio-visual elements—was intentionally developed for hospitalised contexts, where children often face attentional variability, fatigue, and emotional vulnerability. Finally, the prototype underwent expert validation involving both clinical-educational professionals and game design specialists, ensuring its feasibility and relevance in paediatric oncology settings. Taken together, these characteristics position *Yuki's Adventures: Hidden Words* as a context-specific, theory-informed, and clinically sensitive innovation that extends current research in educational technology and provides a foundation for future studies on fluency development in vulnerable paediatric populations.

Building on these, the primary contribution of this work lies in the elaborate design of an adaptable, pedagogically robust serious game that harmonises educational content with engaging gameplay for children facing extraordinary learning barriers. The resulting design patterns offer a valuable reference point for educators, researchers, and developers seeking to create meaningful, skill-transfer-driven serious games in similarly constrained or high-stakes learning environments.

Beyond its practical design value, this study also advances current literature in several key areas. First, it addresses a significant gap by designing and evaluating a serious game tailored to hospitalised preschool children with CNS tumours—an underrepresented population in educational game research. Second, it offers an initial validation of the preliminary prototype's design, grounded in principles of neuroeducation, inclusive pedagogy, and accessible game design, all tailored to a clinical-educational context. This validation serves as a foundational step towards developing a fully functional final version of the game, ensuring pedagogical coherence, contextual relevance, and usability are achieved prior to broader implementation. In this context, if a fully developed final version of the game were pursued, a 24-week implementation period—double the 12-week prototype—would be more suitable to foster deeper engagement and sustained learning. Third, the study makes a methodological contribution by applying a Design-Based Research model within a highly sensitive clinical setting, integrating expert validation with exploratory pilot testing. These elements set the work apart from previous studies and provide a replicable framework for advancing serious games design for vulnerable learners' population.

In light of these contributions, the present study provides both an immediate educational resource and a foundation for future advancements in the design of serious games

for vulnerable learners. Specifically, this study not only offers a validated tool tailored to the needs of children in paediatric oncology but also proposes a methodological model for developing and evaluating serious games in complex learning contexts. In doing so, it highlights the transformative potential of educational games when they are grounded in learner-centred pedagogy, guided by empathy, and developed with scientific rigour. At the same time, it challenges the traditional view of video games as mere entertainment, reframing them instead as powerful cognitive and emotional mediators within fragile and discontinuous educational environments such as hospital classrooms.

Within this scope, the present study did not integrate extended user-centred data such as systematic child feedback or long-term gameplay analytics. While these forms of evidence would certainly enrich final development, their collection lies beyond the objectives of the present exploratory phase. Instead, the prototype offers a conceptual and technical foundation that can serve as a reference point for subsequent projects or collaborations in which user-centred evidence might be incorporated more extensively.

Moreover, although no further iterations are planned at this stage, as the final prototype design has been completed, validated, and the exploratory research line concluded, the study has laid a strong foundation and produced a validated design and design principles that can inform the development of the full-scale serious game in the future, contingent on adequate funding and resources. Looking ahead, research exploring learner motivation and engagement in similar hospital-based serious games could benefit from validated, developmentally appropriate instruments tailored to preschool populations. The Leuven Scale of Involvement [186], for instance, provides a structured observational framework for assessing behavioural engagement and emotional well-being, while the Fun Toolkit [187] offers child-friendly self-report methods such as the Smileyometer and Fun Sorter to capture affective responses to digital learning. These instruments are well aligned with the developmental and cognitive profiles of hospitalised preschoolers and could be applied effectively in clinical-educational contexts in future studies within the field.

Evaluating motivation and engagement in this highly specific population—five-year-old children hospitalised with CNS tumours—would also require an ethically sensitive and contextually appropriate design. A complementary concurrent triangulation mixed-methods approach [188], prioritising quantitative measures while drawing on qualitative observations for additional nuance, could offer a robust framework. Such an approach would acknowledge the multifactorial nature of engagement, which is influenced not only by the game itself but also by mediators such as teacher facilitation [189], environmental conditions, and the impact of illness [190].

Ultimately, this study marks an important step forward while also opening several avenues for future research. Building on the present findings will require studies with larger samples, control groups, and both baseline and post-intervention assessments using validated, age-appropriate instruments, complemented by longitudinal follow-up. In parallel, integrating in-game analytics, AI-based scaffolding, and speech recognition will be crucial to enable real-time adaptation and support personalised learning trajectories. Finally, examining the transferability of this framework to other sensitive educational contexts could extend its social impact and enhance its broader systemic relevance.

Author Contributions: Conceptualization, J.P.T.-S. and M.Á.P.-H.; methodology, J.P.T.-S. and M.Á.P.-H.; software, J.P.T.-S.; validation, J.P.T.-S. and M.Á.P.-H.; formal analysis, J.P.T.-S. and M.Á.P.-H.; investigation, J.P.T.-S.; data curation, J.P.T.-S. and M.Á.P.-H.; writing—original draft preparation, J.P.T.-S.; writing—review and editing, J.P.T.-S. and M.Á.P.-H.; visualization, M.Á.P.-H. and J.P.T.-S.; supervision, M.Á.P.-H.; Project administration, M.Á.P.-H. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding. This research is a part of PhD scholarship at an international programme by Iberoamerican International University.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki and approved by Ethics Committee of Iberoamerican International University (protocol code CR-177 approved on 11/11/22) with additional institutional approval granted by the University of Jaén.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study prior to participation.

Data Availability Statement: The data presented in this study are not available in accordance with Regulation (EU) of the European Parliament and of the Council 2016/679 of 27 April 2016 regarding the protection of natural persons with regard to the processing of personal data and the free circulation of these data (RGPD) and due to ethical committee conditions, user research data is only accessible to authors and supervisors.

Acknowledgments: The authors wish to thank the Iberoamerican International University for the opportunity to carry out this investigation.

Conflicts of Interest: The authors declare no conflicts of interest.

Appendix A

Table A1. Objectives and Game Mechanics by World in *Yuki's Adventures: Hidden Words*.

World	Instructional Objective	Learning Dynamics and Game Mechanics	Reading Content and Targets
World 1: Rurópolis	As the target users are five-year-old children—many of whom may have limited or no prior exposure to video games—the objective is to familiarise them with the game interface and mechanics. The world introduces rapid naming, recognition of high-frequency Spanish words by semantic categories, and initial reading segmentation through mini-games focused on words, syllables, the alphabetic principle, and phonemes.	Players interact with the environment to identify and remember the names of common objects. Items must be collected and stored, either for survival within the forest or for future trade in World 3. The mechanics promote visual recognition, categorisation, and associative memory.	Complete, categorised, high-frequency words from the child's immediate environment containing all 27 Spanish graphemes and basic syllabic combinations. Letter identification in both upper and lower case.
World 2: Bosque Luminaria	This stage focuses on developing fluency in decoding increasingly complex Spanish syllables: direct (e.g., *pa*, *ma*), inverse (e.g., *al*, *as*), and mixed or clenched syllables (e.g., *pal*, *cas*), along with simple word decoding.	Players explore the forest, locating hidden objects and participating in challenges that require decoding syllables embedded within words. The mini-games progressively cover 22 graphemes and introduce inverse and mixed syllables. Tasks rely on matching, sequencing, and timed recognition.	Direct syllables and simple words. Introduction to mixed and inverse syllables, progressing toward more complex syllabic structures.
World 3: Educacity	The goal is to automate decoding and promote initial fluency in reading common words and simple sentences. Activities are designed to reinforce the grapheme–syllable–word–sentence hierarchy and support transfer to everyday reading situations.	Each in-game day is divided into two sessions: **Morning (Virtual Classroom):** The teacher introduces a new grapheme, associated syllables, and target vocabulary. **Afternoon (Virtual Reading Club):** Players practise sentence reading using the new grapheme and the one introduced the day before. Every four graphemes, players embark on a field trip to a recreational location: book café, basketball court, football field, volleyball court, hockey court, tennis court, rugby field, beach, and finally, an amusement park.	Mixed syllables, inverse syllables, clenched syllables, and short, meaningful sentences. Emphasis on decoding automatisation and fluency with cumulative content.

Appendix B

Table A2. Correspondence between the eclectic method activities and game-based tasks in *Yuki's Adventures: Hidden Words*.

Instructional Activity (Eclectic Method)	Corresponding In-Game Task
Rapid naming of visual images	Display of objects on screen upon entering the castle and forest in World 1 and 2.
Visual word recognition	Exploration inside the house in World 1.
Word–object association	Exploration in the antechamber and dungeon of World 1, and in the forest of World 2. Also practiced in mini-games: association, hockey, and rugby goal-shooting.
Phoneme recognition (phonological awareness)	Mini-games in the basketball pavilion and on the beach (volleyball).
Initial letter–sound identification	Mini basketball and volleyball games in the school hall.
Auditory discrimination: select the correct word from three oral options	Frog mini-game.
Letter selection: complete the word by choosing the missing letters (alphabetic principle)	Bubble mini-game.
Direct reading of syllables	Association mini-game, interaction with the teacher in World 3, and all mini-games in the amusement park.
Syllable recognition within words (uppercase/lowercase)	Forest exploration in World 2 with object interaction showing visual and written forms (syllables highlighted in red). Fishing mini-game included.
Syllable arrangement to form words (syllables scrambled)	Association mini-game in the antechamber and dungeon.
Image-based syllable selection (scrambled syllables to match picture)	Football and volleyball mini-games in the school pavilion.
Syllable selection to complete words (syllabic awareness)	Target and rocket mini-games.
Select syllables to build complete words	Multiple mini-games: association, dartboards, basketball, bumper cars, boat ride, Ferris wheel, balloon popping, clown tent, crazy cups, rockets, and karaoke.
Syllable recall reading (spaced repetition and active recall)	Interaction with two girls in the forest of World 2 and with the computer inside the forest hut.
Word reading	Witch challenge in World 2, maze challenge, and teacher interactions in World 3.
Word–image matching	Card and virtual reality mini-games in the forest.
Image–word matching	Mini-game on the beach.
Sentence reading	Reading Club in World 3; interaction with the girl in the school cafeteria; mini-games: football, volleyball, hockey, tennis, rugby, clown tent, and crazy cups.
Sentence–image matching (highlighted word in red)	Football goal-shooting mini-game in the pavilion.
Select the image that represents the full sentence	Roller coaster, rockets, and karaoke mini-games.
Lexical awareness: arrange words to form a sentence	Bumper cars and boat mini-games.
Sentence–image correspondence (choose the sentence that best matches the image)	Ferris wheel mini-game.

Appendix C

Table A3. Mapping of eclectic method activities to their pedagogical purpose.

Eclectic Method Activity Adapted in the Prototype	Pedagogical Purpose
Rapid naming of visual images	Enhance speed of lexical access and visual-verbal retrieval, supporting rapid automatized naming (RAN) development.
Visual word recognition	Foster the ability to identify and internalise high-frequency words through visual exposure.
Associating words with elements	Strengthen the connection between written language and concrete referents to support semantic development.
Phoneme recognition (phonological awareness)	Enhance phonemic awareness by improving the discrimination and articulation of individual sounds.
Recognition of the first letter of the word and its sound	Support grapheme–phoneme correspondence and phonological decoding.
Listen to three options and select the corresponding word	Improve auditory discrimination, listening comprehension, and lexical retrieval.
Select missing letters in a word	Reinforce the alphabetic principle and orthographic processing.
Direct reading of syllables	Promote early fluency and decoding accuracy at the syllabic level.
Select elements associated with target syllables and visual cues	Reinforce syllabic learning through image–syllable associations.
Reconstruct a word from scrambled syllables	Develop lexical assembly skills and understanding of syllabic structure.
Select syllables in order to form a word matching an image	Strengthen image–word mapping and sequential syllabic decoding.
Discriminate and select missing syllables	Support word construction and analysis through syllabic awareness.
Select syllables to form a complete word	Consolidate word synthesis and phonological construction
Syllable recall reading (spaced repetition and active recall)	Enhance retention and fluency through repetition and memory activation
Word reading	Foster lexical fluency and reading comprehension.
Select the correct word or verb to match an image	Strengthen semantic understanding and visual-word association.
Select the corresponding image for a written word	Reinforce the link between orthographic input and semantic representation.
Simple sentence reading	Initiate fluency and syntactic processing at very basic sentence level.
Select the image corresponding to a word highlighted in a sentence	Improve lexical focus and sentence comprehension through selective attention.
Select the image that best represents a sentence	Support inferential understanding and global sentence meaning.
Order words to form a sentence	Develop syntactic awareness and sentence construction skills.
Select the sentence that best matches an image	Foster comprehension, fluency, and integration of text with visual context.

Appendix D

Table A4. Overview of article structure and key contributions.

Section	Summary
Introduction	Introduces the dual challenges faced by children with central nervous system (CNS) conditions in oncology settings. Highlights overlap between brain regions and reading fluency skills, underscoring the need for targeted interventions.
Background and Related Works	Reviews literature on paediatric oncology pedagogy, reading fluency (rapid automatized naming, decoding, fluency), serious games in reading fluency, paediatric and healthcare education and evaluation. Identifies not only the lack of literacy-focused serious games for paediatric oncology classrooms but the need for cognitively adapted tools targeting pre-reading processes, as children with CNS tumours often experience disruptions in neural circuits associated with fluency development due to both tumour location and treatment-related sequelae, as highlighted in recent neurocognitive research.
Materials and Methods	Describes the design of a serious game *Yuki's Adventures: Hidden Words* using a design-based research approach. Details pedagogical features (scaffolding, spaced repetition, explicit instruction, accessibility) and evaluation methods (focus group, interviews, validated questionnaire, expert review).
Results	A focus group with hospital-based teachers (N = 6) and in-depth interviews with 30 experts across fields such as educational game design; educational technology; psycholinguistics and reading processes applied to digital environments; computer engineering, programming, and development; hospital education and health-related technologies; and accessibility in reading-focused games provided qualitative evidence supporting the prototype's contextual adequacy, pedagogical relevance, and usability. Participants emphasised its alignment with the emotional and cognitive profiles of children with CNS tumours and recognised its potential to stimulate motivation and foundational pre-reading fluency and literacy skills, although certain modifications were deemed necessary and subsequently implemented. Expert ratings (N = 274) on contextual, pedagogical, and technical quality were consistently high (all > 4.0/5). Mixed-methods integration confirmed strong cross-construct coherence and developmental convergence. Additionally, a hospital-based pilot with four preschool children in an oncology classroom (ages 5–6) indicated promising trends in stimulating pre-reading processes.
Discussion, Conclusions and Future Work	Validates the prototype's pedagogical, contextual, and technical design quality, establishing design guidelines that can inform the creation of full-scale serious games in similar contexts. Highlights methodological contributions and points to the need for future research with larger samples, validated instruments, and advanced features (e.g., in-game analytics, AI scaffolding, speech recognition), as well as examining transferability to other educational settings.

References

1. Manzano-León, A.; Rodríguez-Ferrer, J.M.; Aguilar-Parra, J.M.; Fernández-Campoy, J.M.; Trigueros, R.; Martínez-Martínez, A.M. Juega y aprende: Influencia de la gamificación y aprendizaje basado en juego en los procesos lectores de alumnado de secundaria. *Rev. Psicodidáct.* **2022**, *27*, 38–46. [CrossRef]
2. Manzano-León, A.; Rodríguez-Ferrer, J.M.; Collado-Soler, R. Play and Reading Process of Secondary Students in Areas of Need for Social Transformation. *Bellaterra J. Teach. Learn. Lang. Lit.* **2023**, *16*, e1099. [CrossRef]
3. Rodríguez-Ferrer, J.M.; Manzano-León, A.; Aguilar-Parra, J.M.; Cangas, A. Effectiveness of Gamification and Game-Based Learning in Spanish Adolescents with Dyslexia: A Longitudinal Quasi-Experimental Research. *Res. Dev. Disabil.* **2023**, *141*, 104603. [CrossRef]
4. Ait-Abdellah-Sefian, H.; Rodríguez-Ferrer, J.-M.; Rodríguez-Moreno, J.; Aguilar-Parra, J.M.; Manzano-León, A. Promoción de funciones ejecutivas y fluencia mediante juegos en estudiantes en situación de incorporación tardía al sistema educativo. *Ocnos* **2025**, *24*, 1. [CrossRef]

5. Lee, M.M.; Stoodley, C.J. Neural Bases of Reading Fluency: A Systematic Review and Meta-Analysis. *Neuropsychologia* **2024**, *202*, 108947. [CrossRef] [PubMed]
6. Cavalheiro, S.; Valsechi, L.C.; Dastoli, P.A.; Nicácio, J.M.; Cappellano, A.M.; Saba Da Silva, N.; Silva Da Costa, M.D. Outcomes and Surgical Approaches for Pineal Region Tumors in Children: 30 Years' Experience. *J. Neurosurg. Pediatr.* **2023**, *32*, 184–193. [CrossRef] [PubMed]
7. Seco Pozo, A.C.; Mederos Rodríguez, A.; García Fariña, M.; González Cruz, M.; Caparrós Nieto, A.B.; Palenzuela Afonso, B.; García Conde, M.; Febles García, P. Tumores del sistema nervioso central en la edad pediátrica. *Rev. Cas. Can. Pediátr.* **2023**, *47*, 264–273.
8. Vásquez Gómez, F.; Carceller Ortega, E.; Lassaletta Atienza, Á. Tumores cerebrales en niños. *Pediatr. Integr.* **2021**, *15*, 357–366.
9. Peña Hita, M.Á.; Tacoronte Sosa, J.P. Videojuegos, fluidez lectora y aulas hospitalarias: Una revisión multivocal de la literatura. *Profesorado* **2024**, *28*, 223–245. [CrossRef]
10. Dinç, F.; Kurt, A.; Yıldız, D. Hospital Classroom Experiences of Hospitalized School-Age Pediatric Oncology Patients: A Qualitative Study. *J. Pediatr. Nurs.* **2023**, *73*, e220–e226. [CrossRef]
11. Wang, H.; Liu, F.; Yu, D. Complex Network of Eye Movements during Rapid Automatized Naming. *Front. Neurosci.* **2023**, *17*, 1024881. [CrossRef]
12. Norton, E.S.; Wolf, M. Rapid Automatized Naming (RAN) and Reading Fluency: Implications for Understanding and Treatment of Reading Disabilities. *Annu. Rev. Psychol.* **2012**, *63*, 427–452. [CrossRef]
13. Kendeou, P.; Van Den Broek, P.; Helder, A.; Karlsson, J. A Cognitive View of Reading Comprehension: Implications for Reading Difficulties. *Learn. Disabil. Res. Pract.* **2014**, *29*, 10–16. [CrossRef]
14. Adlof, S.M.; Catts, H.W.; Little, T.D. Should the Simple View of Reading Include a Fluency Component? *Read. Writ.* **2006**, *19*, 933–958. [CrossRef]
15. Archer, A.L.; Gleason, M.M.; Vachon, V.L. Decoding and Fluency: Foundation Skills for Struggling Older Readers. *Learn. Disabil. Q.* **2003**, *26*, 89–101. [CrossRef]
16. Calet, N.; Gutiérrez-Palma, N.; Defior, S. Effects of Fluency Training on Reading Competence in Primary School Children: The Role of Prosody. *Learn. Instr.* **2017**, *52*, 59–68. [CrossRef]
17. Ecalle, J.; Dujardin, E.; Gomes, C.; Cros, L.; Magnan, A. Decoding, Fluency and Reading Comprehension: Examining the Nature of Their Relationships in a Large-Scale Study with First Graders. *Read. Writ. Q.* **2021**, *37*, 444–461. [CrossRef]
18. Foorman, B.R.; Wu, Y.-C.; Quinn, J.M.; Petscher, Y. How Do Latent Decoding and Language Predict Latent Reading Comprehension: Across Two Years in Grades 5, 7, and 9? *Read. Writ.* **2020**, *33*, 2281–2309. [CrossRef]
19. Hjetland, H.N.; Brinchmann, E.I.; Scherer, R.; Hulme, C.; Melby-Lervåg, M. Preschool Pathways to Reading Comprehension: A Systematic Meta-Analytic Review. *Educ. Res. Rev.* **2020**, *30*, 100323. [CrossRef]
20. Hudson, A.; Koh, P.W.; Moore, K.A.; Binks-Cantrell, E. Fluency Interventions for Elementary Students with Reading Difficulties: A Synthesis of Research from 2000–2019. *Educ. Sci.* **2020**, *10*, 52. [CrossRef]
21. Kim, Y.-S.; Wagner, R.K.; Foster, E. Relations Among Oral Reading Fluency, Silent Reading Fluency, and Reading Comprehension: A Latent Variable Study of First-Grade Readers. *Sci. Stud. Read.* **2011**, *15*, 338–362. [CrossRef]
22. Kuhn, M.R.; Schwanenflugel, P.J.; Meisinger, E.B. Aligning Theory and Assessment of Reading Fluency: Automaticity, Prosody, and Definitions of Fluency. *Read. Res. Q.* **2010**, *45*, 230–251. [CrossRef]
23. Steinle, P.K.; Stevens, E.; Vaughn, S. Fluency Interventions for Struggling Readers in Grades 6 to 12: A Research Synthesis. *J. Learn. Disabil.* **2022**, *55*, 3–21. [CrossRef]
24. Caravolas, M.; Lervåg, A.; Mousikou, P.; Efrim, C.; Litavský, M.; Onochie-Quintanilla, E.; Salas, N.; Schöffelová, M.; Defior, S.; Mikulajová, M.; et al. Common Patterns of Prediction of Literacy Development in Different Alphabetic Orthographies. *Psychol. Sci.* **2012**, *23*, 678–686. [CrossRef] [PubMed]
25. Cutting, L.E.; Denckla, M.B. The Relationship of Rapid Serial Naming and Word Reading in Normally Developing Readers: An Exploratory Model. *Read. Writ.* **2001**, *14*, 673–705. [CrossRef]
26. Christopher, M.E.; Miyake, A.; Keenan, J.M.; Pennington, B.; DeFries, J.C.; Wadsworth, S.J.; Willcutt, E.; Olson, R.K. Predicting Word Reading and Comprehension with Executive Function and Speed Measures across Development: A Latent Variable Analysis. *J. Exp. Psychol. Gen.* **2012**, *141*, 470–488. [CrossRef]
27. Johnston, T.C.; Kirby, J.R. The Contribution of Naming Speed to the Simple View of Reading. *Read. Writ.* **2006**, *19*, 339–361. [CrossRef]
28. Wimmer, H.; Mayringer, H.; Landerl, K. The Double-Deficit Hypothesis and Difficulties in Learning to Read a Regular Orthography. *J. Educ. Psychol.* **2000**, *92*, 668–680. [CrossRef]
29. Marín Suelves, D.; Vidal Esteve, M.I.; Donato, D.; Granados Saiz, J. Análisis del estado del arte sobre el uso de los videojuegos en Educación Infantil y Primaria. *Innoeduca Int. J. Technol. Educ. Innov.* **2021**, *7*, 4–18. [CrossRef]

30. Puspitasari, C.; Subiyanto, S. A New Tool to Facilitate Learning Reading for Early Childhood. *Malays. Online J. Educ. Technol.* **2017**, *5*, 1–15.
31. De Souza, G.N.; Brito, Y.P.D.S.; Tsutsumi, M.M.A.; Marques, L.B.; Goulart, P.R.K.; Monteiro, D.C.; Santana, Á.L.D. The Adventures of Amaru: Integrating Learning Tasks into a Digital Game for Teaching Children in Early Phases of Literacy. *Front. Psychol.* **2018**, *9*, 2531. [CrossRef]
32. Rodríguez Rodríguez, C.; Jiménez González, J.E.; Díaz Megolla, A.; González Martín, D. Tradislexia: Un videojuego para la mejora de la lectura en niños con dyslexia. In *Nuevas Tendencias en TIC y Educación*; González González, C.S., Ed.; Bubok Publishing: Madrid, Spain, 2011; Volume 1, pp. 122–134.
33. Koval-Saifi, N.; Plass, J. *Antura and the Letters: Impact and Technical Evaluation*; World Vision and Foundation for Information Technology Education and Development: Washington, DC, USA, 2018.
34. Steenbeek-Planting, E.G.; Boot, M.; De Boer, J.C.; Van De Ven, M.; Swart, N.M.; Van Der Hout, D. Evidence-Based Psycholinguistic Principles to Remediate Reading Problems Applied in the Playful App Letterprins: A Perspective of Quality of Healthcare on Learning to Read. In *Games for Health; Springer Fachmedien Wiesbaden*; Springer Fachmedien: Wiesbaden, Germany, 2013; pp. 281–291. [CrossRef]
35. Van Gorp, K.; Segers, E.; Verhoeven, L. Enhancing Decoding Efficiency in Poor Readers via a Word Identification Game. *Read. Res. Q.* **2017**, *52*, 105–123. [CrossRef]
36. Potier Watkins, C.; Caporal, J.; Merville, C.; Kouider, S.; Dehaene, S. Accelerating Reading Acquisition and Boosting Comprehension with a Cognitive Science-Based Tablet Training. *J. Comput. Educ.* **2020**, *7*, 183–212. [CrossRef]
37. Pérez-Quichimbo, S.-M.; Navas-Moya, M.-P.; Montes-León, S.-R.; Sambachi-Chilig, P.-A.; Barrera-Quimbita, E.-D. Therapy Using Serious Games to Improve Phonological Awareness in Children with Functional Dyslexia. In *Communications in Computer and Information Science*; Springer International Publishing: Cham, Switzerland, 2022; pp. 121–134. [CrossRef]
38. Nugraha, M.R.; Rosmansyah, Y.; Mulyani, E.D.S.; Mufizar, T.; Anwar, D.S.; Hidayat, C.R. A Serious Game for Improving Reading Skill of Preschoolers. *Elinvo* **2022**, *7*, 84–91. [CrossRef]
39. Amorim, A.N.; Jeon, L.; Abel, Y.; Albuquerque, E.X.S.; Soares, M.; Silva, V.C.; Oliveira Neto, J.R. Escribo play Learning Games Can Foster Early Reading and Writing for Low-Income Kindergarten Children. *Comput. Educ.* **2022**, *177*, 104364. [CrossRef]
40. Gaytán-Lugo, L.S.; Santana-Mancilla, P.C.; Santarrosa-García, A.; Medina-Anguiano, A.; Hernández-Gallardo, S.C.; Garcia-Ruiz, M.A. Developing a Serious Game to Improve Reading Comprehension Skills in Third Graders. *Res. Comput. Sci.* **2015**, *89*, 71–79. [CrossRef]
41. Casillas, J.C.L.; Valdez, K.M.; Lopez, S.R.F.; Canto, D.; Carranza, D.B.; Negron, A.P.P. Saving the Word: A Serious Game to Improve Kids Reading Skills. In Proceedings of the 2021 Mexican International Conference on Computer Science (ENC), Morelia, Mexico, 9–11 August 2021; pp. 1–4. [CrossRef]
42. Palma, C.A.M.; González, R.M.R. Un software educativo como una herramienta pedagógica en la mejora de las habilidades de lectoescritura utilizando el método ecléctico. *Rev. Tecnol. Cienc. Educ.* **2019**, *2019*, 172–186. [CrossRef]
43. Serrano, F.; Bravo Sánchez, J.F.; Gómez Olmedo, M. Galexia: Evidence-Based Software for Intervention in Reading Fluency and Comprehension. In Proceedings of the 10th International Technology, Education and Development Conference, Valencia, Spain, 7–9 March 2016. [CrossRef]
44. López-Olóriz, J.; Pina, V.; Ballesta, S.; Bordoy, S.; Pérez-Zapata, L. Proyecto Petit UBinding: Método de adquisición y mejora de la lectura en primero de primaria. Estudio de eficacia. *Rev. Logop. Foniatr. Audiol.* **2020**, *40*, 12–22. [CrossRef]
45. Sánchez Castro, S.; Pascual Sevillano, M.Á. Eficacia de un juego serio digital para la mejora de la comprensión lectora y el rendimiento académico. *Investig. Sobre Lect.* **2022**, *17*, 40–66. [CrossRef]
46. Bruckman, A. Can Educational Be Fun? In Proceedings of the Game Developers Conference, San Jose, CA, USA, 15–19 March 1999.
47. González González, C.S. Uso de las TIC en Aulas Hospitalarias: El Caso de los Videojuegos. In Proceedings of the Primer Congreso Internacional De Pedagogía Hospitalaria En Colombia, Bogotá, Colombia, 13–14 March 2015. Available online: https://silo.tips/downloadFile/primer-congreso-internacional-de-pedagogia-hospitalaria-en-colombia-cphcol-intro (accessed on 17 January 2023).
48. González González, C.S. Tecnologías para la e-inclusión de niños y niñas hospitalizados(as). In *Educación y Salud. La Pedagogía Hospitalaria Como un Puente Entre Ambas*; Hergué Editorial: Huelva, Spain, 2019; pp. 103–123.
49. González González, C.S. *Uso de Las TIC Para la Atención Educativa Hospitalaria y Domiciliaria*; McGraw-Hill Interamericana de España: Madrid, Spain, 2015; pp. 145–163.
50. Moriconi, V.; Maroto, C.; Cantero-García, M. Efectividad de la realidad virtual (RV) en la disminución del distrés de niños y adolescentes con cáncer: Revisión sistemática. *Psicooncología* **2022**, *19*, 45–61. [CrossRef]

51. Guerra Antequera, J.; Revuelta Domínguez, F.I. Videojuegos precursores de emociones positivas: Propuesta metodológica con Minecraft en el aula hospitalaria. *Int. J. Educ. Res. Innov.* **2015**, *3*, 105–120.
52. Perandones, E. Creando videojuegos hospitarios: Taller "Aventura en el hospital". *Arte Individ. Soc.* **2012**, *23*, 41–53. [CrossRef]
53. Tacoronte Sosa, J.P.; Peña Hita, M.D.L.Á. Juegos serios para la iniciación lectora en pacientes pediátricos hospitalizados en oncología. *Rev. Tecnol. Cienc. Y Educ.* **2024**, *2024*, 105–132. [CrossRef]
54. Arnab, S.; Clarke, S. Towards a Trans-disciplinary Methodology for a Game-based Intervention Development Process. *Br. J. Educ. Technol.* **2017**, *48*, 279–312. [CrossRef]
55. Bellotti, F.; Berta, R.; de Gloria, A.; Ott, M.; Arnab, S.; de Freitas, S.; Kiili, K. Designing Serious Games for Education: From Pedagogical Principles to Game Mechanisms. In Proceedings of the 5th European Conference on Game-Based Learning, Athens, Greece, 20–21 October 2011; Academic Publishing Ltd.: Reading, UK; pp. 26–34.
56. Marne, B.; Wisdom, J.; Huynh-Kim-Bang, B.; Labat, J.-M. The Six Facets of Serious Game Design: A Methodology Enhanced by Our Design Pattern Library. In *Lecture Notes in Computer Science*; Springer: Berlin/Heidelberg, Germany, 2012; pp. 208–221. [CrossRef]
57. Ávila-Pesántez, D.; Rivera, L.A.; Alban, M.S. Approaches for Serious Game Design: A Systematic Literature Review. *Comput. Educ. J.* **2017**, *8*, 3.
58. Ravyse, W.S.; Seugnet Blignaut, A.; Leendertz, V.; Woolner, A. Success Factors for Serious Games to Enhance Learning: A Systematic Review. *Virtual Real.* **2017**, *21*, 31–58. [CrossRef]
59. Wang, C.-M.; Lee, B.-T.; Lo, T.-Y. The Design of a Novel Digital Puzzle Gaming System for Young Children's Learning by Interactive Multi-Sensing and Tangible User Interfacing Techniques. *Sustainability* **2023**, *15*, 3036. [CrossRef]
60. Abbasi, S.; Kazi, H.; Kazi, A.W.; Khowaja, K.; Baloch, A. Gauge Object Oriented Programming in Student's Learning Performance, Normalized Learning Gains and Perceived Motivation with Serious Games. *Information* **2021**, *12*, 101. [CrossRef]
61. Pesantez, A.L.; Rodriguez, C.A.; Martínez, E.C.; Carvajal, F. Mobile Application Prototype for the Development of Phonological Awareness in Children's School. In Proceedings of the International Conference on Industrial Engineering and Operations Management, Rome, Italy, 26–28 July 2022; pp. 266–277. [CrossRef]
62. Sim, G.; Cassidy, B. Investigating the Fidelity Effect When Evaluating Game Prototypes with Children. In *Electronic Workshops in Computing*; BCS Learning & Development: Swindon, UK, 2013. [CrossRef]
63. Calderón, A.; Ruiz, M. A Systematic Literature Review on Serious Games Evaluation: An Application to Software Project Management. *Comput. Educ.* **2015**, *87*, 396–422. [CrossRef]
64. Desurvire, H.; Caplan, M.; Toth, J.A. Using Heuristics to Evaluate the Playability of Games. In *CHI '04 Extended Abstracts on Human Factors in Computing Systems*; ACM: Vienna, Austria, 2004; pp. 1509–1512. [CrossRef]
65. Escudeiro, P.; Escudeiro, N. Evaluation of Serious Games in Mobile Platforms with QEF: QEF (Quantitative Evaluation Framework). In Proceedings of the 2012 IEEE Seventh International Conference on Wireless, Mobile and Ubiquitous Technology in Education, Takamatsu, Kagawa, Japan, 27–30 March 2012; IEEE: Takamatsu, Kagawa, Japan; pp. 268–271. [CrossRef]
66. Escudeiro, P.; Escudeiro, N. Evaluating Educational Games in Mobile Platforms. *Int. J. Mob. Learn. Organ.* **2013**, *7*, 14. [CrossRef]
67. Estrada-Molina, O.; Fuentes-Cancell, D.R.; Morales, A.A. The Assessment of the Usability of Digital Educational Resources: An Interdisciplinary Analysis from Two Systematic Reviews. *Educ. Inf. Technol.* **2022**, *27*, 4037–4063. [CrossRef]
68. Federoff, M.A. Heuristics and Usability Guidelines for the Creation and Evaluation of Fun in Video Games. Ph.D. Thesis, Indiana University, Bloomington, IN, USA, 2002.
69. Frazão, K.; Costa, J.; Viana, D.; Rivero, L. Analyzing App Store Comments and Quality Attributes for Defining an Inspection Checklist for Mobile Educational Games. In Proceedings of the XXXIV Brazilian Symposium on Software Engineering, Natal, Brazil, 21–23 October 2020; pp. 854–859. [CrossRef]
70. Harpur, P.-A.; De Villiers, R. MUUX-E, a Framework of Criteria for Evaluating the Usability, User Experience and Educational Features of m-Learning Environments. *S. Afr. Comput. J.* **2015**, *56*, 34–51. [CrossRef]
71. Hermawati, S.; Lawson, G. Establishing Usability Heuristics for Heuristics Evaluation in a Specific Domain: Is There a Consensus? *Appl. Ergon.* **2016**, *56*, 34–51. [CrossRef]
72. Koeffel, C.; Hochleitner, W.; Leitner, J.; Haller, M.; Geven, A.; Tscheligi, M. Using Heuristics to Evaluate the Overall User Experience of Video Games and Advanced Interaction Games. In *Evaluating User Experience in Games*; Human-Computer Interaction Series; Springer: London, UK, 2010; pp. 233–256. [CrossRef]
73. Korhonen, H.; Koivisto, E.M.I. Playability Heuristics for Mobile Games. In Proceedings of the 8th Conference on Human-Computer Interaction with Mobile Devices and Services, Helsinki, Finland, 12–15 September 2006; ACM: Helsinki, Finland; pp. 9–16. [CrossRef]

74. Korhonen, H.; Koivisto, E.M.I. Playability Heuristics for Mobile Multi-Player Games. In Proceedings of the 2nd International Conference on Digital Interactive Media in Entertainment and Arts, Perth, Australia, 19–21 September 2007; ACM: Perth, Australia; pp. 28–35. [CrossRef]
75. Nunes Marciano, J.; Cunha De Miranda, L.; Esteves Cunha De Miranda, E. Evaluating Multiple Aspects of Educational Computer Games: Literature Review and Case Study. *Int. J. Comput. Games Technol.* **2014**, *2014*, 139205. [CrossRef]
76. Mitgutsch, K.; Alvarado, N. Purposeful by Design? A Serious Game Design Assessment Framework. In Proceedings of the International Conference on the Foundations of Digital Games, Raleigh, NC, USA, 29 May–1 June 2012; ACM: Raleigh, NC, USA; pp. 121–128. [CrossRef]
77. Mohamed, H.; Jaafar, A. Development and Potential Analysis of Heuristic Evaluation for Educational Computer Game (PHEG). In Proceedings of the 5th International Conference on Computer Sciences and Convergence Information Technology, Seoul, Republic of Korea, 30 November–2 December 2010; IEEE: Seoul, Republic of Korea; pp. 222–227. [CrossRef]
78. Oprins, E.; Visschedijk, G.; Roozeboom, M.B.; Dankbaar, M.; Trooster, W.; Schuit, S.C.E. The Game-Based Learning Evaluation Model (GEM): Measuring the Effectiveness of Serious Games Using a Standardised Method. *Int. J. Technol. Enhanc. Learn.* **2015**, *7*, 326. [CrossRef]
79. Papaloukas, S.; Patriarcheas, K.; Xenos, M. Usability Assessment Heuristics in New Genre Videogames. In Proceedings of the 2009 13th Panhellenic Conference on Informatics, Corfu, Greece, 10–12 September 2009; IEEE: Corfu, Greece; pp. 202–206. [CrossRef]
80. Phan, M.H.; Keebler, J.R.; Chaparro, B.S. The Development and Validation of the Game User Experience Satisfaction Scale (GUESS). *Hum. Factors* **2016**, *58*, 1217–1247. [CrossRef]
81. Pinelle, D.; Wong, N.; Stach, T. Heuristic Evaluation for Games: Usability Principles for Video Game Design. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Florence, Italy, 5–10 April 2008; ACM: Florence, Italy; pp. 1453–1462. [CrossRef]
82. Trisnadoli, A.; Hendradjaya, B.; Danar Sunindyo, W. A Proposal of Quality Model for Mobile Games. In Proceedings of the 2015 International Conference on Electrical Engineering and Informatics (ICEEI), Denpasar, Indonesia, 10–11 August 2015; IEEE: Denpasar, Indonesia; pp. 377–381. [CrossRef]
83. Yanez-Gomez, R.; Font, J.L.; Cascado-Caballero, D.; Sevillano, J.-L. Heuristic Usability Evaluation on Games: A Modular Approach. *Multimed. Tools Appl.* **2019**, *78*, 4937–4964. [CrossRef]
84. Zaibon, S.B.; Shiratuddin, N. Mobile Game-Based Learning (mGBL): Application Development and Heuristics Evaluation Strategy. *Malays. J. Learn. Instr.* **2010**, *7*, 37–73. [CrossRef]
85. Amory, A. Game Object Model Version II: A Theoretical Framework for Educational Game Development. *Educ. Tech. Res. Dev.* **2007**, *55*, 51–77. [CrossRef]
86. Fokides, E.; Atsikpasi, P.; Kaimara, P.; Deliyannis, I. Let Players Evaluate Serious Games. Design and Validation of the Serious Games Evaluation Scale. *ICGA J.* **2019**, *41*, 116–137. [CrossRef]
87. Fu, F.-L.; Su, R.-C.; Yu, S.-C. EGameFlow: A Scale to Measure Learners' Enjoyment of e-Learning Games. *Comput. Educ.* **2009**, *52*, 101–112. [CrossRef]
88. Martinez, K.; Menéndez-Menéndez, M.I.; Bustillo, A. A New Measure for Serious Games Evaluation: Gaming Educational Balanced (GEB) Model. *Appl. Sci.* **2022**, *12*, 11757. [CrossRef]
89. Paliokas, I.; Arapidis, C.; Mpimpitsos, M. PlayLOGO 3D: A 3D Interactive Video Game for Early Programming Education: Let LOGO Be a Game. In Proceedings of the 2011 Third International Conference on Games and Virtual Worlds for Serious Applications, Athens, Greece, 4–6 May 2011; IEEE: Athens, Greece; pp. 24–31. [CrossRef]
90. Papadakis, S.; Kalogiannakis, M.; Zaranis, N. Designing and Creating an Educational App Rubric for Preschool Teachers. *Educ. Inf. Technol.* **2017**, *22*, 3147–3165. [CrossRef]
91. Petri, G.; Gresse Von Wangenheim, C.; Borgatto, A.F. MEEGA+, Systematic Model to Evaluate Educational Games. In *Encyclopedia of Computer Graphics and Games*; Springer International Publishing: Cham, Switzerland, 2018; pp. 1–7. [CrossRef]
92. Barbosa, M.B.; Rêgo, A.B.; de Medeiros, I. HEEG: Heuristic Evaluation for Educational Games. In Proceedings of the SBGames, Piauí, Brazil, 11–13 November 2015; pp. 224–227.
93. Rooney, P. A Theoretical Framework for Serious Game Design: Exploring Pedagogy, Play and Fidelity and their Implications for the Design Process. *Int. J. Game-Based Learn.* **2012**, *2*, 41–60. [CrossRef]
94. Savi, R.; Wangenheim, C.G.V.; Borgatto, A.F. A Model for the Evaluation of Educational Games for Teaching Software Engineering. In Proceedings of the 2011 25th Brazilian Symposium on Software Engineering, Sao Paulo, Brazil, 28–30 September 2011; IEEE: Sao Paulo, Brazil; pp. 194–203. [CrossRef]
95. Zhang, H.; Fan, X.; Xing, H. Research on the Design and Evaluation of Educational Games Based on the RETAIN Model. In Proceedings of the 2010 Third International Symposium on Knowledge Acquisition and Modeling, Wuhan, China, 20–21 October 2010; IEEE: Wuhan, China; pp. 375–378. [CrossRef]

96. Bakhtiari, R.; Habibzadeh, Z. Designing a Framework and Validating a Tool for Evaluating the Educational Quality of Serious Games: A Meta-Synthesis. *Int. J. Serious Games* **2023**, *10*, 61–83. [CrossRef]
97. Hainey, T.; Connolly, T.M.; Boyle, E.A.; Wilson, A.; Razak, A. A Systematic Literature Review of Games-Based Learning Empirical Evidence in Primary Education. *Comput. Educ.* **2016**, *102*, 202–223. [CrossRef]
98. Missen, M.M.S.; Javed, A.; Asmat, H.; Nosheen, M.; Coustaty, M.; Salamat, N.; Prasath, V.B.S. Systematic Review and Usability Evaluation of Writing Mobile Apps for Children. *New Rev. Hypermedia Multimed.* **2019**, *25*, 137–160. [CrossRef]
99. Petri, G.; Gresse von Wangenheim, C. How to Evaluate Educational Games: A Systematic Literature Review. *J. Univers. Comput. Sci.* **2016**, *22*, 992–1021.
100. Silveira, A.C.; Martins, R.X.; Vieira, E.A.O. E-Guess: Usability Evaluation for Educational Games. *RIED Rev. Iberoam. Educ. Distancia* **2020**, *24*, 245. [CrossRef]
101. Tahir, R.; Wang, A.I. State of the Art in Game Based Learning: Dimensions for Evaluating Educational Games. In Proceedings of the 11th European Conference on Games Based Learning, Graz, Austria, 5–6 October 2017; pp. 641–650.
102. Vee Senap, N.M.; Ibrahim, R. A Review of Heuristics Evaluation Component for Mobile Educational Games. *Procedia Comput. Sci.* **2019**, *161*, 1028–1035. [CrossRef]
103. Vieira, E.A.O.; Silveira, A.C.D.; Martins, R.X. Heuristic Evaluation on Usability of Educational Games: A Systematic Review. *Inform. Educ.* **2019**, *18*, 427–442. [CrossRef]
104. Wang, X.; Tan, J. A Systematic Literature Review on How to Evaluate Educational Games in K-12. *Adv. Educ. Technol. Psychol.* **2021**, *5*, 82–92. Available online: https://clausiuspress.com/article/2435 (accessed on 15 March 2024).
105. Tahir, R.; Wang, A.I. Codifying Game-Based Learning: The LEAGUE Framework for Evaluation. In Proceedings of the 12th European Conference on Game-Based Learning, Trondheim, Norway, 4–5 October 2018; pp. 677–686.
106. Mohd Khir, N.H.B.; Ismail, M. Review on Gamification in Children Computer Interaction (CCI) for Persona Modelling. *Bull. Electr. Eng. Inform.* **2019**, *8*, 1411–1417. [CrossRef]
107. Badkook, N.; Sinnari, D.; Almakky, A. Developing Comprehensive E-Game Design Guidelines to Support Children with Language Delay: A Step-by-Step Approach with Initial Validation. *Multimodal Technol. Interact.* **2025**, *9*, 68. [CrossRef]
108. Cova, Á.; Arrieta, X.; Reveros, V. Análisis y Comparación de diversos modelos de evaluación de software educativo. *Enlace* **2008**, *5*, 45–67.
109. Hinojo-Lucena, F.-J.; Aznar-Díaz, I.; Cáceres-Reche, M.-P.; Romero-Rodríguez, J.-M. Artificial Intelligence in Higher Education: A Bibliometric Study on Its Impact in the Scientific Literature. *Educ. Sci.* **2019**, *9*, 51. [CrossRef]
110. Chanchí, G.E.G.; Álvarez, M.C.G.; Campo, W.Y.M. Criterios de usabilidad para el diseño e implementación de videojuegos. *Rev. Ibér. Sist. Tecnol. Inf.* **2020**, *E26*, 461–474.
111. Software Engineering—Product Quality—Part 1: Quality Model. 2001. Available online: https://www.iso.org/standard/22749.html (accessed on 21 October 2023).
112. Ergonomics of Human-System Interaction—Part 11: Usability: Definitions and Concepts. 2018. Available online: https://www.iso.org/standard/63500.html (accessed on 21 October 2023).
113. Hitchens, M.; Tulloch, R. A Gamification Design for the Classroom. *Interact. Technol. Smart Educ.* **2018**, *15*, 28–45. [CrossRef]
114. Adams, M.J. *Beginning to Read: Thinking and Learning About Print*; 10. print.; A Bradford Book; MIT Press: Cambridge, MA, USA, 1998.
115. Bravo Valdivieso, L. La conciencia fonológica como una posible "zona de desarrollo próximo" para el aprendizaje de la lectura inicial. *Rev. Latinoam. Psicol.* **2004**, *36*, 21–32.
116. Bus, A.G.; Van IJzendoorn, M.H. Phonological Awareness and Early Reading: A Meta-Analysis of Experimental Training Studies. *J. Educ. Psychol.* **1999**, *91*, 403–414. [CrossRef]
117. Chambers, B.; Cheung, A.C.K.; Slavin, R.E. Literacy and Language Outcomes of Comprehensive and Developmental-Constructivist Approaches to Early Childhood Education: A Systematic Review. *Educ. Res. Rev.* **2016**, *18*, 88–111. [CrossRef]
118. Ehri, L.C.; Nunes, S.R.; Willows, D.M.; Schuster, B.V.; Yaghoub-Zadeh, Z.; Shanahan, T. Phonemic Awareness Instruction Helps Children Learn to Read: Evidence From the National Reading Panel's Meta-Analysis. *Read. Res. Q.* **2001**, *36*, 250–287. [CrossRef]
119. Fernández-Otoya, F.A.; Raposo-Rivas, M.; Halabi-Echeverry, A.X. A Qualitative Systematic Literature Review on Phonological Awareness in Preschoolers Supported by Information and Communication Technologies. *Educ. Sci.* **2022**, *12*, 382. [CrossRef]
120. Guarneros-Reyes, E.; Vega-Pérez, L. Patrón de desarrollo de la conciencia fonológica en niños preescolares. *Rev. Digit. Int. Psicol. Cienc. Soc.* **2015**, *1*, 70–82. [CrossRef]
121. Hipfner-Boucher, K.; Milburn, T.; Weitzman, E.; Greenberg, J.; Pelletier, J.; Girolametto, L. Relationships between Preschoolers' Oral Language and Phonological Awareness. *First Lang.* **2014**, *34*, 178–197. [CrossRef]
122. Snowling, M.J.; Hulme, C. Evidence-Based Interventions for Reading and Language Difficulties: Creating a Virtuous Circle. *Br. J. Educ. Psychol.* **2011**, *81*, 1–23. [CrossRef]

123. Stanovich, K.E.; Cunningham, A.E.; Cramer, B.B. Assessing Phonological Awareness in Kindergarten Children: Issues of Task Comparability. *J. Exp. Child. Psychol.* **1984**, *38*, 175–190. [CrossRef]
124. Chai, Z.; Vail, C.O.; Ayres, K.M. Using an iPad Application to Promote Early Literacy Development in Young Children with Disabilities. *J. Spec. Educ.* **2015**, *48*, 268–278. [CrossRef]
125. Goo, M.; Myers, D.; Maurer, A.L.; Serwetz, R. Effects of Using an iPad to Teach Early Literacy Skills to Elementary Students with Intellectual Disability. *Intellect. Dev. Disabil.* **2020**, *58*, 34–48. [CrossRef]
126. Martinez-Figueira, M.E.; Borda-Valderrama, J.; Montes-Betancourt, M.; Falla-Ortiz, J. ICT in the Development of Phonological Awareness: A Critical Literature Review. *Digit. Educ. Rev.* **2023**, *44*, 125–132. [CrossRef]
127. Parry, C.-R.L.; Kumar, S.; Gallingane, C. Integration of Tablets in Phonological Awareness and Phonics Instruction: A Systematic Review. *J. Res. Child. Educ.* **2024**, *38*, 559–576. [CrossRef]
128. Beck, I.L.; McKeown, M.G. Increasing Young Low-Income Children's Oral Vocabulary Repertoires through Rich and Focused Instruction. *Elem. Sch. J.* **2007**, *107*, 251–271. [CrossRef]
129. Biemiller, A.; Boote, C. An Effective Method for Building Meaning Vocabulary in Primary Grades. *J. Educ. Psychol.* **2006**, *98*, 44–62. [CrossRef]
130. Feustel, T.C.; Shiffrin, R.M.; Salasoo, A. Episodic and Lexical Contributions to the Repetition Effect in Word Identification. *J. Exp. Psychol. Gen.* **1983**, *112*, 309–346. [CrossRef]
131. Marulis, L.M.; Neuman, S.B. The Effects of Vocabulary Intervention on Young Children's Word Learning: A Meta-Analysis. *Rev. Educ. Res.* **2010**, *80*, 300–335. [CrossRef]
132. Roskos, K.; Burstein, K. Assessment of the Design Efficacy of a Preschool Vocabulary Instruction Technique. *J. Res. Child. Educ.* **2011**, *25*, 268–287. [CrossRef]
133. Samudra, P.G.; Wong, K.M.; Neuman, S.B. Promoting Low-Income Preschoolers' Vocabulary Learning from Educational Media: Does Repetition Support Memory for Learned Word Knowledge? *J. Cogn. Educ. Psych.* **2019**, *18*, 160–173. [CrossRef]
134. Kim, M.K.; Bryant, D.P.; Bryant, B.R.; Park, Y. A Synthesis of Interventions for Improving Oral Reading Fluency of Elementary Students with Learning Disabilities. *Prev. Sch. Fail. Altern. Educ. Child. Youth* **2017**, *61*, 116–125. [CrossRef]
135. Kuhn, M.R.; Stahl, S.A. Fluency: A Review of Developmental and Remedial Practices. *J. Educ. Psychol.* **2003**, *95*, 3–21. [CrossRef]
136. Lee, J.; Yoon, S.Y. The Effects of Repeated Reading on Reading Fluency for Students with Reading Disabilities: A Meta-Analysis. *J. Learn. Disabil.* **2017**, *50*, 213–224. [CrossRef]
137. Massler, U.; Müller, W.; Iurgel, I.; Haake, S.; Gantikow, A.; Hadzilacos, T. Meaningful, Gamified Training of Reading Fluency. *Front. Comput. Sci.* **2022**, *4*, 968137. [CrossRef]
138. Padeliadu, S.; Giazitzidou, S. A Synthesis of Research on Reading Fluency Develompent: Study of Eight Meta-Analyses. *Eur. J. Spec. Educ. Res.* **2018**, *3*, 4. [CrossRef]
139. Stevens, E.A.; Walker, M.A.; Vaughn, S. The Effects of Reading Fluency Interventions on the Reading Fluency and Reading Comprehension Performance of Elementary Students with Learning Disabilities: A Synthesis of the Research from 2001 to 2014. *J. Learn. Disabil.* **2017**, *50*, 576–590. [CrossRef] [PubMed]
140. Soriano, M.; Miranda, A.; Soriano, E.; Nievas, F.; Félix, V. Examining the Efficacy of an Intervention to Improve Fluency and Reading Comprehension in Spanish Children with Reading Disabilities. *Int. J. Disabil. Dev. Educ.* **2011**, *58*, 47–59. [CrossRef]
141. Therrien, W.J. Fluency and Comprehension Gains as a Result of Repeated Reading: A Meta-Analysis. *Remedial Spec. Educ.* **2004**, *25*, 252–261. [CrossRef]
142. Wolf, M.; Katzir-Cohen, T. Reading Fluency and Its Intervention. *Sci. Stud. Read.* **2001**, *5*, 211–239. [CrossRef]
143. Nousia, A. The Integration of New Technologies and Video Games in Preschool Education. *Eur. J. Open Educ. E-Learn. Stud.* **2023**, *8*, 1. [CrossRef]
144. Piquette, N.A.; Savage, R.S.; Abrami, P.C. A Cluster Randomized Control Field Trial of the ABRACADABRA Web-Based Reading Technology: Replication and Extension of Basic Findings. *Front. Psychol.* **2014**, *5*, 1413. [CrossRef]
145. Rakimahwati, R.; Hanifa, N.; Aryani, N. Android Based Educational Game Development to Improve Early Childhood Reading Ability. *Al-Ishlah J. Pendidik.* **2022**, *14*, 1123–1134. [CrossRef]
146. Segers, E.; Verhoeven, L. Effects of Vocabulary Training by Computer in Kindergarten. *Comput. Assist. Learn.* **2003**, *19*, 557–566. [CrossRef]
147. Burchinal, M.; Krowka, S.; Newman-Gonchar, R.; Jayanthi, M.; Gersten, R.; Wavell, S.; Lyskawa, J.; Haymond, K.; Bierman, K.; Gonzalez, J.E.; et al. *Preparing Young Children for School*; WWC 2022009; National Center for Education Evaluation and Regional Assistance (NCEE), Institute of Education Sciences, U.S. Department of Education: Washington, DC, USA, 2022. Available online: https://ies.ed.gov/ncee/wwc/Docs/PracticeGuide/TO4_PRACTICE_GUIDE_Preparing-for-School_07222022_v6.pdf (accessed on 15 June 2024).

148. Sargiani, R.D.A.; Ehri, L.C.; Maluf, M.R. Teaching Beginners to Decode Consonant–Vowel Syllables Using Grapheme–Phoneme Subunits Facilitates Reading and Spelling as Compared with Teaching Whole-Syllable Decoding. *Read. Res. Q.* **2022**, *57*, 629–648. [CrossRef]
149. Vazeux, M.; Doignon-Camus, N.; Bosse, M.-L.; Mahé, G.; Guo, T.; Zagar, D. Syllable-First Rather than Letter-First to Improve Phonemic Awareness. *Sci. Rep.* **2020**, *10*, 22130. [CrossRef]
150. Piasta, S.B.; Hudson, A.K. Key Knowledge to Support Phonological Awareness and Phonics Instruction. *Read. Teach.* **2022**, *76*, 201–210. [CrossRef]
151. Gholami, M.; Shirazi, S.; Nikookar, N.; Rahimi, A. Comparison of Effectiveness of Phonetic Intervention, Holistic Intervention and Mixed Intervention on the Rate and Accuracy of Reading Skills of Children with Reading Difficulty in Grade 2. *IIOAB J.* **2016**, *7*, 270–276.
152. Gutiérrez, R. Habilidades favorecedoras del aprendizaje de la lectura en alumnos de 5 y 6 años. *Rev. Signos* **2018**, *51*, 45–60. [CrossRef]
153. Liekum, L.B. Probing the Impact of the Eclectic Approach in the Reading Comprehension Competence of Basic Schools Pupils in Ghana: With Kuka Primary School in Focus. *Am. J. Multidis. Res. Innov.* **2022**, *1*, 41–46. [CrossRef]
154. Van De Walle De Ghelcke, A.; Rossion, B.; Schiltz, C.; Lochy, A. Impact of Learning to Read in a Mixed Approach on Neural Tuning to Words in Beginning Readers. *Front. Psychol.* **2020**, *10*, 3043. [CrossRef]
155. Hughes, C.A.; Morris, J.R.; Therrien, W.J.; Benson, S.K. Explicit Instruction: Historical and Contemporary Contexts. *Learn. Disabil. Res. Pract.* **2017**, *32*, 140–148. [CrossRef]
156. Kamil, M.L.; Borman, G.D.; Dole, J.; Kral, C.C.; Salinger, T.; Torgesen, J. *Improving Adolescent Literacy: Effective Classroom and Intervention Practices: A Practice Guide*; NCEE #2008-4027; National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education: Washington, DC, USA, 2008. Available online: https://ies.ed.gov/ncee/wwc/Docs/practiceguide/adlit_pg_082608.pdf (accessed on 21 July 2024).
157. Kirschner, P.A.; Sweller, J.; Clark, R.E. Why Minimal Guidance During Instruction Does Not Work: An Analysis of the Failure of Constructivist, Discovery, Problem-Based, Experiential, and Inquiry-Based Teaching. *Educ. Psychol.* **2006**, *41*, 75–86. [CrossRef]
158. Kirschner, P.A.; Hendrick, C.; Caviglioli, O. *How Learning Happens: Seminal Works in Educational Psychology and What They Mean in Practice*; A David Fulton Book; Routledge, Taylor & Francis Group: London, UK; New York, NY, USA, 2020. [CrossRef]
159. Rosenshine, B.; Stevens, R. Teaching Functions. In *Handbook of Research on Teaching*; Macmillan: New York, NY, USA, 1986; pp. 376–391.
160. Stockard, J.; Wood, T.W.; Coughlin, C.; Rasplica Khoury, C. The Effectiveness of Direct Instruction Curricula: A Meta-Analysis of a Half Century of Research. *Rev. Educ. Res.* **2018**, *88*, 479–507. [CrossRef]
161. Vaughn, S.; Fletcher, J. Explicit Instruction as the Essential Tool for Executing the Science of Reading. *Read. Leag. J.* **2021**, *2*, 4–11.
162. Brame, C.J.; Biel, R. Test-Enhanced Learning: The Potential for Testing to Promote Greater Learning in Undergraduate Science Courses. *CBE—Life Sci. Educ.* **2015**, *14*, es4. [CrossRef]
163. Bjork, E.L.; Bjork, R.A. Making Things Hard on Yourself, but in a Good Way: Creating Desirable Difficulties to Enhance Learning. In *Psychology and the Real World: Essays Illustrating Fundamental Contributions to Society*; Worth Publishers: New York, NY, USA, 2011; pp. 56–64.
164. Karpicke, J.D. Retrieval-Based Learning: A Decade of Progress. In *Learning and Memory: A Comprehensive Reference*; Elsevier: Amsterdam, The Netherlands, 2017; pp. 487–514. [CrossRef]
165. Sweller, J.; Ayres, P.L.; Kalyuga, S. Explorations in the learning sciences, instructional systems and performance technologies. In *Cognitive Load Theory*; Springer: New York, NY, USA, 2011.
166. Marinelli, C.V.; Nardacchione, G.; Trotta, E.; Di Fuccio, R.; Palladino, P.; Traetta, L.; Limone, P. The Effectiveness of Serious Games for Enhancing Literacy Skills in Children with Learning Disabilities or Difficulties: A Systematic Review. *Appl. Sci.* **2023**, *13*, 4512. [CrossRef]
167. De Gloria, A.; Bellotti, F.; Berta, R. Serious Games for Education and Training. *Int. J. Serious Games* **2014**, *1*, 1. [CrossRef]
168. Mayer, R.E. *Multimedia Learning*, 2nd ed.; Cambridge University Press: Cambridge, UK, 2009. [CrossRef]
169. Dunlosky, J.; Rawson, K.A.; Marsh, E.J.; Nathan, M.J.; Willingham, D.T. Improving Students' Learning with Effective Learning Techniques: Promising Directions From Cognitive and Educational Psychology. *Psychol. Sci. Public Interest* **2013**, *14*, 4–58. [CrossRef]
170. Carpenter, S.K.; Agarwal, P.K. How to Use Spaced Retrieval Practice to Boost Learning. 2020. Available online: https://pdf.retrievalpractice.org/SpacingGuide.pdf (accessed on 19 November 2023).
171. Wimmer, G.E.; Poldrack, R.A. Reward Learning and Working Memory: Effects of Massed versus Spaced Training and Post-Learning Delay Period. *Mem. Cogn.* **2022**, *50*, 312–324. [CrossRef]

172. Blumberg, F.C.; Fisch, S.M. Introduction: Digital Games as a Context for Cognitive Development, Learning, and Developmental Research. *New Dir. Child Adolesc. Dev.* **2013**, *2013*, 1–9. [CrossRef]
173. Boyan, A.; Sherry, J.L. The Challenge in Creating Games for Education: Aligning Mental Models With Game Models: Challenge and Educational Games. *Child Dev. Perspect.* **2011**, *5*, 82–87. [CrossRef]
174. Lipina, S.; Sigman, M.; Fernández Slezak, D.; Battro, A. *Pensar Las Tic Desde La Ciencia Cognitiva Y La Neurociencia*; Gedisa: Barcelona, Spain, 2017.
175. Martí-Parreño, J.; Galbis-Córdova, A.; Miquel-Romero, M.J. Students' Attitude towards the Use of Educational Video Games to Develop Competencies. *Comput. Hum. Behav.* **2018**, *81*, 366–377. [CrossRef]
176. Howard-Jones, P. Videojuegos, Neurociencias y Educación. In *Pensar Las Tic Desde La Ciencia Cognitiva Y La Neurociencia*; Gedisa: Barcelona, Spain, 2017; pp. 69–92.
177. Sierra-Martínez, L.M.; Chanchí-Golondrino, G.E.; Gómez-Álvarez, M.C. Directrices para el diseño y la construcción de videojuegos serios educativos. *Rev. Colomb. Educ.* **2022**, *84*, 1–22. [CrossRef]
178. O'Brien, D.; Lawless, K.A.; Schrader, P.G. A Taxonomy of Educational Games. In *Advances in Game-Based Learning*; IGI Global: Hershey, PA, USA, 2010; pp. 1–23. [CrossRef]
179. Kickmeier-Rust, M. Talking Digital Educational Games. In Proceedings of the 1st International Open Workshop on Intelligent Personalization and Adaptation in Digital Educational Games, Graz, Austria, 14 October 2009. Available online: http://www.eightydays.eu/files/80Days_Workshop_Proceedings.pdf (accessed on 11 January 2024).
180. Teixeira, J.S.F.; de Jesus, V.; Sá, E.; Torres Fernandes, C. Taxonomy of Educational Games Compatible with the LOM-IEEE Data Model. In Proceedings of the Workshop Brasileiro em Web Semântica e Educação, São Paulo, Brasil, 22 October 2007; pp. 37–47.
181. Londoño, L.M.; Rojas, M.D. Determinación de criterios generales para el diseño de juegos serios: Modelo metodológico integrador. *Inf. Tecnol.* **2021**, *32*, 123–132. [CrossRef]
182. Cabero-Almenara, J.; Barragán-Sánchez, R.; Palacios-Rodríguez, A.; Martín-Párraga, L. Design and Validation of T-MOOC for the Development of the Digital Competence of Non-University Teachers. *Technologies* **2021**, *9*, 84. [CrossRef]
183. Cabero Almenara, J.; Barroso Osuna, J. La utilización del juicio de experto para la evaluación de TIC: El coeficiente de competencia experta. *Bordón Rev. Pedagog.* **2013**, *65*, 25–38. [CrossRef]
184. Lam, S.K.; Ahearne, M.; Mullins, R.; Hayati, B.; Schillewaert, N. Exploring the Dynamics of Antecedents to Consumer–Brand Identification with a New Brand. *J. Acad. Mark. Sci.* **2013**, *41*, 234–252. [CrossRef]
185. González-Trujillo, M.C.; Calet, N.; Defior, S.; Gutiérrez-Palma, N. Scale of Reading Fluency in Spanish: Measuring the Components of Fluency/Escala de Fluidez Lectora En Español: Midiendo Los Componentes de La Fluidez. *Estud. Psicol.* **2014**, *35*, 104–136. [CrossRef]
186. Laevers, F. Assessing the Quality of Childcare Provision: "Involvement" as Criterion. *Res. Early Child.* **1997**, *3*, 151–165.
187. Read, J.C. Validating the Fun Toolkit: An Instrument for Measuring Children's Opinions of Technology. *Cogn. Technol. Work* **2008**, *10*, 119–128. [CrossRef]
188. Creswell, J.W.; Plano Clark, V.L. *Designing and Conducting Mixed Methods Research*, 3rd ed.; SAGE: Thousand Oaks, CA, USA, 2018.
189. Reeve, J.; Jang, H.; Carrell, D.; Jeon, S.; Barch, J. Enhancing Students' Engagement by Increasing Teachers' Autonomy Support. *Motiv. Emot.* **2004**, *28*, 147–169. [CrossRef]
190. Mombaers, T.; Donche, V. Hospital School Students' Academic Motivation and Support Needs: A Self-Determination Perspective. *Front. Educ.* **2020**, *5*, 106. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

Multimodal Technologies and Interaction

MDPI

Article

Evaluating a UDL-Based Digital Escape Room for Bullying Awareness in Secondary Education

Yosra Achab-Moukram [1], Paula Rodríguez-Rivera [2], José Manuel Aguilar-Parra [3,*], José M. Rodríguez-Ferrer [4], Javier Rodríguez-Moreno [1] and Ana Manzano-León [3,*]

[1] Department of Education, University of Jaén, 23071 Jaén, Spain; yosra.achab@gmail.com (Y.A.-M.); jrmoreno@ujaen.es (J.R.-M.)
[2] Faculty of Education and Social Work, Department of Psychosocioeducational Analysis and Intervention, University of Vigo, 32004 Ourense, Spain; paula.rodriguez.rivera@uvigo.es
[3] Health Research Centre, Department of Psychology, University of Almería, 04120 Almería, Spain
[4] Faculty of Health Sciences, Department of Health Sciences, University of Burgos, 09001 Burgos, Spain; jmrferrer@ubu.es
* Correspondence: jmaguilar@ual.es (J.M.A.-P.); aml570@ual.es (A.M.-L.)

Abstract: This study evaluates the effectiveness of a digital escape room designed according to the Universal Design for Learning (UDL) principles as an educational tool to raise awareness about school bullying and cyberbullying in lower secondary education. Using a quasi-experimental design with 137 students (66 in the control group and 71 in the experimental group), the gamified activity was compared to a traditional expository intervention. Empathy was measured using the Brief Basic Empathy Scale (BES-B), and bullying-related knowledge and experiences were assessed with the SPB-A scale. The escape room integrated key UDL principles, including multiple means of representation (e.g., video, images, text), action and expression (interactive tasks and puzzles), and engagement (narrative-based missions and collaborative challenges). The results indicate significant improvements in knowledge about behaviors related to school bullying among participants who completed the escape room. Although changes in affective and cognitive empathy were not statistically significant, positive trends were observed in the experimental group. The immersive, collaborative, and accessible nature of the activity positions the escape room as a promising tool within broader strategies for bullying prevention.

Keywords: digital escape room; school bullying; empathy; gamification; Universal Design for Learning (UDL); secondary education

1. Introduction

1.1. Bullying and Cyberbullying in School Contexts

The prevalence of school bullying has become a growing concern in recent years, with significant impacts on the academic and social well-being of students [1,2]. This phenomenon disproportionately affects students from marginalized groups, including racial, ethnic, and sexual minorities, who are often subjected to bullying rooted in systemic inequalities and sociocultural narratives of otherness [3,4]. The construction of "the other" (identities deemed non-normative or deviant by dominant social hierarchies) fosters environments where discrimination is normalized, particularly in schools lacking inclusive frameworks [5]. For instance, students with behavioral disabilities experience heightened bullying due to intersecting prejudices against disability. At the same time, LGBTQ+ youth

report bullying rates 2–3 times higher than their heterosexual peers, often linked to gender expression or sexual orientation [6,7]. Such dynamics amplify in digital spaces, where cyberbullying leverages anonymity and the permanence of online content to perpetuate harassment, further alienating marginalized individuals from the safe participation in virtual communities [8]. This lack of digital safety poses severe risks, as marginalized youth may withdraw from online engagement, affecting their academic, social, and personal development [9].

Empathy emerges as an essential socioemotional competence for understanding and addressing the complexity of bullying and cyberbullying in school settings. Numerous studies indicate that the ability to recognize and share the emotions of others decreases the likelihood of developing aggressive behaviors and strengthens social cohesion [10,11]. Likewise, higher levels of empathy have been identified as a protective factor, mitigating the incidence of victimization and reducing the frequency of perpetration [12,13].

Beyond its direct impact on aggression, empathy also plays a key role in shaping bystander responses to bullying incidents. When students are encouraged to empathize with victims, they become more likely to intervene or seek help, disrupting bullying trajectories and contributing to a safer school climate [14,15]. In digital contexts, this "digital empathy" can manifest as an increased awareness of the emotional weight of online messages and interactions, which helps prevent normalizing harmful behaviors [16,17]. In the context of school interventions, fostering empathy through shared-perspective activities, cooperative learning, and emotional education has proven effective in sensitizing students to diversity and injustice and promoting prosocial behaviors that bolster the educational community's cohesion [18]. Such approaches encourage students to actively consider others' feelings and experiences, laying the groundwork for peer support networks and inclusive school cultures. Moreover, systematically implementing empathy-based training can enhance students' self-regulatory skills, reducing tendencies toward bullying and facilitating long-term behavioral change [19,20].

1.2. The Role of Escape Rooms in Pedagogy: Innovative Pedagogical Tools and Processes

Digital technologies play a paradoxical role in this context. On the one hand, they can exacerbate harm. Around 15–20% of U.S. adolescents experience cyberbullying, with higher rates among girls, LGBTQ+ youth, and students with disabilities [21]. Features like anonymity, rapid dissemination, and the absence of immediate adult oversight create fertile ground for continuous aggression. Perpetrators target victims through hate speech, image-based abuse, doxing, or exclusion from social networks, often causing long-lasting psychological trauma, including anxiety, depression, and suicidal ideation [22].

The integration of escape rooms into pedagogical strategies against bullying and cyberbullying represents a significant step forward toward evidence-based approaches that link experiential learning with the prevention of discriminatory behaviors. Recent studies have highlighted that these gamified environments, by combining structured narratives and collaborative problem-solving, strengthen critical socioemotional skills such as empathy and self-regulation, which are essential for dismantling bullying dynamics [23]. Furthermore, adapting escape rooms to digital contexts enables the replication of online risk scenarios, providing students with opportunities to practice ethical and technical responses to cyberbullying, ranging from managing privacy to reporting abusive content [24].

Although the literature supports the effectiveness of these tools in reducing both victimization and perpetration behaviors [25], challenges remain in their systemic implementation. The lack of teacher training in gamified design and the technological gap in under-resourced educational settings limit the scalability of these approaches [26]. How-

ever, collaborative initiatives promoting modular templates and open-source resources mitigate these barriers, facilitating the adoption of escape rooms as accessible and adaptable educational tools [27,28].

UDL shifts the paradigm from reactive to proactive solutions. Rather than retrofitting classrooms to address bullying, the framework embeds flexibility into curriculum design by anticipating learner variability [29,30]. Its three principles—multiple means of engagement, representation, and expression—enable educators to create environments that value diverse identities [31]. By incorporating digital tools such as interactive narratives or gamified scenarios, teachers can address sensitive topics like discrimination through modalities that resonate with varied learning styles [32]. For instance, a student struggling with text-based content might engage more deeply with a video narrative on inclusion, while another might process ethical dilemmas through a collaborative digital simulation [33].

A recent systematic review demonstrates how escape-room mechanics align with the three core principles of the Universal Design for Learning (UDL), mapping typical features such as branching narratives and adaptive puzzles onto the principle of engagement, while multimodal clues and embedded audiovisual prompts address the principle of representation [34]. Likewise, varied challenge formats, such as code-breaking activities and role-playing, support the principle of action and expression [35]. This structural alignment highlights the intrinsic inclusivity of escape rooms: they offer multiple, flexible pathways for students to engage with content, to access information through diverse sensory modalities, and to demonstrate their understanding using different formats [36]. As such, escape rooms represent a pedagogical strategy that inherently operationalizes UDL guidelines, making them particularly suitable for educational interventions targeting diverse student populations.

Extending UDL's emphasis on variability and flexibility, recent research has explored other innovative digital tools that foster inclusion and socio-emotional learning. For instance, a quasi-experimental study found that a gamified interactive e-book significantly enhanced motivation and anti-bullying knowledge among Chinese primary school students compared to traditional lecture-based approaches [37]. In parallel, intelligent virtual assistants developed for educational contexts have shown promise in detecting bullying patterns through conversational analysis and providing students with confidential support and reporting options, thereby increasing their self-efficacy in help-seeking and bystander intervention [38]. Similarly, immersive virtual reality (VR) simulations have been used to create controlled, scenario-based environments where students interact through avatars in harassment situations, resulting in measurable gains in empathy and decision-making skills in pre-post assessments [39].

While such technologies expand the toolkit of anti-bullying interventions, playful strategies like escape rooms offer a unique combination of experiential, narrative-driven, and collaborative elements within a single, cohesive framework. This integration makes them particularly effective for reinforcing socio-emotional competencies in physical and digital learning environments. Emerging evidence suggests that these gamified formats, combining structured storytelling with collective problem-solving, can foster critical skills such as empathy, emotional regulation, and cooperative decision-making, all essential for disrupting bullying dynamics [40]. Furthermore, adapting escape rooms to digital contexts enables the replication of online risk scenarios, providing students with opportunities to practice ethical and technical responses to cyberbullying, ranging from managing privacy to reporting abusive content [41].

Although the literature supports the effectiveness of these tools in reducing both victimization and perpetration behaviors, challenges remain in their systemic implementation.

The lack of teacher training in gamified design and the technological gap in under-resourced educational settings limit the scalability of these approaches [42]. However, collaborative initiatives promoting modular templates and open-source resources mitigate these barriers, facilitating the adoption of escape rooms as accessible and adaptable educational tools. This study contributes to the growing body of research on gamified educational interventions by providing empirical evidence on the effectiveness of a digital escape room designed to address bullying and cyberbullying in secondary education. For this reason, the present study was guided by the following research questions:

(Q1) Which are the differences in affective and cognitive empathy levels after the intervention between adolescents who participate in a digital escape room and those who receive an expository intervention?

(Q2) Which are the differences in the perception of behaviors related to bullying (identification, witnessing, and experiencing) between students who participate in a digital escape room and those who receive a traditional intervention?

2. Materials and Methods

2.1. Participants

The study involved 137 students enrolled in the first cycle of Compulsory Secondary Education (CSE) in Spain, with an average age of 13.32 years (*SD* = 0.88). Participants were recruited from six classrooms across two public secondary schools in Andalusia and selected based on accessibility and willingness to participate. High schools were contacted through existing collaborative networks with the research team, and participation was voluntary for both institutions and students.

Eligibility criteria for inclusion in the study were: (1) being enrolled in a participating school and present during the scheduled intervention session; (2) being in the 12–14 age range, which corresponds to early adolescence and the first two years of CSE in the Spanish education system; and (3) having obtained signed informed consent from a parent or legal guardian. No exclusion criteria related to sex, academic performance, or prior experiences with bullying were applied to ensure representativeness and ecological validity.

To assign participants to conditions, entire classes were randomly assigned to either the experimental group (digital escape room; n = 71) or the control group (traditional anti-bullying session; n = 66), maintaining the natural classroom structure to respect school dynamics. Table 1 outlines the sample distribution by group, gender, and age.

Table 1. Study participants.

	Sex		**Age**		***N***
	Female	**Male**	***M***	***DT***	
Control	32	34	13.41	0.86	66
Experimental	39	32	13.24	0.89	71
Total	71	66	13.32	0.88	137

Before data collection, students were informed about the study's aims and procedures and assured that their participation would remain anonymous. Informed consent was also obtained from their legal guardians. The digital escape room activity (experimental group) and the traditional anti-bullying session (control group) were integrated into the students' regular tutoring schedules to ensure ecological validity.

This study followed the ethical principles outlined in the Declaration of Helsinki. Ethical approval was obtained from the University of Almería Ethics Committee (UALBIO2023/019).

2.2. Instruments

2.2.1. Empathy

Empathy was measured using the Brief Basic Empathy Scale (BES-B), a psychometric instrument composed of 9 items designed to assess empathy in children and adolescents by distinguishing between two main dimensions: affective empathy and cognitive empathy. In the study by Merino-Soto, López-Fernández, and Grimaldo-Muchotrigo [43], the scale demonstrated an adequate bifactorial structure and acceptable levels of reliability (α and $\omega > 0.70$), as well as metric and scalar invariance across sex and educational level. These findings support the conclusion that the scale measures empathy equivalently among Spanish-speaking children and adolescents, regardless of gender.

2.2.2. Bullying-Related Knowledge and Experience

The Self-Perception and Perception of Bullying in Adolescents Scale (SPB-A) [44] was employed to evaluate students' understanding of bullying and their related experiences. This instrument comprehensively assesses the experience of school bullying among students in Compulsory Secondary Education. It consists of 27 items distributed across four factors: knowledge of behaviors that constitute bullying, witnessing bullying behaviors, experiencing bullying behaviors, and the impact of bullying behaviors. The scale was validated through exploratory and confirmatory factor analyses, demonstrating a robust factorial structure and high internal consistency ($\alpha = 0.86$). In addition, multigroup factorial invariance analyses confirmed that the instrument is invariant across sex, grade level, and school type, supporting its reliability and validity for use in diverse educational settings.

2.3. Escape Room Design

The digital escape room was developed through a multi-phase design process grounded in pedagogical and technical criteria aligned with Universal Design for Learning (UDL) principles. The objective was to create an engaging, accessible, and educational experience that increases students' awareness of school bullying and promotes socio-emotional competencies.

2.3.1. Platform Selection and Design Phases

The activity was created using Genially, a web-based platform selected for its interactive capabilities, ease of use, and compatibility with multimedia resources. Genially allowed the integration of dynamic visual elements, audio, videos, and interactive components without requiring advanced programming skills, thus ensuring replicability in diverse educational settings. Its responsive design enabled implementation across devices (computers, tablets, or interactive whiteboards).

The development process followed three main stages:

- Narrative conceptualization: A fictional storyline was constructed around a school experiencing a bullying crisis, where superhero characters enlist students to help identify and resolve harmful behaviors. This narrative structure was chosen to foster emotional engagement and intrinsic motivation through role-play and mission-based tasks.
- Content and challenge design: Educational content was structured into three thematic environments: the classroom, the library, and the cafeteria. Each contains multiple interactive tasks. These tasks included decoding puzzles, completing video-based reflections, solving word searches, and participating in a comic-based scenario. Activities aligned with bullying-related learning objectives.
- Visual and multimedia development: All visual content was created through the AI-based generative tools DALL·E and ChatGPT 4o with an image generator GPT

or sourced from the royalty-free image platform Pixabay. The resulting materials were curated to ensure cultural sensitivity, visual clarity, and thematic coherence with the narrative.

The research team pilot-tested the escape room internally to ensure its functionality, clarity of instructions, and age-appropriateness.

2.3.2. Escape Room Structure

The escape room was structured around a narrative [45] set in a fictional school environment. In the storyline, the school principal expresses concern about a rise in bullying cases and activates the "red alert," summoning several superheroes—fictional characters familiar to young audiences [46] to help address the problem. These superheroes request the students' collaboration to identify inappropriate behaviors and propose ways to improve coexistence within the school.

Once the mission is accepted, students—working in small groups of 3 to 4 players—must overcome a series of challenges distributed across three virtual school settings: the classroom, the library, and the cafeteria.

In the first phase, located in the classroom, the challenges include:

- A spiral with a hidden message: Participants must interact with various objects in the virtual environment to find a clue that helps decode a message related to bullying.
- A video activity: Students watch a YouTube video about a bullied child. After viewing the video, they must answer a comprehension question related to it and solve an interactive puzzle.

In the second phase, taking place in the library, students face two activities:

- A word search: Students must identify six concepts related to bullying. Based on the terms they find, a reflective question is then posed.
- An interactive comic: A bullying scenario is illustrated ("Extraña como yo" by Gavin Aun Thang) to foster empathy and recognize inappropriate behaviors.

The third phase occurs in the cafeteria, where students encounter reflective messages and practical advice on preventing bullying. They must also find a hidden password to complete the activity there.

Upon entering the password correctly, students are directed to a final screen. This screen displays the date of the International Day Against Bullying and congratulates the students as a symbolic recognition of their engagement with the proposed mission (See Figure 1).

2.3.3. Application of Universal Design for Learning (UDL) Principles in the Escape Room Design

The escape room was designed in accordance with the principles of the Universal Design for Learning (UDL) [47] to provide an educational experience that is accessible, inclusive, and meaningful for all students. Specifically, the following aspects were considered:

- Providing multiple means of representation: Content is presented through diverse formats (text, video, images, audio, comics, interactive activities), allowing students to access information according to their preferences and learning styles.
- The escape room provides multiple means of action and expression: It includes a range of challenges, from puzzle-solving to reflective tasks, allowing students to demonstrate their understanding through various cognitive and creative approaches.
- Providing multiple means of engagement: The gamified narrative, use of fictional characters, and progression through phases foster intrinsic motivation and active par-

ticipation. Furthermore, the reflective messages and connection to a socially relevant issue (bullying) encourage emotional engagement with the activity.



Figure 1. Sample screenshots from the escape room. Note: English translation: Congratulations!! Now you have superpowers against bullying!

In addition, the design prioritizes intuitive and autonomous navigation, enabling implementation in various contexts (individual, group, classroom, or remote) and accommodating different learning paces. This approach fosters an inclusive learning experience tailored to the diverse needs and abilities of all students.

2.4. Procedure

This study followed a quasi-experimental pre- and post-test design with a control and an experimental group. The study was conducted during regular school hours in public secondary education centers in [removed for revision], Spain. After obtaining the necessary approvals from the schools, participants were assigned to either the experimental or control group based on their existing class groups, as random assignment was not feasible in the natural school setting.

Before the intervention, all participants completed the battery of instruments. These pre-test assessments were administered collectively in the classroom under the supervision of one of the researchers and the classroom teacher.

Following the pre-test, the interventions were implemented:

- Experimental group: Students participated in a digital escape room developed with the Genially platform, chosen for its interactive and engaging features, designed according to the Universal Design for Learning (UDL) principles to ensure accessibility and engagement. The activity lasted approximately one hour and was completed in small groups of 3 to 4 students. The escape room included multimedia challenges, such as videos, comics, interactive puzzles, and reflection questions, distributed across three virtual environments: the classroom, library, and cafeteria. All were related to the theme of bullying and school coexistence.
- The control group received a one-hour standard session with their classroom tutor, which was carefully designed to align with the duration and topic of the escape room. This ensured that both groups were exposed to comparable levels of the subject matter, enhancing the study's balance and impartiality.

To enhance comparability, the control session covered the same core themes: recognizing bullying behaviors, understanding the consequences of bullying, and strategies for intervention. It was delivered in a lecture-based format using a slideshow presentation accompanied by group discussion facilitated by the classroom tutor. No interactive or gamified elements were included. While both conditions addressed the same learning objectives, the instructional strategies differed notably in terms of engagement and modality: the experimental group participated in a collaborative, multimodal, narrative-based activity, whereas the control group engaged in a conventional didactic session with limited student interaction.

One week after the intervention, both groups completed the same set of instruments used in the pre-test and post-test under the same conditions. Data collection was anonymous and coded to allow for pre–post comparisons. This study complied with the recommendations of the Declaration of Helsinki. Ethical approval was obtained from the University of Almería Ethics Committee (UALBIO2023/019).

2.5. Data Analysis

The data for this research study were processed using SPSS Version 26. To ensure coherence with the two research questions, the analysis was structured accordingly:

For RQ1 (Which are the differences in affective and cognitive empathy levels after the intervention between adolescents who participate in a digital escape room and those who receive an expository intervention?), the following steps were followed:

- Initial equivalence between groups was assessed using independent samples *t*-tests on pre-test scores for affective and cognitive empathy.
- Subsequently, individual pre–post differences (Δ) were calculated for each participant by subtracting post-test scores from pre-test scores.
- These difference scores were compared between groups using independent samples *t*-tests to evaluate.
- Effect sizes were computed using Cohen's d to estimate the magnitude of the differences beyond statistical significance.

For RQ 2 (Which are the differences in the perception of behaviors related to bullying—identification, witnessing, and experiencing—between students who participate in a digital escape room and those who receive a traditional intervention?), a similar analytic strategy was used:

- Pre-test equivalence between groups was verified using independent samples *t*-tests for each bullying-related indicator.

- Post–pre difference scores (Δ) were calculated for each participant on the three indicators: knowledge about bullying behaviors, witnessing bullying, and experiencing bullying.
- Group differences in these change scores were assessed using *t*-tests, and Cohen's d was again used to interpret effect sizes (small ≈ 0.20, moderate ≈ 0.50, large ≥ 0.80).

Results were reported with means, standard deviations, t-values, *p*-values, and effect sizes for all comparisons.

3. Results

Regarding the first research question (Which are the differences in affective and cognitive empathy levels after the intervention between adolescents who participate in a digital escape room and those who receive an expository intervention?), an analysis was conducted to compare changes in affective and cognitive empathy between the experimental group (digital escape room) and the control group (traditional expository session).

First, the groups' initial equivalence was assessed using an independent samples *t*-test applied to the pre-test scores. No statistically significant differences were found in empathy dimensions, indicating that both groups started from similar levels before the intervention (See Table 2).

Table 2. Means (M) and standard deviations (SD) of affective and cognitive empathy by group and moment.

Variable	Moment	Experimental Group (M ± DT)	Control Group (M ± DT)	*p* (Pre-Test)
Affective empathy	Pre	12.35 ± 3.86	11.35 ± 4.32	$p = 0.153$
	Post	16.43 ± 2.15	12.89 ± 3.37	$p < 0.001$
Cognitive empathy	Pre	18.52 ± 3.84	17.76 ± 4.66	$p = 0.295$
	Post	21.27 ± 2.40	19.18 ± 3.58	$p < 0.001$

Subsequently, the pre–post differences (Δ) were calculated, and *t*-tests were applied to compare independent samples. Although the differences did not reach statistical significance, the experimental group showed more significant improvements than the control group in both dimensions of empathy. The effect size was small, especially in the cognitive dimension (See Table 3).

Table 3. Comparison of pre–post differences in affective and cognitive empathy (Δ).

Variable	Δ Experimental Group (M ± DT)	Δ Control Group (M ± DT)	t	*p*	d Cohen
Affective empathy	4.08 ± 3.03	1.54 ± 4.97	3.63	$p < 0.001$	0.62
Cognitive empathy	2.75 ± 2.92	1.44 ± 5.91	1.65	0.100	0.28

For the second research question (Which are the differences in the perception of behaviors related to bullying (identification, witnessing, and experiencing) between students who participate in a digital escape room and those who receive a traditional intervention?), three dimensions were analyzed: knowledge about behaviors that involve bullying, bullying behaviors witnessed, and bullying behaviors suffered.

In this case, the equivalence between groups was rigorously checked in the pre-test, and no significant differences were found, providing reassurance about the study's methodology (See Table 4).

Table 4. Means (M) and standard deviations (SD) of the perception of behaviors related to bullying by group and moment.

Variable	Moment	Experimental (M ± DT)	Control (M ± DT)	*p* (Pre-Test)
Knowledge about bullying behaviors	Pre	44.84 ± 13.50	43.45 ± 14.53	$p = 0.574$
	Post	50.19 ± 15.59	39.92 ± 16.46	$p < 0.001$
Bullying behaviors witnessed	Pre	36.03 ± 18.28	35.15 ± 15.39	$p = 0.767$
	Post	37.14 ± 15.50	32.55 ± 16.25	$p = 0.102$
Bullying behaviors suffered	Pre	19.31 ± 11.95	19.97 ± 14.82	$p = 0.782$
	Post	18.16 ± 10.86	19.28 ± 13.59	$p = 0.613$

Subsequently, pre–post differences (Δ) were calculated, and independent samples *t*-tests were conducted. Although the differences did not reach statistical significance, the experimental group demonstrated more substantial improvements than the control group in both dimensions of empathy. The effect size for cognitive empathy was small, indicating no significant changes occurred following the program's implementation. However, for affective empathy, the observed effect size was medium [48], which also suggests the presence of statistically meaningful changes. According to Cohen, a medium effect size may already indicate a noticeable change in this field (See Table 5).

Table 5. Comparison of pre–post differences in perception of behaviors related to bullying (Δ).

Variable	Experimental Δ (M ± DT)	Control Δ (M ± DT)	t	*p*	d Cohen
Knowledge about bullying behaviors	5.34 ± 20.41	−3.53 ± 21.14	−2.43	0.016	−0.427
Bullying behaviors witnessed	1.10± 25.06	−2.60 ± 22.47	−0.891	0.375	−0.156
Bullying behaviors suffered	−1.15 ± 16.93	−0.74 ± 16.93	0.135	0.893	0.024

4. Discussion

This study demonstrates that a digital escape room based on Universal Design for Learning (UDL) principles can be an effective introductory strategy for enhancing knowledge and critical awareness about bullying among secondary school students.

The results highlight two key findings: first, the intervention's capacity to enhance students' knowledge about bullying behaviors, and second, its potential as a motivational entry point within broader socio-educational programs aimed at promoting empathy. Although increases were observed in both affective and cognitive empathy, only the affective component showed a statistically significant change, and even then, effect sizes were in the small to medium range. These modest results suggest that while the intervention may trigger initial emotional engagement, it is unlikely to generate more profound or lasting changes in empathic abilities without further reinforcement. This finding is consistent with prior research, which shows that short, stand-alone interventions can yield transient empathy gains. However, sustained exposure is generally required to consolidate these effects [49,50]. Recent thematic reviews also highlight that effective empathy development often requires multi-week interventions that include structured perspective-taking and opportunities for reflection, to foster sustained behavioral change [51]. Therefore, brief activities like digital escape rooms may be best conceptualized as complementary tools that spark interest and raise awareness, rather than as definitive solutions for empathy training. This limitation is significant when addressing bullying linked to systemic exclusion, such as that experienced by LGBTQ+ students or learners with disabilities, where long-term, inclusive pedagogical frameworks are necessary to challenge entrenched social biases [52].

The most significant result was an improvement in knowledge about bullying behaviors, supported by the power of gamified narratives to enhance the acquisition of

socio-emotional competencies [53]. Escape rooms, through their immersive and collaborative structure, enable students to identify and articulate behaviors such as exclusion, cyberbullying, and verbal microaggressions [54]. For marginalized populations, this explicit training in recognizing patterns of discrimination represents a critical first step toward breaking cycles of victimization [55].

Moreover, the reduction in self-reported bullying experiences and observation by bystanders suggests that increased awareness may precede behavioral changes [56,57]. These findings are supported by previous studies, which have shown that gamified interventions increase awareness of bullying, laying the groundwork for gradual transformations [58,59]. Based on the Universal Design for Learning (UDL) principles, the design amplified these effects by ensuring equitable accessibility. Multiple formats (videos, comics, puzzles) and collaborative tasks removed barriers for diverse learning styles [60], demonstrating how pedagogical inclusion enhances the impact of interventions [61].

These results position digital escape rooms as versatile tools within comprehensive anti-bullying strategies. Their ability to generate initial engagement and critical awareness makes them ideal allies for long-term initiatives, such as peer mentoring programs or restorative justice curricula [62,63]. For instance, integrating these activities with modules on intersectional diversity or digital literacy workshops could amplify their relevance in contexts marked by structural inequalities [64].

While the study's results provide promising evidence regarding the potential of digital escape rooms designed following UDL principles to promote empathy and awareness of bullying among adolescents, several limitations must be acknowledged to contextualize the findings appropriately.

First, the sample was selected using a non-probabilistic convenience method from only two public schools. This restricts the external validity of the results and limits the ability to generalize findings to broader populations. Future research should aim to include more diverse and representative samples, spanning different regions, school types, and socioeconomic backgrounds, to enhance generalizability and evaluate the impact of contextual variables.

Second, the intervention consisted of a single session lasting approximately one hour. While short-term positive outcomes were observed, particularly in recognizing bullying behaviors, it is essential to acknowledge the challenges of achieving and sustaining meaningful empathy gains through such brief interventions. Therefore, future studies should adopt longitudinal designs or include follow-up assessments to examine whether initial improvements in empathy translate into durable behavioral changes over time.

Third, the study relied exclusively on self-report instruments. Although these tools demonstrated sound psychometric properties, self-report data may be subject to limitations such as social desirability bias or limited self-awareness. Incorporating behavioral observations, peer nominations, or teacher assessments in future research would provide more robust and triangulated measures of empathy-related change.

Finally, considerations related to digital equity and facilitation expertise also emerged. Variability in device quality, internet access, and familiarity with digital tools may have impacted students' ability to fully engage with the escape room, particularly in under-resourced settings. These challenges intersect with language proficiency, which can influence how students interpret and respond to text- and audio-based content. Grounded in the UDL framework, future iterations of the escape room should include customizable and inclusive features that actively support linguistic and cognitive diversity. These may involve multilingual interfaces, audio narration to assist students with reading difficulties, and simplified or adjustable text formats [65]. Moreover, incorporating accessibility tools

such as haptic feedback or voice-guided navigation, which have been successfully tested in accessible VR escape rooms, could ensure broader usability for students with visual impairments [66].

5. Conclusions

Digital escape rooms based on UDL principles offer a promising tool for enhancing secondary students' knowledge and critical awareness of bullying behaviors. The intervention's immersive, gamified structure proved particularly effective in helping students identify and articulate subtle forms of bullying, such as exclusion, cyberbullying, and microaggressions—behaviors often normalized in school settings.

While empathy gains were modest, this aligns with its socially constructed nature. Sustained engagement is typically needed to deepen empathetic skills. This suggests that brief interventions, such as escape rooms, are better positioned as catalysts for long-term programs rather than standalone solutions [67]. Their ability to generate initial engagement and foundational awareness makes them ideal for integration into comprehensive strategies, such as peer mentoring or restorative justice curricula, which can foster more profound empathy and promote lasting behavioral change over time. The UDL framework was crucial in ensuring accessibility and inclusivity, critical to the intervention's success. By incorporating multiple modalities (e.g., videos, puzzles, collaborative tasks), the escape room accommodated diverse learning needs, reinforcing the importance of pedagogical flexibility in addressing systemic inequities, especially for marginalized groups like LGBTQ+ students or those with disabilities who face disproportionate bullying risks.

Moreover, these findings have clear practical implications for educators and policymakers. Incorporating digital escape rooms into existing anti-bullying programs can engage students early and foster critical awareness. Schools should provide facilitator training in UDL principles and basic technical troubleshooting to ensure consistent, immersive experiences. Future research should explore the longitudinal implementation of escape rooms, integrating them into multi-session modules, and assessing their impact on behavioral outcomes over time. It would also be valuable to compare different gamified formats (e.g., VR simulations, mobile apps) and to investigate scalability across diverse cultural and resource contexts.

Author Contributions: Conceptualization, Y.A.-M. and A.M.-L.; methodology, Y.A.-M.; software, A.M.-L.; validation, A.M.-L., J.M.A.-P. and J.R.-M.; formal analysis, A.M.-L.; investigation, Y.A.-M., A.M.-L. and P.R.-R.; resources, J.M.A.-P.; data curation, Y.A.-M.; writing—original draft preparation, Y.A.-M., A.M.-L. and P.R.-R.; writing—review and editing, A.M.-L. and P.R.-R.; visualization, A.M.-L., J.M.A.-P. and J.R.-M.; supervision, A.M.-L., J.M.A.-P., J.R.-M. and J.M.R.-F.; project administration, A.M.-L.; funding acquisition, J.M.R.-F. All authors have read and agreed to the published version of the manuscript.

Funding: The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: Projects to Strengthen Research Groups FEDER-UAL (Ref: P_FORT_GRUPOS_2023/05).

Institutional Review Board Statement: The study was conducted following the Declaration of Helsinki and approved by the University of Almería Ethics Committee (UALBIO2023/019), approval date 1 October 2023.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The dataset of this study is available in Zenodo at the following link: https://doi.org/10.5281/zenodo.15292394 (accessed on 6 June 2025).

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Bourou, A.; Papageorgiou, E. Prevalence of Aggressive Behavior in Greek Elementary School Settings from Teachers' Perspectives. *Behav. Sci.* **2023**, *13*, 390. [CrossRef] [PubMed]
2. Ahmad, W.; Siddique, M.S.; Jahangir, R.; Rahim, R. Prevalence of Bullying and Perceived Social Support Among Students in Peshawar. *J. Gandhara Med. Dent. Sci.* **2020**, *7*, 40–45. [CrossRef]
3. Yasmini, N.W.S. Integrasi satua bali dalam konseling postmodern untuk meminimalisasi perilaku bullying siswa. *Indones. J. Educ. Dev. IJED* **2020**, *1*, 190–198.
4. Gage, N.A.; Katsiyannis, A.; Rose, C.; Adams, S.E. Disproportionate Bullying Victimization and Perpetration by Disability Status, Race, and Gender: A National Analysis. *Adv. Neurodev. Disord.* **2021**, *5*, 256–268. [CrossRef]
5. O'Malley, M.; Brock, S.E. Introduction to Special Issue: Addressing Aggression and Violence in Diverse School Settings. *Contemp. Sch. Psychol.* **2020**, *24*, 249–251. [CrossRef]
6. Schmitz, R.M.; Tabler, J.; Charak, R.; Gomez, G.; Cole, R.E.; Schmitz, J.J. Navigating Digital Geographies and Trauma Contexts: Conceptions of Online Communities and Experiences Among LGBTQ+ People During COVID-19. *Int. J. Environ. Res. Public Health* **2025**, *22*, 443. [CrossRef]
7. Lee, J.M.; Park, J.; Lee, H.; Lee, J.; Mallonee, J. The Impact of Cyberbullying Victimization on Academic Satisfaction among Sexual Minority College Students: The Indirect Effect of Flourishing. *Int. J. Environ. Res. Public Health* **2023**, *20*, 6248. [CrossRef]
8. Kala, E.M. Influence of Online Platforms on Criminal Behavior. *Int. J. Res. Stud. Comput. Sci. Eng. IJRSCSE* **2024**, *10*, 25–37. [CrossRef]
9. Alturif, G.; Alsand, H. The repercussions of cyberbullying on students on educational platforms. *Front. Psychol.* **2024**, *15*, 1426467. [CrossRef]
10. Denervaud, S.; Mumenthaler, C.; Gentaz, E.; Sander, D. Emotion recognition development: Preliminary evidence for an effect of school pedagogical practices. *Learn. Instr.* **2020**, *69*, 101353. [CrossRef]
11. Daly, B.P.; Resnikoff, A.; Litke, S. Effective School Leadership for Supporting Students' Mental Health: Findings from a Narrative Literature Review. *Behav. Sci.* **2025**, *15*, 36. [CrossRef] [PubMed]
12. Iswinarti, I.; Hidayah, N. Empathy enhancement through group play therapy to reduce aggressive behavior. *J. Konseling Dan Pendidikan* **2020**, *8*, 151–158. [CrossRef]
13. León-del-Barco, B.; Lázaro, S.M.; Polo-del-Río, M.-I.; López-Ramos, V.-M. Emotional Intelligence as a Protective Factor against Victimization in School Bullying. *Int. J. Environ. Res. Public Health* **2020**, *17*, 9406. [CrossRef] [PubMed]
14. Huang, L.; Li, W.; Xu, Z.; Sun, H.; Ai, D.; Hu, Y.; Wang, S.; Li, Y.; Zhou, Y. The severity of cyberbullying affects bystander intervention among college students: The roles of feelings of responsibility and empathy. *Psychol. Res. Behav. Manag.* **2023**, *16*, 893–903. [CrossRef]
15. Salem, A.A.M.S.; Al-Huwailah, A.H.; Abdelsattar, M.; Al-Hamdan, N.A.H.; Derar, E.; Alazmi, S.; Abu Al-Diyar, M.; Griffiths, M.D. Empathic Skills Training As a Means of Reducing Cyberbullying among Adolescents: An Empirical Evaluation. *Int. J. Environ. Res. Public Health* **2023**, *20*, 1846. [CrossRef]
16. Gupta, C.; Jogdand, D.S.; Kumar, M. Reviewing the Impact of Social Media on the Mental Health of Adolescents and Young Adults. *Cureus* **2022**, *14*, e30143. [CrossRef]
17. Canonigo, J.; Uy, L.K.; Culajara, C.L. Social Media Usage, Body Image, and Psychological Well-Being among College Students. *Asian J. Educ. Soc. Stud.* **2024**, *50*, 451–470. [CrossRef]
18. Rentzi, A. Social Justice in Education: Creating an Inclusive Culture at Schools Through Critical Multicultural Theory. *Eur. J. Educ. Pedagog.* **2024**, *5*, 94–97. [CrossRef]
19. Lin, J.C.; Shih, Y.-H. Strategies for preventing school bullying—A life education perspective. *Front. Psychol.* **2024**, *15*, 1429215. [CrossRef]
20. Ducato, A. Building Classroom Culture and Community with an Emphasis on Inclusive Practices. Master's Thesis, Dominican University of California, San Rafael, CA, USA, 2022.
21. Eisenberg, M.E.; Lawrence, S.E.; Eadeh, H.-M.; Suresh, M.; Rider, G.N.; Gower, A.L. Emotional Distress Disparities Across Multiple Intersecting Social Positions: The Role of Bias-Based Bullying. *Pediatrics* **2024**, *153*, e2023061647. [CrossRef]
22. Sorrentino, A.; Esposito, A.; Acunzo, D.; Santamato, M.; Aquino, A. Onset risk factors for youth involvement in cyberbullying and cybervictimization: A longitudinal study. *Front. Psychol.* **2023**, *13*, 1090047. [CrossRef] [PubMed]
23. Wang, Z.; Harun, J.; Yuan, Y. Enhancing Reading Instruction Through Gamification: A Systematic Review of Theoretical Models, Implementation Strategies, and Measurable Outcomes (2020–2024). *J. Inf. Technol. Educ. Res.* **2024**, *23*, 028. [CrossRef] [PubMed]

24. Williams, C.; Griffin, K.W.; Botvin, C.M.; Sousa, S.; Botvin, G.J. Effectiveness of Digital Health Tools to Prevent Bullying among Middle School Students. *Adolescents* **2023**, *3*, 110–130. [CrossRef]
25. Rodríguez-Rivera, P.; Manzano-León, A.; Dapía Conde, M.D.; Rodríguez-Ferrer, J.M. Unleashing creativity and cooperation: A qualitative case study on designing digital breakouts for social education degrees. *Societies* **2023**, *13*, 203. [CrossRef]
26. Acosta Yela, M.T.; Manrique Toro, R.M.; Rodriguez Triviño, J.A.; Viteri Chiriboga, E.A. The Role of Gamification in Teacher Training for Educational Inclusion. *Evol. Stud. Imaginative Cult.* **2024**, *8*, 1442–1453. [CrossRef]
27. Alieto, E.; Abequibel-Encarnacion, B.; Estigoy, E.; Balasa, K.; Eijansantos, A.; Torres-Toukoumidis, A. Teaching inside a digital classroom: A quantitative analysis of attitude, technological competence and access among teachers across subject disciplines. *Heliyon* **2024**, *10*, e24282. [CrossRef]
28. Leung, L.Y.Y.; Zou, T.X.P. The role of community in research-based learning: Undergraduate perspectives. *Educ. Res.* **2025**, *67*, 60–78. [CrossRef]
29. Ismailov, M.; Chiu, T.K.F. Catering to Inclusion and Diversity With Universal Design for Learning in Asynchronous Online Education: A Self-Determination Theory Perspective. *Front. Psychol.* **2022**, *13*, 819884. [CrossRef]
30. Lim, I. A physical neuroscience-themed escape room: Design, implementation, and students' perceptions. *Educ. Inf. Technol.* **2024**, *29*, 8725–8740. [CrossRef]
31. Hernández, L.E.; Darling-Hammond, L. Creating Identity-Safe Schools and Classrooms. *Learn. Policy Inst.* **2022**.
32. Breakstone, J.; Smith, M.; Wineburg, S.; Rapaport, A.; Carle, J.; Garland, M.; Saavedra, A. Students' Civic Online Reasoning: A National Portrait. *Educ. Res.* **2021**, *50*, 505–515. [CrossRef]
33. Lvina, E. Teaching ethics and ethical leadership with the Harvard Everest simulation. *J. Leadersh. Educ.* **2024**. ahead-of-print. [CrossRef]
34. Makri, A.; Vlachopoulos, D.; Martina, R.A. Digital Escape Rooms as Innovative Pedagogical Tools in Education: A Systematic Literature Review. *Sustainability* **2021**, *13*, 4587. [CrossRef]
35. Manzano-León, A.; Rodríguez-Ferrer, J.M.; Aguilar-Parra, J.M.; Martínez Martínez, A.M.; Luque de la Rosa, A.; Salguero García, D.; Fernández Campoy, J.M. Escape Rooms as a Learning Strategy for Special Education Master's Degree Students. *Int. J. Environ. Res. Public Health* **2021**, *18*, 7304. [CrossRef]
36. Sánchez, A.M. Using digital educational escape rooms as a motivational review tool for Economics. *Int. J. Manag. Educ.* **2023**, *21*, 100852. [CrossRef]
37. Shao, J.; Abdul Rabu, S.N.; Chen, C. Gamified interactive e-books for bullying prevention: Enhancing knowledge and motivation in Chinese primary schools. *Front. Psychol.* **2025**, *16*, 1509549. [CrossRef]
38. Sá, C. Designing an Intelligent Virtual Assistant for Preventing Bullying and Cyberbullying in Education. *Open J. Appl. Sci.* **2025**, *15*, 874–884.
39. Liu, Y.-L.; Chang, C.-Y.; Wang, C.-Y. Using VR to investigate bystander behavior and the motivational factors in school bullying. *Comput. Educ.* **2023**, *194*, 104696. [CrossRef]
40. Moukram, Y.; Manzano-León, A.; Rodríguez-Ferrer, J.; Rodríguez-Moreno, J.; Aguilar-Parra, J. A systematic review of Gamification as a playful strategy to prevent bullying. *Environ. Soc. Psychol.* **2022**, *7*, 38–50.
41. Alismaiel, O.A. Digital Media Used in Education: The Influence on Cyberbullying Behaviors among Youth Students. *Int. J. Environ. Res. Public Health* **2023**, *20*, 1370. [CrossRef]
42. Santos, J.; Andrade, E.; Benevides, K.; Silva, K.; Nascimento, J.; Bittencourt, I.; Pereira, M.; Fernandes, S.; Isotani, S. Does gender stereotype threat affects the levels of aggressiveness, learning and flow in gamified learning environments?: An experimental study. *Educ. Inf. Technol.* **2023**, *28*, 1637–1662. [CrossRef] [PubMed]
43. Merino-Soto, C.; López-Fernández, V.; Grimaldo-Muchotrigo, M. Invarianza de medición y estructural de la Escala Básica de Empatía Breve (bes-b) en niños y adolescentes peruanos. *Rev. Colomb. Psicol.* **2019**, *28*, 15–32. [CrossRef]
44. Álvarez Blanco, L.; Iglesias García, M.T.; Urbano Contreras, A.; García Díaz, V. Diseño y validación de la Escala de Autopercepción y Percepción del Acoso Escolar en Adolescentes (APAE-A). *Rev. Educ.* **2022**, *397*, 15–42.
45. Nicholson, S. Creating Engaging Escape Rooms for the Classroom. *Child. Educ.* **2018**, *94*, 44–49. [CrossRef]
46. Grimmer, T. *Jugamos a Ser Superhéroes: Aprendizaje Efectivo, Valores Positivos Y Competencias Para La Vida*; Narcea Ediciones: Madrid, Spain, 2022; Volume 41.
47. Pastor, C.A. Diseño Universal para el Aprendizaje: Un modelo teórico-práctico para una educación inclusiva de calidad. *Particip. Educ.* **2019**, *6*, 55–68.
48. Cohen, J. *Statistical Power Analysis for the Behavioral Sciences*; Routledge: New York, NY, USA, 2013.
49. Levett-Jones, T.; Cant, R.; Lapkin, S. A systematic review of the effectiveness of empathy education for undergraduate nursing students. *Nurse Educ. Today* **2019**, *75*, 80–94. [CrossRef]

50. Winter, R.; Issa, E.; Roberts, N.; Norman, R.I.; Howick, J. Assessing the effect of empathy-enhancing interventions in health education and training: A systematic review of randomised controlled trials. *BMJ Open* **2020**, *10*, e036471. [CrossRef]
51. Makal, N.; Çepni, S. Increasing Students' Empathy Levels in Science Education: A Thematic Review. *J. Balt. Sci. Educ.* **2024**, *23*, 914–930. [CrossRef]
52. Miller, R.A.; Wynn, R.D.; Stare, B.; Williamson, J.; Guo, L.; Stare, B.G.; Williamson, J.N. Mental Health and Resilience Among LGBTQ+ College Students with Disabilities. *Curr. J. Divers. Scholarsh. Soc. Change* **2022**, *2*. [CrossRef]
53. Rončević Zubković, B.; Svjetlana, K.-V.; Sanja, S.-A.; Tamara, M.D.; Pahljina-Reinić, R. The role of experience during playing bullying prevention serious game: Effects on knowledge and compassion. *Behav. Inf. Technol.* **2022**, *41*, 401–415. [CrossRef]
54. Bakkum, M.J.; Richir, M.C.; Sultan, R.; de la Court, J.R.; Lambooij, A.C.; van Agtmael, M.A.; Tichelaar, J. Can Students Create Their Own Educational Escape Room? Lessons Learned from the Opioid Crisis Escape Room. *Med. Sci. Educ.* **2021**, *31*, 1739–1745. [CrossRef] [PubMed]
55. Beagan, B.L.; Tameera, M.; Kim, B.; Bea, W.; Weinberg, M. Microaggressions experienced by LGBTQ academics in Canada: "just not fitting in. . . it does take a toll". *Int. J. Qual. Stud. Educ.* **2021**, *34*, 197–212. [CrossRef]
56. Peng, Z.; Li, L.; Su, X.; Lu, Y. A pilot intervention study on bullying prevention among junior high school students in Shantou, China. *BMC Public Health* **2022**, *22*, 262. [CrossRef] [PubMed]
57. Wu, Q.; Jia, F. Empowering Students against Ethnic Bullying: Review and Recommendations of Innovative School Programs. *Children* **2023**, *10*, 1632. [CrossRef]
58. Oyekoya, O.; Urbanski, J.; Shynkar, Y.; Baksh, A.; Etsaghara, M. Exploring First-Person Perspectives in Designing a Role-Playing VR Simulation for Bullying Prevention: A Focus Group Study. *Front. Virtual Real.* **2021**, *2*, 672003. [CrossRef]
59. Dailey, S.F.; Roche, R.R. The SHIELD Framework: Advancing Strength-Based Resilience Strategies to Combat Bullying and Cyberbullying in Youth. *Int. J. Environ. Res. Public Health* **2025**, *22*, 66. [CrossRef]
60. Engaging potential: Using the UDL principle of engagement to motivate year 7–8 learners with autism in writing tasks. *Kairaranga* **2023**, *24*, 57–70. [CrossRef]
61. Adom, D.; Abroampa, W.K.; Amoako, R.; Toquero, C.M.; Kquofi, S. Towards an Inclusive Pedagogy: Applying the Universal Design for Learning in an Introduction to History of Global Art Course in Ghana. *J. Contemp. Issues Educ.* **2023**, *18*, 79–97. [CrossRef]
62. Hendry, B.P.; Hellsten, L.-a.M.; McIntyre, L.J.; Smith, B.R.R. Recommendations for cyberbullying prevention and intervention: A Western Canadian perspective from key stakeholders. *Front. Psychol.* **2023**, *14*, 1067484. [CrossRef]
63. Saleh Alabdulaziz, M. Escape rooms technology as a way of teaching mathematics to secondary school students. *Educ. Inf. Technol.* **2023**, *28*, 13459–13484. [CrossRef]
64. González, T.; Epstein, R. Critical race feminism, health, and restorative practices in schools: Centering the experiences of Black and Latina girls. *Mich. J. Gend. Law* **2022**, *29*, 409. [CrossRef]
65. Coronado, A.; Carvalho, S.; Berretta, L. Game accessibility: Adaptation of a digital escape room game to improve spatial cognitive skills on blind people. In Proceedings of the 25th Symposium on Virtual and Augmented Reality, Rio Grande, Brazil, 6–9 November 2024; pp. 174–182.
66. Coronado, A.; Carvalho, S.T.; de Oliveira Berretta, L. Escape-INF-VR: An Accessible VR Escape Game Proposal for Blind Individuals. In Proceedings of the Entertainment Computing—ICEC 2024, Cham, Switzerland, 30 September–3 October 2025; pp. 337–341.
67. Kasperski, R.; Levin, O.; Hemi, M.E. Systematic Literature Review of Simulation-Based Learning for Developing Teacher SEL. *Educ. Sci.* **2025**, *15*, 129. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

Multimodal Technologies and Interaction

MDPI

Article

Space Medicine Meets Serious Games: Boosting Engagement with the Medimon Creature Collector

Martin Hundrup [1], Jessi Holte [2], Ciara Bordeaux [2], Emma Ferguson [2], Joscelyn Coad [2], Terence Soule [3] and Tyler Bland [4,*]

[1] Computer Science Department, Washington State University, Pullman, WA 99164, USA
[2] Virtual Technology and Design Program, University of Idaho, Moscow, ID 83844, USA
[3] Computer Science Department, University of Idaho, Moscow, ID 83844, USA
[4] WWAMI Medical Education Department, University of Idaho, Moscow, ID 83844, USA
* Correspondence: tbland@uidaho.edu

Abstract: Serious games that integrate educational content with engaging gameplay mechanics hold promise for reducing cognitive load and increasing student motivation in STEM and health science education. This preliminary study presents the development and evaluation of the Medimon NASA Demo, a game-based learning prototype designed to teach undergraduate students about the musculoskeletal and visual systems—two critical domains in space medicine. Participants (n = 23) engaged with the game over a two-week self-regulated learning period. The game employed mnemonic-based characters, visual storytelling, and turn-based battle mechanics to reinforce medical concepts. Quantitative results demonstrated significant learning gains, with posttest scores increasing by an average of 23% and a normalized change of c = 0.4. Engagement levels were high across multiple dimensions of situational interest, and 74% of participants preferred the game over traditional formats. Qualitative analysis of open-ended responses revealed themes related to intrinsic appeal, perceived learning efficacy, interaction design, and cognitive resource management. While the game had minimal impact on short-term STEM career interest, its educational potential was clearly supported. These findings suggest that mnemonic-driven serious games like Medimon can effectively enhance engagement and learning in health science education, especially when aligned with real-world contexts such as space medicine.

Keywords: serious games; game-based learning; mnemonics; STEM education

1. Introduction

Students in STEM and health science disciplines often face cognitive load due to the volume and complexity of required content [1–3]. This challenge, combined with decreasing interest in traditional educational approaches, contributes to disengagement and difficulty retaining knowledge [4–7]. As a result, there is growing interest in innovative teaching strategies that can both reduce cognitive load and increase motivation to learn—particularly among younger learners. Game-based learning (GBL) has emerged as a promising approach that promotes active learning, motivation, and knowledge retention through interactive and immersive experiences [8–11]. Serious games, which are games whose primary purpose is for education or training rather than entertainment, have been increasingly adopted in health professions education for their ability to blend learning objectives with engaging gameplay [12,13]. Targeted examples underscore the breadth of serious-game applications in health care training. Empiric, a tabletop card game for

guideline-based antibiotic prescribing, engaged practicing clinicians at two continuing medical education conferences, who rated the session highly and reported intent to change prescribing behavior [14]. Likewise, AntibioGame®, a digital case-based simulation on rational antibiotic use, was perceived as attractive, usable, and fun by medical students [15]. Together, these studies illustrate how well-designed GBL interventions can foster both knowledge acquisition and behavior change across the spectrum of medical education.

Another well-established method for supporting learning in high cognitive load fields is the use of mnemonics [16,17]. Mnemonics leverage visual, linguistic, and spatial cues to aid in memory encoding and recall, making them useful for students managing large quantities of information common in STEM and health science fields [18–21]. Some common examples of visual mnemonics used in medical education are Sketchy Medical and Picmonic [22,23]. While both GBL and mnemonics have been individually validated, few studies meaningfully combine the two into a unified educational experience.

To address this gap, we developed Medimon, an innovative GBL platform that merges mnemonic-based learning with the mechanics of a creature-collector role-playing game [24]. In Medimon, players explore fictional environments, capture characters based on body systems, and engage in turn-based battles using moves and abilities grounded in medical mnemonics. The platform is designed to reduce cognitive load, enhance engagement, and improve knowledge retention through multisensory learning.

Our current preliminary study centers on the Medimon NASA Demo prototype, which focuses on two body systems of critical relevance to NASA and the future of space exploration: the musculoskeletal and visual systems. Astronauts operating in microgravity environments experience significant health challenges, including skeletal tissue degeneration, muscle atrophy, cardiovascular deconditioning, and Spaceflight-Associated Neuro-Ocular Syndrome (SANS) [25,26]. These issues are not only clinically important but also serve as compelling educational topics with real-world relevance that can capture student interest.

The Medimon NASA Demo was designed as a prototype serious game to teach undergraduate students about these systems through a self-regulated learning (SRL) environment [27,28]. We hypothesize that this preliminary demo will increase student knowledge of these body systems, foster engagement with health science education, and potentially boost interest in pursuing STEM-related careers. Though piloted with undergraduates, the demo's content aligns with advanced high-school biology standards, positioning this as an early-stage feasibility study.

2. Materials and Methods

2.1. Participants

Undergraduate students at the University of Idaho were recruited via institution email invitations and in-person presentations delivered by the authors. Because a central research aim was to examine whether the intervention could stimulate interest in pursuing STEM careers among students not already committed to those fields, we deliberately focused our in-person recruitment efforts on courses outside the traditional STEM disciplines. Brief 15 min study invitation presentations were delivered in two settings that met this criterion yet still showed a natural affinity for game-based learning: (i) students in the Virtual Technology and Design (VTD) program—whose curriculum includes video-game production—and (ii) an entrepreneurship course in the College of Business and Economics. While our primary intent was to engage non-STEM students, we did ultimately enroll a small number (n = 3) of participants majoring in STEM-related fields, including computer science, cybersecurity, medicine, and nutrition science (Table 1).

Table 1. Participant demographics. Avg: Average, SD: Standard deviation.

Characteristic	Count (Avg)	SD
Age	(21)	2.2
Gender		
Male	11	
Female	9	
Non-binary/third gender	3	
Ethnicity		
White	23	
Hispanic	1	
Asian	3	
American Indian/Alaskan Native	1	
Major		
Virtual Technology and Design	12	
Creative Writing	1	
Conservation Biology	1	
Computer Science	1	
Cybersecurity	1	
Studio Art and Design	1	
Medicine	1	
Nutrition Science	1	
Marketing	2	
English	1	
Human Resources Management	1	
Business Management	1	
Finance	1	

Recruitment occurred on a rolling basis throughout the Spring 2025 semester. Overall, 92 undergraduates provided informed consent to take part in the study (Figure 1). Of these, 83 students (90%) completed the baseline (pre-test) survey. During the subsequent two-week intervention window, 61 students discontinued participation—by never launching the game or by not completing the post-test or the Situational Interest Survey–Multimedia (SIS-M)—leaving 23 students (28% of those who completed the pre-test) who (a) engaged with the game, and (b) submitted both the post-test and the SIS-M. These 23 participants constituted the final analytic sample for all achievement, engagement, and interest analyses reported below. Demographic information is summarized in Table 1. Some participants indicated multiple ethnicities and Majors.

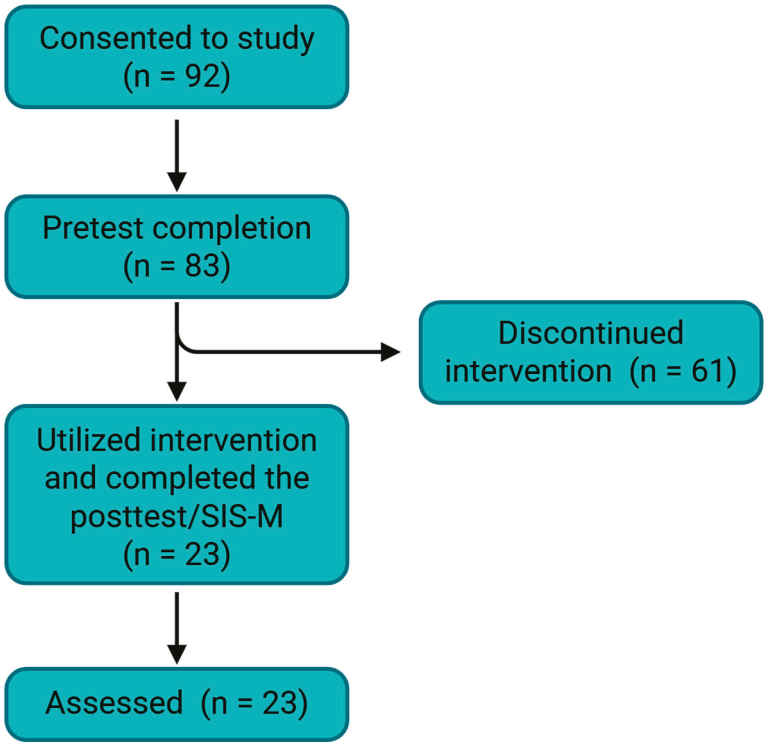


Figure 1. Flow diagram showing recruitment, exclusions, and final analytic sample (n = 23).

2.2. Design

The prototype was designed as a top-down, role-playing game (RPG) with creature-collector mechanics, drawing inspiration from early-generation titles such as *Pokémon Red, Blue, and Yellow* for the Game Boy. Players navigate a 2D overworld from a bird's-eye perspective, encounter and recruit educational creatures (called "Medimon"), and engage in turn-based battles that integrate medical content with gameplay. This familiar format was selected to evoke nostalgia, lower the learning curve for non-gamers, and scaffold engagement through recognizable game systems. The game was developed using the Unity game engine and compiled for both Windows (PC) and macOS platforms. Installation files were distributed to participants via the digital game-hosting platform itch.io, allowing easy and secure access for download.

The game prototype taught concepts from the musculoskeletal and visual systems using mnemonic-based characters and environments. Each character was part of a Medimon family representing a specific body system. Characters were designed across three evolutionary stages—Stage 1 (baby), Stage 2 (adolescent), and Stage 3 (adult)—with both healthy and diseased versions. Healthy Medimon represented normal anatomy and physiology, while diseased Medimon embodied related medical conditions. Diseased versions could be transformed into their healthy counterparts using in-game therapy items modeled after real-world treatments. Families and characters included in the prototype are listed in Appendix Table A1.

Each character featured visual mnemonics, which were visual designs that reflect the biological function or pathology of the real-world item they represent (Figure 2). For example, the Osteoclast character visually breaks down bone, symbolizing bone resorption. Locations within the game also supported mnemonic learning, such as the Bone family's building, "Skeleton Crew Framing," a hardware store where players accessed visual aids and encountered related wild Medimon (Figure 3).

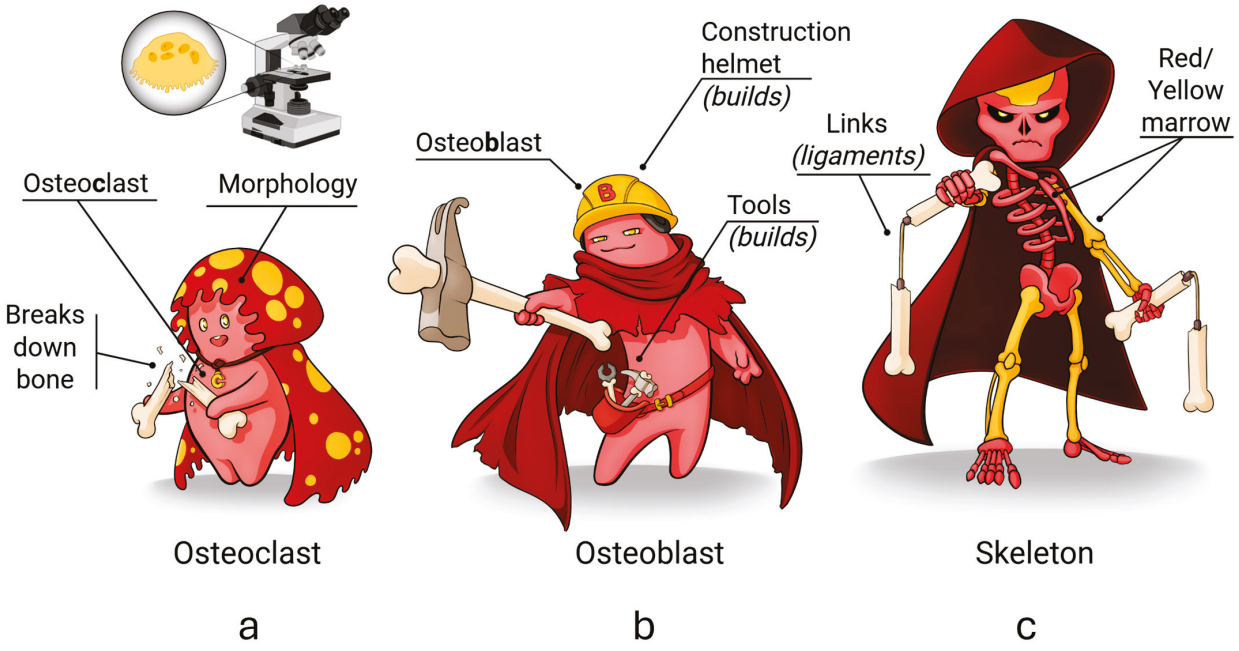


Figure 2. Bone family with labeled visual mnemonics. (**a**) Osteoclast is the stage 1 (baby) of the Bone family. It represents osteoclast cells in bone tissue that resorb bone. (**b**) Osteoblast is the stage 2 (adolescent) of the Bone family. It represents osteoblast cells in bone tissue, which promote the buildup of bone. (**c**) Skeleton is the stage 3 (adult) of the Bone family. It represents the skeleton and marrow of the skeletal system.



Figure 3. The interior of the Bone family location, Skeleton Crew Framing (building exterior shown in the bottom right corner). Five main types of bones are represented: long, short, flat, irregular, and sesamoid. Tendons are also represented with a visual connection to the Muscle character, representing the function of tendons connecting muscle to bone.

The turn-based battle system enabled players to battle and capture wild Medimon. Victories awarded experience points that allowed Medimon on the player's team to level up, enhance stats, and evolve. Key game statistics included:

- **HP (Health Points):** Measures life; reaching zero causes the Medimon to "flat line," meaning it cannot be used in battle until healed.
- **ATP (Adenosine Triphosphate):** Energy required to perform moves in battle, representing cellular energy.
- **STS (Stress):** Increases with each move during battle; high STS may cause healthy Medimon to evolve into diseased versions or diseased Medimon to self-destruct.
- **Attack and Defense:** Determine battle effectiveness and resistance to damage.

Moves within battles contained linguistic mnemonics (Figure 4). For example, the Osteoclast's "Bone Loss" move reinforces its biological role in bone resorption.

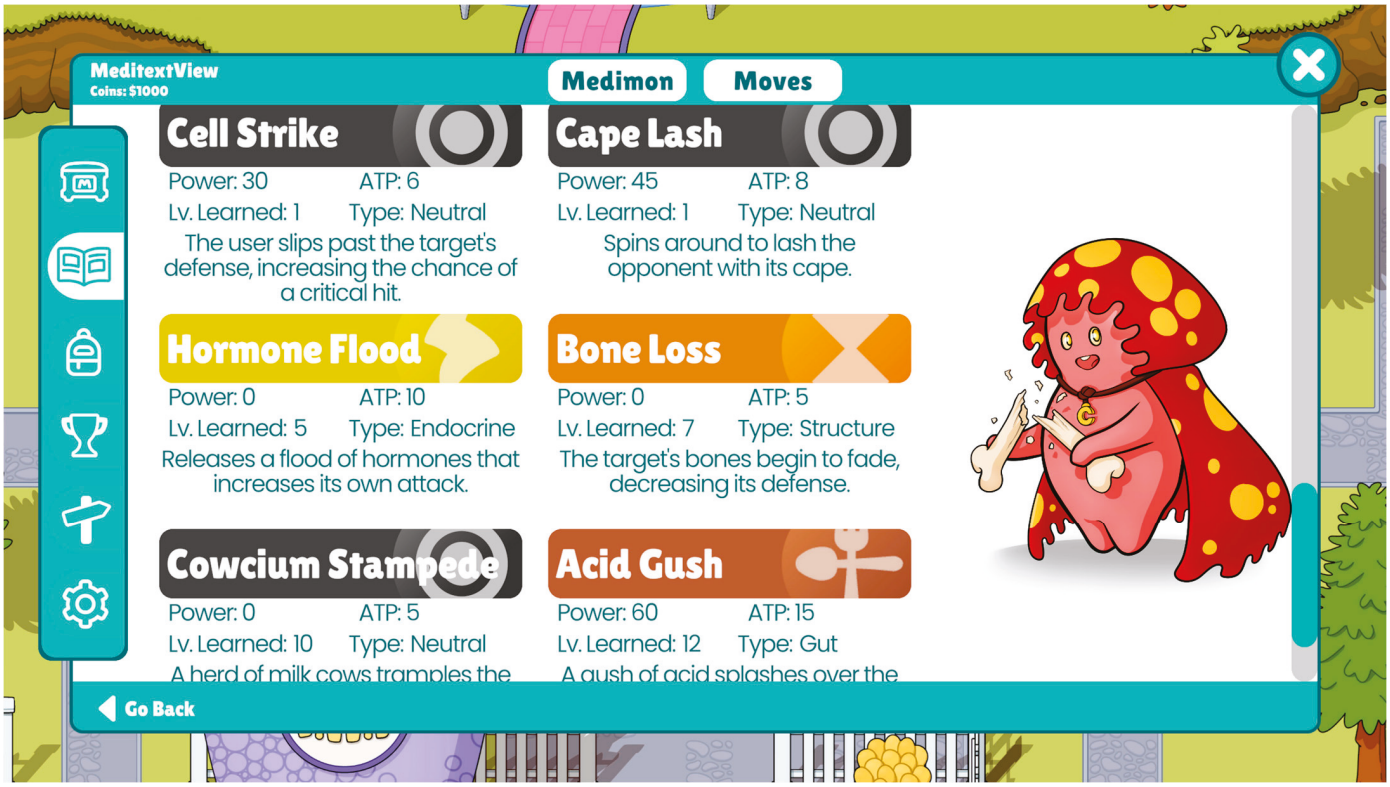


Figure 4. Osteoclast battle moves menu screen. Lv: Level.

Players could also battle non-playable characters (NPCs) to level up their Medimon and visit a hospital to restore HP, ATP, and reduce STS. A comprehensive tutorial guided players through core mechanics, and the in-game menu provided explanations of educational content and mnemonics.

Cognitive Load Theory (CLT) distinguishes between intrinsic load (the complexity inherent to the material), extraneous load (processing demands imposed by the learning environment), and germane load (mental effort devoted to schema construction) [3]. With these principles in mind, we deliberately incorporated several game mechanics that either lighten unnecessary load or channel effort toward meaningful learning. First, visual mnemonics—for example, depicting an Osteoclast physically breaking a bone to signify bone resorption—provide concrete, image-based cues that "chunk" complex physiological information, thereby lowering intrinsic load while simultaneously exploiting dual-coding

effects [29]. Second, context-sensitive tool-tips and inventory pop-ups supply just-in-time explanations of battle moves or character functions at the moment players need them; this eliminates the split-attention problem and trims extraneous load. Third, a guided tutorial sequence introduces mechanics in graduated steps rather than all at once, keeping the element interactivity within the learner's working-memory capacity and managing intrinsic load through scaffolding. Finally, spaced-retrieval Medimon battles require players to revisit earlier content under slightly varied conditions; by prompting effortful recall at optimally spaced intervals, these encounters shift mental resources toward germane load, fostering robust schema formation and long-term retention. Collectively, these intertwined design decisions operationalize CLT within the gameplay loop, ensuring that the cognitive effort students expend is both efficient and educationally productive.

2.3. Intervention

Participants were invited to download and play the Medimon NASA Demo, a serious game prototype integrating health science content related to the musculoskeletal and visual systems. They were instructed to engage with the game over a two-week period in an SRL manner, allowing them to control when and how often they played. The game is freely available online at: https://panacea-interactive.itch.io/medimon-nasa-project (accessed on 3 August 2025).

2.4. Achievement

To evaluate learning, participants completed a short multiple-choice questionnaire before (pretest) and after (posttest) the two-week game period. Questions assessed knowledge of cellular functions, anatomical structures, and associated diseases covered in the game. Participants completed the pre-/posttest through the Qualtrics platform.

2.5. STEM Career Interest

As part of the pretest and posttest, participants rated their interest in pursuing a STEM career using a 5-point Likert scale (1 = Very interested; 5 = Very not interested). Changes in response were used to evaluate whether the game influenced career aspirations.

2.6. Data Collection

After completing the posttest, participants also completed the Situational Interest Survey of Multimedia (SIS-M), a validated instrument for evaluating multimedia engagement [30,31]. The SIS-M evaluates triggered interest, maintained interest, maintained feeling, and value interest using a 5-point Likert scale (1 = Strongly disagree; 5 = Strongly agree) (Appendix Table A2), and has been used in prior medical education studies [32–35]. The 12-item SIS-M was administered with two additional items: (1) a preference ranking between different health science learning modalities (Medimon game, traditional materials, or no preference), and (2) an open-ended explanation of that preference. Participants completed the SIS-M through the Qualtrics platform.

2.7. Data Analysis

Quantitative data analysis was conducted using Microsoft Excel, GraphPad Prism (v 9.5.1 (733)), and GPT o3 [36]. Student achievement was assessed by calculating the mean pretest and posttest scores for each participant group, with statistical comparisons conducted using a paired two-tailed *t*-test ($\alpha = 0.5$) and a Wilcoxon signed-rank test. To further evaluate learning gains, normalized change calculations were applied, providing a standardized metric for interpreting improvement relative to baseline performance [37].

For the Situational Interest Survey in Multimedia (SIS-M), responses were analyzed across four validated dimensions of situational interest: triggered interest, maintained interest, maintained-feeling, and maintained-value. We classified mean SIS-M scores < 2.5 = low, 2.5–4 = moderate, >4 = high engagement.

The open-ended SIS-M survey responses were analyzed using a large language model (LLM)-driven qualitative workflow informed by recent applications in medical education research [32–35]. An agentic analysis process was utilized [38,39], which was powered by the "thinking" Google Gemini 2.5 Pro Experimental 03-25 LLM. The process involved three specialized generative artificial intelligent (genAI) agents, each guided by distinct system-level prompts tailored to their roles: Principal Investigator (PI), Qualitative Researcher (QR), and Subject Matter Expert (SME) (Figure 5a). The workflow operated in a sequential, iterative format in which the output from one agent was refined by the next (Figure 5b). Prompt engineering strategies, including Persona Prompting [40,41], were used to optimize agent performance. A detailed account of the full workflow—including all prompts, agent instructions, and resulting outputs—is provided in the Supplemental Materials (Methods S1).

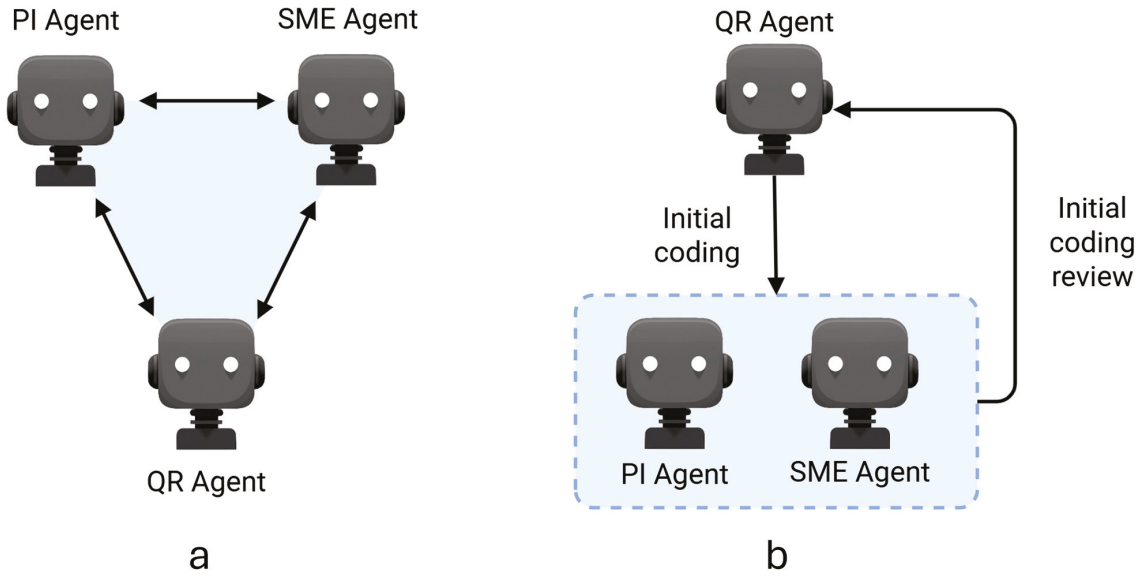


Figure 5. Agentic genAI thematic analysis. (**a**) genAI agents and workflow of the thematic analysis of the open-ended survey responses. (**b**) Example workflow for a single step of the agentic thematic analysis. This represents the step of initial coding of the survey responses by the QR agent and review by the PI and SME agents. PI: Principal Investigator, SME: Subject Matter Expert, QR: Qualitative Researcher.

2.8. Ethical Considerations

The study protocol was reviewed and approved as exempt by the Institutional Review Board (23-042) at the University of Idaho. Informed consent was obtained from all participants prior to data collection.

3. Results

3.1. Achievement

All participants completed a multiple-choice questionnaire before and after engaging with the Medimon NASA Demo. Analysis revealed that posttest scores were significantly higher than pretest scores (pretest: 42% ± 13%; posttest: 65% ± 24%; $p < 0.001$; Wilcoxon signed-rank test: W = 11, $n = 20$, $p = 0.00043$, r = 0.78), indicating a learning gain (Figure 6a,b).

To account for baseline knowledge variability and ceiling effects, we applied the normalized change method [37]. The average normalized change across participants was c = 0.4 (Figure 6c), representing a moderate improvement in knowledge acquisition during the two-week self-regulated learning period.

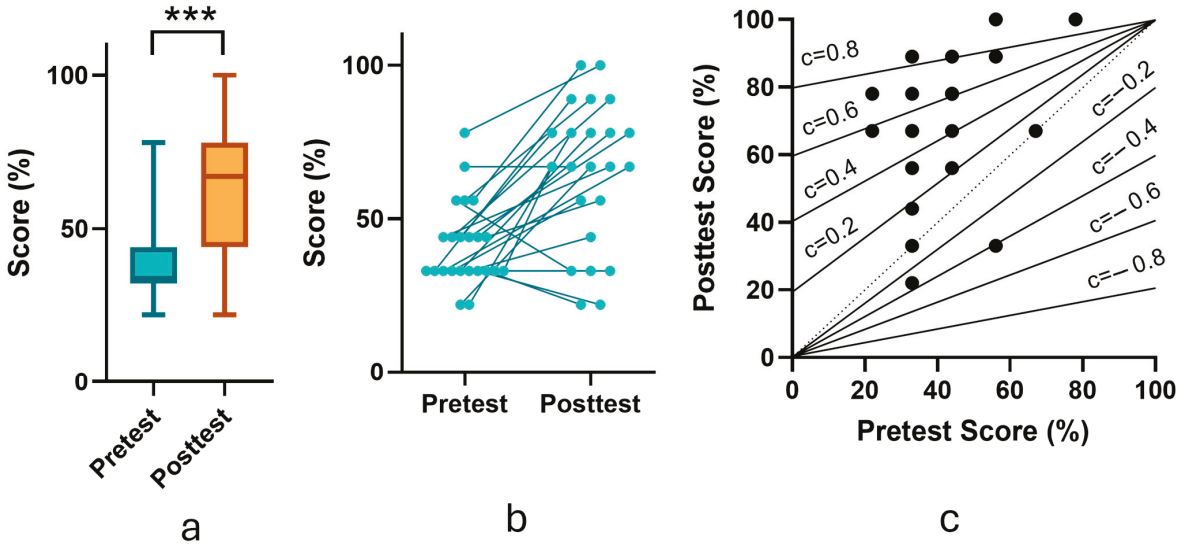


Figure 6. Achievement results. (**a**) Pretest scores before accessing the game and posttest scores after playing the game for two weeks. (**b**) Individual changes in scores between the pretest and posttest. (**c**) Normalized changes in scores between the pretest and posttest. *** $p < 0.001$.

3.2. Engagement

The SIS-M provided insights into participants' levels of interest and engagement with the Medimon NASA Demo. The participants' had a high average triggered situational interest (M = 4.15, SD = 0.78) and moderate maintained feeling (M = 3.92, SD = 0.79) in the game (Figure 7a).

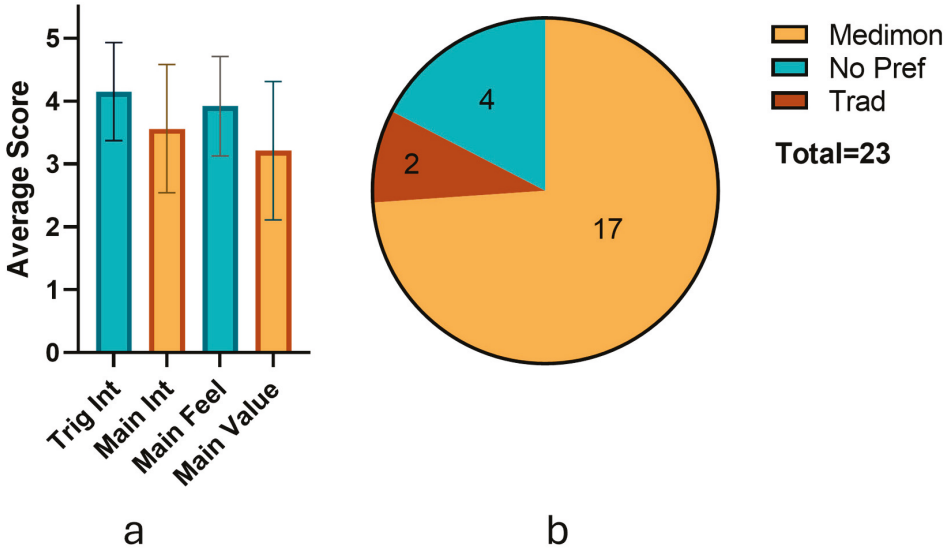


Figure 7. SIS-M results. (**a**) Average scores (error bars = standard deviation) of the SI types. (**b**) Responses to the question "Which type of health science education do you prefer?". Trig Int: Triggered Interest, Main Int: Maintained Interest, Main Feel: Maintained Feeling, Main Value: Maintained Value, No Pref: No Preference, Trad: Traditional.

The results for maintained interest (M = 3.56, SD = 1.02) and maintained value (M = 3.21, SD = 1.10) were lower but still moderate, indicating a more neutral stance with the game. Overall, the majority of participants (74%, 17/23) preferred the game for learning the health sciences, compared to the minority of participants who preferred traditional types of materials such as lectures and textbook readings (9%, 2/23) and those who had no preference in learning material format (17%, 4/23) (Figure 7b).

To understand the underlying reasons for the learning modality preferences reported in the SIS-M, thematic analysis was conducted on participants' open-ended explanations. This analysis yielded four key themes illuminating the factors driving participants' choices: (1) Intrinsic Appeal and Affective Engagement, (2) Perceived Learning Efficacy and Mechanisms, (3) Interaction Design and Learner Experience, and (4) Cognitive Factors and Resource Management (Table 2).

Table 2. Thematic analysis of the SIS-M open-ended responses. Trad: Traditional.

Theme 1: Intrinsic Appeal and Affective Engagement	
Description	This theme captures the enjoyment, motivation, and emotional engagement of learners. It reflects how fun, esthetically pleasing, and attention-grabbing each modality was perceived to be.
Prefer Medimon	Participants frequently highlighted the game's engaging and entertaining nature as central to their preference. Many linked fun to increased focus, emotional reward, and effective learning. Multimodal features (e.g., art, music) were especially impactful.
Prefer Trad	Enjoyment was viewed as secondary. These participants emphasized learning efficacy and saw games as supplemental.
No Preference	Acknowledged the game's fun and appeal but doubted its learning value. Some felt enjoyment did not equate to educational effectiveness.
Theme 2: Perceived Learning Efficacy and Mechanisms Description	
Description	This theme focuses on how well participants believed each method supported understanding and memory, and which features helped or hindered learning.
Prefer Medimon	Reported that the game made content easier to understand and remember. Valued mnemonics, interactivity, and integration of content into gameplay.
Prefer Trad	Believed textbooks provided better depth and clarity. Viewed games as better for review, not initial learning.
No Preference	Questioned the game's educational depth. Noted retention challenges and issues with recalling in-game content.
Theme 3: Interaction Design and Learner Experience Description	
Description	This theme explores how features like interactivity, structure, and user control shaped learner experience.
Prefer Medimon	Valued the game's interactive nature and familiarity with game mechanics. Highlighted active engagement.
Prefer Trad	Appreciated the control and structure textbooks provided. Emphasized ease of navigation and review.
No Preference	Criticized the game's lack of structure and goals. Wanted more learning guidance and incentives. Still, some praised its accessibility.
Theme 4: Cognitive Factors and Resource Management Description	
Description	This theme involves perceived mental effort and efficiency. It covers how learners managed cognitive demands and time investment.
Prefer Medimon	Some felt the game reduced cognitive load, making complex ideas easier to process through engaging formats.
Prefer Trad	Prioritized time efficiency and directness of textbooks for faster learning.
No Preference	Found the game cognitively demanding without clear structure. Highlighted the need for active attention to extract learning.

3.3. STEM Career Interest

To measure the game's potential impact on career interests toward pursuing a career in a STEM field, participants rated their desire to pursue a STEM career before and after gameplay. Most participants showed no change in their interest in pursuing a STEM career over the two-week period. However, five students reported increased interest in STEM careers, while six reported decreased interest (Figure 8). These results suggest that while the demo had a positive educational impact, its effect on career intentions was limited during the short intervention window. It should also be noted that this data reflects short-term interest, not career commitment.

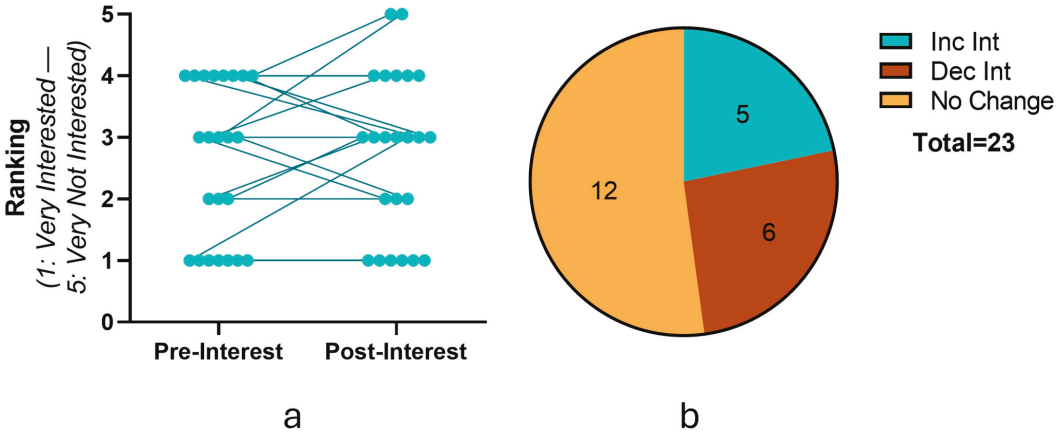


Figure 8. STEM career interest results. (**a**) Individual changes in pursuing a career in a STEM field before (pre-interest) and after (post-interest) playing the game. (**b**) Distribution of responses that either increased interest, decreased interest, or had no change in interest in pursuing a STEM field career. Inc Int: Increase Interest, Dec Int: Decrease Interest.

4. Discussion

The Medimon NASA Demo was designed to merge GBL with mnemonic strategies to address two persistent challenges in STEM and health science education: cognitive overload [1,2] and student disengagement [42]. Our preliminary findings suggest that this approach can improve student knowledge in key biomedical areas while maintaining high levels of learner engagement in an SRL environment.

Participants demonstrated significant learning gains, as evidenced by a 67% average improvement in posttest scores and a normalized change score of c = 0.4. This supports other studies also showing increased efficacy of serious games to promote learning and retention [43]. These results are particularly notable given the brief, two-week exposure to the intervention and the lack of structured instructional support, allowing for higher SRL [27,28]. The results affirm that the game's specific mechanisms, including its interactive nature and mnemonic-rich design (key elements highlighted in Theme 2: Perceived Learning Efficacy and Mechanisms), can serve as effective learning supports even without traditional scaffolding. The finding lends further support to multimedia-learning theory, which posits that interactive, learner-controlled exploration can substitute for external scaffolds when instructional materials are well-designed [44]. In particular, Medimon's mnemonic-rich visuals and creature narratives harness dual-coding and imagery principles, mechanisms repeatedly shown to strengthen recall and transfer in biomedical domains [29]. Taken together, the results indicate that Medimon can function as a stand-alone learning tool, a feature that is especially attractive for content areas such as space medicine where traditional classroom resources are limited.

The SIS-M results further reinforce the educational potential of Medimon. Participants reported high levels of triggered situational interest and maintained feeling. These findings align with previous studies in medical and STEM education that have shown serious games can foster motivation and sustained engagement [45]. These quantitative findings align strongly with Theme 1 (Intrinsic Appeal and Affective Engagement) identified in the qualitative analysis, where participants emphasized the game's fun, engaging nature, and esthetic appeal. Importantly, qualitative feedback highlighted that the integration of visual and linguistic mnemonics helped students grasp and retain difficult concepts, such as the function of osteoclasts in bone resorption or the etiology of conditions like retinitis pigmentosa and myasthenia gravis. This speaks directly to Theme 2 (Perceived Learning Efficacy and Mechanisms), as participants explicitly valued how mnemonics, interactivity, and the integration of content into gameplay made learning easier and more memorable. The presence of spatially located mnemonic-rich environments and NPCs provided contextual reinforcement, turning complex biological and pathological processes into accessible and memorable experiences [46]. Participants reported more moderate levels of maintained interest and value-related interest. This outcome is consistent with the participant demographics, as the majority of students were not enrolled in STEM-related majors and, therefore, may not have perceived a health science educational game as highly relevant or essential to their academic studies.

Interestingly, while the game succeeded in improving knowledge and engagement, it had minimal impact on students' STEM career aspirations. Only five participants reported increased interest in pursuing a STEM field, while five others reported a decline. This suggests that while GBL tools like Medimon can effectively enhance content understanding and immediate engagement, influencing long-term identity formation and career trajectories may require more prolonged, repeated exposure or the integration of mentorship and narrative career exploration within the game environment [47–50]. Career interest is often shaped by a broader range of psychosocial factors, including role modeling [51,52], perceived self-efficacy [50,53], and exposure to professional pathways [54,55]. Thus, future iterations of Medimon could include embedded role-play elements, career profiles, or narrative arcs that highlight the real-world applications of biomedical knowledge in various STEM professions, particularly in space medicine.

Evidence suggests that stable personality traits may moderate both engagement with serious games and downstream educational outcomes. For example, individuals high in *Openness to Experience* tend to prefer novel, visually rich learning environments, whereas those high in *Conscientiousness* often exhibit greater persistence and goal-directed behavior—qualities that would translate into deeper interaction with game mechanics and higher knowledge gains [56,57]. Recent work using game-based assessment platforms further demonstrates that fine-grained Big Five facets can be reliably inferred from in-game behaviors such as exploration patterns, risk taking, and time-on-task [58]. Although personality was not measured in the present study, future iterations should incorporate brief, validated inventories (e.g., NEO-FFI) and log-data analytics to parse how individual differences shape situational interest, persistence, and ultimately academic achievement. Accounting for these moderating variables will clarify for whom—and under what dispositional conditions—Medimon-style interventions are most effective.

4.1. Limitations

Several limitations of the preliminary current study should be noted. First, the sample size was relatively small (n = 23) and limited to students from a single institution, which may affect the generalizability of the results. One potential contributor to the high participant dropout

rate may have been the complexity of the game access process. Participants were required to create an itch.io account, download the itch.io app, save the game to their account, and then install and play the game through the app. Future studies will explore more streamlined and user-friendly methods for distributing and accessing the game to reduce barriers to participation. Second, we did not track in-game behavior metrics such as time played, specific characters interacted with, or the frequency of battles completed. These metrics could provide a deeper understanding of how specific game mechanics influence learning and engagement. Third, the two-week duration may have been too short to observe meaningful changes in career interest. Fourth, the qualitative findings highlight specific design limitations perceived by some participants, particularly the lack of pedagogical structure, clear goals, and review mechanisms (Theme 3). These factors may have dampened learning outcomes or preference for students who thrive on more guided learning, suggesting the current prototype may be more appealing to intrinsically motivated or experienced gamers. Furthermore, concerns about learning depth and recall difficulty (Theme 2) identified by some participants suggest areas where the game's instructional design could be strengthened. Longitudinal studies are needed to assess how sustained interaction with GBL platforms like Medimon influences educational outcomes and career development over time.

Despite these limitations, this study provides compelling preliminary evidence that serious games integrating mnemonic strategies can enhance learning and engagement in undergraduate health science education. The thematic emphasis on musculoskeletal and visual systems within the context of space medicine also offers a novel interdisciplinary entry point for students who may not initially be drawn to STEM fields but are intrigued by the adventure and problem-solving aspects of space exploration. The self-paced nature of the demo aligns well with modern learner preferences [59,60] and demonstrates the value of autonomy and GBL feedback in promoting knowledge acquisition [61].

4.2. Future Directions

Future research should expand the scope of Medimon to include additional body systems and pathologies, incorporate real-time data analytics, and assess its impact across diverse learner populations. Future iterations of Medimon should directly address the design feedback revealed through the qualitative analysis (Theme 3). This includes incorporating clearer learning objectives, progress tracking, feedback mechanisms, structured challenges or 'learning checks', and potentially dedicated review modes to enhance learner control and address recall concerns (Theme 2). These would incorporate problem-based tasks with appropriate feedback to promote extended learning. There is also potential to explore collaborative modes of gameplay and integration into flipped or hybrid classroom models. As space agencies and health care systems prepare the next generation of scientists and clinicians, tools like Medimon offer a scalable, engaging, and pedagogically sound platform to inspire and educate tomorrow's innovators. Finally, future work could explicitly embed career exploration elements or narratives to potentially influence STEM interest more effectively.

5. Conclusions

Within the limits of a small convenience sample, these findings provide preliminary evidence that the Medimon NASA Demo represents a promising fusion of game-based learning and mnemonic strategies tailored to the unique needs of STEM and health science education. By embedding visually and linguistically rich mnemonics into an engaging, interactive gaming experience, the prototype succeeded in increasing student knowledge of the musculoskeletal and visual systems—two critical domains in space medicine. Partic-

ipants not only demonstrated significant learning gains but also reported high levels of engagement and enjoyment in the game.

While the demo did not lead to a marked increase in STEM career interest over the short duration of the preliminary study, the strong indicators of cognitive and emotional engagement suggest that platforms like Medimon can serve as effective entry points into complex biomedical content. These results support the growing body of evidence that serious games, when thoughtfully designed, can bridge the gap between passive content delivery and active, meaningful learning.

Future work should focus on expanding content coverage, enhancing career exploration elements, integrating in-game analytics, and testing the platform across diverse learner populations and institutional contexts. With further development and longitudinal evaluation, Medimon has the potential to play a transformative role in health science education, helping cultivate a more scientifically literate and inspired generation of learners—both on Earth and beyond.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/mti9080080/s1, Methods S1: Agentic workflow, prompts, and outputs of the thematic analysis.

Author Contributions: Conceptualization, M.H. and T.B.; data curation, T.B.; formal analysis, T.B.; investigation, M.H. and T.B.; methodology, M.H., J.H. and T.B.; visualization, J.H., C.B., E.F., J.C. and T.B.; writing—original draft, T.B.; writing—review and editing, T.S. and T.B.; funding acquisition, T.S. and T.B. All authors have read and agreed to the published version of the manuscript.

Funding: Support for this research was provided by the NASA Idaho Space Grant Consortium, a NASA funded program under Federal Award 80NSSC20M0108. The content of this manuscript is solely the responsibility of the authors and does not necessarily represent the official views of the funding agencies. Publication of this article was funded by the University of Idaho—Open Access Publishing Fund.

Institutional Review Board Statement: This study was approved as exempt by the Institutional Review Board of the University of Idaho (23-042) on 15 February 2023.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The datasets presented in this article are not publicly available due to the potential risk of re-identification of individual participants. Given the sensitivity of the data and the possibility of linking responses to specific students, requests to access the datasets should be directed to Tyler Bland (tbland@uidaho.edu).

Acknowledgments: We would like to thank all the University of Idaho undergraduate students who participated in this study for their time and feedback for enhancing the Medimon platform.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

GBL	Game-Based Learning
LLM	Large Language Model
SANS	Spaceflight-Associated Neuro-Ocular Syndrome
SIS-M	Situational Interest Survey in Multimedia
SRL	Self-Regulated Learning
STEM	Science, Technology, Engineering, and Mathematics
genAI	Generative Artificial Intelligence

Appendix A

Table A1. Families and characters in the game. 1 = stage 1, 2 = stage 2, 3 = stage 3.

Family	Healthy Characters	Diseased Characters
Bone	1. Osteoclast	1. Osteoporosis
	2. Osteoblast	2. Postmenopausal Osteoporosis
	3. Skeleton	3. Paget Disease
Muscle	1. Myofibril	1. Myasthenia Gravis
	2. Muscle Fiber	2. Epilepsy
	3. Muscle	3. Rhabdomyolysis
Eye	1. Rod	1. Retinitis Pigmentosa
	2. Cone	2. Color Blindness
	3. Eye	3. Glaucoma

Table A2. SIS items.

SIS Type	Survey Item
SI-triggered	The Medimon NASA demo was interesting.
	The Medimon NASA demo grabbed my attention.
	The Medimon NASA demo was often entertaining.
	The Medimon NASA demo was so exciting, it was easy to pay attention.
SI-maintained-feeling	What I learned from the Medimon NASA demo is fascinating to me.
	I am excited about what I learned from the Medimon NASA demo.
	I like what I learned from the Medimon NASA demo.
SI-maintained-value	I found the information from the Medimon NASA demo interesting.
	What I studied in the Medimon NASA demo is useful for me to know.
	The things I studied in the Medimon NASA demo are important to me.
	What I learned from the Medimon NASA demo can be applied to my major/career.
	I learned valuable things from the Medimon NASA demo.

References

1. Van Merriënboer, J.J.G.; Sweller, J. Cognitive load theory in health professional education: Design principles and strategies. *Med. Educ.* **2010**, *44*, 85–93. [CrossRef]
2. Young, J.Q.; Van Merrienboer, J.; Durning, S.; Cate, O.T. Cognitive Load Theory: Implications for medical education: AMEE Guide No. 86. *Med. Teach.* **2014**, *36*, 371–384. [CrossRef]
3. Sweller, J. Cognitive load theory, learning difficulty, and instructional design. *Learn. Instr.* **1994**, *4*, 295–312. [CrossRef]
4. Teschers, C.; Neuhaus, T.; Vogt, M. Troubling the boundaries of traditional schooling for a rapidly changing future–Looking back and looking forward. *Educ. Philos. Theory* **2024**, *56*, 873–884. [CrossRef]
5. Kuh, G.D.; Schneider, C.G. *High-Impact Educational Practices: What They Are, Who Has Access to Them, and Why They Matter*; Assn of Amer Colleges: Washington, DC, USA, 2008; p. 35.
6. Stains, M.; Harshman, J.; Barker, M.K.; Chasteen, S.V.; Cole, R.; DeChenne-Peters, S.E.; Eagan, M.K., Jr.; Esson, J.M.; Knight, J.M.; Laski, F.A.; et al. Anatomy of STEM teaching in North American universities. *Science* **2018**, *359*, 1468–1470. [CrossRef] [PubMed]
7. Freeman, S.; Eddy, S.L.; McDonough, M.; Smith, M.K.; Okoroafor, O.N.; Jordt, H.; Wenderoth, M.P. Active learning increases student performance in science, engineering, and mathematics. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 8410–8415. [CrossRef]
8. Games2train, M.P.; York, N. Digital game-based learning. *Comput. Entertain. (CIE)* **2003**, *1*, 21.
9. Edwards, S.L.; Gantwerker, E.; Cosimini, M.; Christy, A.L.; Kaur, A.W.; Helms, A.M.; Stiver, A.K.; Mikaela, L.; Zachary, L. Game-Based Learning in Neuroscience. *Neurol. Educ.* **2023**, *2*, e200103. [CrossRef]

10. Hamari, J.; Shernoff, D.J.; Rowe, E.; Coller, B.; Asbell-Clarke, J.; Edwards, T. Challenging games help students learn: An empirical study on engagement, flow and immersion in game-based learning. *Comput. Human Behav.* **2016**, *54*, 170–179. [CrossRef]
11. Nadeem, M.; Oroszlanyova, M.; Farag, W. Effect of Digital Game-Based Learning on Student Engagement and Motivation. *Computers* **2023**, *12*, 177. [CrossRef]
12. Xu, M.; Luo, Y.; Zhang, Y.; Xia, R.; Qian, H.; Zou, X. Game-based learning in medical education. *Front. Public Health* **2023**, *11*, 1113682. [CrossRef]
13. Graafland, M.; Schraagen, J.M.; Schijven, M.P. Systematic review of serious games for medical education and surgical skills training. *Br. J. Surg.* **2012**, *99*, 1322–1330. [CrossRef] [PubMed]
14. Cosimini, M.; Ochoa, D.M.; Yu, D.; Chiang, A. Empiric: A Card Game for Guideline-Based Antibiotic Prescribing Used for Continuing Medical Education. *MedEdPORTAL* **2025**, *21*, 11533. [CrossRef] [PubMed]
15. Tsopra, R.; Courtine, M.; Sedki, K.; Eap, D.; Cabal, M.; Cohen, S.; Bouchaud, O.; Mechaï, F.; Lamy, J.-B. AntibioGame®: A serious game for teaching medical students about antibiotic use. *Int. J. Med. Inform.* **2020**, *136*, 104074. [CrossRef]
16. Putnam, A.L. Mnemonics in education: Current research and applications. *Transl. Issues Psychol. Sci.* **2015**, *1*, 130–139. [CrossRef]
17. Akpan, J.; Notar, C.E.; Beard, L. The Impact of Mnemonics as Instructional Tool. *J. Educ. Hum. Dev.* **2021**, *10*, 20–28. [CrossRef]
18. Lewis, J.B.; Mulligan, R.; Kraus, N. The Importance of Medical Mnemonics in Medicine. *Pharos* **2018**, *2018*, 30–35.
19. Assistant Professor, M.S.; Mary Zachariah Associate Professor, A.; Balakrishnan Assistant Professor, S. Effectiveness of mnemonics based teaching in medical education. *Int. J. Health Sci.* **2022**, *6*, 9635–9640.
20. Olu-Ajayi, F.E. The Upshot of Mnemonics on Gender and Other Learning Outcomes of Senior Secondary School Students in Biology. Available online: https://www.eajournals.org/ (accessed on 5 April 2023).
21. Al-Maqbali, F.; Ambusaidi, A.; Shahat, M.A.; Alkharusi, H. The effect of teaching science based on mnemonics in reducing the sixth-grade female students' cognitive load according to their imagery style. *J. Posit. Psychol.* **2022**, *6*, 2069–2084.
22. Yang, A.; Goel, H.; Bryan, M.; Robertson, R.; Lim, J.; Islam, S.; Speicher, M.R. The picmonic® learning system: Enhancing memory retention of medical sciences, using an audiovisual mnemonic web-based learning platform. *Adv. Med. Educ. Pract.* **2014**, *5*, 125–132.
23. Sketchy-Learning Made Unforgettable. Available online: https://www.sketchy.com/ (accessed on 19 June 2025).
24. Bland, T.; Guo, M. Visual Mnemonics and Gamification: A New Approach to Teaching Muscle Physiology. *J. Technol.-Integr. Lessons Teach.* **2024**, *3*, 73–82. [CrossRef]
25. Wolfe, J.W.; Rummel, J.D. Long-term effects of microgravity and possible countermeasures. *Adv. Space Res.* **1992**, *12*, 281–284. [CrossRef] [PubMed]
26. Lee, A.G.; Mader, T.H.; Gibson, C.R.; Tarver, W.; Rabiei, P.; Riascos, R.F.; Galdamez, L.A.; Brunstetter, T. Spaceflight associated neuro-ocular syndrome (SANS) and the neuro-ophthalmologic effects of microgravity: A review and an update. *Npj Microgravity* **2020**, *6*, 7. [CrossRef]
27. Winne, P.H.; Azevedo, R. Metacognition and Self-Regulated Learning. *Camb. Handb. Learn. Sci.* **2022**, *10*, 93–113.
28. Zheng, B.; Sun, T. Self-Regulated Learning and Learning Outcomes in Undergraduate and Graduate Medical Education: A Meta-Analysis. *Eval Health Prof.* **2024**, *3*, 01632787241288849. [CrossRef]
29. Clark, J.M.; Paivio, A. Dual coding theory and education. *Educ. Psychol. Rev.* **1991**, *3*, 149–210. [CrossRef]
30. Dousay, T.A. Effects of redundancy and modality on the situational interest of adult learners in multimedia learning. *Educ. Technol. Res. Dev.* **2016**, *64*, 1251–1271. [CrossRef]
31. Dousay, T.A.; Trujillo, N.P. An examination of gender and situational interest in multimedia learning environments. *Br. J. Educ. Technol.* **2019**, *50*, 876–887. [CrossRef]
32. Bland, T. Enhancing Medical Student Engagement Through Cinematic Clinical Narratives: Multimodal Generative AI–Based Mixed Methods Study. *JMIR Med. Educ.* **2025**, *11*, e63865. [CrossRef]
33. Bland, T.; Guo, M.; Dousay, T.A. Multimedia design for learner interest and achievement: A visual guide to pharmacology. *BMC Med. Educ.* **2024**, *24*, 113. [CrossRef]
34. Worthley, B.; Guo, M.; Sheneman, L.; Bland, T. Antiparasitic Pharmacology Goes to the Movies: Leveraging Generative AI to Create Educational Short Films. *AI* **2025**, *6*, 60. [CrossRef]
35. Alomar, Z.; Guo, M.; Bland, T. AI-Generated Mnemonic Images Improve Long-Term Retention of Coronary Artery Occlusions in STEMI: A Comparative Study. *Technologies* **2025**, *13*, 217. [CrossRef]
36. Singleton, V.; Bordeaux, C.; Ferguson, E.; Bland, T. An Educational Trading Card Game for a Medical Immunology Course. *Educ. Sci.* **2025**, *15*, 768. [CrossRef]
37. Marx, J.D.; Cummings, K.; Phys, A.J. Normalized change. *Am. J. Phys.* **2007**, *75*, 87–91. [CrossRef]

38. Tian, J.; Wang, L.; Fard, P.; Junior, V.M.; Blacker, D.; Haas, J.S.; Patel, C.; Murphy, S.N.; Moura, L.M.V.R.; Estiri, H. An Agentic AI Workflow for Detecting Cognitive Concerns in Real-World Data. Available online: https://arxiv.org/abs/2502.01789v1 (accessed on 9 April 2025).
39. Li, X. A Review of Prominent Paradigms for LLM-Based Agents: Tool Use (Including RAG), Planning, and Feedback Learning. Available online: https://arxiv.org/abs/2406.05804 (accessed on 9 April 2025).
40. Wang, J.; Liu, Z.; Zhao, L.; Wu, Z.; Ma, C.; Yu, S.; Dai, H.; Yang, Q.; Liu, Y.; Zhang, S.; et al. Review of Large Vision Models and Visual Prompt Engineering. *Meta-Radiol.* **2023**, *1*, 100047. [CrossRef]
41. White, J.; Fu, Q.; Hays, S.; Sandborn, M.; Olea, C.; Gilbert, H.; Elnashar, A.; Spencer-Smith, J.; Schmidt, D.C. A Prompt Pattern Catalog to Enhance Prompt Engineering with ChatGPT. Available online: https://arxiv.org/abs/2302.11382v1 (accessed on 23 September 2024).
42. Schnitzler, K.; Holzberger, D.; Seidel, T. All better than being disengaged: Student engagement patterns and their relations to academic self-concept and achievement. *Eur. J. Psychol. Educ.* **2021**, *36*, 627–652. [CrossRef]
43. Haoran, G.; Bazakidi, E.; Zary, N. Serious Games in Health Professions Education: Review of Trends and Learning Efficacy. *Yearb. Med. Inform.* **2019**, *28*, 240. [CrossRef]
44. Mayer, R.E. Cognitive Theory of Multimedia Learning. *Camb. Handb. Multimed. Learn.* **2012**, *41*, 31–48.
45. Watsjold, B.K.; Cosimini, M.; Mui, P.; Chan, T.M. Much ado about gaming: An educator's guide to serious games and gamification in medical education. *AEM Educ. Train.* **2022**, *6*, e10794. [CrossRef]
46. Qureshi, A.; Rizvi, F.; Syed, A.; Shahid, A.; Manzoor, H. The method of loci as a mnemonic device to facilitate learning in endocrinology leads to improvement in student performance as measured by assessments. *Adv. Physiol. Educ.* **2014**, *38*, 140. [CrossRef]
47. Eccles, J. Who Am I and What Am I Going to Do with My Life? Personal and Collective Identities as Motivators of Action. *Educ. Psychol.* **2009**, *44*, 78–89. [CrossRef]
48. Gee, J.P. What video games have to teach us about learning and literacy. *Comput. Entertain. (CIE)* **2003**, *1*, 20. [CrossRef]
49. Shaffer, D.W.; Gee, J.P. *How Computer Games Help Children Learn*; Palgrave Macmillan: London, UK, 2007; pp. 1–242. [CrossRef]
50. Lent, R.W.; Brown, S.D.; Hackett, G. Toward a Unifying Social Cognitive Theory of Career and Academic Interest, Choice, and Performance. *J. Vocat. Behav.* **1994**, *45*, 79–122. [CrossRef]
51. Zhang, H.; Couch, S.; Estabrooks, L.; Perry, A.; Kalainoff, M.Z. Role models' influence on student interest in and awareness of career opportunities in life sciences. *Int. J. Sci. Educ. Part B* **2023**, *13*, 381–399. [CrossRef]
52. Gibson, D.E. Role models in career development: New directions for theory and research. *J. Vocat. Behav.* **2004**, *65*, 134–156. [CrossRef]
53. Pham, M.; Lam, B.Q.; Tuan Ngoc Bui, A. Career exploration and its influence on the relationship between self-efficacy and career choice: The moderating role of social support. *Heliyon* **2024**, *10*, e31808. [CrossRef] [PubMed]
54. Mathur, A.; Chow, C.S.; Feig, A.L.; Kenaga, H.; Moldenhauer, J.A.; Muthunayake, N.S.; Ouellett, M.L.; Pence, L.E.; Straub, V. Exposure to multiple career pathways by biomedical doctoral students at a public research university. *PLoS ONE* **2018**, *13*, e0199720. [CrossRef]
55. Blank, R.; Daniels, R.J.; Gilliland, G.; Gutmann, A.; Hawgood, S.; Hrabowski, F.A.; Pollack, M.E.; Price, V.; Reif, L.R.; Schlissel, M.S. A new data effort to inform career choices in biomedicine. *Science* **2017**, *358*, 1388–1389. [CrossRef]
56. Costa, P.T.; McCrae, R.R. The revised NEO personality inventory (NEO-PI-R). In *The SAGE Handbook of Personality Theory and Assessment: Volume 2—Personality Measurement and Testing*; Sage Publications, Inc.: Thousand Oaks, CA, USA, 2008; pp. 179–198. [CrossRef]
57. Costa, P.T.; Mccrae, R.R.; Kay, G.G. Persons, Places, and Personality: Career Assessment Using the Revised NEO Personality Inventory. *J. Career Assess.* **1995**, *3*, 123–139. [CrossRef]
58. Wu, F.Y.; Mulfinger, E.; Alexander, L.; Sinclair, A.L.; McCloy, R.A.; Oswald, F.L. Individual differences at play: An investigation into measuring Big Five personality facets with game-based assessments. *Int. J. Sel. Assess.* **2022**, *30*, 62–81. [CrossRef]
59. Tullis, J.G.; Benjamin, A.S. On the effectiveness of self-paced learning. *J. Mem. Lang.* **2011**, *64*, 109. [CrossRef]
60. Rhode, J.F. Interaction Equivalency in Self-Paced Online Learning Environments: An Exploration of Learner Preferences. *Int. Rev. Res. Open Distrib. Learn.* **2009**, *10*, 1–23. [CrossRef]
61. Tavares, N. The use and impact of game-based learning on the learning experience and knowledge retention of nursing undergraduate students: A systematic literature review. *Nurse Educ. Today* **2022**, *117*, 105484. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

Multimodal Technologies and Interaction

MDPI

Article

Evaluating Educational Game Design Through Human–Machine Pair Inspection: Case Studies in Adaptive Learning Environments

Ioannis Sarlis [1], Dimitrios Kotsifakos [2,*] and Christos Douligeris [1,*]

[1] Department of Informatics, University of Piraeus, 18534 Piraeus, Greece; sarlisj@unipi.gr
[2] Department of Business Administration, University of West Attica, 12241 Athens, Greece
* Correspondence: dkotsifakos@uniwa.gr (D.K.); cdoulig@unipi.gr (C.D.); Tel.: +30-6938416632 (D.K.)

Abstract: Educational games often fail to effectively merge game mechanics with educational goals, lacking adaptive feedback and real-time performance monitoring. This study explores how Human–Computer Interaction principles and adaptive feedback can enhance educational game design to improve learning outcomes and user experience. Four educational games were analyzed using a mixed-methods approach and evaluated through established frameworks, such as the Serious Educational Games Evaluation Framework, the Assessment of Learning and Motivation Software, the Learning Object Evaluation Scale for Students, and Universal Design for Learning guidelines. In addition, a novel Human–Machine Pair Inspection protocol was employed to gather real-time data on adaptive feedback, cognitive load, and interactive behavior. Findings suggest that Human–Machine Pair Inspection-based adaptive mechanisms significantly boost personalized learning, knowledge retention, and student motivation by better aligning games with learning objectives. Although the sample size is small, this research provides practical insights for educators and designers, highlighting the effectiveness of adaptive Game-Based Learning. The study proposes the Human–Machine Pair Inspection methodology as a valuable tool for creating educational games that successfully balance user experience with learning goals, warranting further empirical validation with larger groups.

Keywords: game-based learning effectiveness; human–computer interaction; interaction design in educational software; adaptive personalized feedback mechanisms; learner experience; student engagement metrics; human–machine pair inspection; educational game evaluation

1. Introduction

Educational games are powerful tools for boosting student engagement and improving learning outcomes. They play a vital role in transforming traditional learning environments, turning static classrooms into dynamic, interactive spaces [1]. By blending educational content with captivating gameplay, these tools harness interactivity to spark curiosity, keep students engaged, and promote deeper learning. As a result, Game-Based Learning (GBL) has grown in popularity and effectiveness in recent years [2]. GBL provides students with opportunities to tackle complex and abstract concepts through hands-on, experiential activities and interactive engagement. Techniques borrowed from video games—such as achievement badges, leaderboards, and contextual guidance—are essential for maintaining motivation [3] and for creating environments where effort is visibly rewarded [4]. Engaging

in tasks like solving challenging problems (e.g., decoding) or breaking down complex, multi-level problems into simpler ones helps students solidify concepts and reinforce learning objectives, often within an enjoyable context [5].

Although the terms Game-Based Learning (GBL) and gamification are often used interchangeably, the literature draws a clear distinction: GBL involves the use of complete games or serious games where rules, narrative, and mechanics are intrinsically aligned with learning objectives to create immersive, contextualized experiences, whereas gamification applies individual game design elements—such as points, badges, leaderboards, and progress bars—to non-game contexts to boost motivation without delivering the holistic experience of a game. Game-Based Learning is widely recognized as an effective strategy to foster motivation and engagement [6], particularly when embedded within constructivist and socio-cultural paradigms [7]. This distinction is grounded in Deterding et al. [8], who define gamification as the application of "gameful" elements in non-game settings, while critics like Bogost [9] warn that superficial gamification can result in shallow engagement when not paired with meaningful content or narrative. Effective GBL design, therefore, requires embedding learning deeply within game mechanics rather than relying solely on extrinsic reward structures. Additionally, contemporary Science, Technology, Engineering, and Mathematics (STEM) education has increasingly embraced the broader STEAM (Science, Technology, Engineering, Arts, and Mathematics) approach by incorporating the arts to foster creativity, design thinking, and human-centered innovation. Integrating the arts amplifies the capacity of scientific and technical learning experiences to produce multidimensional, empathetic solutions, and recent work shows that STEAM promotes engagement, interdisciplinary understanding, and creative problem-solving in Game-Based Learning environments [10].

However, the success of educational games is not solely dependent on students' motivation to engage. The ongoing interaction between the player and the game system plays a significant role, particularly when it is supported by timely and adaptive feedback mechanisms [11]. Such feedback not only sustains engagement but also supports learning by helping learners monitor and regulate their progress [12]. In this regard, principles from Human–Computer Interaction (HCI) is highly relevant, as it emphasizes usability, personalization, and the optimization of user experience in interactive systems [13]. Recent advances in adaptive gamification further highlight the importance of tailoring in-game responses to individual learners through dynamic feedback loops and user modeling. Core HCI principles include usability, consistency, feedback responsiveness, error prevention, and personalization—each directly influencing educational game engagement and effectiveness [14]. Games need to react dynamically to what the player does. Features like consistent real-time feedback challenge that scale appropriately, and smart progress tracking works with such features to create a truly personal learning journey [15]. This journey adjusts based on the student's performance, maintaining an optimal balance between challenge and engagement. Such adaptive systems do not just pull players deeper into the game world (enhancing immersion); they actively help players acquire knowledge and skills by ensuring the game mechanics respect cognitive load principles and align with effective problem-solving approaches [16].

Knowing whether these games work is, naturally, a major focus in education today. Rigorous evaluation is a key, not just to seeing immediate results, but also to understanding the long-term effects these games have on learning [17]. Structured approaches to development, like the traditional waterfall model (Figure 1), enrich quality control from start to finish. Specific evaluation frameworks, such as the Assessment of Learning and Motivation Software (ALMAS), the Learning Object Evaluation Scale for Students (LOES-S), the Seri-

ous Educational Games Evaluation Framework (SELEAG), the Educational Context-Based Assessment (ECBA), the GameFlow model, Universal Design for Learning (UDL), and the Model for the Evaluation of Educational Games (MEEGA+), provide tools to examine usability, player motivation, and learning effectiveness, checking that games achieve their objectives both technically and educationally [18]. These evaluation approaches often highlight the importance of easy-to-use interfaces, engaging narratives, and smart adaptive mechanics in keeping players engaged and improving learning retention. Going a step further, methods focusing on player interaction allow for collecting real-time data about user engagement and thinking processes. This information can then feed back into refining the game's design and tailoring the educational content dynamically. The precise role of HCI principles is to apply them in a way that ensures educational games are flexible, truly responsive to individual learners, and welcoming to all (inclusive); this ultimately serves to bridge the gap between engaging entertainment and effective, structured learning [19].

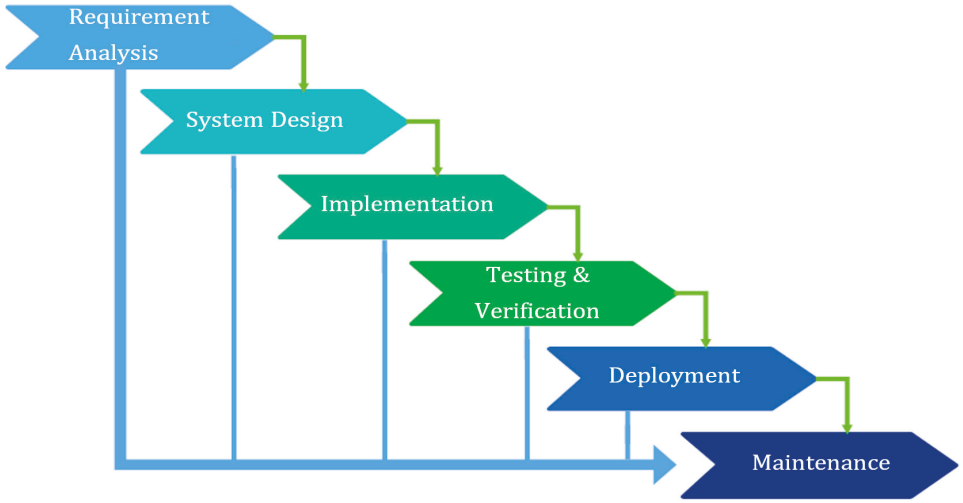


Figure 1. The phases of educational software development according to the waterfall model.

Building on these foundations, this study aims to answer the following research questions:

1. How can game design principles and interactive mechanisms be effectively aligned with educational objectives to enhance student engagement and learning outcomes?
2. What are the advantages and limitations of Human–Machine Pair Inspection (HMPI) in evaluating educational games, and how do these affect learning outcomes and player engagement?
3. In what ways do different evaluation frameworks—such as ALMAS, ECBA, GameFlow, and SELEAG—differ in how they measure the effectiveness of educational games in terms of player motivation, cognitive skill development, and knowledge retention?

To explore these questions, this paper delves into how specific game elements—challenges, puzzles, engaging narratives—can foster crucial skills like critical thinking, teamwork, and problem-solving. The paper also examines the potential of newer technologies like micro:bit, Arduino, and Virtual/Augmented Reality (VR/AR) to create learning environments that are both immersive ('experiential') and safe places to experiment.

1.1. Game Design Principles

Grounded in research, educational game design balances engagement and learning value by leveraging STEAM to combine analytical rigor with creativity [20]. Games built around adventure or treasure hunts, for example, can create particularly immersive and

meaningful learning journeys [21]. Crafting these games requires carefully linking game mechanics to specific learning goals while encouraging active participation through challenges that push for deeper understanding. Storylines are vital, too; they provide context that boosts engagement and ties game activities to believable situations. Well-designed games also mirror effective real-world teaching, weaving in practical tasks and interactive methods to boost learning [22]. Logically, evaluation must ascertain the actual impact of these games, confirming that design elements yield meaningful improvements in learning. As exemplified in Section 1.7, the case studies instantiate these STEAM-based, transdisciplinary approaches through narrative, puzzles, and creative challenges.

Within educational games, "Game Design Principles" refers to the core strategies for weaving learning content and teaching approaches into the gameplay itself. Key principles of game design include:

- Alignment of Mechanics with Learning Objectives: Ensuring that gameplay directly reinforces the intended learning outcomes.
- Embedded Assessment: Integrating evaluation of player learning within the game experience.
- Adaptive Difficulty: Adjusting the challenge level based on individual player performance.
- Meaningful Feedback: Providing feedback that is both informative and relevant to the learning goals.
- Motivation and Engagement: Designing gameplay that sustains player interest and encourages continued learning.
- Interdisciplinary Integration: Combining scientific, technical, and artistic perspectives to foster both analytical and creative learning.

1.2. Aligning Game Mechanics with Learning Objectives

A core challenge in educational game design is ensuring every game feature serves the learning objectives. Research backs this up, showing a strong link between well-integrated educational content and engaging mechanics, which ultimately leads to better learning results [23]. Studies also confirm that weaving learning directly into the gameplay helps students acquire and remember knowledge more effectively [24]. Game challenges, mechanics, and rewards should be designed to build skills like problem-solving, critical thinking, and teamwork [25]. For instance, math games might use logic puzzles to sharpen analytical abilities. Clear in-game instructions are also important, so players understand the goals and how to play, letting them focus on the learning itself.

While educational game design is often approached from an instructional or cognitive lens, perspectives from game studies enrich our understanding of how rules, mechanics, and narrative framing shape player engagement and meaning-making. For instance, Bogost [9], Juul [26], and Salen and Zimmerman [27] highlight that games function as semiotic systems—where rules do not merely regulate but express pedagogical intent. Integrating this perspective supports the design of mechanics that not only challenge the player cognitively but also position the learner within an expressive narrative world with educational consequences.

Information and Communication Technologies (ICT) support the creation of engaging adventure and treasure-hunt games. These tools help to build interactive, immersive worlds where learners actively participate and gain hands-on experience. Examples include

- Virtual Reality (VR): Immerses students in simulated scenarios, like navigating a rainforest, designing virtual prototypes, or conducting virtual experiments safely [28].
- Augmented Reality (AR) Games: Enhances physical spaces with digital overlays [29], allowing activities like identifying virtual plants on a nature walk or solving puzzles using virtual objects in the real world [30].

- Programming Platforms (Hardware): Students can combine platforms like Micro:bit and Arduino for complex game setups (e.g., escape rooms with physical locks controlled by Arduino, managed by Micro:bit). In a treasure hunt, devices might communicate to navigate a robot or trigger effects upon finding solutions [31].
- Gamification Platforms: Tools like *Minecraft Education Edition* allow exploration through interactive environments, perhaps tasking learners to build functional systems to understand processes [32].
- Environmental Puzzles: Can require applying understanding of systems, like simulating food webs, modeling environmental impacts, or using physics to build structures in a virtual treasure hunt. The game *The Legend of the Dead Knight's Spirit* provides a prime example: players needed to apply the Pythagorean Theorem to calculate the rope length required to reach a tower, using the moat width and tower height.

Elements like these are common in immersive adventure and treasure-hunt formats By using these technologies well, educators can create rich worlds where learning happens through exploration and tackling problems. Location-based AR games, for instance, have successfully boosted student motivation and spatial skills by letting learners interact with virtual elements in their actual surroundings.

1.3. Designing Interactive Challenges for Deeper Learning

Captivating puzzles and challenges are the heart of successful adventure and treasure-hunt games, demanding careful design to ensure real educational value. Research confirms that interactive simulations, for example, significantly boost student engagement and deepen understanding [33]. So, when these challenges are carefully crafted to promote deeper learning, they typically incorporate several essential elements, such as the following:

- **Immediate and constructive feedback:** Providing players with real-time guidance that identifies their weaknesses and areas needing further attention.
- **Scenario-based problems:** Using realistic dilemmas to spur critical thinking and problem-solving [34].
- **Collaborative tasks:** Encouraging teamwork and communication through shared goals, which enhances problem-solving skills.
- **Adaptive difficulty:** Adjusting challenge levels to match individual players' progress and learning style, keeping motivation high [35].

There are games like *Super Turtle Adventures*, which introduces players to Logo programming [36], or *Musical Mastermind*, which cleverly combines music theory with computational thinking [37], clearly demonstrating how complex topics can be integrated into an interactive game and transformed into an enjoyable learning experience. Similarly, the *Digital Little Music Makers* project embodies constructionist principles of "making" to teach music composition, further proving the power of creative, hands-on learning in digital environments [38].

The educational value of these games is further enhanced by high-quality digital learning objects. When games include built-in tools that provide immediate feedback or present real-world problems (such as resource management or understanding recycling systems, as shown in Figure 2), learning becomes much more dynamic and meaningful to students. In *Super Turtle Adventures*, students use Logo to program a turtle robot to sort litter into recycling bins. Moreover, such learning objects are vital as they help students connect theoretical knowledge with practical application. For example, students could practice interpreting various types of data [39], conduct virtual experiments in chemistry or genetics [40]. Customizable environments also help cater to different learning styles.

By weaving these interactive challenges and objects into the game's narrative, educational content becomes a natural part of the experience. Mixing disciplines within puzzles, like combining physics and engineering, can also be very effective at maintaining learner interest [41].



Figure 2. Learning to code for a greener planet.

1.4. Crafting Compelling Narratives

A good story is often a key to keeping players engaged and giving context to the learning in educational games. Seamlessly integrating learning goals within the game's narrative tends to enhance player motivation for exploration and learning. For example, a game narrative centered on finding sustainable energy sources can teach environmental concepts within an exciting adventure. Drawing on game studies, narrative framing in educational games should be seen not just as context, but as a mechanism for shaping learner subjectivity and agency. The story becomes a framework through which players interpret challenges, inhabit roles, and internalize pedagogical meaning, making narratives essential components of game-based epistemology.

Effective storylines often use immersive themes and believable challenges, such as

- Sustainability Missions: A scientific expedition to discover renewable energy sources educates players about sustainability while entertaining them.
- Role-Playing Scenarios: Players take on various roles, immersing themselves in real-world challenges.

Narrative framing in educational games should be seen not just as context, but as a mechanism for shaping learner subjectivity and agency. The story becomes a framework through which players interpret challenges, inhabit roles, and internalize pedagogical meaning.

1.5. Research-Backed Design Strategies

Research points to several core strategies for creating impactful educational games:

- Alignment with Learning Objectives: Every game element should support measurable educational goals [42].
- Design and Aesthetics: Visual appeal and overall design significantly influence how usable and credible the game feels [43].

- Usability: The game must be easy to navigate and interact with.
- Engaging Gameplay: Variety in challenges, immediate feedback, and a sense of accomplishment are crucial for motivation [44].
- Appropriate Difficulty: Challenges should adapt to different learners to prevent frustration or boredom; balancing difficulty with adaptive feedback and real-time analytics deepens engagement [45].
- Constructivist/Problem-Based Learning: Games should encourage active learning through exploration, experimentation, and tackling problems.
- Accessibility: Games need to be usable by all learners, including those with disabilities, incorporating features like adjustable difficulty, alternative inputs, or text-to-speech options. Technological accessibility must also account for content formats such as video captions, audio descriptions, and compatibility with screen readers—especially in multimedia-rich educational environments [46].

1.6. Real-World Examples and Resources

Real-world examples highlight the power of these approaches in education:

- *DragonBox*: Uses engaging card games and puzzles to teach algebraic concepts [47].
- Labster: Offers interactive simulations of lab experiments across various science fields [48].
- Game Development Tools: Platforms like Scratch, Unity, and GameMaker Studio provide user-friendly interfaces and resources, empowering educators and students alike to create their own interactive learning experiences [49].

The principles and strategies discussed above—from aligning game mechanics with learning goals to crafting engaging challenges and narratives—provide a foundation for effective educational game design. To see how these concepts are translated into practice, this section examines four distinct case studies. These examples illustrate different approaches to leveraging gamification and digital tools across various disciplines, highlighting the real-world application and potential impact of thoughtfully designed educational games.

1.7. Case Studies and Applications

Presented here are four case studies of educational games combining gamification and digital technologies to support learning across subjects. The selected titles, (1) *Super Turtle Adventures*, (2) *Musical Mastermind*, (3) *The Legend of the Dead Knight's Spirit*, and (4) *Urban History of Drapetsona*, incorporate thematic or historical narratives alongside puzzle-solving and problem-based learning activities. All four games are rooted in STEAM approaches to foster creativity, cultural awareness, and deeper engagement.

Their thoughtfully developed scenarios are designed to engage learners while promoting both skill development and conceptual understanding. Through the integration of artistic, cultural, or musical dimensions into technical content, the games demonstrate how immersive mechanics can foster transdisciplinary learning. They are carefully aligned with specific educational goals, merging disciplines like programming, music theory, history, and cultural heritage into coherent and meaningful learning experiences.

1.7.1. *Super Turtle Adventures*

Super Turtle Adventures is a novice programming environment (NPE) designed to introduce students to the Logo programming language through block-based coding puzzles [50]. Beyond developing computational thinking skills, the game also aims to promote environmental awareness. Players guide a virtual turtle through a series of challenges, learning core programming concepts such as movement commands and loops.

In the game *Super Turtle Adventures*, developed by Sarlis et al. (2019) [50], one of the central tasks involves writing code to control a robot that assists a turtle in collecting and recycling waste. This coding component is seamlessly integrated into gameplay, encouraging students to use logic and planning to solve real-world inspired problems. As players proceed, they are required to compose increasingly complex instructions to achieve their recycling goals, making code-writing itself a meaningful part of the problem-solving experience. The game features a visual interface that supports learners as they progress from simple instructions to more advanced procedural programming. This design encourages the development of logical reasoning and creativity in a format that is both engaging and interactive (see Figure 3). In *Super Turtle Adventures*, students use block-based coding to guide a turtle through puzzles related to recycling. The effectiveness of the game was also assessed using the SELEAG (Serious Educational Games Evaluation Framework), focusing on aspects such as learning outcomes, player motivation, engagement, usability, and overall user experience.



Figure 3. The interface of *Super Turtle Adventures*.

According to the authors, the results of the evaluation pointed to significant gains in programming skills and environmental understanding (see Tables A1–A5 in Appendix A). This evaluation confirmed that *Super Turtle Adventures* effectively boosts students' grasp of programming concepts while keeping them motivated through interactive gameplay tied to real-world environmental themes. To ensure wide access, *Super Turtle Adventures* is available online where students can practice Logo programming anytime, tackling coding puzzles and honing problem-solving skills enjoyably.

1.7.2. *Musical Mastermind*

Musical Mastermind is an inventive educational tool blending music theory with computational thinking. It offers students an interactive way to decode and program musical sequences using micro:bit v2.0-v2.2 devices. Taking a STEAM-based approach, the game cultivates creativity, logical reasoning, and collaborative learning as students experiment with sound and LED interactions. By combining coding with music composition, *Musical Mastermind* sharpens both technical and artistic skills within a dynamic, engaging experience (Figure 4).

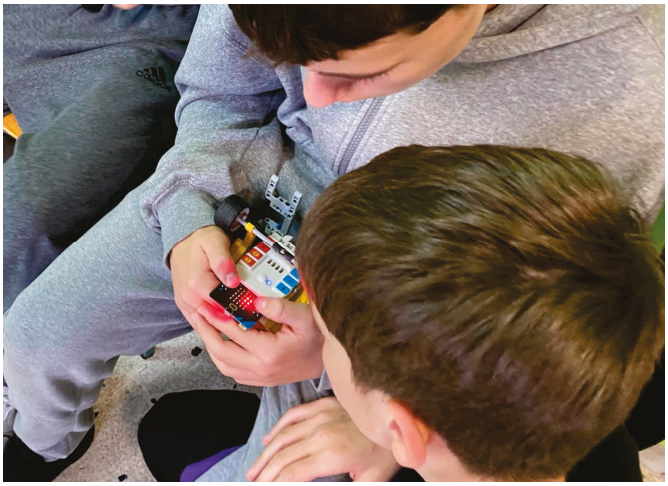


Figure 4. *Musical Mastermind:* Students program a micro:bit to decode the correct musical sequence.

The game's evaluation used the UDL framework, which focuses on accessibility, adaptability, and inclusivity to ensure it meets diverse learner needs. The results of a study by Sarlis et al. (2025) [37] showed that students acquired skills remarkably well, especially in problem-solving, pattern recognition, and computational thinking. A key factor was the high motivation driven by the challenge of cracking and programming musical codes; students showed more interest in exploring music-based puzzles afterward (see Tables A1–A5 in Appendix A). Some students wanted to tackle further coding challenges involving music, underscoring the game's success in building sustained interest and deeper learning across both domains.

1.7.3. *The Legend of the Dead Knight's Spirit*

The *Immersive Gamified Learning in the Old City of Rhodes* project [51] turns a UNESCO World Heritage site into an interactive classroom, using geolocation, QR codes, and story-driven challenges. This engaging setup lets students explore local history while boosting their digital literacy, spatial awareness, and problem-solving abilities. By merging learning with real-world discovery, the project takes a multidisciplinary approach, making history feel more real and relevant through tech-powered storytelling and immersive play (Figure 5).

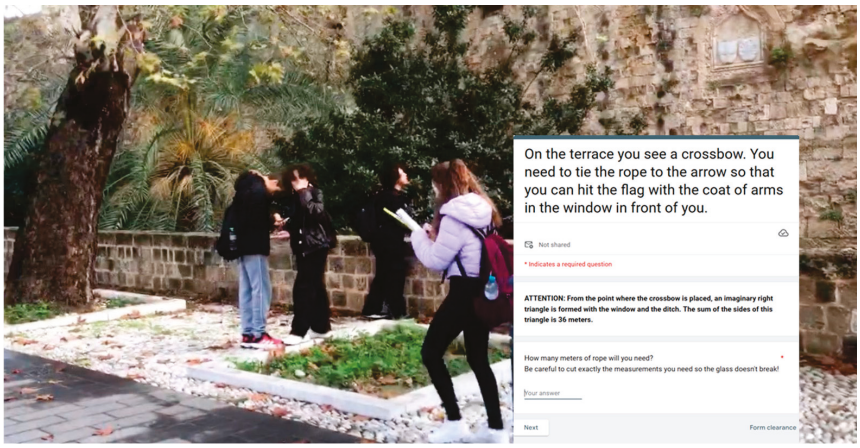


Figure 5. "Riddle of the Rope": Treasure hunt participants calculate the exact length for the crossbow challenge.

Evaluation using the LOES-S framework assessed the pedagogical value, content quality, and interactivity of the game. The findings, as revealed by the study of Sarlis et al. (2024) [51] indicated that students effectively learned about the Knights of Rhodes and other historical aspects of the medieval city of Rhodes. The game also motivated students to apply math and physics skills in solving complex puzzles required to advance, which involved calculations, logic, and spatial reasoning grounded in interdisciplinary approaches to create a meaningful learning context (see Tables A1–A5 in Appendix A). Moreover, the evaluation showed that students capably used digital tools, navigating QR codes and Google Maps to find historical spots and access materials on Google Drive. Their skill in interpreting location data and using technology for exploration confirmed the game's power to improve both historical knowledge and vital 21st-century digital literacy [52].

1.7.4. *Educational Cultural Game for Drapetsona's Industrial Heritage*

The *Educational Cultural Game for Drapetsona's Industrial Heritage* is a point-and-click puzzle adventure drawing players into the city's industrial past [53]. Using realistic settings and historical storylines, students explore local heritage, tackle interactive challenges, and cultivate logical reasoning, storytelling skills, and civic awareness. By connecting with cultural themes via problem-solving mechanics, the game aims to build a sense of place and appreciation for industrial history, making the learning experience relevant and engaging. It features mini challenges testing problem-solving and historical knowledge, like unlocking a padlock with the city's founding date or solving physics puzzles related to factory equipment (Figure 6). These interactive parts make for an immersive experience, merging education and exploration.

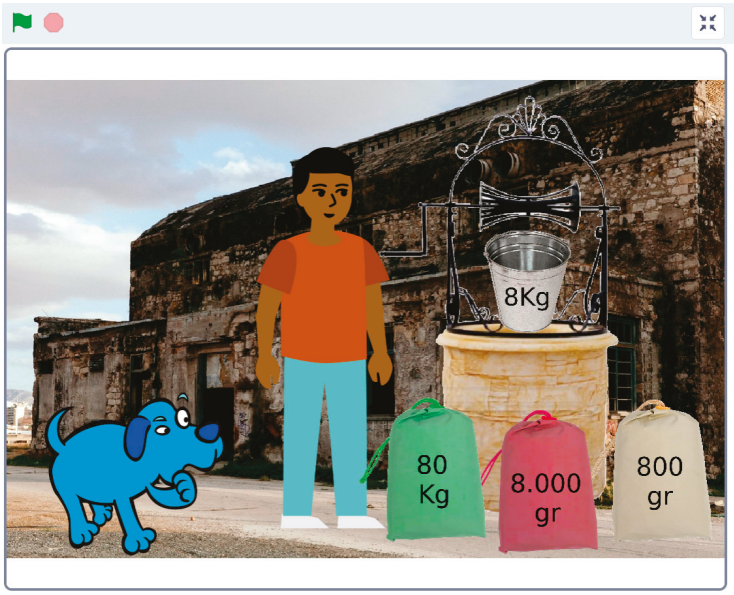


Figure 6. "Math Challenge in an Abandoned Factory": The player uses math to balance the pulley and lift an 8 kg bucket.

Due to logistical issues ("unfavorable conditions"), the game has not yet had a formal classroom evaluation. However, its designers tested it during development using the AL-MAS framework, measuring engagement, motivation, and learning (see Tables A1–A5 in Appendix A). The plan is to introduce the game soon in local schools to teach Drapetsona's

urban history. Following implementation, student and educator feedback will be gathered to assess its educational impact, engagement levels, and success in fostering historical awareness.

These four case studies (*Super Turtle Adventures*, *Musical Mastermind*, *The Legend of the Dead Knight's Spirit*, and *Urban History of Drapetsona*) showcase practical applications of the design principles discussed earlier. Evaluating their impact on learning, engagement, and user experience necessitates a robust methodological framework. The following section outlines the mixed-methods approach adopted for this research, detailing how established evaluation frameworks (such as SELEAG, ALMAS, LOES-S, and UDL), alongside the HMPI protocol, were employed to systematically analyze these educational games.

2. Materials and Methods

To address the research questions thoroughly, this study utilized a mixed-methods approach, combining quantitative and descriptive (qualitative) data. Central to this approach was the Human–Machine Pair Inspection (HMPI) framework, which facilitates the analysis of how human cognition interacts with game mechanics during gameplay. HMPI was employed in conjunction with other established evaluation tools to provide a comprehensive assessment of the educational games' design and effectiveness. These included

- The Serious Educational Games Evaluation Framework (SELEAG), which assesses learning outcomes, motivation, usability, and feedback mechanisms.
- The Assessment of Learning and Motivation Software (ALMAS), which measures engagement, learning outcomes, and motivational aspects.
- The Learning Object Evaluation Scale for Students (LOES-S), which evaluates digital tools from the student perspective in terms of usability, content quality, and interactivity.
- The Universal Design for Learning (UDL) framework, which promotes accessibility and personalization for diverse learner needs [54].

Combining these different evaluation methods allows for a deeper assessment of how educational games can be improved for maximum impact.

2.1. Qualitative Analysis

The qualitative component of the study—focusing on game design principles, the strengths and limitations of HMPI, and the comparison of different evaluation frameworks—employed a case-study methodology. Participants included small cohorts (n = 10–25) of secondary students (ages 12–16) from three urban schools. All participants and guardians provided informed consent. Ethics approval was secured through institutional review boards aligned with GDPR [55]. Interviews and observations, though qualitative, complemented the quantitative tracking of engagement and learning performance to triangulate findings. This method was chosen to gain both a broad and in-depth understanding of how different educational games perform across dimensions such as engagement, usability, and pedagogical effectiveness. Four educational games were selected for analysis: *Super Turtle Adventures*, *Musical Mastermind*, *The Legend of the Dead Knight's Spirit*, and *Urban History of Drapetsona*. Each game was examined using established frameworks (SELEAG, ALMAS, LOES-S, UDL) alongside the HMPI protocol. As previously mentioned, HMPI focuses on the interplay between player cognition and game mechanics. Utilizing HMPI provided a better understanding of how players interact with the learning environment, respond to real-time feedback, and engage with adaptive features. This deeper perspective informs recommendations for the optimal design and assessment of educational games. It should be noted that the game *Urban History of Drapetsona* was evaluated by a pilot group of educators rather than students. Their expert feedback was instrumental in identifying design limitations and pedagogical affordances within the game's structure.

2.2. Quantitative Analysis

For the quantitative component, numerical data were collected to measure aspects such as player actions, engagement levels, and learning outcomes. Data analysis was conducted using standard Python tools (https://www.python.org/ (accessed on 12 July 2025)). The combination of quantitative and qualitative findings yields a solid and detailed understanding pertinent to the research questions, supporting a thorough analysis of educational game design and evaluation [56].

2.3. Evaluation Strategies and Frameworks

To ensure the effectiveness of educational games, a combination of general evaluation strategies and specific, established frameworks is often used throughout the design process. Common evaluation strategies include

- **Pre- and Post-Tests:** Utilized to assess knowledge acquisition and cognitive gains [57].
- **Observational Studies:** Involving direct player observation, video analysis, or gameplay recordings to analyze engagement, interaction patterns, and behaviors [58].
- **Surveys and Focus Groups:** Employed to gather qualitative feedback from students and educators for design refinement [59].
- **Embedded Assessment:** Integrating real-time evaluation mechanisms to monitor player performance and provide adaptive feedback seamlessly within gameplay [60].

Alongside these general strategies, a range of established frameworks provides deeper, more structured insights into specific aspects of Game-Based Learning [61]. Prominent examples include

ALMAS (Assessment of Learning and Motivation Software)
Measures engagement, learning outcomes, and motivation in educational games.

To improve the assessment of motivation, the ALMAS (Assessment of Learning and Motivation Software) framework is utilized. ALMAS is an evaluation tool that measures student engagement, learning outcomes, and motivation in educational games [62]. Its innovation lies in its ability to integrate affective factors, such as engagement, with cognitive outcomes like understanding and performance. The assessment of motivation is conducted through a set of key dimensions that include

- Learning Effectiveness: The software's ability to improve learning.
- Mobilization: The degree to which the game sustains user interest.
- Engagement: How the game holds the student's attention during the activity.
- Long-term Impact: The effects of the game on students' skills and knowledge over time.
- Feedback Quality: The quality and usefulness of the feedback provided.

Specifically, motivation is assessed using pre- and post-game surveys, observational metrics, and self-reporting scales that capture the user's level of interest and persistence in completing tasks.

ECBA (Emerging Technology-based Competence-Based Assessment)
Evaluates how emerging technologies enhance competence development.

This framework evaluates how well games develop specific user competencies using emerging technologies. Dimensions assessed include competence development via technology, performance analysis against objectives, software adaptability, student proficiency with new technologies, and quality of interaction.

GameFlow

Ensures a balance between challenge and enjoyment in educational games.

The goal of GameFlow is to optimize the gaming experience by balancing enjoyment and learning, grounded in flow theory [63]. Player's engagement is maintained by appropriately calibrating the difficulty to ensure that it remains challenging but not discouraging. Key dimensions that influence the experience include appropriate challenge, player control, and clear objectives. Providing immediate and useful feedback is also a critical factor. Immersion can be achieved through narrative elements or gameplay challenges. Additionally, opportunities for social interaction enhance participation. Clarity in presentation and usability of the interface are equally important. Finally, to maintain flow, balancing player skill with the challenge they face is of critical importance. This subtle balance, after all, is fundamental for keeping players engaged and fully immersed in the flow experience of the game.

LOES-S (Learning Object Evaluation Scale for Students)
Assessing the effectiveness and usability of digital learning tools from the student perspective.

Developed to gather students' perspectives on the effectiveness, design, and usability of digital learning tools, the Learning Object Evaluation Scale for Students (LOES-S) [64] is widely used in adaptive systems and multimedia environments. A key focus of LOES-S is evaluating educational usefulness—specifically, how effectively a resource facilitates knowledge acquisition and understanding. Features like real-time personalization and adaptivity, known to enhance learning outcomes are central to its assessment. Beyond usefulness, LOES-S scrutinizes content quality, emphasizing accuracy and alignment with curriculum standards. It also prioritizes inclusivity and accessibility, evaluating aspects such as intuitive interfaces and compatibility with assistive technologies. Visual design elements, including multimodal interaction options, clear layouts, and simple navigation, are examined for their role in reducing cognitive load and promoting student engagement. Finally, LOES-S evaluates interactivity and adaptability, highlighting the importance of student autonomy and personalized learning paths in modern, learner-centered environments. This learner-centered approach makes LOES-S a highly valuable framework for assessing digital educational experiences.

MEEGA+ (Motivation, Engagement, and Enjoyment in Games Assessment)
Assessing Motivation, Engagement, and Enjoyment in Educational Games.

The MEEGA+ framework evaluates how educational games influence student motivation, engagement, and enjoyment. Motivation refers to the sustained interest students maintain throughout gameplay [65]. Engagement captures the degree of active participation and immersion during the game experience. Enjoyment—understood as the extent to which students find the drill pleasurable and rewarding—is also a central indicator of a game's overall appeal.

SciLOET (Science Learning Object Evaluation Tool)
Evaluates the quality and scientific validity of digital learning objects.

This framework is employed to evaluate digital learning objects specifically within science education [66]. It assesses multiple dimensions to ensure comprehensive quality. The Pedagogical Dimension examines aspects of instructional design and the support provided to learners. Concurrently, the Scientific Dimension focuses on the accuracy of the scientific content and its alignment with inquiry-based learning approaches. Finally, the Technological Dimension evaluates the object's usability and its compatibility across different technological platforms.

SELEAG (Serious Educational Learning Evaluation Assessment Grid)

Assesses interaction, engagement, and goal alignment in serious games.

This framework assesses user interaction, engagement, and the alignment of educational games with learning goals. Its evaluation criteria encompass several key dimensions. Learning outcomes are evaluated to measure the educational effectiveness of the game. The framework also examines motivation, specifically how well the game sustains player interest. Usability is another critical factor, referring to the game's accessibility and ease of use for diverse users Finally, the quality, clarity, and usefulness of in-game feedback and assessment mechanisms are assessed [67].

Universal Design for Learning (UDL)

Design for Learner Variability and Inclusion.

Universal Design for Learning (UDL) is an educational framework that promotes accessibility and inclusion by addressing the diverse needs of learners. One of its main principles is providing multiple means of engagement to support motivation, along with presenting content in varied formats to match different learning styles [68]. It also encourages flexible ways for students to demonstrate their understanding. Evidence shows that applying UDL in practice contributes to more inclusive learning environments and helps improve student motivation. As Mygdanis [69] notes, the goal of UDL is to support the creation of adaptable educational contexts capable of meeting a broad range of student needs.

Each framework offers unique insights into learner experience, instructional quality, and technological affordances, making them essential for a holistic evaluation approach. Additionally, recent research has highlighted the importance of embedding ethical foresight into the evaluation of educational technologies. As Efthymiou et al. (2023) [55] argue, educational tools—especially those employing real-time tracking, AI-based feedback, or adaptive learning mechanisms—must be designed with societal responsibility, equity, and inclusivity in mind. Ethical considerations such as data privacy, algorithmic transparency, and the potential for bias in adaptive systems should be factored into the development and evaluation process. By integrating ethical principles into established evaluation frameworks, educational game developers can ensure not only effective learning outcomes but also alignment with responsible and inclusive technological practices.

2.4. Integrating Real-Time Evaluation into Educational Game Design

The Human–Machine Pair Inspection (HMPI) protocol is a formative evaluation methodology that monitors real-time interaction between a user and an adaptive system. It draws upon HCI principles, learning analytics, and cognitive workload theory to analyze how game mechanics affect learner engagement, understanding, and progression [70]. Understanding student interaction within educational games—alongside the cognitive effort required and the system's real-time adaptability—calls for a structured and responsive evaluation method. The HMPI framework effectively addresses these needs. Instead of relying only on post-release evaluations, HMPI emphasizes the value of ongoing feedback throughout the development process. This steady stream of insights enables designers to refine usability, ensure pedagogical value, and maintain player engagement across all stages of game creation.

2.4.1. The Role of HMPI in Educational Game Evaluation

HMPI plays an important role in improving educational games by examining players' behavior and cognitive responses during gameplay. When user feedback is incorporated in real time, it becomes easier to fine-tune game mechanics for deeper immersion and more effective learning experiences. Moreover, this continuous exchange of information also supports the personalization of learning, allowing the game to adjust aspects like difficulty or provide more

targeted guidance to individual players. Furthermore, it helps measure how gameplay affects skills like logical reasoning, teamwork, and critical thinking. By checking user interactions continuously, HMPI helps developers ensure their games are engaging, adaptive, and educationally effective throughout the design process, not just at the end.

Key functions of HMPI that help game evaluation:

- Tracking Engagement: Monitors how players interact and how adaptive features affect motivation.
- Addressing Learning Effectiveness: Analyzes thinking effort, problem-solving approaches, and knowledge retention to ensure meaningful learning.
- Evaluating Personalized Experiences: Checks how well AI or other systems adapt game difficulty, feedback, and pacing.
- Measuring Skill Development: Gauges the game's impact on logic, teamwork, and critical thinking.

Additionally, ethical considerations—such as informed consent, data anonymization, and transparency in adaptive feedback systems—were carefully observed during gameplay evaluation. These precautions ensure that real-time data collection and analysis respect student autonomy and privacy in educational contexts.

2.4.2. HMPI Across the Educational Game Development Lifecycle

Integrating HMPI at various stages of the educational software development lifecycle assists in minimizing usability issues proactively and sustaining the effectiveness of adaptive learning models. This approach complements iterative playtesting methodo-logies, permitting refinements before significant problems escalate.

Application of HMPI is relevant across key development stages:

- Design Phase: Informing interface design and initial usability considerations through HCI principles.
- Development Phase: Refining difficulty scaling and user feedback mechanisms via iterative playtesting.
- Evaluation Phase: Utilizing real-time analytics to assess cognitive workload, engagement levels, and learning efficacy.
- Post-Deployment Analysis: Monitoring player interaction trends after release using adaptive tracking and embedded assessment data.

Figure 7 illustrates how HMPI mechanisms integrate across the four core phases of development—design, implementation, evaluation, and post-deployment—providing actionable data at each stage.

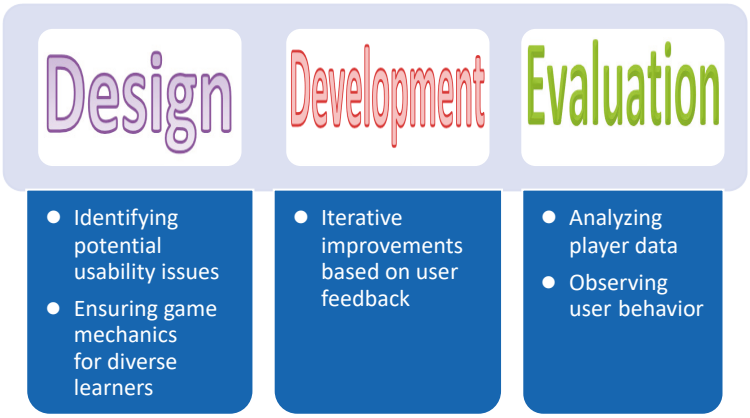


Figure 7. A model for applying Human–Machine Pair Inspection throughout game development.

2.5. Integrating Evaluation Throughout Development

Effective evaluation should be integrated throughout the development process, following a structured model such as the waterfall methodology (Figure 1). Recent implementations in educational game development, such as the work by Helda, Sanusi, and Suhendra [71], demonstrate how the waterfall model facilitates clarity and systematic progress across design phases:

- Early Stages (Design/Requirements): Emphasis on usability heuristics and accessibility standards (e.g., applying UDL guidelines).
- Middle Stages (Implementation/Testing): Direct usability testing, assessment of engagement and motivation (e.g., using SELEAG, ALMAS, LOES-S), analysis of cognitive load, and tracking of learning analytics are critical. Validation of content accuracy (e.g., via SciLOET) and assessment of competence development through technology (e.g., via ECBA) are also relevant here.
- Later Stages (Deployment/Maintenance): Long-term monitoring of engagement, motivation, and enjoyment (e.g., using MEEGA+) provides ongoing insights.

The systematic application of suitable evaluation methods and frameworks at each phase enables continuous refinement, ensuring educational games remain accessible, motivating, and pedagogically sound. Figure 8 maps evaluation tools (e.g., SELEAG, UDL, LOES-S) to the waterfall stages, clarifying when and how to apply them for maximal insight. This alignment helps ensure that evaluation is not isolated but embedded meaningfully throughout the development cycle.

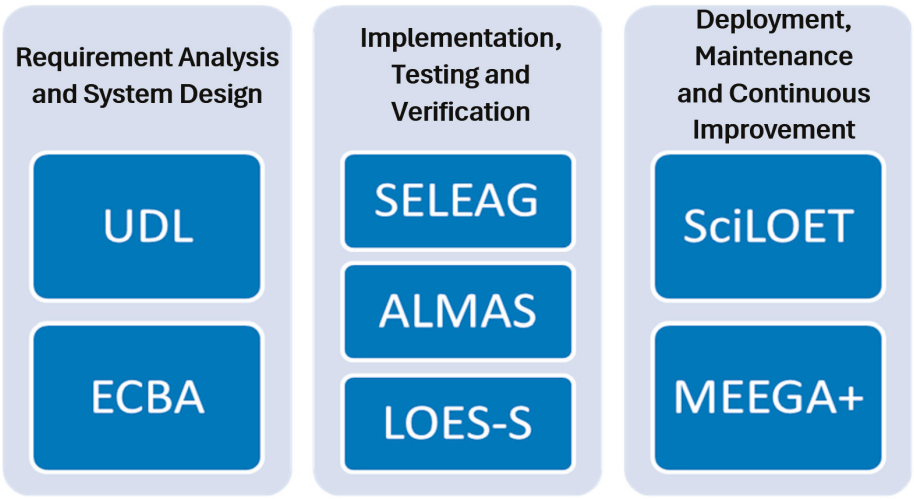


Figure 8. Alignment of evaluation strategies with the phases of educational software development following the waterfall model.

2.6. Advancing HMPI in Educational Game Development

Looking ahead, analytics methodologies derived from HMPI are anticipated to be pivotal in advancing personalized and adaptive learning experiences within educational games. Future innovations may concentrate on AI-based real-time assessments, evaluation models incorporating immersive technologies, and broader cross-disciplinary applications.

Emerging Trends in HMPI-Driven Evaluation Include

- AI-driven Adaptive Learning Analytics: Real-time customization of gameplay based on behavioral data.
- AR and VR for Immersive Learning: Utilizing immersive environments not only for learning delivery but also for assessment.

- Personalized Learning Pathways: Employing AI to adapt challenge progression based on learner performance.
- Cross-Disciplinary Applications: Extending HMPI and similar frameworks to evaluate educational tools in diverse fields like STEAM, humanities, and vocational training.
- Gamified Assessment: Embedding evaluation within authentic, contextualized game challenges.

As these evaluation approaches evolve, they are poised to enhance the design of educational games, making them more personalized, engaging, and pedagogically effective. Notably, models like GEB [72] exemplify the integration of pedagogy with interactive technology.

The synthesis of these diverse evaluation methodologies and frameworks facilitates a comprehensive assessment of educational game performance. This multifaceted approach also facilitates the identification of specific strengths and weaknesses, yielding valuable insights for iterative improvement and future design. Accordingly, the following section details the empirical results of the selected cases obtained from applying these methods.

3. Results

This section presents the findings of the study, detailing the application of the Human–Machine Pair Inspection (HMPI) framework across four educational games. The results are organized around the key research questions, with a focus on interactive game mechanics, engagement patterns, and the effectiveness of HMPI-driven learning models. Table 1 summarizes the core characteristics of each game, including technologies used and human–machine interaction patterns, while Table 2 presents the comparative results of evaluations from multiple frameworks. This approach, which was based on the triangulation of metrics, qualitative feedback, and gameplay data analysis, ensured the reliability and depth in interpreting each game's educational effectiveness.

Table 1. The game mechanics, evaluation, and Human–Machine Pair Inspection.

Game's Name	***Super Turtle Adventures***	***Urban History of Drapetsona***	***Musical Mastermind***	***The Legend of the Dead Knight's Spirit***
Programming Language	Scratch	Scratch	Micro:bit (Microcode)	Google Forms, Maps and G-Drive objects
Human Interaction	❖ Structured environment to learn Logo	❖ Point-and-click puzzle adventure	❖ Hands-on micro:bit melody decoding	❖ GPS + QR codes, location-based puzzles
Machine Interaction	❖ Intelligent guide	❖ Narrative-driven feedback	❖ Interactive feedback Validates inputs	❖ Real-time geolocation feedback
Evaluation	SELEAG	ALMAS	UDL	LOES-S
Phases of Construction	Design → Maintenance	Design → Implementation	Design → Deployment	Design → Deployment

Table 2. Comparative evaluation results by framework and game.

Evaluation Method/Game	***Super Turtle Adventures***	***The Legend of the Dead Knight's Spirit***	***Musical Mastermind***	***Urban History of Drapetsona***
ALMAS (Motivation)	Moderate—Novelty appreciated, but repetitive tasks reduced motivation.	High—Real-world exploration was highly engaging.	High—Decoding music engaged auditory learners.	Moderate—Educator reviewers noted narrative complexity but appreciated puzzle-based engagement.
SELEAG (Usability and Feedback)	High—Intuitive interface, responsive guide.	Moderate—Some learners struggled with QR interactions.	Moderate—Input validation helpful but not always accurate.	Low—Feedback was inconsistent, and interface usability raised concerns among educators.

Table 2. *Cont.*

Evaluation Method/Game	*Super Turtle Adventures*	*The Legend of the Dead Knight's Spirit*	*Musical Mastermind*	*Urban History of Drapetsona*
LOES-S (Student Perspective)	Positive—Students liked structured progression.	Mixed—Appreciated the novelty but wanted clearer instructions.	Positive–Enjoyable physical interaction and sound-based feedback.	-
UDL (Accessibility and Flexibility)	High—Multiple pathways and task supports.	Moderate—GPS-based puzzles limit accessibility.	High—Supported diverse learners with different sensory input.	-

3.1. HMPI Case Study: The Legend of the Dead Knight's Spirit

This location-based game, set in the Old Town of Rhodes, combines historical exploration with puzzle-solving. The HMPI analysis showed high motivation from its immersive narrative (ALMAS framework) but moderate usability (SELEAG framework), as some learners struggled with QR code interactions. From the student perspective (LOES-S), results were mixed, with many appreciating the novelty but requesting clearer instructions. The game received a moderate rating for accessibility (UDL framework) because its GPS-based puzzles limited inclusivity.

3.2. HMPI Case Study: Urban History of Drapetsona

Developed as a point-and-click puzzle game, this project immerses players in Drapetsona's industrial history. HMPI analysis for this game was derived exclusively from evaluations by the scientific team, as it has not yet undergone full classroom implementation. ALMAS scores were moderate, reflecting engaging puzzles but complex narratives. SELEAG scores were low due to inconsistent feedback and interface usability issues. Neither the LOES-S nor UDL frameworks were applied because the game was not evaluated directly by students.

3.3. HMPI Case Study: Super Turtle Adventures

This game provides a structured programming environment for learning Logo through block-based coding. HMPI evaluation showed moderate ALMAS scores—novelty was appreciated, but repetitive tasks slightly reduced motivation. SELEAG results were high, reflecting an intuitive interface and responsive feedback. LOES-S ratings were positive, with students enjoying the structured progression. UDL scores were high, acknowledging the multiple pathways and supports for diverse learning needs.

3.4. HMPI Case Study: Musical Mastermind

This game blends music theory and computational thinking using micro:bit devices. The HMPI findings showed high ALMAS motivation scores, as decoding music engages auditory learners. SELEAG scores were moderate, with input validation being helpful but not always accurate. LOES-S results were positive, reflecting enjoyment of the physical interaction and sound-based feedback. UDL evaluation rated the game highly for accommodating diverse learner needs.

3.5. Cross-Case Analysis and Framework Comparison

Looking ahead, analytics methodologies derived from HMPI are anticipated to be pivotal in advancing personalized and adaptive learning experiences within educational games. Future innovations may concentrate on AI-based real-time assessments, evaluation models incorporating immersive technologies, and broader cross-disciplinary approaches.

Across the four case studies, HMPI proved to be a dynamic learning model that fostered interdisciplinary skill development. Key aspects included

- Interaction and Engagement: Students interacted with both digital and physical tools (e.g., GPS, QR codes, micro:bit). Games promoting hands-on engagement and exploratory problem-solving lead to deeper participation.
- Multidisciplinary Skill Development: The games successfully combine elements of STEAM, history, and language, providing a holistic educational experience.
- Inclusivity: The gender analysis indicated that a balance among different gameplay elements (competition, exploration, structured learning) is essential for ensuring inclusivity.
- The cross-case comparison, as detailed in Table 2, demonstrates how each evaluation framework provides complementary insights. ALMAS primarily measures motivation, SELEAG assesses usability, LOES-S captures the student experience, and UDL evaluates accessibility and adaptability. By integrating these frameworks, this study provides a more holistic and actionable assessment of each game's effectiveness.

The results from these four case studies demonstrate the adaptability of the HMPI framework across varied educational contexts. They also highlight the added value of integrating multiple evaluation frameworks, as seen in the complementary insights of ALMAS, SELEAG, LOES-S, and UDL. Building on these empirical findings, the next section discusses their broader implications for educational game design and the advancement of human–machine collaborative learning environments. By integrating these frameworks, this study provides a more holistic and actionable assessment of each game's effectiveness [73].

4. Discussion

Building on the results, this section examines how the findings align with educational objectives, evaluates the strengths and limitations of HMPI, and compares multiple evaluation frameworks. It also considers emerging trends in educational game development and offers forward-looking recommendations.

4.1. Emerging Trends in Educational Game Development

As technology advances, several key trends appear to be shaping the future of Game-Based Learning:

- AI-Driven Adaptive Learning: A prominent trend involves AI-powered educational games designed to dynamically adjust content, difficulty, and feedback based on player behavior and performance.
- Immersive Technologies (VR and AR): Virtual and augmented reality show significant potential to enhance hands-on exploration, making learning more experiential and engaging.
- Cross-Disciplinary Educational Games: There is increasing integration of STEAM and humanities within single games, aiming to create more holistic learning experiences [74].
- Gamification of Assessment: Traditional assessment methods are progressively incorporating gamification elements, leading to more interactive and potentially more student-driven evaluations.

Human–Machine Pair Inspection (HMPI) Expansion: The application of HMPI is expected to extend beyond educational games into related areas such as virtual classrooms, AI tutors, and interactive simulations.

4.2. Broader Implications of HMPI in Education

Gamified learning environments enriched by human–machine collaboration can increase student motivation and cultivate essential 21st-century skills. By integrating digital tools with educational frameworks, HMPI ensures that learning remains engaging, accessible, and relevant in an increasingly technology-driven world. The comparative analysis of these four games shows the flexibility of HMPI, demonstrating its effectiveness across varied learning contexts (Table 1). Each game is evaluated based on programming language, human–machine interaction (HMI), evaluation frameworks, and development phases, showcasing unique technologies, interaction models, and assessment methodologies designed to create meaningful and engaging learning experiences. This diversity underscores the versatility of HMPI, demonstrating its potential to substantially impact education by making learning more interactive, interdisciplinary, and impactful.

4.3. Limitations of the Study

While this study provides valuable insights into the role of human–machine interaction and adaptive feedback in educational games, it is subject to certain limitations:

- Sample Size and Scope: The study analyzed a limited number of case studies, which may not fully capture broader trends in educational game design and effectiveness.
- Generalizability: The findings are based on specific Game-Based Learning contexts and technologies, which may not be directly applicable to all educational settings.
- Methodological Constraints: Although the study employed both qualitative and quantitative analysis, further refinement of evaluation criteria (e.g., comparative assessments across multiple game types) is necessary to enhance validity.
- Real-World Testing: The study's conclusions relied on controlled environments rather than large-scale classroom implementations, requiring further empirical validation with diverse participants.

These limitations are further elaborated in Section 5.1, along with suggestions for overcoming them in future research.

4.4. Addressing Research Question 1: Alignment of Game Design with Educational Objectives

RQ1: How can game design principles and interactive mechanics align effectively with educational objectives to enhance student engagement and learning outcomes?

Game design principles and interactive mechanics must be strategically aligned with educational objectives to create meaningful learning experiences that maximize engagement, knowledge retention, and skill acquisition. A fundamental strategy is to structure challenges so that difficulty gradually increases, allowing learners to build their abilities step by step. Complementing this design approach with adaptive feedback systems provides immediate, personalized responses that reinforce learning and maintain engagement. Additionally, features such as leaderboards, rewards, and time-based challenges help sustain motivation. Research also highlights the value of interactive storytelling, challenge-based progression, and collaborative learning in enhancing conceptual understanding and long-term retention. This was observed in the design of *The Legend of the Dead Knight's Spirit* and *Musical Mastermind*, where the integration of structured challenges with responsive, narrative-driven feedback improved students' conceptual grasp and intrinsic motivation. Ensuring accessibility and adaptability is equally important, meaning that tasks should be offered in multiple formats—visual, auditory, and hands-on—while accommodating varying levels of difficulty. As technology advances, innovations such as AI customization,

augmented reality, and immersive storytelling can further align game mechanics with authentic, real-world contexts, enhancing both engagement and learning effectiveness.

4.5. Addressing Research Question 2: Benefits and Limitations of HMPI

RQ2: What are the benefits and limitations of Human–Machine Pair Inspection (HMPI) in evaluating educational games, and how do they influence learning outcomes and player engagement?

A. Key Benefits of HMPI

- Personalized Learning and Adaptive Feedback: HMPI supports real-time adaptation of game difficulty and feedback based on user performance, ensuring balanced challenges that sustain engagement and accommodate individual learning needs.
- Enhanced Learning Outcomes: Immediate feedback fosters faster knowledge acquisition and retention. Research indicates that students interacting with HMPI-driven educational games exhibit higher retention rates and improved problem-solving skills.
- Increased Engagement: The dynamic nature of HMPI enables immersive, interactive learning experiences, keeping students motivated through personalized challenges and continuous progression tracking.
- Data-Driven Insights for Game Improvement: Analytics on player behavior, cognitive load, and engagement patterns allow developers and educators to refine game design and optimize learning experiences.
- Interdisciplinary Skill Development: Educational games that incorporate HMPI promote cross-disciplinary learning by blending elements of STEAM and humanities, fostering creativity, computational thinking, and critical reasoning [75].

B. Key Limitations of HMPI

- Data Interpretation Challenges: HMPI systems produce a large volume of data. Making sense of this data effectively requires a good understanding of both how students learn (educational psychology) and data analysis techniques (data science). If the data are not interpreted correctly, any changes made to educational designs might not help students improve.
- Scalability and Resource Intensity: Implementing HMPI requires appropriate technological infrastructure, posing challenges for underfunded institutions.
- Balancing Automation with Human Input: While automated feedback from HMPI can be a great help to learners, relying too much on these automated systems might reduce important opportunities for human interaction. This includes valuable moments like discussions led by teachers, students working together, and direct guidance from educators.
- Connecting with Teaching Strategies: HMPI primarily looks at how a learner interacts with a specific system or game. For HMPI to be truly effective in an educational setting, it must align well with broader teaching methods and overall educational goals. Its success depends on how seamlessly it fits into the existing pedagogical framework.

C. Impact of HMPI on Learning Outcomes and Player Engagement

- Improved Knowledge Retention: Personalized feedback and adaptive difficulty levels help students retain concepts more effectively.

- Higher Motivation and Immersion: Real-time interaction and progress tracking maintain student interest
- Balanced Learning Experience: HMPI's mix of challenge and support prevents frustration or disengagement
- Inclusive Game Design: Analyzing gameplay data enables targeted refinements to address diverse cognitive abilities and engagement preferences.

In conclusion, while HMPI offers substantial advantages in enhancing learning experiences and engagement in educational games, its overall success relies on careful implementation, balanced integration with human instruction, and accurate interpretation of data.

4.6. Evaluating Educational Games Through Multiple Frameworks

RQ3: In what ways do different evaluation frameworks—such as ALMAS, ECBA, GameFlow, and SELEAG—differ in how they measure the effectiveness of educational games in terms of player motivation, cognitive skill development, and knowledge retention?

Assessing the effectiveness of educational games requires multidimensional frameworks that can capture the complexity of learning experiences. Each framework offers distinct lenses through which to evaluate core components such as engagement, skill acquisition, and retention [76]. A comparative summary of evaluation results across the four games examined is provided in Table 2, which illustrates how each framework highlights different strengths and weaknesses per game.

Based on the application of the different frameworks, several key insights were gained. The ALMAS framework, which focuses on motivation and long-term engagement showed that *The Legend of the Dead Knight's Spirit* and *Musical Mastermind* were highly rated due to their engaging narrative and real-world exploration. *Super Turtle Adventures* received a moderate score because repetitive tasks reduced motivation, while *Urban History of Drapetsona* also scored moderately due to narrative complexity noted by educator reviewers. The SELEAG framework, which concentrates on usability and pedagogical alignment, found *Super Turtle Adventures* to be highly effective due to its intuitive interface. *Urban History of Drapetsona* received a low score because of inconsistent system feedback and interface issues. *The Legend of the Dead Knight's Spirit* and *Musical Mastermind* were rated as moderate due to some learners struggling with QR interactions and inconsistent input validation, respectively.

The LOES-S framework, which reflects the student perspective highlighted that both *Super Turtle Adventures* and *Musical Mastermind* were positively received for their structured progression and enjoyable interaction. *The Legend of the Dead Knight's Spirit* received mixed results, with students appreciating the novelty but requesting clearer instructions. LOES-S was not applied to *Urban History of Drapetsona*, as it was not evaluated by students. Finally, the UDL framework, which promotes accessibility and personalization, rated *Super Turtle Adventures* and *Musical Mastermind* highly for supporting diverse learner needs. *The Legend of the Dead Knight's Spirit* scored moderately, as its GPS-based puzzles limited accessibility. UDL was not applied to *Urban History of Drapetsona* for the same reason as LOES-S.

4.7. Synthesizing Insights and Looking Forward

The synthesis of the findings reveals a dynamic interplay among game mechanics, pedagogical goals, learner engagement, and evaluative methodologies. At the center of this ecosystem lies the principle of alignment, ensuring that game mechanics meaningfully correspond to intended learning outcomes. This alignment becomes more powerful when it is supported by adaptive feedback systems, such as those utilized within the HMPI

framework. As shown in previous sections, such systems enhance personalization, sustain engagement, and promote deeper learning. The HMPI approach offers valuable contributions to educational game evaluation, particularly in its ability to generate actionable data on learner behavior and performance. However, as the findings suggest, the successful implementation of HMPI is contingent upon a few factors. These include the capacity to interpret complex data outputs, the availability of institutional resources, and the alignment of automated systems with broader pedagogical strategies. In addition, learner diversity remains a critical consideration. Variations in engagement across demographic lines, such as gender or cognitive style, highlight the need for inclusive and adaptable game design that supports equity in learning experience.

Moreover, the evaluation frameworks discussed in Section 4.3 reinforce the importance of multidimensional assessment. Tools such as ALMAS, ECBA, GameFlow, and SELEAG offer complementary strengths, and their combined use allows for more comprehensive and context-sensitive analysis. The synergy between HMPI and these frameworks presents a compelling direction for future research and practice. Moving forward, efforts should concentrate on refining the integration of HMPI and hybrid evaluation models across diverse educational settings. This includes testing scalability, optimizing real-time analytics, and supporting educators with practical tools for interpretation and implementation. Technological innovation, particularly in areas such as AI-driven feedback, adaptive storytelling, and multimodal interaction—can further enhance the alignment between game mechanics and authentic learning needs. Ultimately, the findings underscore a critical insight: the effectiveness of serious games depends not only on innovative design or powerful technologies, but on how these elements are meaningfully combined, evaluated, and adapted to diverse learner populations.

5. Conclusions

Rigorous evaluation of educational games is essential for validating their pedagogical relevance, ensuring high design quality, and confirming their sustained impact on student learning. The integration of structured evaluation models—such as SELEAG, GameFlow, and ECBA—equips educators and developers with essential tools to understand what drives learner motivation, fosters engagement, and facilitates cognitive development. These frameworks provide both analytic rigor and design-oriented insights, enabling the creation of personalized and inclusive learning experiences that respond to diverse learner profiles. A central contribution of this study lies in the demonstrated potential of the HMPI approach in strengthening the educational effectiveness of serious games. HMPI supports critical components of Game-Based Learning, including adaptive feedback, interactive storytelling, and cognitively engaging activities. Its interdisciplinary foundation—linking STEAM, humanities, and cognitive science—positions it as a powerful mechanism for promoting creativity, problem-solving, and collaborative learning.

Moreover, HMPI-generated insights offer meaningful support for inclusive instructional design, highlighting the importance of addressing individual learning differences and gender-based engagement patterns [77]. The combination of gamification and artificial intelligence for personalizing gameplay has also shown promising results, positively influencing academic achievement, learner motivation, and depth of engagement [78]. In addition, the implementation of multimodal learning strategies—such as integrating hands-on experimentation with problem-solving scenarios—has been found to enhance memory retention and critical thinking by activating multiple sensory and cognitive channels [79]. These approaches make learning more inclusive and impactful, especially when applied across different educational levels. Advancing the field of educational game design

thus requires a balanced integration of technological innovation, pedagogical insight, and inclusive design principles. The findings of this study form a solid foundation for the ongoing development of serious games that are not only engaging and adaptive but also pedagogically robust and socially responsive.

5.1. Challenges and Areas for Further Research

While HMPI and AI-driven educational game models offer significant advantages, several challenges must be addressed to optimize their effectiveness and scalability.

5.1.1. Interpreting Data from HMPI and AI-Driven Feedback

One of the most pressing challenges in the deployment of HMPI and AI-powered educational games is the effective interpretation of the big data that they generate. These systems produce continuous streams of learners' behavioral data feeds, ranging from interaction patterns and response times to cognitive load indicators and engagement metrics. While data feeds hold considerable promise for informing personalized instruction, their utility is contingent on the availability of robust analytic frameworks and user-friendly visualization tools. Educators and developers often lack the technical expertise or infrastructure required to transform raw data into actionable insights. Without advanced analytics and clearly defined pedagogical models for interpretation, there is a risk of misalignment between data-driven recommendations and actual learner needs. In such cases, adaptive feedback mechanisms may become superficial or even counterproductive, failing to support meaningful learning progress. Furthermore, ethical considerations surrounding data privacy, consent, and algorithmic bias must also be considered, particularly when systems are used in formal educational settings. Future research should explore how data from HMPI systems can be meaningfully contextualized within broader educational goals and used to support teachers rather than replace them.

5.1.2. Scalability of Personalized Learning Approaches

As previously demonstrated using HMPI, personalized learning strategies have the potential to enhance student engagement and adaptation. However, implementing such strategies across many students without compromising quality demands substantial technological and infrastructural support. While AI has the potential to tailor learning experiences to individual needs, the computational resources required to maintain real-time adaptation can be costly and logistically demanding. Ensuring that these technologies remain accessible to a broad range of educational institutions, including those with limited resources, is essential for promoting equitable learning opportunities. Future research should explore cost-effective cloud-based solutions and optimized AI algorithms that can facilitate scalability while maintaining high levels of personalization.

5.1.3. Balancing Automation with Human-Led Instruction

As previously demonstrated in the context of HMPI-based systems, AI-driven learning environments enhance adaptability and responsiveness. However, over-reliance on automation may inadvertently diminish the role of human-led discussions and instructor intervention. Personalized instruction from educators plays a crucial role in providing emotional support, fostering critical thinking, and encouraging deeper engagement—elements that AI alone cannot fully replicate. Striking the right balance between automated feedback and teacher involvement is essential to maintaining a meaningful and interactive learning experience. Educational game developers should prioritize designs that complement hu-

man instruction rather than replace it, ensuring that AI serves as an enhancement rather than a substitute for traditional teaching methods [80].

5.1.4. Need for Large-Scale Empirical Validation

Controlled case studies have yielded valuable insights into adaptive learning mechanics and Human–Machine Pair Inspection (HMPI), yet they often fail to capture the full complexity of real-world classrooms. To determine how these systems perform in everyday educational settings, large-scale trials are required—engaging more students across a variety of schools and learning levels. Future investigations ought to compare diverse learner populations across multiple institutions and track their progress over time [81]. Such expansive experimentation will reveal how cultural background, infrastructure, and individual learning differences shape the effectiveness of AI-driven, game-based interventions. While pilot studies offer important initial evidence, mixed-methods replications and comparative field trials are essential to determine whether early findings are held on a scale. Anchoring these efforts in established frameworks—such as Juul's rule-based formal system model or Sweetser and Wyeth's GameFlow model—will ensure methodological rigor without sacrificing contextual relevance.

Building this broader empirical foundation would refine evaluation methodologies, strengthen adaptive feedback algorithms, and ensure that AI-driven personalization supports rather than replaces human instruction. Ultimately, this groundwork will enable the development of educational games that are inclusive, scalable, and demonstrably effective at fostering deep learning, skill acquisition, and equitable access.

The following section outlines practical recommendations and future directions, addressing the opportunities and challenges that lie ahead.

6. Recommendations and Future Directions

As previously discussed in the results and discussion sections, adaptive learning systems have shown strong potential for enhancing personalization and engagement. The ongoing evolution of educational games is closely linked to advancements in data analytics, learner-centered design, and emerging technologies. By integrating artificial intelligence AI-driven learning, immersive narrative structures, and evaluation frameworks such as HMPI, educators and developers can design learning experiences that are engaging, inclusive, and pedagogically effective. When coupled with real-world applications, interdisciplinary content, and adaptive learning models, game-based education has the potential to remain a vital and transformative component of contemporary learning environments [82]. To maximize both impact and scalability, future research and development should focus on the following key areas.

6.1. Advancing Adaptive Learning and Personalization

A primary avenue for enhancing educational games lies in the development of AI-powered adaptive learning systems capable of dynamically adjusting difficulty levels in response to learner performance. By analyzing in-game behavior, machine-learning algorithms can identify areas of cognitive struggle and modify challenges in real time to support individual learning trajectories.

Additionally, the implementation of personalized learning pathways allows students to progress at their own pace, engaging with content that aligns with their cognitive styles and preferences. Predictive analytics can further refine game mechanics, ensuring that content remains appropriately scaffolded to sustain motivation and improve knowledge retention.

6.2. Enhancing Interactivity Through Hands-On and Analytical Engagement

Effective educational games should offer a balanced combination of hands-on activities and structured, logic-based problem-solving tasks to address a broad range of learning styles. The inclusion of physical tools—such as micro:bit devices, QR codes, and IoT-enabled resources—enhances the learning experience for kinesthetic learners by linking digital content with tangible interaction. At the same time, strategically designed analytical challenges can stimulate critical thinking and foster deep cognitive engagement. Future developments in Game-Based Learning are expected to integrate physical and digital modalities seamlessly, creating immersive, hybrid environments that foster both practical and analytical learning [83].

6.3. Expanding Gamified Assessments and Alternative Evaluation Methods

As educational games continue to evolve, assessment methodologies should also be adapted to the games' design. Traditional exams and standardized tests are often limited in capturing the depth and dynamism of Game-Based Learning. On the other hand, future assessment systems should incorporate interactive elements such as simulations, scenario-based challenges, and digital storytelling [84].

Embedding real-time performance analytics within games allows for continuous monitoring of cognitive development, engagement levels, and content mastery. AI-driven adaptive quizzes can further personalize assessment by adjusting difficulty and content based on individual learner profiles. They transform evaluation into a fluid and formative component of the learning process.

6.4. Immersive Learning with Augmented Reality (AR) and Virtual Reality (VR)

The integration of AR and VR technologies into educational games offers transformative possibilities for experiential learning [85]. Augmented reality can be used to create location-based storytelling experiences, where students explore historical sites, scientific phenomena, or cultural artifacts through interactive AR overlays. Virtual reality, on the other hand, can provide fully immersive collaborative problem-solving scenarios, enabling students to engage in realistic team-based challenges while developing communication, cooperation, and critical thinking skills. As these technologies become more and more accessible, their application in educational games will provide students with deeper, more interactive learning experiences.

6.5. Scaling up Human–Machine Pair Inspection (HMPI) in Educational Research

The HMPI framework should be expanded to incorporate AI tutors, voice recognition, and multimodal learning environments, further refining interactive educational experiences. As HMPI collects and analyzes data on student interactions, its effectiveness can be enhanced through AI-driven insights that monitor cognitive load, engagement patterns, and learning efficiency. Future research should also explore the integration of emotion recognition technologies to assess affective states such as frustration or confusion, enabling real-time adaptation of content based on learners' emotional and cognitive needs.

6.6. Strengthening Teacher Training and Curriculum Integration

For Game-Based Learning to be effectively implemented, educators must be adequately trained and supported. The successful implementation of Game-Based Learning hinges on strengthening teacher training and facilitating curriculum integration. To this end, professional development programs must equip educators not only with technical skills but also with pedagogical strategies to integrate educational games into their lesson

plans and align them with specific curriculum goals. AI-powered tools can further support educators by offering content customization and differentiated instruction based on student data. Additionally, curricular models must evolve to embed Game-Based Learning principles at a structural level, ensuring that games function as core pedagogical instruments rather than as optional enhancements.

6.7. Ensuring Accessibility, Inclusivity, and Bridging the Digital Divide

Universal Design principles must underpin the development of educational games to ensure accessibility for learners of all abilities, including students with learning disabilities. In parallel, developers and policymakers must address the digital divide by expanding access to educational games in underserved regions and resource-constrained schools. Research should prioritize cost-effective, cloud-based platforms and mobile-friendly technologies to make Game-Based Learning accessible on a global scale. The promotion of open-access development tools will further democratize educational innovation, enabling learners from all socioeconomic backgrounds to benefit.

6.8. Final Thoughts

The application of rigorous evaluation strategies in educational game design holds significant potential to enhance student engagement, motivation, and interdisciplinary learning. These strategies also address critical challenges related to accessibility, personalization, and assessment. Emerging technologies—such as AI, AR/VR, and gamified assessment tools—are rapidly reshaping the educational landscape, enabling learning experiences that are more immersive, adaptive, and interactive. These advancements have shown measurable improvements in learner engagement, motivation, and learning outcomes, paving the way for the next generation of dynamic, impactful Game-Based Learning environments. As technological innovation accelerates, Game-Based Learning is becoming a pivotal medium for engaging students with digital content in ways that are both personalized and inclusive. Ensuring accessibility, ethical AI implementation, and interdisciplinary integration will be crucial in scaling these innovations for global educational impact. The future of education will be driven by intelligent, interactive, and inclusive learning environments, where games serve as powerful tools for knowledge acquisition, skill development, and creative problem-solving.

Author Contributions: As authors, we have organized strict formatting requirements, and all manuscripts contain the required sections: Author Information, Abstract, Keywords, Introduction, Materials and Methods, Results, Conclusions, Figures and Tables with Captions, Funding Information, Author Contributions, Conflict of Interest, and other Ethics Statements. I.S. and D.K.: Writing, Statistics, and Conceptualization. D.K. and C.D.: Conceptualization, Review, and Editing. D.K. and C.D.: Visualization. All authors have read and agreed to the published version of the manuscript.

Funding: This work has been partly supported by COSMOTE under a PEDION24 grant. (University of Piraeus Research Center).

Institutional Review Board Statement: The research was conducted following the Declaration of Helsinki, and the protocol was approved by the Ethics Committee of the University of Piraeus and the University of West Attica as described in the general Code of Ethics and Conduct (2020), the general regulations of international publications, and the official legislation on research (No. 8465 (1) Regulation for the Conduct of Postdoctoral Research at the University of West Attica, dated 8 March 2019, Sheet Number 827).

Informed Consent Statement: The research material (photos or statistical tables) collected and reproduced in the manuscript is about educational and learning activities within the context of daily

teaching in the schools where we are the school's principals. Within the framework of these actions, Informed consent for publication was obtained from all identifiable human participants, students, or teachers. Informed oral consent from the teachers and written consent from the parents of the students for publication were obtained from all identifiable human participants.

Data Availability Statement: There are no other data other than those mentioned in the text and appendices.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript.

Abbreviation	Full Term	Description
AI	Artificial Intelligence	Use of machines and algorithms to simulate human intelligence processes.
ALMAS	Assessment of Learning and Motivation Software	Framework for measuring engagement, motivation, and learning outcomes in educational games.
AR	Augmented Reality	Technology overlaying digital content on the real-world environment.
ECBA	Emerging Technology-based Competence-Based Assessment	Framework assessing competence development using emerging technologies.
GBL	Game-Based Learning	Use of complete games aligned with learning objectives to enhance engagement and understanding.
HCI	Human-Computer Interaction	Field focusing on the design and use of computer technology and user interfaces.
HMPI	Human-Machine Pair Inspection	Protocol for real-time analysis of learner interaction, engagement, and adaptive feedback in educational games.
ICT	Information and Communication Technologies	Technologies used for communication, creation, and sharing of information in education and other fields.
IoT	Internet of Things	A network of interconnected devices that collect and exchange data.
LOES-S	Learning Object Evaluation Scale for Students	A tool for evaluating the usability, content quality, and interactivity of digital learning resources from the student perspective.
MEEGA+	Motivation, Engagement, and Enjoyment in Games Assessment	Framework measuring motivational and engagement aspects of educational games.
SciLOET	Science Learning Object Evaluation Tool	A tool for assessing scientific accuracy, pedagogical design, and technological usability of science-focused learning objects.
SELEAG	Serious Educational Games Evaluation Framework	Evaluation grid focusing on learning outcomes, motivation, usability, and feedback.
STEAM	Science, Technology, Engineering, Arts, and Mathematics	Educational approach integrating arts into STEM for creativity and innovation.
UDL	Universal Design for Learning	Framework for inclusive education that addresses learner variability through multiple means of engagement, representation, and expression.
VR	Virtual Reality	An immersive computer-generated environment simulating real or imagined experiences.

Appendix A

Table A1. Learning outcomes and knowledge acquisition [1].

Game	Achievement of Learning Objectives (%)	Evaluation Method
Super Turtle Adventures	85%	SELEAG
Musical Mastermind	87%	UDL
The Legend of the Dead Knight's Spirit	88%	LOES-S
Urban History of Drapetsona	84%	ALMAS [2]

[1] The effectiveness of the games in promoting learning objectives was assessed based on students' ability to grasp and apply new concepts. [2] The assessment was conducted by the designers during the development and implementation phase of the game.

Table A2. Engagement and player experience [1].

Game	Engagement (Boys) (%)	Engagement (Girls) (%)	Overall Engagement (%)
Super Turtle Adv.	88	89	88.5
Musical Mastermind	90	91	90.5
The Legend of the Dead Knight's Spirit	92	93	92.5
Drapetsona Cult. Game	87	88	87.5

[1] Engagement levels were measured through player participation, motivation, and interaction with game mechanics.

Table A3. Gameplay experience and scenario effectiveness [1].

Game	Gameplay Quality (%)	Scenario Quality (%)
Super Turtle Adventures	86	89
Musical Mastermind	89	91
The Legend of the Dead Knight's Spirit	91	92
Urban History of Drapetsona	85	88

[1] Each game was assessed based on game mechanics, player interaction, and overall scenario design.

Table A4. Gender-based gameplay preferences [1].

Game	Boys' Preferred Element	Girls' Preferred Element
Musical Mastermind	Hands-on micro:bit experimentation	Puzzle-solving and melody decoding
The Legend of the Dead Knight's Spirit	Searching for hotspots (competition)	Historical learning and exploration

[1] The research results were based on student questionnaires completed after their gameplay experience.

Table A5. Overall effectiveness [1].

Game	Overall Effectiveness (%)
Super Turtle Adventures	86
Musical Mastermind	90
The Legend of the Dead Knight's Spirit	92
Urban History of Drapetsona	85
Average	***88.25***

[1] The research results were based on student questionnaires completed after their gameplay experience.

References

1. Ren, J.; Xu, W.; Liu, Z. The impact of educational games on learning outcomes. *Int. J. Game-Based Learn.* **2024**, *14*, 1–25. [CrossRef]
2. Gao, M.; Yunus, M.M.; Rafiq, K.R.M. Educational games and game-based approaches in higher education: A systematic review (2014–2023). *Int. J. Acad. Res. Progress. Educ. Dev.* **2024**, *13*, 72–82. [CrossRef]
3. Oliveira, W.; Pastushenko, O.; Rodrigues, L.; Toda, A.M.; Palomino, P.T.; Hamari, J.; Isotani, S. Does gamification affect flow experience? A systematic literature review. In Proceedings of the 5th International GamiFIN Conference 2021 (GamiFIN 2021), Online, 7–10 April 2021. [CrossRef]
4. Luo, J. Validating the impact of gamified technology-enhanced learning environments on motivation and academic performance: Enhancing TELEs with digital badges. *Front. Educ.* **2024**, *9*, 1429452. [CrossRef]
5. Pinto, M.; Leite, C. Digital technologies in support of students' learning in higher education: A literature review. *Digit. Educ. Rev.* **2020**, *37*, 343–360. [CrossRef]
6. González-Pérez, L.I.; Ramírez-Montoya, M.S. Components of Education 4.0 in 21st Century Skills Frameworks: Systematic Review. *Sustainability* **2022**, *14*, 1493. [CrossRef]
7. Kalogiannakis, M.; Papadakis, S.; Zourmpakis, A.-I. Gamification in Science Education. A Systematic Review of the Literature. *Educ. Sci.* **2021**, *11*, 22. [CrossRef]
8. Deterding, S.; Dixon, D.; Khaled, R.; Nacke, L. From game design elements to gamefulness: Defining "gamification". In Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments, Tampere, Finland, 28–30 September 2011; pp. 9–11. [CrossRef]
9. Bogost, I. *Persuasive Games: The Expressive Power of Videogames*; MIT Press: Cambridge, MA, USA, 2007. [CrossRef]
10. Maričić, M.; Lavicza, Z. Enhancing student engagement through emerging technology integration in STEAM learning environments. *Educ. Inf. Technol.* **2024**, *29*, 23361–23389. [CrossRef]
11. Chiotaki, D.; Poulopoulos, V.; Karpouzis, K. Adaptive game-based learning in education: A systematic review. *Front. Comput. Sci.* **2023**, *5*, 1062350. [CrossRef]
12. Acosta, H.D.; Lee, S.; Mott, B.; Bae, H.; Glazewski, K.; Hmelo-Silver, C.; Lester, J. Multimodal learning analytics for predicting student collaboration satisfaction in collaborative game-based learning. In Proceedings of the 17th International Conference on Educational Data Mining, Atlanta, GA, USA, 14–17 July 2024; pp. 224–235. [CrossRef]
13. Fang, X. *HCI in Games: 4th International Conference, HCI-Games 2022, Held as Part of the 24th HCI International Conference, HCII 2022, Virtual Event, 26 June–1 July 2022, Proceedings*, 1st ed.; Springer: Cham, Switzerland, 2022. [CrossRef]
14. MacDonald, C.M.; Girouard, A.; Granollers, T.; Joshi, A.; Kang, J.; Kharrufa, A.; Slegers, K.; St-Cyr, O. Editorial: Teaching and learning human-computer interaction (HCI): Current and emerging practices. *Front. Comput. Sci.* **2023**, *5*, 1188680. [CrossRef]
15. Sanjaya, K.; Sanjaya, P.K.; Rathika, S.; Rajalakshmi, B. Human-computer interaction: Enhancing user experience in interactive systems. *Kufa J. Eng.* **2023**, *14*, 23–41. [CrossRef]
16. Wu, C.; Su, Y.; Li, Y.; Chen, G. A comprehensive exploration of personalized learning in smart education: From student modeling to personalized recommendations. *arXiv* **2024**, arXiv:2402.01666. Available online: https://arxiv.org/abs/2402.01666 (accessed on 10 July 2025). [CrossRef]
17. OECD. *PISA 2022 Assessment and Analytical Framework*; OECD Publishing: Paris, France, 2022. [CrossRef]
18. Estrada-Molina, O.; Fuentes-Cancell, D.R.; Alvarez Morales, A. The assessment of the usability of digital educational resources: An interdisciplinary analysis from two systematic reviews. *Educ. Inf. Technol.* **2022**, *27*, 4037–4063. [CrossRef]
19. Veytia Bucheli, M.G.; Gómez-Galán, J.; Cáceres Mesa, M.L.; López Catalán, L. Digital technologies as enablers of universal design for learning: Higher education students' perceptions in the context of SDG4. *Discov. Sustain.* **2024**, *5*, 473. [CrossRef]
20. Gui, Y.; Cai, Z.; Yang, Y.; Kong, L.; Fan, X.; Tai, R.H. Effectiveness of digital educational games and game design in STEM learning: A meta-analytic review. *Int. J. STEM Educ.* **2023**, *10*, 36. [CrossRef]
21. Martia, L.A.; Nurafni, N. The impact of treasure hunt games on learning outcomes in time measurement. *Al-Jabar J. Pendidik. Mat.* **2023**, *14*, 515–525. [CrossRef]
22. Ruhama, U.; Yulian, R.; Rangkuti, Z. Blended Learning in the Digital Age: Integrating Augmented Reality (AR) and Gamification for Effective Teaching. *Int. J. Innov. Think.* **2025**, *2*, 51–65. Available online: https://www.researchgate.net/publication/391889934_Blended_Learning_in_the_Digital_Age_Integrating_Augmented_Reality_AR_and_Gamification_for_Effective_Teaching (accessed on 15 July 2025). [CrossRef]
23. Li, Y.; Chen, D.; Deng, X. The impact of digital educational games on students' motivation for learning: The mediating effect of learning engagement and the moderating effect of the digital environment. *PLoS ONE* **2024**, *19*, e0294350. [CrossRef]
24. Lai, C.-H.; Hu, P.-Y. The gaming revolution in history education: The practice and challenges of integrating game-based learning into formal education. *Information* **2025**, *16*, 490. [CrossRef]

25. Haryanti, T.; Destari, D.; Hidayat, N.R.; Nugroho, A.Y.; Waahib, A.N. Simulated learning in STEM education: Bridging theory and practice through AR. *J. Emerg. Technol. Educ.* **2024**, *2*, 72–82. Available online: https://journal.ypidathu.or.id/index.php/jete/article/download/1368/926/23344 (accessed on 12 July 2025). [CrossRef]
26. Juul, J. *Half-Real: Video Games Between Real Rules and Fictional Worlds*; MIT Press: Cambridge, MA, USA, 2005; ISBN 978-0-262-10110-3.
27. Salen Tekinbas, K.; Zimmerman, E. *Rules of Play: Game Design Fundamentals*; MIT Press: Cambridge, MA, USA, 2003; ISBN 9780262240451.
28. Gleichgerrcht, E.; Llorente, J.; Serebrisky, A. Immersive serious games for learning the concept of density: An analysis of the impact of virtual reality on conceptual change in high school students. *arXiv* **2024**, arXiv:2401.01831. Available online: https://arxiv.org/abs/2401.01831 (accessed on 10 July 2025).
29. Puspitasari, N.P.D.; Putro, H.P.; Sari, R.P.P.; Hapsari, A.D.P. Development virtual laboratory based on augmented reality for basic concepts of physics and chemistry courses for elementary teacher students. In *The 9th Annual International Seminar on Trends in Science and Science Education (AISTSSE) 2022*; Sciendo: Boston, MA, USA, 2023; pp. 601–608. [CrossRef]
30. Lampropoulos, G. Augmented Reality, Virtual Reality, and Intelligent Tutoring Systems in Education and Training: A Systematic Literature Review. *Appl. Sci.* **2025**, *15*, 3223. [CrossRef]
31. McLaren, B.M.; Nguyen, H.A. Digital learning games in artificial intelligence in education (AIED): A review. In *Handbook of Artificial Intelligence in Education*; Edward Elgar Publishing: Cheltenham, UK, 2023. [CrossRef]
32. Kharitonov, V.; Kulik, S.; Zhadko, S. Gamification method using Minecraft for training future teachers of computer science and mathematics. In Proceedings of the 4th International Workshop on Interactive Environments and Emerging Technologies for eLearning (IEETeL 2023), Guimaraes, Portugal, 12–14 July 2023; pp. 209–216. Available online: https://ceur-ws.org/Vol-3771/paper26.pdf (accessed on 24 July 2025).
33. Almasri, F. Simulations to teach science subjects: Connections among students' engagement, self-confidence, satisfaction, and learning styles. *Educ. Inf. Technol.* **2022**, *27*, 1587–1610. [CrossRef] [PubMed]
34. Krouska, A.; Troussas, C.; Mylonas, P.; Sgouropoulou, C. Exploring the acceptance and impact of a digital escape room game for environmental education using structural equation modeling. *Information* **2025**, *16*, 528. [CrossRef]
35. Guzmán, D.E.; Rengifo, C.F.; Garcia, C.E. Assessing Game Mechanics of a VR-Based Cognitive Game for Older Adults to Enable Dynamic Difficulty Adjustment. In *Converging Clinical and Engineering Research on Neurorehabilitation V*; Pons, J.L., Tornero, J., Akay, M., Eds.; Springer: Cham, Switzerland, 2024; pp. 390–394. [CrossRef]
36. Sarlis, I. Super Turtle Adventures [Online Game]. Available online: http://superturtleadventures.schl.gr (accessed on 7 August 2025).
37. Sarlis, I.; Kotsifakos, D.; Douligeris, C. The Musical Mastermind: A case study on fostering music theory and computational thinking. In Proceedings of the IEEE Educon 2025, London, UK, 22–25 April 2025. [CrossRef]
38. Mygdanis, Y. "Little music makes & producers": Designing and implementing an innovative educational technological-musical approach in primary education to enhance creativity. In Proceedings of the 13th Panhellenic and International Conference "ICT in Education", Kavala, Greece, 29 September–1 October 2024; pp. 380–387. Available online: https://www.researchgate.net/publication/382248007 (accessed on 17 July 2025).
39. Liu, Y.; Ma, Z.; Ng, J.T.D.; Hu, X. Multimodal learning analytics for game-based assessment of collaborative problem-solving skills among young students. In Proceedings of the 15th International Learning Analytics and Knowledge Conference (LAK '25), Dublin, Ireland, 3–7 March 2025; pp. 963–969. [CrossRef]
40. Tene, T.; Marcatoma Tixi, J.A.; Palacios Robalino, M.d.L.; Mendoza Salazar, M.J.; Vacacela Gomez, C.; Bellucci, S. Integrating immersive technologies with STEM education: A systematic review. *Front. Educ.* **2024**, *9*, 1410163. [CrossRef]
41. Tiwari, R. The Integration of AI and Machine Learning in Education and Its Potential to Personalize and Improve Student Learning Experiences. *Int. J. Sci. Res. Eng. Manag.* **2023**, *7*. [CrossRef]
42. Pacheco-Velazquez, E.; Bester, A.; Rabago-Mayer, L.; Ro, V. What do we evaluate in serious games? A systematic review. In Proceedings of the 17th European Conference on Games Based Learning (ECGBL), Enschede, The Netherlands, 5–6 October 2023; pp. 696–704. [CrossRef]
43. Barbero, G.; Bonsangue, M.M.; Hermans, F.F.J. How to evaluate games in education: A literature review. *arXiv* **2024**, arXiv:2407.03879. Available online: https://arxiv.org/abs/2407.03879 (accessed on 10 July 2025). [CrossRef]
44. Behnamnia, N.; Kamsin, A.; Hayati, S. Impact of digital game-based learning on STEM education in primary schools: A meta-analysis of learning approaches. *Innoeduca Int. J. Technol. Educ. Innov.* **2024**, *10*, 113–140. Available online: https://revistas.uma.es/index.php/innoeduca/article/view/19694/20777 (accessed on 13 July 2025). [CrossRef]
45. Manoharan, A.; Nagulapally, S. Adaptive gamification algorithms for personalized learning experiences in educational platforms. *Int. Res. J. Mod. Eng. Technol. Sci.* **2024**, *6*, 2582–5208. [CrossRef]

46. Almeqdad, Q.I.; Alodat, A.M.; Alquraan, M.F.; Mohaidat, M.A.; Al-Makhzoomy, A.K. The effectiveness of universal design for learning: A systematic review of the literature and meta-analysis. *Cogent Educ.* **2023**, *10*, 2218191. [CrossRef]
47. Reyes-Millán, M.; Vázquez-Villegas, P.; Membrillo-Hernández, J. Using an adaptive learning tool to improve student performance and satisfaction in online and face-to-face education for a more personalized approach. *Smart Learn. Environ.* **2024**, *11*, 6. [CrossRef]
48. Al-Khayat, M.R.; Gargash, M.U.; Atiq, A.F. The Effectiveness of Game-based Learning in Enhancing Students' Motivation and Cognitive Skills. *J. Educ. Teach. Methods* **2023**, *2*, 50–62. Available online: https://gprjournals.org/journals/index.php/jetm/article/view/199/231 (accessed on 26 July 2025).
49. Sarlis, I.; Kotsifakos, D.; Douligeris, C. Modern methods for evaluating educational games: From usage analysis to long-term impact on learning. In Proceedings of the 5th Panhellenic Conference on "Educational Games and Art in Education and Culture", Athens, Greece, 17–19 October 2024; ISBN 978-618-5458-82-9. [CrossRef]
50. Sarlis, I.; Kotsifakos, D.; Zacharias, P. Creating a Novice Programming Environment for Learning LOGO. In Proceedings of the 11th Conference on Informatics in Education, Piraeus, Athens, 29 September–1 October 2019; ISBN 978-960-578-058-6. Available online: https://schl.gr/papers/CIE_2019.pdf (accessed on 22 July 2025).
51. Sarlis, I.; Antonopoulou, S.; Antoniou, M.; Kotsifakos, D.; Douligeris, C. Leveraging digital tools for immersive gamified learning in the Old City of Rhodes. In Proceedings of the IEEE Global Engineering Education Conference (EDUCON), Kos Island, Greece, 8 May 2024. [CrossRef]
52. Petratos, S.; Botha, R.; Mtshabe, M. Digital Skills' Impact on the Use of Digital Learning Tools. In Proceedings of the 23rd European Conference on e-Learning (ECEL 2024), Pretoria, South Africa, 24–25 October 2024. [CrossRef]
53. Kassi, D.; Sarlis, I.; Kotsifakos, D.; Douligeris, C. An Educational Cultural Game for an Industrial Urban Landscape. In Proceedings of the 21st European Conference on e-Learning (ECEL); University of Brighton: Brighton, UK, 2022; Volume 21, pp. 171–180. [CrossRef]
54. Center for Applied Special Technology (CAST). Universal Design for Learning Guidelines. Available online: https://udlguidelines.cast.org/ (accessed on 10 July 2025).
55. Efthymiou, L.; Epaminonda, E.; Ktoridou, D.; Dionysiou, I. Societal and ethical implications of technology in education. In Proceedings of the 2023 IEEE Global Engineering Education Conference (EDUCON), Kuwait City, Kuwait, 1–4 May 2023. [CrossRef]
56. Liu, Y.; Ma, S.; Chen, Y. The impacts of learning motivation, emotional engagement, and psychological capital on academic performance in a blended learning university course. *Front. Psychol.* **2024**, *15*, 1357936. [CrossRef] [PubMed]
57. Sun, X.; Sailer, M. Students' behavioural, cognitive and affective outcomes in gamified flipped classrooms: A meta-analysis. *Rev. Educ.* **2025**, *13*, e70039. [CrossRef]
58. Esterhazy, R.; Kaliisa, R.; Sanchez, D.; Langford, M.; Damsa, C. Advancing Multimodal Collaboration Analytics: A Scoping Review. *J. Learn. Anal.* **2025**, 1–20. [CrossRef]
59. Nadeem, M.; Oroszlanyova, M.; Farag, W. Effect of Digital Game-Based Learning on Student Engagement and Motivation. *Computers* **2023**, *12*, 177. [CrossRef]
60. Johnson, A.; Johnson, B. *Adaptive Learning Technologies and Their Impact on Education*; Academic Press: New York, NY, USA, 2022.
61. Jonas, J.M.; Ogodo, J. Framework as a Process: A User-Centric Conceptual Framework for Game-Based Learning. *J. Res. Sci. Math. Technol. Educ.* **2025**, *8*, 335–355. [CrossRef]
62. Dabbous, M.; Kawtharani, A.; Fahs, I.; Sakr, F. The role of game-based learning in experiential education: Tool validation, motivation assessment, and outcomes evaluation among a sample of pharmacy students. *Educ. Sci.* **2022**, *12*, 434. [CrossRef]
63. Sweetser, P.; Wyeth, P. GameFlow: A model for evaluating player enjoyment in games. *Comput. Entertain.* **2005**, *3*, 3. [CrossRef]
64. Kay, R.H.; Knaack, L. Assessing Learning, Quality and Engagement in Learning Objects: The Learning Object Evaluation Scale for Students (LOES-S). *Educ. Technol. Res. Dev.* **2009**, *57*, 147–168. [CrossRef]
65. Petri, G.; Gresse von Wangenheim, C.; Borgatto, A.F. MEEGA+, a systematic model to evaluate educational games. In *Encyclopedia of Computer Graphics and Games*; Lee, N., Ed.; Springer: Cham, Switzerland, 2024. [CrossRef]
66. Mikropoulos, T.A.; Papachristos, N.M. SciLOET: A Framework for Assessing Digital Learning Objects for Science Education. In *Technology and Innovation in Learning, Teaching and Education*; Reis, A., Barroso, J., Lopes, J.B., Mikropoulos, T., Fan, C.W., Eds.; Springer: Cham, Switzerland, 2021; pp. 339–350. [CrossRef]
67. Zacharis, G.K.; Mikropoulos, T.A. Evaluating digital learning objects in science education with the Science Learning Objects Evaluation Tool. In *Research on E-Learning and ICT in Education*; Bratitsis, T., Ed.; Springer: Cham, Switzerland, 2023; pp. 3–18. [CrossRef]
68. Rivas Ginel, M.I.; Theroine, S. Gender in Video Games: Creativity for Inclusivity. *arXiv* **2024**, arXiv:2407.16536. [CrossRef]

69. Mygdanis, Y. Digital Little Music Makers: Designing and implementing an innovative educational music proposal embodying the concept of "making". In Proceedings of the Constructionism 2025: Building Communities, Bridging Ideas, Zurich, Switzerland, 24–27 June 2025; Volume 8. [CrossRef]
70. Dai, Y.; Liu, S.; Xu, G. Enhancing human-machine pair inspection with risk number and code inspection diagram. *Softw. Qual. J.* **2024**, *32*, 939–959. [CrossRef]
71. Helda, N.; Sanusi, S.; Suhendra, R. Innovative educational game: Designing and developing interactive web-based memory card games as an educational medium for children. *J. Smart Comput. Artif. Intell.* **2024**, *1*, 320–326. [CrossRef]
72. Martinez, K.; Menéndez-Menéndez, M.I.; Bustillo, A. A New Measure for Serious Games Evaluation: Gaming Educational Balanced (GEB) Model. *Appl. Sci.* **2022**, *12*, 11757. [CrossRef]
73. Pietriková, E.; Sobota, B. Game Development and Testing in Education. In *Game Theory: From Idea to Practice*; Sobota, B., Ed.; IntechOpen: London, UK, 2022. [CrossRef]
74. Osipovskaya, E.; Tasi, P.; Green, M.; Lukac, D. The State of Research on Effects of Gamification in the Context of Learning. In Proceedings of the New Perspectives in Science Education (12th Edition), Hybrid Event, 16–17 March 2023; pp. 138–143, ISBN 979-12-80225-55-9. Available online: https://kar.kent.ac.uk/100567/ (accessed on 25 July 2025).
75. Volden, T.; Grbić, D.; Burelli, P. Procedurally generating rules to adapt difficulty for narrative puzzle games. *arXiv* **2023**. [CrossRef]
76. Verville, L.; Côté, P.; Grondin, D.; Mior, S.; Kay, R. The development and evaluation of an online educational tool for the evidence-based management of neck pain by chiropractic teaching faculty. *J. Chiropr. Educ.* **2021**, *35*, 95–105. [CrossRef]
77. Nguyen, H.A.; Else-Quest, N.; Richey, J.E.; Hammer, J.; Di, S.; McLaren, B.M. Gender differences in learning game preferences: Results using a multi-dimensional gender framework. In *Artificial Intelligence in Education. AIED 2023*; Wang, N., Rebolledo-Mendez, G., Matsuda, N., Santos, O.C., Dimitrova, V., Eds.; Springer: Cham, Switzerland, 2023; pp. 533–545. [CrossRef]
78. Li, M.; Ma, S.; Shi, Y. Examining the effectiveness of gamification as a tool promoting teaching and learning in educational settings: A meta-analysis. *Front. Psychol.* **2023**, *14*, 1253549. [CrossRef]
79. Sharma, D.; Das, A. The role of multimodal strategies in enhancing cognitive and meta-cognitive skills in middle school learners: A psychological exploration. *Int. J. Res. Publ. Rev.* **2024**, *5*, 2098–2103. [CrossRef]
80. Shukla, H.; Pandey, K. Human-AI Collaboration in Teaching and Learning. *Int. J. Sci. Soc. Sci. Res. (IJSSSR)* **2025**, *2*, 367–374. Available online: https://www.researchgate.net/publication/391277461_Human-AI_Collaboration_in_Teaching_and_Learning (accessed on 22 July 2025). [CrossRef]
81. Reichert-Schlax, J.; Zlatkin-Troitschanskaia, O.; Frank, K.; Brückner, S.; Schneider, M.; Müller, A. Development and evaluation of digital learning tools promoting applicable knowledge in economics and German teacher education. *Educ. Sci.* **2023**, *13*, 481. [CrossRef]
82. Aragonez, T.; Saur-Amaral, I.; Gouveia, M. Game-based learning in higher education: A systematic literature review. In Proceedings of the EDULEARN21 Conference, Online, 5–6 July 2021. [CrossRef]
83. Tarng, W.; Lin, Y.-J.; Ou, K.-L. A virtual experiment for learning the principle of the Daniell cell based on augmented reality. *Appl. Sci.* **2021**, *11*, 762. [CrossRef]
84. Pacheco-Velazquez, E.; Aguilar-Ávalo, N. Student's experience in using simulators for the development of complex skills in logistics education: Assessing enjoyment of learning. *Smart Learn. Environ.* **2022**, *9*, 18. [CrossRef]
85. Udeozor, C.; Chan, P.; Russo Abegão, F.; Glassey, J. Game-based assessment framework for virtual reality, augmented reality and digital game-based learning. *Int. J. Educ. Technol. High. Educ.* **2023**, *20*, 36. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

Multimodal Technologies and Interaction

MDPI

Article

Designing Gamified Intergenerational Reverse Mentorship Based on Cognitive Aging Theory

Sijie Sun

Department of Philosophy, Autonomous University of Barcelona, 08193 Barcelona, Spain; 1583451@uab.cat

Abstract: With the intensification of population aging, elderly individuals face significant barriers in learning digital skills, leading to a widening digital divide. Based on Cognitive Aging Theory, this study proposes and validates an original paradigm called Structured Gamified Intergenerational Digital Reverse Mentorship to enhance digital literacy among the elderly. Cognitive Aging Theory suggests that due to declines in memory, attention, and executive function, older adults encounter challenges when learning new technologies, while gamified learning combined with intergenerational interaction can help reduce cognitive load and increase learning motivation. This study designed a collaborative gamified digital reverse mentorship application, "Digital Bridge", and employed a randomized controlled trial method, assigning 90 participants aged 60 and above into three groups: the traditional digital mentorship group (Group A), the independent gamified learning group (Group B), and the collaborative gamified digital mentorship group (Group C). Each intervention session lasted 30 min and was conducted in a controlled environment. Experimental results showed that Group C significantly outperformed Groups A and B in digital skill acquisition, user experience, and learning motivation ($p < 0.001$), indicating that the combination of gamified learning and intergenerational interaction effectively enhances learning interest, reduces learning anxiety, and improves skill transferability. This study provides a new approach to elderly digital literacy education and offers theoretical and practical support for the design of future age-friendly digital learning tools.

Keywords: cognitive aging theory; user experience design; age-friendly design; gamification design; digital literacy for older adults

1. Introduction

As population aging intensifies, elderly individuals are facing increasingly severe challenges related to the digital divide, and significant barriers in using smart devices and digital technologies [1,2]. Although the Internet penetration rate among the elderly population has gradually increased, cognitive aging has led to a decline in learning capability, and the lack of age-appropriate learning resources continues to pose significant challenges in digital skill acquisition [1,3,4]. Therefore, it is of great practical significance to narrow the digital divide for the elderly, enabling them to better adapt to the digital society and benefit from the convenience and opportunities offered by digital technology.

Intergenerational digital reverse mentorship fosters the intergenerational transfer of digital skills through interactions between younger and older generations, and serves as a potentially effective intervention strategy. Research has shown that digital games can not only enhance the digital literacy among the elderly, but also improve memory, motor coordination, and cognitive function [5–7]. However, current digital reverse mentorship

models largely rely on the support of family or community, and lack systematic and sustainable learning mechanisms. Therefore, they fail to meet the growing demand for digital skill acquisition among the elderly [8]. Moreover, existing digital skills training methods for the elderly primarily depend on offline courses or informal assistance from family members, lacking systematicity, engagement, and personalization, and mainly focusing on resolving specific operational issues rather than providing comprehensive support for enhancing digital literacy. Although the elderly show a strong interest in leisure and cognitive games [6,9], the market still lacks digital training tools specifically designed for the elderly, especially gamified products. Existing research predominantly focuses on adapting general applications for elderly users, but lacks targeted design tailored to their cognitive characteristics and learning needs. Therefore, it is of significant research value to develop a structured and gamified digital reverse mentorship program based on Cognitive Aging Theory to enhance digital literacy, optimize user experience, and increase learning motivation among the elderly.

Based on the theories of cognitive aging [10,11] and user experience design [12,13], the present study developed a gamified learning app for digital reverse mentorship tailored for elderly users. Cognitive Aging Theory suggests that the elderly face considerable challenges in performing complex tasks due to the decline in memory, attention, and reaction speed. Therefore, gamification design was employed to decompose complex digital learning tasks into simple and progressive micro-tasks, thereby gradually increasing the difficulty through level-based mechanisms. Additionally, instant feedback, virtual rewards, and social interactions were incorporated to reduce cognitive load, enhance engagement, and improve learning participation. Furthermore, based on user experience design theory, interface layout, operational processes, and interaction methods were optimized to align with the cognitive characteristics and usage habits of the elderly. Such system-supported design by theory not only improves the usability and attractiveness of the app, but also enhances the positive emotions and satisfaction of elderly users, providing them with an efficient and enjoyable digital learning experience and promoting digital skill acquisition and practical application.

The design was created using the ADDIE (Analysis, Design, Development, Implementation, and Evaluation) model [14], incorporating survey research, data analysis, prototype design, and experimental design to systematically evaluate the effectiveness of gamified digital reverse mentorship. A total of 90 participants aged 60 and above were recruited and randomly assigned to three different learning model groups. Then, the study evaluated user experience and learning motivation among elderly participants across different learning models using the UEQ-S (User Experience Questionnaire—Short Version) and the ARCS Model of Motivation scale while assessing learning outcomes through repeated task performance tests [15–17]. The primary objective of this study is to design an app and evaluate its effectiveness on gamified digital reverse mentorship interventions in enhancing digital literacy among older adults and to explore its potential in optimizing user experience and increasing learning motivation.

This study merges the application of Cognitive Aging Theory and user experience design theory in enhancing digital literacy among the elderly, offering new academic insights into intergenerational education and digital skills training. Practically, this study develops a gamified digital reverse mentorship learning app, filling the market gap in digital skills training tools for older adults while improving engagement.

The research findings provide scientific evidence for optimizing digital education for older adults and offer valuable references for governments, educational institutions, and

social organizations in formulating digital inclusion policies. This, in turn, helps bridge the digital divide, fosters intergenerational interaction, and promotes social harmony.

2. Literature Review

2.1. Background Dimension: Older Adults' Digital Literacy and Intergenerational Digital Reverse Mentorship in the Context of the Digital Divide

The three-level model of the digital divide [18] reveals the core contradictions in the digital challenges faced by the elderly. After overcoming physical access barriers (first level), they encountered a "usage divide" (second level), characterized by difficulties in digital skill acquisition and insufficient intrinsic motivation, ultimately resulting in significant generational differences in digital resource acquisition outcomes (third level). The current research presents two opposing paradigms to explain this phenomenon: the individual deficiency paradigm emphasizes the limitations imposed by cognitive aging [19] and low self-efficacy [20], advocating for the improvement of cognitive reserves through repetitive training. In contrast, the structural constraint paradigm criticizes the imbalance in social resource allocation [21] and the lack of intergenerational support [22,23], calling for the reconstruction of technology empowerment networks. However, existing studies typically isolate the interaction between individuals and systems, failing to propose an integrated approach that combines cognitive intervention with social support.

Intergenerational Digital Reverse Mentorship offers an innovative solution to this issue. "Mentorship" generally refers to a structured relationship in which a more experienced individual (the mentor) provides guidance and support to a less experienced person (the mentee) in areas such as skills, experience, emotional development, and career advancement [24]. "Digital Reverse Mentorship," on the other hand, specifically refers to a mentorship model in the technological domain where younger, more tech-savvy individuals mentor older adults, thereby reversing traditional generational roles [25]. This mechanism not only facilitates technical learning, but also fosters emotional support and social connection. Research has indicated that technology transfer from younger to elderly individuals not only alleviates cognitive decline [26,27], but also reduces technology-related anxiety through emotional bonding [28], proving to be more effective than unidirectional teaching methods [29]. However, existing studies primarily focus on the social function of mentorship, while overlooking its potential as a systematic intervention tool. Therefore, this study introduced a gamified intergenerational mentorship model that integrates cognitive training within a social support network, simultaneously targeting individual cognitive deficits (e.g., decreased working memory) and structural support gaps (e.g., insufficient intergenerational interaction opportunities), thereby addressing a key gap in the intervention mechanisms of the digital divide.

Research on digital literacy of the elderly stems from a structural reflection on the digital divide, grounded in both Cognitive Aging Theory and social resource distribution theory. Cognitive Aging Theory [3] suggests that due to memory decline, attention deficits, and slower information processing speeds, the elderly experience severe cognitive overload in complex digital tasks, leading to "technological avoidance" [1]. Social resource distribution theory [18] highlights that low-income and low-education elderly populations are trapped in a "digital exclusion" cycle due to systemic resource deprivation, such as the lack of age-friendly tools and training opportunities. The data indicate that 25% of individuals aged 65 and above remain in a state of "digital isolation" [30,31], and digital skill deficiencies not only limit their access to health information [32], but also significantly correlate with worsening depressive symptoms [33]. Although the COVID-19 pandemic has accelerated the improvement of digital literacy among the elderly, skill growth remains

uneven, and some groups still lack adequate technological support, highlighting the necessity for sustained intervention and structured learning. Additionally, the existing research primarily examines digital literacy barriers from physiological and psychological perspectives, neglecting the potential of intergenerational collaboration as an intervention strategy, thereby limiting an in-depth understanding of how collaborative learning enhances digital literacy among the elderly.

Intergenerational digital reverse mentorship, where the younger generation transfers digital skills to the elderly, offers a promising solution to this issue. Studies have shown that involving younger individuals in technology training for the elderly can foster a supportive learning environment and effectively meets their unique needs [34]. Moreover, integrating cultural sensitivity can enhance technology acceptance [35], while collaborative learning models provide personalized educational support and boost the confidence of the elderly in technology use [30]. However, current digital mentorship practices still have several limitations. Cota et al. (2015) and Marston (2013) pointed out that although informal family interactions can boost the technological confidence of the elderly, they rely on sporadic support and lack structured learning frameworks and systematic assistance [6,9,36]. Existing learning tools (such as tutorials and manuals) have dismal interactivity and contextual detachment [37], failing to encourage sustained participation. The elderly tend to prefer relaxed and interactive learning environments [38]. "Gamification" refers to the application of game elements (such as points, challenges, and real-time feedback) in non-game contexts to enhance user engagement, motivation, and learning outcomes [39]. In the context of digital literacy training, gamification can help reduce learning anxiety, increase enjoyment, and promote emotional involvement [40]. Cognitive games have been shown to improve their cognitive capabilities and learning motivation through task segmentation and instant feedback [5]. However, most market products focus on general cognitive training (such as Sudoku) rather than deep integration of digital skill development [41].

Although intergenerational collaboration programs such as Cyber-Seniors have demonstrated a certain success in higher education settings [42], their linear teaching models lack dynamic adaptability, and resource allocation disparities limit scalability [43]. More importantly, existing studies present several disjointed aspects: (1) the analysis of digital literacy barriers and intergenerational interaction theories remain separate, and have not formed an integrated framework; (2) intervention practices rely on informal settings, making it difficult to achieve sustainability [9]; and (3) technological design efforts focus primarily on interface simplification while neglecting cognitive load management and emotional motivation [1]. Despite the importance of addressing individual barriers, integrating intergenerational collaboration not only enhances the effectiveness of digital literacy interventions, but also fosters a more inclusive digital society.

In the present study, the "Structured Gamified Intergenerational Digital Reverse Mentorship" paradigm was proposed to overcome these limitations. Through progressive digital tasks (from basic operations to complex skills), a bidirectional intergenerational collaboration module (real-time voice/text guidance), and a multidimensional incentive system (points rewards, emotional feedback), a threefold intervention framework of "cognitive adaptation–emotional engagement–skill internalization" was constructed, and provided a scientific basis and practical solution for bridging the digital divide.

2.2. Theoretical Dimension: The Dual-Drive Framework of Cognitive Aging Theory and User Experience Design Theory

Cognitive Aging Theory and User Experience Design Theory play complementary roles in the digital skill learning process for the elderly. Cognitive Aging Theory focuses on

the changes in cognitive function that occur with aging, particularly their impact on the execution of complex tasks [10,12]. Research has indicated that the elderly experience a gradual decline in controlled processing, working memory, and cognitive reserve, leading to increasingly challenging learning and operating in digital environments [44]. The changes in controlled processing make it more difficult for the elderly to adapt to new tasks, while slower processing speed and reduced short-term memory capacity increase the likelihood of errors or cognitive overload in attention-intensive tasks [45,46]. Additionally, the elderly have a diminished ability to filter out irrelevant information, making them more susceptible to distractions and thus reducing learning efficiency [47]. As cognitive reserve declines, they are more likely to feel anxious or frustrated when facing new technologies or complex operations [48,49]. Therefore, when designing digital learning tools, it is essential to provide adaptive support tailored to cognitive characteristics of the elderly to reduce cognitive load, enhance learning efficiency, and improve user experience.

User Experience Design (UXD) focuses on user needs, and emphasizes interaction experience, operational convenience, and emotional resonance to optimize product usability and accessibility [9]. Unlike traditional instructional design, UXD places greater emphasis on interactive fluidity, the integration of functionality, and emotional connection, ensuring a positive user experience before, during, and after product usage. For example, a digital reverse mentorship game for the elderly can incorporate personalized tutorials and step-by-step demonstrations to provide a progressive learning path based on users' initial proficiency level, so as to prevent frustration caused by steep learning curves [50]. Additionally, UXD prioritizes real-time feedback mechanisms, such as visual or auditory rewards upon task completion, to enhance users' sense of accomplishment and willingness to continue engaging. Designing a system that dynamically adapts to user needs can not only improve the usability of gamified applications, but also effectively promote the motivation of the elderly to explore digital technology proactively.

The integration of Cognitive Aging Theory and User Experience Design Theory has formed a framework of "Cognitive Adaptation–Emotional Drive–Skill Internalization", providing a dual-compensation pathway for digital skill learning for the elderly. First, in the cognitive adaptation phase, the system reduces initial learning barriers through segmented tasks, thereby decreasing cognitive load to enable the elderly to gradually become familiar with digital operations. Second, in the emotional drive phase, intergenerational social incentives (such as virtual badge sharing and collaborative tasks) are introduced to enhance user motivation by making the learning process more socially valuable and interactive. Finally, in the skill internalization phase, scenario-based simulation tasks (such as online shopping and social media interactions) reinforce skill transfer, allowing the elderly to flexibly apply their acquired digital skills in their daily life. Neuroscientific research supports this framework, as fNIRS (functional near-infrared spectroscopy) studies have shown that gamified learning activates both the dorsolateral prefrontal cortex (responsible for cognitive control) and the ventral striatum (involved in reward processing), indicating that gamification effectively enhances learning outcomes for the elderly [51].

In summary, Cognitive Aging Theory provides a theoretical foundation for cognitive support in digital skill learning, while User Experience Design optimizes interaction methods to enhance learning experiences and emotional engagement. The combination of these two theories overcomes the limitations of single-theory approaches. For instance, pure cognitive training may lead to low compliance due to its tedious nature [52], while isolated emotional incentives lack a physiological basis for cognitive improvement. Building on this theoretical framework, this study posits that a collaborative gamified model, by integrating cognitive adaptation mechanisms, intergenerational guidance, and contextualized

task design, will outperform traditional mentorship and independent learning models in improving digital skill acquisition among older adults. Therefore, we propose Hypothesis 1 (H1) and Hypothesis 2 (H2):

H1: *The Collaborative Gamified Digital Reverse Mentorship model (Model C) is more effective in enhancing digital skill acquisition among older adults than the Traditional Digital Reverse Mentorship model (Model A) and the Independent Gamified Learning model (Model B).*

H2: *The Collaborative Gamified Digital Reverse Mentorship model (Model C) provides a better user experience than the Traditional Digital Reverse Mentorship model (Model A) and the Independent Gamified Learning model (Model B).*

2.3. Technical Dimension: The Integration of Gamification and Aging-Friendly Design

Aging-Friendly Design and Gamification Theory exhibit deep complementarity in the field of digital skills training for the elderly. The aging-friendly design holds a primary goal to reduce cognitive load for elderly users and promote them to adapt to and use digital technology, which is achieved through measures such as high-contrast visual encoding, simplified menu structures, and fewer interaction steps to optimize the user experience [53]. Although these measures have been proven to effectively enhance perceived ease of use (PEU), studies have indicated that optimizing the interface alone is insufficient to improve the learning motivation and skill retention of the elderly [1]. Based on the Technology Acceptance Model (TAM), research has suggested that the sustained learning of elderly users depends not only on perceived ease of use, but also on perceived usefulness (PU) and hedonic motivation (HM) [54] In other words, the elderly require not only easy-to-use interfaces, but also engaging and motivating learning experiences to sustain long-term use.

Compared with traditional aging-friendly optimizations, gamification strategies can reduce cognitive load while enhancing engagement and learning persistence. Gamification Theory enhances user retention by incorporating mechanisms such as task segmentation, instant feedback, and virtual rewards, ensuring that the elderly can maintain cognitive engagement throughout the learning process [7,55]. Task segmentation breaks down complex digital skills into progressive challenges, allowing elderly users to gradually adapt to and master new skills [6]. Instant feedback provides positive reinforcement during the learning process and increases self-efficacy [7,42], while virtual rewards (e.g., points and badges) leverage the goal-gradient effect [56] to boost motivation. However, most existing gamified applications for the elderly primarily focus on cognitive training (e.g., Sudoku, memory training) [57] and lack real-life application scenarios despite their ability to enhance cognition, thereby limiting the ability of the elderly to transfer learned skills into daily digital interactions [55,58]. In contrast, contextualized game design can strengthen the practical relevance of skill acquisition. For example, by simulating tasks such as online medical consultations, social media interactions, and digital payments, the elderly can learn digital skills in familiar environments and improve skill transfer and long-term retention [58]. Furthermore, neuroimaging studies suggest that contextualized game design can enhance functional connectivity between the hippocampus and prefrontal cortex, potentially delaying cognitive decline in the elderly [5,59]. Therefore, embedding gamification mechanisms into aging-friendly design not only lowers learning barriers, but also enhances learning motivation and long-term skill retention.

Based on aging-friendly design and gamification strategies, intergenerational interaction technologies further optimize digital learning experiences of the elderly. Research has indicated that teaching digital skills to the elderly by young individuals not only

reduces their technology anxiety, but also enhances learning experiences through social interaction [27]. However, traditional intergenerational mentorship models typically rely on offline support from family members or volunteers and lack systematic and sustainable technological support, leading to individual learning discrepancies among the elderly [9]. To address this issue, recent studies have explored how technology can facilitate intergenerational interaction, such as real-time collaboration and remote guidance features, so as to enable younger users to provide digital assistance to the elderly at any time [28]. Additionally, gamified task-sharing and collaborative learning models allow younger and older participants to engage in learning tasks together, thereby facilitating knowledge transfer and increasing learning motivation for the elderly. The integration of artificial intelligence (AI) further optimizes the learning experience. For instance, intelligent recommendation systems can tailor learning tasks based on progress, provide personalized teaching suggestions, reduce frustration, and ultimately improve the digital skill mastery rate among elderly users.

In summary, traditional aging-friendly interface optimizations alone cannot fully meet the long-term digital learning needs of the elderly. Instead, integrating gamification mechanisms with intergenerational interaction offers a more effective approach by enhancing motivation, improving skill transferability, and fostering long-term retention of digital skills. In this study, a systematic and sustainable training model for the elderly was explored by incorporating gamification strategies into aging-friendly design, coupled with intergenerational collaboration, with the aim to help the elderly better adapt to digital environments and promote sustainable technology adoption. The combination of gamified learning and intergenerational collaboration is expected to foster both extrinsic and intrinsic learning motivations among elderly users, as supported by motivation theories and prior empirical findings. Hence, we formulate Hypothesis 3 (H3):

H3: *The Collaborative Gamified Digital Reverse Mentorship model (Model C) leads to higher learning motivation compared to the Traditional Digital Reverse Mentorship model (Model A) and the Independent Gamified Learning model (Model B).*

3. Materials and Methods

3.1. Research Design

This study adopts a design-based research methodology, systematically guided by the ADDIE instructional design model, to develop and evaluate a gamified intergenerational mentorship learning application, "Digital Bridge", with the aim of enhancing digital literacy among older adults. Rather than treating the application design itself as the primary research objective, this study focuses on evaluating the effectiveness of a collaborative gamified mentorship mechanism in a real-life learning context, grounded in both cognitive and user-experience theories.

Originally developed in the field of instructional design, the ADDIE model provides a structured process for creating effective and learner-centered educational programs. The ADDIE model consists of five stages, Analysis, Design, Development, Implementation, and Evaluation, as illustrated in Figure 1. This model is widely recognized as an effective methodological framework for developing training programs, ensuring scientific rigor, systematic organization, and practical applicability in instructional content [14].

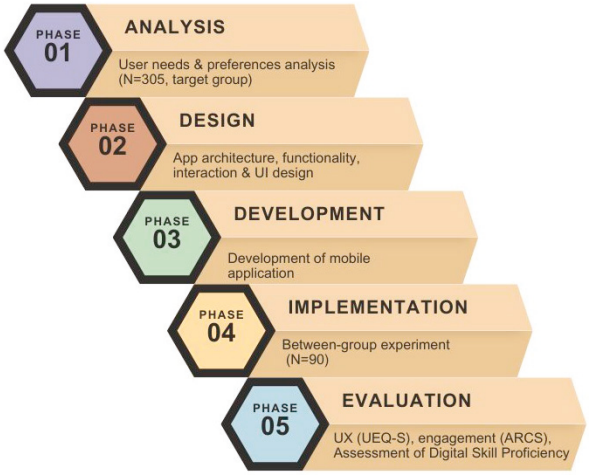


Figure 1. Mobile application development process using the ADDIE model.

3.1.1. Phase One: Analysis Phase

During the analysis phase, this study combined questionnaire surveys and focus group discussions to comprehensively understand older adults' digital skills status, needs, barriers, and preferences, providing data support and practical guidance for the design of the "Digital Reverse Mentorship" gamified application.

First, to ensure the scientific validity and content accuracy of the questionnaire, the research team consulted three gerontology professors (3), two design professors (2), two elderly university instructors (2), and one community center staff member (1). Based on their suggestions, a preference and needs assessment questionnaire was developed, covering six sections: basic information, current digital skills, digital skill needs, learning barriers, willingness for gamified learning, and open-ended suggestions.

Data collection was conducted in Taiyuan, Shanxi Province, China, via the "Wenjuanxing" online survey platform from November 15 to 5 December 2024, with 305 older adults (aged 60+) participating. All respondents provided written informed consent. To enhance comprehension, researchers explained each question during the survey process to ensure older participants could respond accurately.

To further explore user needs and ensure a multidimensional research approach, a focus group discussion was organized at a community center in Taiyuan, Shanxi Province. The discussion involved eight participants from diverse backgrounds, including older adults (2), family members (2), community center staff (1), elderly university instructors (1), experienced aging-friendly designers (1), and gerontology experts specializing in the digital divide (1). The discussion focused on the following key topics:

① How can gamification help older adults overcome learning barriers?

② How can game-based learning content and interaction be optimized to reduce learning anxiety, boost confidence, and promote intergenerational interaction?

③ What is the specific role and impact of younger individuals in assisting older adults with digital skill learning?

④ What is the current state of older adults' digital literacy, and what are the strengths and limitations of existing "digital reverse mentorship" models?

⑤ What is the potential of cognitive games in improving older adults' digital literacy and learning motivation?

The discussion followed an open-ended format, encouraging participants to share real experiences and insights. Sufficient background information was provided to facilitate meaningful dialogue. By incorporating perspectives from different stakeholders, the study

refined the design direction of the gamified intergenerational digital mentorship application "Digital Bridge", ensuring that it meets the practical needs of older users while also achieving broader social and educational objectives, such as enhancing intergenerational interaction and improving the effectiveness and sustainability of digital skills learning.

3.1.2. Phase Two and Three: Design and Development Phase

Based on the findings from the analysis phase, including user needs exploration and focus group interviews, and guided by Cognitive Aging Theory and Aging-Friendly Experience Design, this study aimed to develop a "Digital Reverse Mentorship" application, "Digital Bridge", that balances practicality and engagement. The core design philosophy of this application is to accommodate older users' cognitive, emotional, and social needs by using gamification strategies to lower learning barriers, stimulate intrinsic motivation, and promote active intergenerational interaction. To achieve these goals, the application underwent comprehensive optimization in terms of workflow and structural design, UI (user interface) enhancements, interaction modes, and feedback mechanisms, creating a holistic digital literacy learning platform that combines education, entertainment, and social engagement to help older adults integrate into the digital world.

This study introduced "Intergenerational Digital Reverse Mentorship" as the core concept of the application, emphasizing the use of digital technology to bridge the intergenerational digital divide by enabling younger users to pass on digital skills to older adults while fostering mutual learning and interaction. Based on this concept, the application's functional design includes two user modes: the "Digital Learner" mode (for seniors) and the "Tech Mentor" mode (for juniors). It integrates four key modules: a digital skills learning module, an intergenerational interaction module (e.g., a "Call for Help" feature), a gamified reward system (including points, virtual achievements, and emotional rewards), and an optional AI voice assistant (providing game hints or acting as a virtual junior). The workflow design follows standard operational logic such as onboarding screens for first-time users, registration pages, and senior–junior binding pages, while ensuring a clear and progressive learning path from simple to complex and from basic to advanced content, with logical continuity between topics. The interface is simple and intuitive, tailored to the cognitive characteristics of older users, for example, by reducing the number of interactive elements per page and enhancing their visibility. A "Call for Help" function allows older users to request assistance from their junior counterparts when facing difficulties, who can then provide real-time guidance via voice or text, enhancing interactivity and collaboration. Additionally, a task challenge mechanism is introduced, encouraging older users to complete various digital skill challenges to earn achievement-based rewards, thereby boosting learning engagement and motivation.

In UI (user interface) and interaction design, this study optimized the layout based on Cognitive Aging Theory to make the application more intuitive, simplified, and user-friendly. The interface uses large fonts, high-contrast color schemes, and clear navigation structures to ensure that older users can browse and operate the app effortlessly. Additionally, for critical operations (such as payments and information submissions), confirmation dialogs were introduced to reduce the risk of errors and improve safety and reliability. To enhance learning interest and user experience, gamification-based interactions were further optimized. For instance, upon completing a task, users receive instant feedback (e.g., animated prompts, voice guidance), helping them track their progress and achievements while boosting confidence. At the same time, operation processes were simplified, reducing multi-level menus and improving touch interactions to ensure that tasks can be completed with minimal steps.

In the development phase, Android Studio was used as the primary development environment. Given the operational habits of older users, the application's interface interactions adopted a "tap-to-enter" navigation method, allowing users to easily access different functional modules without complex operations. Moreover, this study employed an iterative feedback loop model, collecting continuous user feedback throughout the development process to optimize operational fluency, interaction experience, and visual design. For example, early user testing revealed that some older adults were unfamiliar with swipe gestures, leading to an adjustment to a "single-tap confirmation" mode to reduce learning difficulty and error rates. Additionally, voice-assisted functions were incorporated into key learning tasks, further reducing cognitive load and improving interactive accessibility.

At present, the application has only completed the initial stage of basic development, with five levels created for testing purposes. Core features such as account registration, senior–junior pairing, interaction, and real-time communication have not yet been implemented. Therefore, during the testing phase of this study, each group used an Android device with the test version of the app installed. To simulate the guidance from juniors to seniors, a volunteer was assigned to observe the senior's gameplay in person. When the senior encountered difficulties and pressed the app's "Request Help" button, the on-site volunteer would provide prompts and guidance, simulating the app's intended intergenerational interaction functionality.

3.1.3. Phases Four and Five: Implementation and Evaluation

This study adopted a between-group experimental design to evaluate the effectiveness of the collaborative gamified digital mentorship application in improving digital literacy among older adults. The research focused on measuring digital skill acquisition and examining the impact of collaborative gamification on user experience and learning motivation.

First, a total of 90 older adults aged 60 and above were recruited from a senior university in Taiyuan, Shanxi Province, China. Participants were recruited after we obtained permission from the university administration. Recruitment was conducted through both online announcements and offline flyers distributed on campus. Inclusion criteria included: age $\geq$ 60 years, basic literacy skills, voluntary participation with signed informed consent, and confirmation in the pre-test that they had not yet mastered the seven digital skills involved in the experiment. To ensure the representativeness of the sample, recruitment considered factors such as educational background, digital skill levels, and smartphone usage experience. Exclusion criteria included: (1) individuals with severe cognitive impairment, mental disorders, or physical conditions that hinder learning, as well as those unable to complete a continuous two-hour learning task; (2) older adults who had participated in similar digital skills training in the past year. No participants dropped out during the experiment. To encourage participation, each participant received a small gift valued at approximately 30 RMB as a token of appreciation after completing the sessions.

Second, this study adopted a three-arm parallel randomized controlled trial (RCT) design, with participants randomly assigned to one experimental group and two control groups (N = 30 per group), each receiving a different learning intervention (Figure 2).

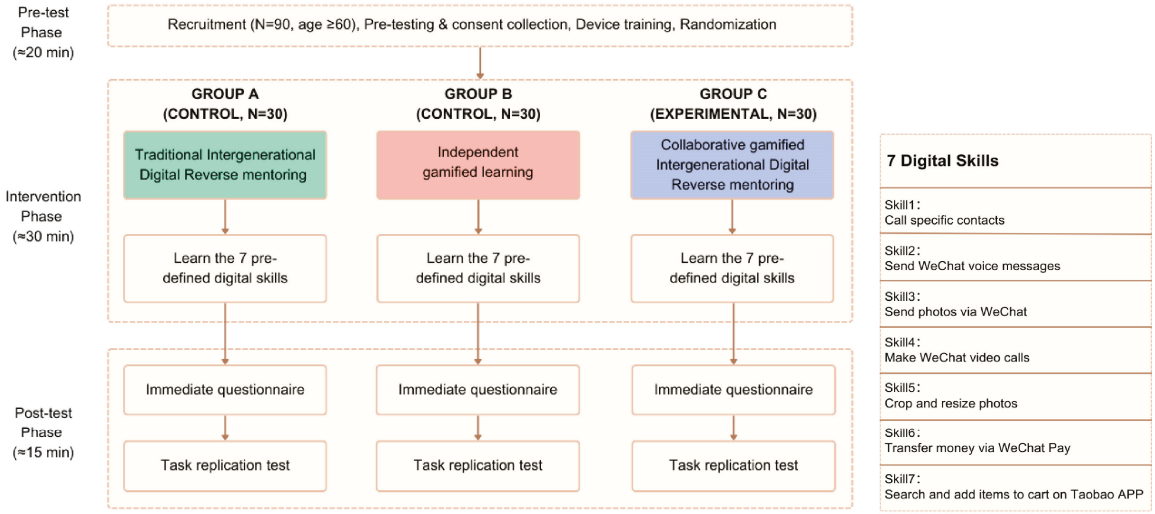


Figure 2. Experimental procedure diagram and seven digital skills.

Since intergenerational interaction was involved in the experiment, the "tech mentor mode" was carried out by trained staff members. All mentors completed a two-day standardized training program, which covered the digital skills instructional script, communication techniques, and strategies for supporting older adults in learning. After the training, mentors were required to pass a teaching simulation assessment to ensure consistency in instructional content and methods, as well as the ability to adapt flexibly to learners' progress.

The independent variable in this study was the three learning intervention modes (A/B/C models), characterized by differences in knowledge transfer medium (offline instruction/human–computer interaction/human–computer collaboration) and interaction dimension (one-way teaching/self-directed exploration/intergenerational collaboration). The dependent variables included:

① Digital Skill Acquisition: seven core digital skills (e.g., mobile payments, social media use) were selected based on the needs assessment and were evaluated using a dichotomous scoring system (1 = success, 0 = failure). Participants were required to perform skill-related tasks independently in a standardized environment, with double-blind observers recording accuracy.

② User Experience: measured using the User Experience Questionnaire—Short (UEQ-S), an eight-item scale divided into pragmatic quality (e.g., efficiency, clarity) and hedonic quality (e.g., attractiveness, innovativeness). Each item was rated on a 7-point semantic differential scale (e.g., "inefficient—efficient"), with scores ranging from 1 (negative experience) to 7 (positive experience). Mean scores for both dimensions were calculated separately to reflect user experience levels.

③ Learning Motivation: measured using the ARCS Model of Motivation (Attention, Relevance, Confidence, Satisfaction), assessed through a 16-item scale (4 items per dimension) on a 5-point Likert scale (1 = strongly disagree, 5 = strongly agree), with a total score ranging from 16 to 80.

Experimental tasks were selected based on the needs assessment, covering high-frequency digital skills commonly required by older adults, including online shopping (76.7%), calling/messaging (73.4%), social media use (68.2%), photo/video editing (20.7%), and video chat (20.3%). Skills were selected based on a progressive difficulty principle, ensuring fairness and scientific rigor in comparing different learning models. All selected skills were confirmed through pre-testing to be unmastered by participants before the

experiment, ensuring the reliability of results. The experiment was conducted in a standardized environment, with the venue set up as a quiet, comfortable, independent space to maintain consistent environmental conditions and lighting. All participants used the same model of Android smartphone, pre-installed with the standardized test application "Digital Bridge", and were provided with a uniform network configuration to minimize external factors influencing the results.

The experiment was divided into the pre-test phase, intervention phase, and post-test phase (as illustrated in Figure 2), with the following procedures:

① Pre-Test Phase: participants' baseline digital skills were assessed to ensure that none had prior knowledge of the seven experimental tasks. Additionally, this phase included device operation training to minimize the impact of unfamiliarity with the equipment.

② Intervention Phase: participants engaged in their assigned learning tasks according to group allocation. Researchers used a standardized observation record sheet to document learning duration, types and frequency of encountered issues, help-seeking behaviors, emotional responses, and learning strategies, ensuring comprehensive data collection.

③ Post-Test Phase: immediately after learning completion, the following assessments were conducted:

a. User Experience Measurement (UEQ-S).
b. Learning Motivation Measurement (ARCS Model of Motivation Scale).
c. Digital Skill Acquisition Assessment: participants performed skill-related tasks independently in a standardized environment, with double-blind observers recording their performance.

This study adopted an immediate testing approach to minimize the effects of short-term memory decay on test results. Additionally, all tests strictly followed a double-blind design, ensuring that both assessors and participants were unaware of their specific group allocation, thus maintaining data objectivity and experimental validity. Furthermore, all experiment staff underwent standardized training to ensure uniform instructional content and interaction methods across groups, adhering to a predefined teaching script to guarantee consistency in the experiment.

3.2. Data Analysis

To ensure the reliability and validity of the measurement instruments, this study conducted reliability and validity analyses on the User Experience Questionnaire—Short (UEQ-S), ARCS Motivation Model Scale, and Digital Skills Test. Internal consistency was assessed using Cronbach's α coefficient, while construct validity was evaluated using Composite Reliability (CR) and Average Variance Extracted (AVE). Prior to analysis, all variables were preprocessed to ensure data suitability and robustness.

For data analysis, this study used R version 4.4.1 (2024-06-14 ucrt) and adopted non-parametric tests to ensure the robustness of the results. Since the data exhibited non-normal distribution (Shapiro–Wilk normality test $p < 0.05$), the Kruskal–Wallis H test was employed to compare differences among the three learning modes (Group A: Traditional Digital Mentorship, Group B: Independent Gamified Learning, Group C: Collaborative Gamified Digital Mentorship) in terms of user experience, learning motivation, and digital skill acquisition. For variables showing significant differences, Dunn's post hoc test was used for pairwise comparisons.

4. Results

This study conducted a systematic data analysis based on the ADDIE model, which encompasses five phases, Analysis, Design, Development, Implementation, and Evaluation, ensuring the scientific rigor and effectiveness of the developed digital learning intervention.

4.1. Findings from the Analysis Phase

During the needs assessment phase, a total of 305 older adults aged 60 and above participated in the questionnaire survey. The sample consisted of 52.46% male and 47.54% female respondents, with 68.2% aged between 61 and 75 years. The educational background of the participants was generally low, with 46.89% having only completed primary education or below and 28.53% having completed junior high school education. Regarding living arrangements, the highest proportion of participants lived with their children or relatives (33.77%), followed by those living with a spouse (30.49%), while 15.74% lived alone. The survey revealed that 69.5% of older adults regularly use smart devices, with smartphones being the most commonly used device (37.7%). However, their overall digital skills remained weak, indicating a significant demand for learning and notable learning barriers.

The findings showed that older adults' most frequently used digital skills were related to daily services and social functions, including online shopping (76.7%), calling/messaging (73.4%), and social media use (68.2%). However, only 15.1% of respondents considered themselves "highly proficient" in digital skills. In terms of learning demands, the most desired skills were photo/video editing (20.7%) and video chat techniques (20.3%), reflecting their high interest in recording life experiences and social interactions. Additionally, the demand for file transfer and printing (13.8%) highlighted practical operational needs in their daily routines (See Figure 3).

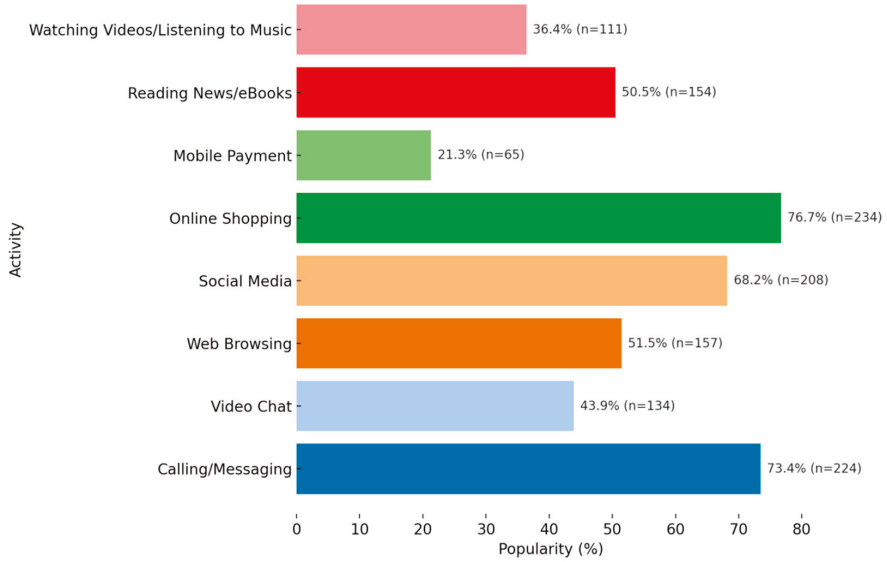


Figure 3. Primary activities using smart devices by the elderly.

However, older adults faced significant barriers when learning digital skills. The most prominent concern was the fear of making mistakes that could damage devices or lead to financial losses (88.5%), followed by difficulty understanding technical terms (58.0%) and lack of patient guidance from family or friends (46.6%). These barriers indicate that older adults commonly experience psychological anxiety and a lack of environmental support when learning digital skills, necessitating tailored learning tools to lower entry barriers (See Figure 4).

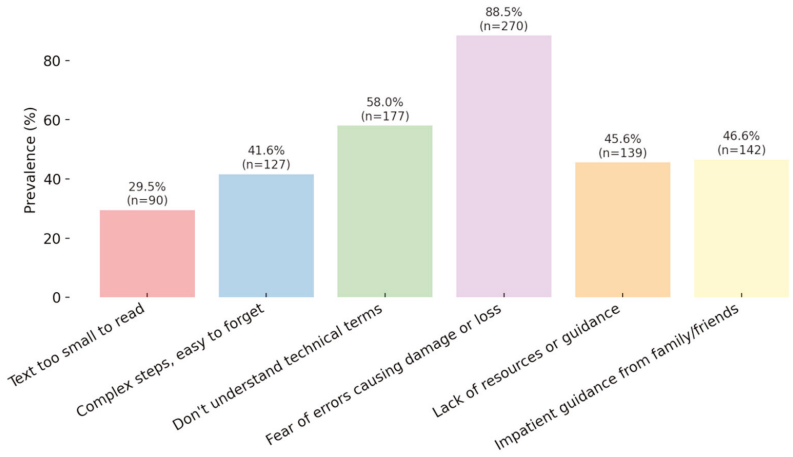


Figure 4. Main difficulties older adults face in learning to use smart devices.

Additionally, survey data showed that older adults had a high acceptance of gamified learning products, with reward mechanisms (89.5%), interactivity (73.4%), and personalized settings (71.5%) being key motivators (See Figure 5). Open-ended suggestions further revealed that older adults prefer digital learning tools with a simple and intuitive interface, instructional videos or image-based tutorials, and real-life scenario-based training. These insights provided valuable references for subsequent gamification design.

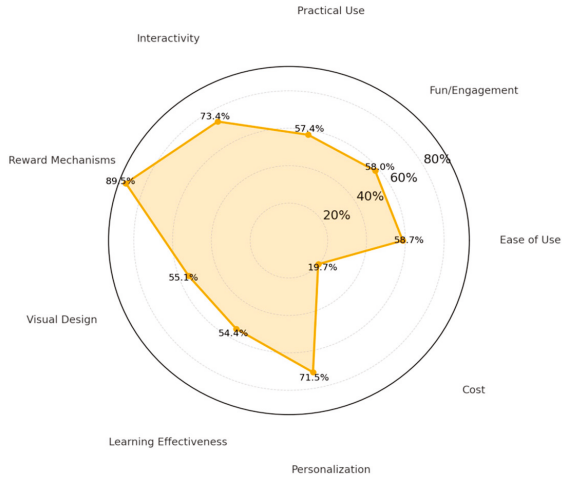


Figure 5. Factors influencing the elderly's willingness to use digital learning products.

The focus group discussions with participants provided valuable insights into the core functions and design strategies of the "Digital Reverse Mentorship" gamified learning application. Participants emphasized that gamification elements could significantly reduce learning barriers for older adults. One participant remarked, "Breaking down complex tasks into smaller, progressive steps makes learning less overwhelming and more manageable." Others highlighted the importance of cognitive load reduction, suggesting that a structured, step-by-step learning approach would help older users retain information more effectively.

Regarding motivation, several participants pointed out that point-based rewards, virtual achievements, and emotional reinforcement mechanisms could make the learning experience more engaging. One participant shared, "Seeing progress through points and badges would keep me motivated to continue learning."

For instructional content, the discussions underscored the need for real-life scenarios to enhance practical understanding. A participant explained, "If I see how these digital

skills apply to everyday tasks, like online banking or video calls with family, I'd be more confident in using them." Additionally, participants stressed the importance of a simple and intuitive interface, with optimized navigation and confirmation dialogs for critical actions. As one participant put it, "I worry about making mistakes, so having a clear confirmation before taking an action would make me feel much safer using the app.".

These discussions confirmed that designing with older adults' needs in mind—by simplifying tasks, providing motivational reinforcement, and ensuring usability—would enhance the effectiveness of the gamified learning experience.

In summary, during the needs assessment phase, this study conducted a comprehensive analysis of older adults' digital learning needs, barriers, and expectations through questionnaire surveys and focus group discussions. By integrating participant's feedback, this study clearly defined the design direction for the gamified digital mentorship application. The findings provided strong data support and theoretical foundations for designing and optimizing gamified learning tools, ensuring that the final product not only meets the practical needs of older adults, but also enhances intergenerational interaction and promotes long-term digital skill development.

4.2. Results from the Design and Development Phases

In the design phase, this study, based on findings from the needs assessment and focus group discussions, and guided by Cognitive Aging Theory and Aging-Friendly Experience Design, developed a Collaborative Gamified Digital Mentorship Application "Digital Bridge". This application aims to enhance older adults' digital literacy, promote intergenerational interaction, and reduce learning barriers through gamification strategies, thereby increasing user motivation and digital skill acquisition. The following sections summarize the design research findings in four key areas: game framework, functionality design, UI design, and interaction design.

In terms of game framework and functionality design, the application consists of four core modules: Digital Skills Learning Module, Intergenerational Interaction Module, Reward Mechanism, and AI Voice Assistant. The Digital Skills Learning Module covers 12 fundamental skills, such as making phone calls, sending text messages, and using social media. Task design is closely aligned with the daily life scenarios of older users, helping them understand the practical significance and application of these operations. The Intergenerational Interaction Module innovatively introduces the "Call for fresh Help" feature, allowing younger users to guide older users in real-time via voice or text, facilitating intergenerational knowledge transfer and collaboration. The Reward Mechanism employs a point-based system and virtual achievements (such as badge collections) to provide multi-dimensional incentives, enhancing user motivation and a sense of accomplishment. Additionally, the AI Voice Assistant offers voice navigation and real-time operational prompts, further reducing learning complexity and enabling older users to complete digital skill training more independently.

In terms of UI design, this study strictly adhered to aging-friendly design principles, optimizing visual accessibility and operational convenience. The visual design employs font sizes of at least 18 pt and high-contrast color schemes, ensuring that interface elements are clearly visible, thus reducing accessibility barriers caused by age-related vision decline. The interface layout, while ensuring functionality completeness, adopts a hierarchical navigation structure, ensuring that core functions (such as digital skills learning and intergenerational interaction) are placed in top-level menus, allowing users to quickly and intuitively locate target functions, minimizing confusion and frustration caused by complex pathways. For example, a secondary confirmation mechanism is implemented in key task steps to prevent accidental errors, thereby improving interaction security. Additionally, the

system provides real-time visual and auditory feedback, ensuring operational visibility and predictability, thus enhancing users' learning experience and confidence.

In terms of interaction design, this study focused on the core objectives of reducing cognitive load and optimizing user experience, employing four key design strategies. First, simplifying navigation paths, using an intuitive menu structure, and clear task flow to reduce operational difficulty and cognitive burden. Second, an instant feedback mechanism, where the system immediately provides success prompts accompanied by dynamic celebratory animations after task completion, enhancing the sense of achievement and learning motivation. Third, progressive task difficulty, where tasks follow an easy-to-difficult learning path—initial tasks focus on basic operations (e.g., answering a phone call), while more complex tasks (e.g., using mobile payments) are gradually introduced as users gain proficiency. This ensures users stay within an optimal challenge zone, preventing frustration due to excessive difficulty. Lastly, the "Call for Help" function, which optimizes the guidance process for younger users, supporting both voice intercom and text prompts to ensure efficient and engaging intergenerational collaboration. When older users encounter operational difficulties, they can instantly request help from their paired younger user, who can provide real-time assistance through voice or text instructions, effectively solving problems while strengthening emotional bonds through intergenerational interaction (see Figure 6).

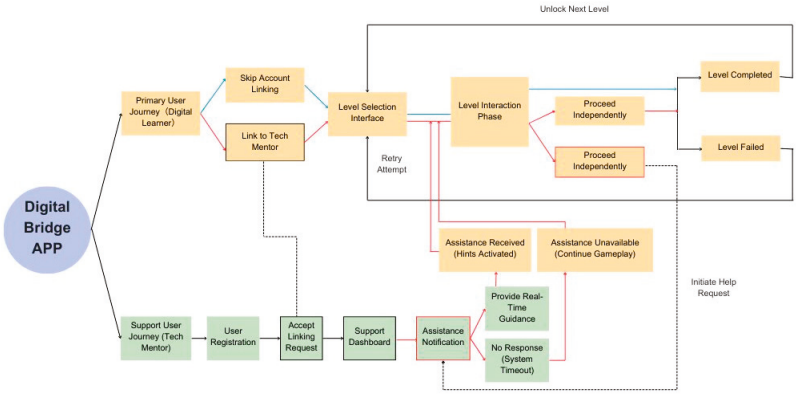


Figure 6. "Digital Bridge" Intergenerational Digital Mentorship Gamified Application process diagram. This game consists of digital learner mode (A) and tech mentor mode (B), with elders as the primary players while younger users assist them upon receiving a help request. The game flow and key interaction design are as follows: After entering the game, users must select either Elder or Younger identity. Elders can choose to play independently or bind with a younger user, enabling them to send help requests during the game. Once bound, when a younger user receives a help request, they can choose to assist the elder. If they choose to help, they will receive a prompt guiding them through the task. If the elder successfully completes a level, they unlock the next one, but if they fail, the next level remains locked.

Through these design strategies, the application is expected to effectively reduce cognitive load, improve operational ease and the learning experience, and establish an interactive and practical digital skills learning environment, enabling older users to integrate more smoothly into the digital society. Ultimately, through systematic design and development, this study aims to build a digital learning platform that aligns with older users' cognitive characteristics, meets their emotional needs, and promotes intergenerational interaction. (see Figures 7 and 8).

Through these design strategies, the application effectively reduces cognitive load, enhances ease of operation and learning experience, and establishes an interactive and practical digital skills learning environment, enabling older users to integrate more easily into the digital

society. Ultimately, through systematic design and development, this study successfully built a digital learning platform that aligns with older users' cognitive characteristics, meets their emotional needs, and promotes intergenerational interaction. At the same time, the application's design fully integrates aging-friendly principles and gamification strategies, offering an intuitive, user-friendly, and emotionally engaging learning experience, laying a solid foundation for subsequent user testing and effectiveness evaluation.



Figure 7. "Digital Bridge" Application Primary User Flow (A) Elder Mode Interface, from left to right: Character Selection Page, where users choose between "I am an elder" and "I am a junior" to enter the game; Bind Younger User Page, prompting the elder to bind with a junior by providing a unique binding code ("Please tell the junior the binding number: 123321"); Home Page (Level Selection Page), where users begin learning tasks from the beginner level and select from Level 1 to Level 3, with a prominent "Start" button to proceed; Task Execution Page (Skill Task: "Call specific contacts"), displaying a task goal ("Goal: Save contact number") with instructions like "Open dial page, click 'Add Contact'" and an option to "Call the junior" for help; and Level Completion Page, which shows feedback upon successfully completing a task ("Score + 1", "Level passed successfully") with options to return home or continue to the next level.



Figure 8. "Digital Bridge" Application Support User Flow (B) Tech Mentor Mode Interface. From left to right: Bind Elderly User Page, where the junior user enters the elder's binding number and sets their own name, confirming the connection to assist the elder. The labels read "I am a junior," "Enter elder's binding number," and "Confirm binding"; Home Page (Always Ready to Help the Elderly), showing that the junior is ready to assist the elder in learning. Levels such as Level 1, 2, and 3 are displayed, along with a button labeled "View Assistance Records"; and Task Execution Page, where the junior sees the task goal ("Goal: Save contact number") and prepares to provide real-time guidance. A banner at the bottom reads: "You are currently in observer mode, assisting the elder to complete the task.".

4.3. Results from the Implementation and Evaluation Phases

In the implementation and evaluation phase, this study conducted a randomized controlled trial (RCT) to assess the effectiveness of the Collaborative Gamified Digital Mentorship Application. A total of 90 older adults aged 60 and above were recruited and randomly assigned to three groups: the Collaborative Gamified Digital Mentorship Group (C Group), the Traditional Digital Mentorship Group (A Group), and the Independent Gamified Learning Group (B Group). Data collection focused on three core indicators: digital skill acquisition, user experience (UEQ-S), and learning motivation (ARCS) (see Figure 9).



Figure 9. Trial scenarios for older adult users. This image presents a sequence of four real-life photos showing elderly users interacting with the "Digital Bridge" application. In the first photo, the user is on the role selection screen, choosing "I am an elder" to begin. The second photo shows the user navigating the level selection page, preparing to start a digital skill training task. In the third photo, the user is actively performing the task, such as making a phone call, as part of the guided activity. The final photo displays the completion screen with a "Level Passed" message, highlighting the user's success and reinforcing a sense of accomplishment. This series visually demonstrates the practical engagement of older adults with the app in real-world settings.

First, in the correlation analysis of group variables, gender, education level, and smartphone proficiency showed no significant differences ($p > 0.05$). Additionally, there was no significant association between group assignment and these variables, indicating that there were no systematic differences in baseline demographic characteristics among the groups. The random assignment was reasonable, making it suitable for further group comparisons (see Table 1). This is also supported by the association analysis results presented in Table 1, where all *p*-values were greater than 0.05 and effect sizes (Cramer's V) were small, indicating no systematic bias across gender, education, or smartphone proficiency among the groups.

Table 1. Results of categorical variable association analysis.

Variable 1	Variable 2	Statistic	Value	*p*-Value	Effect Size
Group	Gender	$\chi^2(2)$	1.71	0.425	0.138
Group	Education Level	Fisher's Exact		0.296	0.231
Group	Smartphone Usage Proficiency	Fisher's Exact		0.226	0.240
Gender	Education Level	Fisher's Exact		0.853	0.128
Gender	Smartphone Usage Proficiency	Fisher's Exact		0.462	0.206
Education Level	Smartphone Usage Proficiency	Fisher's Exact		0.200	0.256

Note: effect size is measured using Cramer's V.

Second, experimental results showed that different learning modes had significant differences in user experience, learning motivation, and digital skill acquisition. Specifically, the Collaborative Gamified Digital Mentorship Model (C Group) outperformed both the Traditional Digital Mentorship Model (A Group) and the Independent Gamified Learning Model (B Group), confirming its effectiveness in improving digital literacy among older adults. As shown in Table 2, Group C not only had the highest average scores across all dimensions, but also the smallest standard deviations, indicating more consistent and reliable outcomes. C Group had the highest scores across all evaluation metrics: User Experience (UEQ-S): 5.59 ± 0.33, Learning Motivation (ARCS): 5.75 ± 0.19, Skill Acquisition (Skills): 6.23 ± 0.94. A Group and B Group showed significantly lower scores: A Group (UEQ-S: 3.83 ± 0.55, ARCS: 3.99 ± 0.30, Skills: 2.97 ± 2.39), B Group (UEQ-S: 3.90 ± 0.44, ARCS: 4.14 ± 0.28, Skills: 2.87 ± 2.05). The small difference between A Group and B Group suggests that traditional teaching and independent game-based learning had comparable effects (see Table 2).

Table 2. Descriptive statistics of anxiety scale scores by group (mean ± SD).

Group	UEQ-S	ARCS	Skills
A	3.83 (±0.55)	2.98 (±0.24)	2.97 (±2.39)
B	3.90 (±0.44)	3.11 (±0.23)	2.87 (±2.05)
C	5.59 (±0.33)	3.66 (±0.52)	6.23 (±0.94)

Note. Data are presented as mean ± SD.

Further analysis of scale reliability and validity confirmed the measurement instruments' robustness. Cronbach's α values: UEQ-S: 0.78, ARCS: 0.91, Skills: 0.85. These values indicate high internal consistency. Construct validity tests (CFI > 0.90, TLI > 0.90, RMSEA < 0.08) showed a good model fit, supporting the individual validity of the measurement tools. Additionally, the Shapiro–Wilk normality test results showed that UEQ-S, ARCS, and Skills scores in C Group failed the normality test ($p < 0.05$). The Levene's test for homogeneity of variance indicated that ARCS and Skills scales showed significant variance differences among groups ($p < 0.05$). Therefore, non-parametric tests were used to ensure robustness of the results (see Table 3).

Table 3. Homogeneity of variance and normality tests.

Variable	Group	Shapiro–Wilk W	Shapiro *p*	Levene Test	Levene *p*
ARCS	A	0.9658199	0.432	F(2, 87) = 5.02	0.009
ARCS	B	0.9456768	0.129	F(2, 87) = 5.02	0.009
ARCS	C	0.8959594	0.007	F(2, 87) = 5.02	0.009
Skills	A	0.8656595	0.001	F(2, 87) = 7.48	0.001
Skills	B	0.9346811	0.065	F(2, 87) = 7.48	0.001
Skills	C	0.7777873	<0.001	F(2, 87) = 7.48	0.001
UEQ-S	A	0.9715425	0.582	F(2, 87) = 2.83	0.064
UEQ-S	B	0.9573003	0.264	F(2, 87) = 2.83	0.064
UEQ-S	C	0.9259038	0.038	F(2, 87) = 2.83	0.064

Note: The Shapiro–Wilk test reports the W statistic and exact *p*-value, while the Levene test reports the F-value and degrees of freedom.

To further evaluate the central tendency and dispersion of the data, this study calculated the median and interquartile range (IQR) for each group (see Table 4). The analysis results indicate that Group C had the highest median scores in UEQ-S (5.6250), ARCS (3.9375), and Skills (6.5), with the smallest IQR, suggesting that the learning outcomes in

Group C were the most stable and consistent. In contrast, Groups A and B had lower median scores, and the IQR of Skills scores was larger (Group A: 4.00, Group B: 2.75), reflecting greater individual variability in adaptation under traditional teaching and independent game-based learning modes. This indicates that the collaborative gamified mentorship model not only led to better outcomes on average, but also produced more stable learning results with less variation, which is critical when designing inclusive interventions for elderly learners.

Table 4. Median and interquartile range (IQR) for each group.

Group	Median_ UEQ-S	IQR_ UEQ-S	Median_ ARCS	IQR_ ARCS	Median_ Skills	IQR_ Skills
A	3.6875	0.625	2.96875	0.296875	2	4
B	3.875	0.375	3.0625	0.359375	3	2.75
C	5.625	0.34375	3.9375	1.046875	6.5	1

The Kruskal–Wallis H test results (see Table 5) further confirmed significant differences among the three groups across all measured indicators (UEQ-S: $\chi^2(2) = 59.39$, $p < 0.001$; ARCS: $\chi^2(2) = 60.99$, $p < 0.001$; Skills: $\chi^2(2) = 36.51$, $p < 0.001$). Dunn's post hoc test (see Tables 6–8) indicated that Group C scored significantly higher than Groups A and B ($p < 0.001$), while there was no significant difference between Groups A and B ($p > 0.05$). This finding highlights that simply using gamified content without intergenerational support may not be sufficient for improving outcomes. The interactivity and social reinforcement in Group C appear to be key differentiators. Additionally, Group C had the highest median and the smallest IQR, indicating a more concentrated data distribution and more stable learning outcomes. In contrast, the Skills scores in Groups A and B showed greater dispersion (IQR = 4.00 and 2.75, respectively), suggesting that older adults exhibited higher variability in adaptation under independent learning and traditional teaching modes.

Table 5. Kruskal–Wallis test results.

Variable	Chi_Squared	df	*p*_Value
UEQ-S_total	59.39197	2	<0.001
ARCS_total	28.23301	2	<0.001
Skills_total	36.50690	2	<0.001

Table 6. The results of Dunn's post hoc test—UEQ-S_total.

Comparison	Z_Value	*p*_unadj	*p*_adj	Significance
A—B	−0.556889	0.578	1.000	ns
A—C	−6.935124	<0.001	<0.001	***
B—C	−6.378235	<0.001	<0.001	***

Note: *** $p < 0.001$; ns indicates non-significant (no statistical significance).

Table 7. The results of Dunn's post hoc test—ARCS_total.

Comparison	Z_Value	*p*_unadj	*p*_adj	Significance
A—B	0.1175083	0.906	1.000	ns
A—C	−5.1728633	<0.001	<0.001	***
B—C	−5.2903715	<0.001	<0.001	***

Note: *** $p < 0.001$; ns indicates non-significant (no statistical significance).

Table 8. The results of Dunn's post hoc test—Skills_total.

Comparison	Z_Value	*p*_unadj	*p*_adj	Significance
A—B	−1.491985	0.136	0.407	ns
A—C	−5.162467	<0.001	0.000	***
B—C	−3.670482	<0.001	0.001	***

Note: *** $p < 0.001$; ns indicates non-significant (no statistical significance).

Moreover, the results suggest that there may be a strong relationship among user experience, learning motivation, and digital skill acquisition. In the Collaborative Gamified Digital Mentorship group (Group C), participants not only achieved significantly higher scores on the UEQ-S and ARCS scales, but also demonstrated better mastery of digital skills. This indicates that a positive user experience and stronger learning motivation may jointly contribute to improved learning outcomes. Specifically, a good user experience such as a clear interface, intuitive operation, and timely feedback can help reduce cognitive load during the learning process, thereby enhancing older adults' willingness to participate and their engagement. Higher learning motivation reflected in interest in the content, confidence, and satisfaction further drives active task completion and continuous practice, leading to better skill acquisition. In contrast, Groups A and B showed relatively lower scores in both user experience and learning motivation, with poorer overall performance in skill acquisition and more dispersed data distribution. This suggests that in the absence of interaction or motivational mechanisms, learning outcomes among older adults are more susceptible to individual differences, resulting in greater variability. Therefore, optimizing user experience and enhancing learning motivation are key strategies for improving the effectiveness of digital skills training for older adults.

5. Discussion

5.1. Effectiveness of the Intergenerational Digital Mentorship Model

Based on Cognitive Aging Theory, this study developed a Collaborative Gamified Digital Mentorship Application and evaluated its effectiveness in improving digital skill acquisition, user experience, and learning motivation among the elderly. The experimental results demonstrate that the Collaborative Gamified Digital Mentorship Model (C Group) significantly outperforms the Traditional Teaching Group (A Group) and the Independent Gamified Learning Group (B Group) across all three dimensions ($p < 0.001$), indicating that a model combining intergenerational interaction and gamified learning can effectively reduce learning anxiety, enhance skill acquisition, and increase learning motivation and engagement among the elderly.

In the present study, intergenerational interaction played a crucial role, as interactions between older and younger users provided real-time feedback and support, so as to reduce technical barriers and boost learning confidence [60]. Particularly in a gamified environment, elderly users actively participated in tasks, received immediate feedback, and engaged with the reward mechanism, which helps them stay motivated and enhance their digital skills [7]. Additionally, the introduction of gamified learning provides an immersive learning experience to make the learning process more interactive and enjoyable, ultimately increasing learning motivation. These findings confirm that combining intergenerational interaction with gamified learning not only leads to more effective skill acquisition, but also improves attitudes toward digital learning of the elderly.

5.2. Skill Acquisition and Learning Experience

In this study, the elderly in C Group exhibited superior skill acquisition, significantly outperforming the other two groups. The combination of intergenerational interaction and gamified learning provided more effective learning support. Through intergenerational collaboration, the elderly received timely assistance and guidance from younger users, which reduced cognitive load during learning and enabled them to acquire digital skills more quickly and accurately. In contrast, the Traditional Teaching Model (A Group) and Independent Gamified Learning Model (B Group) lacked interactive support, making older users feel more isolated and confused during the learning process, ultimately impacting their skill acquisition.

Regarding user experience, the elderly in C Group scored significantly higher in both usability and enjoyment compared to A Group and B Group. Collaborative learning and aging-friendly game design effectively enhanced user experience, making the learning process more seamless and enjoyable. Intergenerational interaction not only provided social support, but also fostered emotional connections, giving the elderly a greater sense of belonging and emotional engagement during learning. Additionally, gamified design elements (such as task breakdown, immediate feedback, and reward mechanisms) increased learning enjoyment, allowing the elderly to better appreciate the learning process, and thereby improving their overall experience.

Another key finding of this study is the significant improvement in learning motivation. C Group achieved the highest learning motivation scores, indicating that the combination of gamification and intergenerational interaction effectively stimulates the elderly's interest in learning. The elderly often struggle with low learning motivation, but gamification elements, task feedback, and intergenerational support in this study successfully enhanced their interest and engagement, providing valuable insights for the design of future educational programs for the elderly.

5.3. Comparison with Existing Research

The findings of this study align with the existing literature and further expand the application of gamified learning in the education of the elderly. For example, Cota et al. (2015) and Marston (2013) found that the elderly show strong interest in puzzle-based games, and this study further confirmed that gamification elements (such as task segmentation, instant feedback, and reward mechanisms) can enhance learning motivation and optimize learning experience. Additionally, the elderly engage more actively in digital skill learning in the interactive learning environment with immediate feedback and rewards [6,9].

Xie et al. (2022) suggested that intergenerational interaction effectively reduces technology anxiety in the elderly, and this study further found that intergenerational collaboration not only improves skill acquisition, but also enhances emotional belonging and social connectedness [27]. Intergenerational interaction helps the elderly overcome challenges in the learning process, thereby strengthening emotional bonds between younger and older generations [61,62]. This social support not only improves learning outcomes, but also increases the social participation of the elderly, allowing them to exert greater social value and achieve emotional fulfillment in digital learning [63,64], which aligns with the core principles of the Ethics of Care with the emphasis on interpersonal relationships, emotional connections, and mutual dependence [65]. From the perspective of the Ethics of Care, digital reverse mentorship is not merely a transfer of knowledge, but also an intergenerational responsibility, a means of fulfilling needs, and a form of emotional support.

This study also validates the positive impact of gamified learning on the elderly, demonstrating that gamified design provides continuous learning incentives, which are

challenging to achieve in traditional teaching and independent game-based learning models. Additionally, this study provides empirical evidence that gamified learning not only enhances skill acquisition, but also improves overall learning experiences by increasing motivation to make learning more engaging, and fostering emotional connections.

In addition, this study conducted a comparative analysis of the UEQ-S and ARCS scores with findings from the related literature, further validating the effectiveness of the system while identifying areas for potential improvement. The UEQ-S results indicated high scores in the pragmatic dimensions (such as efficiency and clarity), suggesting that the system interface was smooth and functionally sound. However, scores in the hedonic dimensions (such as novelty and stimulation) were relatively lower. Compared to the study by Maskeliūnas et al. (2023), where the FGPE+ platform reported high scores in stimulation and attractiveness (M = 4.11, SD = 0.51), our results suggest that enhancing visual engagement and interactivity could further improve the emotional experience [66]. Similarly, Lampropoulos et al. (2020) emphasized that combining augmented reality with gamification significantly enhances the hedonic aspects of learning, offering a promising direction for future design enhancements [67].

In terms of learning motivation, this study recorded high scores in the "Relevance" and "Satisfaction" dimensions of the ARCS scale, which aligns with the findings of Kumar et al. (2023), who reported a strong positive correlation between relevance and learner engagement (r = 0.727) in a biomedical VR education system [68]. However, the "Confidence" dimension received comparatively moderate ratings, which is consistent with the ARCS+G model proposed by Fazamin et al. (2023) [69]. That model recommends integrating gamified motivational strategies—such as goal setting, progress tracking, and achievement badges—to reinforce learners' self-efficacy. These comparisons suggest that the present system demonstrates notable strengths in usability and relevance, but still holds room for improvement in emotional stimulation and confidence-building mechanisms.

5.4. Research Limitations and Future Directions

First, the research sample was primarily drawn from a senior university in Taiyuan, a specific region in China, which may affect the external validity of the findings. Moreover, this study only involved 90 participants. Future research should expand the sample size and include elderly individuals from diverse social strata, cultural backgrounds, and residential areas to enhance the generalizability of the results. For example, older adults from different communities, urban and rural settings, and various cultural environments may exhibit significant differences in their acceptance of and response to gamified learning. Therefore, further studies are needed to verify the broader applicability of intergenerational interaction and gamified learning.

Second, this study primarily focused on short-term learning outcomes and did not track long-term skill retention and cognitive function changes among the elderly. While this study demonstrated that the Collaborative Gamified Digital Mentorship Model can effectively improve digital skill acquisition in the short term, future research can extend the study period to observe long-term effects, such as the sustainability of learning outcomes and skill retention. Since Group C outperformed the other groups, future studies can further analyze the specific characteristics or intervention measures within Group C to identify the key factors contributing to the significant learning improvement. These could include the frequency of intergenerational interactions or the optimization of game tasks, which is conducive to refining the core mechanisms and optimizing intervention strategies, so as to make them applicable to a broader population of the elderly. In addition, future research is recommended to comprehensively consider the impact of multiple factors, such

as cognitive abilities, health status, emotional state, learning motivation, and habits of using digital devices in individuals. It is also advisable to expand the application scenarios by implementing this intervention model in settings such as home environments, community education centers, senior activity venues, or online learning platforms, so as to enhance its practicality and scalability.

Additionally, Group A (traditional teaching) and Group B (independent game-based learning) share relatively similar learning outcomes in this study, suggesting that these two intervention methods may have limited effectiveness in improving digital literacy among the elderly. Future research may consider adjusting experimental designs or intervention approaches, such as introducing more interactive teaching models, optimizing learning content, or incorporating enhanced motivational mechanisms to improve learning outcomes and increase engagement and persistence among the elderly.

Moreover, it is important to consider the potential limitations of the measurement tools used in this study, such as the UEQ-S (User Experience Questionnaire-Short Version) and the ARCS (Attention, Relevance, Confidence, Satisfaction) model. These scales may have inherent constraints in assessing user experience and motivation, particularly when the elderly are engaging with gamified learning applications. Future research should critically evaluate whether these frameworks fully capture the cognitive and emotional responses of elderly learners. Additionally, this study used the Chinese versions of these scales, and it is necessary to assess whether the translation has been appropriately validated to ensure conceptual equivalence with the original instruments. Translation and cultural adaptation processes may introduce subtle variations in meaning, thereby potentially impacting the validity and reliability of the results. Future studies should conduct validation studies on these localized versions to confirm their suitability for assessing user experience and learning motivation in the Chinese-speaking elderly population.

Finally, this study mainly examined skill acquisition and learning experience, while it did not explore the impact of gamified learning on memory, attention, and executive function in the elderly. To comprehensively assess the effectiveness of gamified learning, future studies should integrate neuropsychological assessment tools to systematically evaluate the impact of gamified learning on cognitive capabilities such as memory, attention, and executive function in the elderly. Additionally, this study has a relatively small sample size. Future research should incorporate effect size analysis (e.g., Eta-squared or Epsilon-squared) to assess the practical significance of intergroup differences and provide more precise statistical conclusions, which will contribute to a more comprehensive understanding of the role of gamified learning in enhancing digital literacy among the elderly, and providing a stronger theoretical and practical foundation for subsequent research and applications.

6. Conclusions

This study demonstrates that the Collaborative Gamified Digital Mentorship Model is an effective strategy for improving digital literacy among the elderly, enhancing skill acquisition, optimizing user experience, and increasing learning motivation. The experimental results indicate that learning models combining gamification and intergenerational interaction are more effective than traditional teaching and independent game-based learning in enhancing digital skills, optimizing user experience, and boosting learning motivation. Gamification elements, such as task segmentation, instant feedback, and reward mechanisms, not only optimize user experience, but also increase learning engagement, while intergenerational interaction provides emotional support and a sense of social belonging, thereby achieving more positive and enjoyable learning.

The present study suggests that the combination of gamified learning and intergenerational interaction provides multidimensional support, which not only facilitates skill acquisition, but also enhances the interest and enthusiasm of the elderly for digital learning. Future educational product designs should leverage these findings by integrating gamification strategies and intergenerational collaboration to improve the effectiveness of digital skills training for the elderly.

The theoretical significance of this study lies in providing a new perspective for the application of Cognitive Aging Theory in digital education and validating the effectiveness of integrating gamified learning with intergenerational interaction. Thes findings expand the boundaries of research on education and digital literacy among the elderly, and provide theoretical support for the future development of more interactive and aging-friendly digital education products. In practical terms, this study proposes a feasible intervention strategy that not only helps bridge the digital divide among the elderly, but also promotes intergenerational interaction and provides greater social support for their learning. Additionally, this study provides practical guidance for optimizing digital skills training programs for the elderly, particularly in incorporating gamification and social interaction to enhance learning outcomes and satisfaction.

Future research should further explore long-term intervention effects, cultural adaptability, and more refined personalized learning strategies to develop a more sustainable digital literacy education system for the elderly. Moreover, it should incorporate virtual reality (VR), voice assistants, and other intelligent technologies to optimize digital training programs for the elderly, so as to enhance learning immersion and interactivity. By leveraging artificial intelligence and big data analysis, learning tasks can be dynamically adjusted in terms of difficulty and content, thereby enabling personalized learning support and optimization tailored to the individual needs and progress of elderly learners, finally improving learning effectiveness and continuity.

In conclusion, this study provides practical guidance for enhancing digital literacy among the elderly, improving the access to technology design, and developing future educational products, as well as lays a foundation for building a more inclusive and sustainable digital education system for the elderly. Future research should further explore personalized gamified learning, long-term intervention effects, and the integration of social support networks to drive innovation in digital education for the elderly, thereby ensuring that a wider range of such advancements benefit the elderly population worldwide.

Funding: The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. This research was funded by the China Scholarship Council (File No. 202208390037).

Institutional Review Board Statement: The studies involving humans were approved by Research Ethics Committee (CERec) of the Autonomous University of Barcelona (20241115CSH) on 15 November 2024.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to privacy reasons.

Acknowledgments: I sincerely thank Elena Bartomeu Magaña and Haoran Li, Ruijie Zhang, and Xi Chen for their guidance and support during the data analysis process. Their insights and resources were invaluable in helping me face this challenge and ultimately complete the work independently. I would also like to thank all participants for their support.

Conflicts of Interest: The author declares no conflicts of interest.

References

1. Escourrou, E.; Laurent, S.; Leroux, J.; Oustric, S.; Gardette, V. The shift from old age to very old age: An analysis of the perception of aging among older people. *BMC Prim. Care* **2022**, *23*, 3. [CrossRef] [PubMed]
2. Wang, J.; Zhou, Y.; Zhang, Q.; Li, J.; Zhai, D.; Li, J.; Han, B.; Liu, Z. Loneliness among older Chinese individuals: The status quo and relationships with activity-related factors. *BMC Geriatr.* **2024**, *24*, 42.
3. Schroeder, T.; Dodds, L.; Georgiou, A.; Gewald, H.; Siette, J. Older adults and new technology: Mapping review of the factors associated with older adults' intention to adopt digital technologies. *JMIR Aging* **2023**, *6*, e44564. [CrossRef] [PubMed]
4. Tomczyk, Ł.; Mascia, M.L.; Gierszewski, D.; Walker, C. Barriers to digital inclusion among older people: An intergenerational reflection on the need to develop digital competences for the group with the highest level of digital exclusion. *Soc. Incl.* **2023**, *11*, 44–54. [CrossRef]
5. Buiza, C.; Gonzalez, M.F.; Facal, D.; Martinez, V.; Diaz, U.; Etxaniz, A.; Urdaneta, E.; Yanguas, J. Efficacy of cognitive training experiences in the elderly: Can technology help? In *Universal Access in Human-Computer Interaction. Addressing Diversity, Proceedings of the 5th International Conference, UAHCI 2009, San Diego, CA, USA, 19–24 July 2009. Proceedings, Part I*; Springer: Berlin/Heidelberg, Germany, 2009; pp. 324–333.
6. Cota, T.T.; Ishitani, L.; Vieira, N., Jr. Mobile game design for the elderly: A study with focus on the motivation to play. *Comput. Hum. Behav.* **2015**, *51*, 96–105. [CrossRef]
7. Tahmassebi, S. Digital Game Design for Elderly People. Master's Thesis, Malmö University, Malmö, Sweden, 2018. Available online: https://www.diva-portal.org/smash/record.jsf?pid=diva2:1480001&dswid=-7698 (accessed on 12 May 2025).
8. Breck, B.M.; Dennis, C.B.; Leedahl, S.N. Implementing reverse mentoring to address social isolation among older adults. *J. Gerontol. Soc. Work* **2018**, *61*, 513–525. [CrossRef]
9. Marston, H.R. Design recommendations for digital game design within an ageing society. *Educ. Gerontol.* **2013**, *39*, 103–118. [CrossRef]
10. Francis, J.; Ball, C.; Kadylak, T.; Cotten, S.R. Aging in the digital age: Conceptualizing technology adoption and digital inequalities. In *Ageing and Digital Technology: Designing and Evaluating Emerging Technologies for Older Adults*; Springer: Singapore, 2019; pp. 35–49.
11. Salthouse, T.A. A theory of cognitive aging. In *Handbook of Aging and Cognition*; Academic Press: Cambridge, MA, USA, 2000; Volume 2, pp. 221–276.
12. Hackett, K.; Giovannetti, T. Capturing cognitive aging in vivo: Application of a neuropsychological framework for emerging digital tools. *JMIR Aging* **2022**, *5*, e38130. [CrossRef]
13. Hassenzahl, M. User experience and experience design. In *The Encyclopedia of Human-Computer Interaction*, 2nd ed.; The Interaction Design Foundation: Aarhus, Denmark, 2013; pp. 1–14.
14. Allen, W.C. Overview and evolution of the ADDIE training system. *Adv. Dev. Hum. Resour.* **2006**, *8*, 430–441. [CrossRef]
15. Keller, J.M.; Keller, J.M. The ARCS model of motivational design. In *Motivational Design for Learning and Performance: The ARCS Model Approach*; Springer: Boston, MA, USA, 2010; pp. 43–74.
16. Schrepp, M.; Hinderks, A. Design and Evaluation of a Short Version of the User Experience Questionnaire (UEQ-S). 2017. Available online: https://reunir.unir.net/handle/123456789/11836 (accessed on 12 May 2025).
17. Schrepp, M.; Kollmorgen, J.; Thomaschewski, J. A comparison of SUS, UMUX-LITE, and UEQ-S. *J. User Exp.* **2023**, *18*, 86–104.
18. Van Dijk, J. *The Digital Divide*; John Wiley & Sons: Hoboken, NJ, USA, 2020.
19. Salthouse, T.A. Continuity of cognitive change across adulthood. *Psychon. Bull. Rev.* **2016**, *23*, 932–939. [CrossRef]
20. Bandura, A.; Wessels, S. *Self-Efficacy*; Cambridge University Press: Cambridge, UK, 1997.
21. Neves, B.J.; Braga, R.C.; Melo-Filho, C.C.; Moreira-Filho, J.T.; Muratov, E.N.; Andrade, C.H. QSAR-based virtual screening: Advances and applications in drug discovery. *Front. Pharmacol.* **2018**, *9*, 1275. [CrossRef] [PubMed]
22. Helsper, E. *The Digital Disconnect: The Social Causes and Consequences of Digital Inequalities*; SAGE: Newcastle upon Tyne, UK, 2021.
23. Levin, D.; Arafeh, S.; Rainie, L. *The Digital Disconnect: The Widening Gap Between Internet-Savvy Students and Their Schools*; Pew Internet & American Life Project: Washington, DC, USA, 2002.
24. Bozeman, B.; Feeney, M.K. Toward a useful theory of mentoring: A conceptual analysis and critique. *Adm. Soc.* **2007**, *39*, 719–739. [CrossRef]
25. Chen, Y.C. Effect of reverse mentoring on traditional mentoring functions. *Leadersh. Manag. Eng.* **2013**, *13*, 199–208. [CrossRef]
26. von Preußen, P.W.; Beimborn, D. Turning mentoring around: A case-based analysis of the outcomes of digital reverse mentoring. *Electron. Mark.* **2019**, *29*, 351–369.
27. Xie, X.; Han, Y.; Anderson, A.; Ribeiro-Navarrete, S. Digital platforms and SMEs' business model innovation: Exploring the mediating mechanisms of capability reconfiguration. *Int. J. Inf. Manag.* **2022**, *65*, 102513. [CrossRef]

28. Choi, N.G.; DiNitto, D.M.; Marti, C.N.; Choi, B.Y. Telehealth use among older adults during COVID-19: Associations with sociodemographic and health characteristics, technology device ownership, and technology learning. *J. Appl. Gerontol.* **2022**, *41*, 600–609. [CrossRef]
29. Leist, A.K. Social media use of older adults: A mini-review. *Gerontology* **2013**, *59*, 378–384. [CrossRef]
30. Hladek, M.D.; Wilson, D.; Krasnansky, K.; McDaniel, K.; Shanbhag, M.; McAdams-DeMarco, M.; Crews, D.C.; Brennan, D.C.; Taylor, J.; Segev, D.; et al. Using Photovoice to explore the lived environment and experience of older adults with frailty on their kidney transplant journey. *Kidney360* **2024**, *5*, 589–598. [CrossRef]
31. Jang, S.-B.; Kim, M. Digital fitness revolution: User perspectives on Fitbit's role in health management. *Behav. Sci.* **2025**, *15*, 231. [CrossRef]
32. Nutakor, J.A.; Zhou, L.; Larnyo, E.; Addai-Dansoh, S.; Cui, Y. Impact of health information seeking behavior and digital health literacy on self-perceived health and depression symptoms among older adults in the United States. *Sci. Rep.* **2024**, *14*, 31080. [CrossRef] [PubMed]
33. Kang, H.; Kim, D.H. Socioeconomic, health, and social connectedness factors associated with self-rated health of octogenarians and nonagenarians in South Korea: Urban and rural comparison. *BMC Public Health* **2024**, *24*, 3477. [CrossRef] [PubMed]
34. Stamate, A.N.; Denis, P.L.; Sauvé, G. How cognitive ability shapes personality differentiation in real job candidates: Insights from a large-scale study. *J. Intell.* **2024**, *12*, 34. [CrossRef] [PubMed]
35. Chanquin, E.; Navarro, S.S.; Falzarano, F. Bridging the digital divide: A user-centered design approach to enhancing social connectivity among older adults. *Innov. Aging* **2024**, *8* (Suppl. 1), 1294. [CrossRef]
36. Mannheim, I.; Schwartz, E.; Xi, W.; Buttigieg, S.C.; McDonnell-Naughton, M.; Wouters, E.J.; Van Zaalen, Y. Inclusion of older adults in the research and design of digital technology. *Int. J. Environ. Res. Public Health* **2019**, *16*, 3718. [CrossRef]
37. Tsai, M.C.; Yang, W.S. Familial exchange and intergenerational contact in East Asian societies. In *Family, Work and Wellbeing in Asia*; Springer: Singapore, 2017; pp. 21–42.
38. Punchoojit, L.; Hongwarittorrn, N. Usability studies on mobile user interface design patterns: A systematic literature review. *Adv. Hum.-Comput. Interact.* **2017**, *2017*, 6787504. [CrossRef]
39. Deterding, S.; Dixon, D.; Khaled, R.; Nacke, L. From game design elements to gamefulness: Defining" gamification". In Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments, Tampere, Finland, 28–30 September 2011; pp. 9–15.
40. Hamari, J.; Koivisto, J.; Sarsa, H. Does gamification work?—A literature review of empirical studies on gamification. In Proceedings of the 2014 47th Hawaii International Conference on System Sciences, Waikoloa, HI, USA, 6–9 January 2014; IEEE: Piscataway, NJ, USA, 2014; pp. 3025–3034.
41. Juris, J.; Zvonkovic, A.; Allen, K.R.; Few-Demo, A.; Crawford, A. Grandparent couples' experiences of family leisure. *J. Leis. Res.* **2024**, 1–20. [CrossRef]
42. Leedahl, S.N.; Brasher, M.S.; LoBuono, D.L.; Wood, B.M.; Estus, E.L. Reducing ageism: Changes in students' attitudes after participation in an intergenerational reverse mentoring program. *Sustainability* **2020**, *12*, 6870. [CrossRef]
43. Newsham, N. Understanding the Extent, Demographic Causes, Consequences and Future of Population Decline in Europe. Doctoral Thesis, University of Liverpool, Liverpool, UK, 2024.
44. Alain, C.; McDonald, K.L.; Ostroff, J.M.; Schneider, B. Aging: A switch from automatic to controlled processing of sounds? *Psychol. Aging* **2004**, *19*, 125. [CrossRef]
45. Plass, J.L.; Moreno, R.; Brünken, R. (Eds.) *Cognitive Load Theory*; Cambridge University Press: Cambridge, UK, 2010.
46. Sweller, J. Cognitive load theory. In *Psychology of Learning and Motivation*; Academic Press: Cambridge, MA, USA, 2011; Volume 55, pp. 37–76.
47. Paas, F.; Van Merrienboer, J.J. Cognitive-load theory: Methods to manage working memory load in the learning of complex tasks. *Curr. Dir. Psychol. Sci.* **2020**, *29*, 394–398. [CrossRef]
48. Stern, Y. Cognitive reserve. *Neuropsychologia* **2009**, *47*, 2015–2028. [CrossRef] [PubMed]
49. Song, H.; Kim, M.; Park, D.; Shin, Y.; Lee, J.G. Learning from noisy labels with deep neural networks: A survey. *IEEE Trans. Neural Netw. Learn. Syst.* **2022**, *34*, 8135–8153. [CrossRef] [PubMed]
50. DiNicolantonio, J.J.; O'Keefe, J.H. The importance of marine omega-3s for brain development and the prevention and treatment of behavior, mood, and other brain disorders. *Nutrients* **2020**, *12*, 2333. [CrossRef] [PubMed]
51. Howard-Jones, P.A.; Varma, S.; Ansari, D.; Butterworth, B.; De Smedt, B.; Goswami, U.; Laurillard, D.; Thomas, M.S. The principles and practices of educational neuroscience: Comment on Bowers (2016). *Psychol. Rev.* **2016**, *123*, 620–627. [CrossRef] [PubMed]
52. Lampit, A.; Hallock, H.; Valenzuela, M. Computerized cognitive training in cognitively healthy older adults: A systematic review and meta-analysis of effect modifiers. *PLoS Med.* **2014**, *11*, e1001756. [CrossRef]

53. Fisk, A.D.; Czaja, S.J.; Rogers, W.A.; Charness, N.; Sharit, J. *Designing for Older Adults: Principles and Creative Human Factors Approaches*; CRC Press: Boca Raton, FL, USA, 2020.
54. Venkatesh, V.; Thong, J.Y.; Xu, X. Consumer acceptance and use of information technology: Extending the unified theory of acceptance and use of technology. *MIS Q.* **2012**, *36*, 157–178. [CrossRef]
55. Deterding, S.; Sicart, M.; Nacke, L.; O'Hara, K.; Dixon, D. Gamification: Using game-design elements in non-gaming contexts. In Proceedings of the CHI'11 Extended Abstracts on Human Factors in Computing Systems, Vancouver, BC, Canada, 7–12 May 2011; pp. 2425–2428.
56. Kivetz, R.; Urminsky, O.; Zheng, Y. The goal-gradient hypothesis resurrected: Purchase acceleration, illusionary goal progress, and customer retention. *J. Mark. Res.* **2006**, *43*, 39–58. [CrossRef]
57. Kappen, C. Developmental patterning as a quantitative trait: Genetic modulation of the Hoxb6 mutant skeletal phenotype. *PLoS ONE* **2016**, *11*, e0146019. [CrossRef]
58. Tobias, E.S. Learning with digital media and technology in hybrid music classrooms. In *Teaching General Music: Approaches, Issues, and Viewpoints*; Oxford University Press: Oxford, UK, 2016; pp. 112–140.
59. Kühn, S.; Gallinat, J. Segregating cognitive functions within hippocampal formation: A quantitative meta-analysis on spatial navigation and episodic memory. *Hum. Brain Mapp.* **2014**, *35*, 1129–1142. [CrossRef]
60. Sun, S. Unlocking engagement: Exploring the drivers of elderly participation in digital backfeeding through community education. *Front. Psychol.* **2025**, *16*, 1524373. [CrossRef]
61. Sun, S.; Li, H.; Wei, X. Creative mutual aid community in the context of "Internet+". In Proceedings of the 8th International Conference on E-Society, E-Learning and E-Technologies, Rome, Italy, 10–12 June 2022; pp. 1–6.
62. Zhang, R.; Martí Casanovas, M.; Bosch González, M.; Sun, S. Public value in historic environment regeneration in China: A public perception perspective on spatial form, urban governance, and people's experience (2000–2020). *Land* **2025**, *14*, 267. [CrossRef]
63. Bartomeu, E.; Ventura, O. Participatory practices for co-designing a multipurpose family space in a children's hospital. *Des. Health* **2021**, *5*, 26–38. [CrossRef]
64. Zhang, R.; Martí Casanovas, M.; Bosch González, M.; Sun, S. Revitalizing heritage: The role of urban morphology in creating public value in China's historic districts. *Land* **2024**, *13*, 1919. [CrossRef]
65. Richardson, L.R.; Canham, S.L.; Weldrick, R.; Hoselton, J.; Grittner, A.; Walsh, C.A. "I feel like a somebody again": Ethics of care at a shelter for older adults fleeing abuse in Canada. *J. Gerontol. Soc. Work* **2025**, 1–23. [CrossRef]
66. Maskeliūnas, R.; Damaševičius, R.; Blažauskas, T.; Swacha, J.; Queirós, R.; Paiva, J.C. FGPE+: The mobile FGPE environment and the Pareto-optimized gamified programming exercise selection model—An empirical evaluation. *Computers* **2023**, *12*, 144. [CrossRef]
67. Lampropoulos, G.; Keramopoulos, E.; Diamantaras, K.; Evangelidis, G. Augmented reality and gamification in education: A systematic literature review of research, applications, and empirical studies. *Appl. Sci.* **2022**, *12*, 6809. [CrossRef]
68. Kumar, M.S.; Park, J.H.; Ryu, H.J. Application of AIoMT and virtual reality in medical education: A study based on the ARCS model. *Electronics* **2023**, *12*, 835. [CrossRef]
69. Fazamin, F.N.; Zakaria, M.H.; Ghazali, A.S.; Ahmad, A. Design and evaluation of gamified ARCS-based e-learning systems: An ARCS+G framework. *Interact. Learn. Environ.* **2023**, *32*, 6868–6885. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

Multimodal Technologies and Interaction

MDPI

Article

A Digital Model-Based Serious Game for PID-Controller Education: One-Axis Drone Model, Analytics, and Student Study

Raul Brumar [1,*], Stelian Nicola [1] and Horia Ciocârlie [2]

[1] Department of Automation and Applied Informatics, Polytechnic University of Timisoara, 300006 Timisoara, Romania; stelian.nicola@upt.ro
[2] Department of Computer and Information Technology, Polytechnic University of Timisoara, 300006 Timisoara, Romania; horia.ciocarlie@cs.upt.ro
* Correspondence: raul.brumar@upt.ro

Abstract: This paper presents a serious game designed to support the teaching of PID controllers. The game couples a visually clear Unity scene with a physics-accurate digital model of a drone with a single degree of freedom (called a one-axis drone) and helps prepare students to meet the demands of Industry 4.0 and 5.0. An analytics back-end logs system error at 10 Hz and interaction metrics, enabling instructors to diagnose common tuning issues from a plot and to provide actionable hints to students. The design process that led to choosing the one-axis drone and turbulence application via "turbulence balls" is explained, after which the implementation is described. The proposed solution is evaluated in a within-subjects study performed with 21 students from mixed technical backgrounds across two short, unsupervised tinkering sessions of up to 10 min framed by four quizzes of both general and theoretical content. Three questions shaped the analysis: (i) whether error traces can be visualized by instructors to generate actionable hints for students; (ii) whether brief, unsupervised play sessions yield measurable gains in knowledge or stability; and (iii) whether efficiency of tuning improves without measurable changes in tune performance. Results show that analysis of plotted error values exposes recognizable issues with PID tunes that map to concrete hints provided by the instructor. When it comes to unsupervised play sessions, no systematic pre/post improvement in quiz scores or normalized area under absolute error was observed. However, it required significantly less effort from students in the second session to reach the same tune performance, indicating improved tuning efficiency. Overall, the proposed serious game with the digital twin-inspired one-axis drone and custom analytics back-end has emerged as a practical, safe, and low-cost auxiliary tool for teaching PID controllers, helping bridge the gap between theory and practice.

Keywords: serious game; digital games; digital game-based learning; digital model; digital twin; gamification; automation; regulator

1. Introduction

In the current educational context, technology plays an important role in transforming both teaching and learning. Integrating it into the teaching process offers opportunities to deepen the understanding of both theoretical and practical processes [1–4]. Technology has vast advantages, such as offering unlimited access to resources, personalized content

through dedicated platforms, and increased engagement through interactive serious games. The games can work using traditional technologies, or they can put players in a virtual or augmented reality to increase interactivity even further.

While theoretical knowledge is fundamental, lacking a clear grasp of how it is applied in practice defeats the purpose of learning. Theory gains real value when it can be applied in practice [5], and education should prepare students for the everyday challenges of the labor market [6,7]. Students should develop the required skills by applying their knowledge in concrete settings.

Technology allows us to develop innovative ways to introduce students to practical examples such as simulations and virtual laboratories. Problem-based projects, serious games situated in VR/AR, online collaboration platforms, and industry partnerships are some of the ways technology enables students to improve their practical skills [8,9].

Serious educational games, along with gamification, have become powerful tools for improving learning [4,10]. Although they are both inspired by traditional entertainment games, they represent two distinct approaches. Serious games are, in fact, games. However, in contrast to entertainment games, they are designed primarily for education, training, simulation, and problem solving. They integrate game mechanics into specific contexts so that learners can acquire knowledge through experience and direct interaction [11,12]. Gamification, on the other hand, is a design approach used in application development, not games, that applies elements and techniques with the goal of increasing user engagement, retention, and motivation [4,10]. In education, gamification reframes learning tasks as game-like progression [13,14].

The digital twin is an increasingly present concept when real-time virtual representations of a physical objects are discussed [15–17]. Kritzinger et al. [18] define a digital twin as a virtual counterpart that maintains bidirectional, real-time synchronization with its physical twin and distinguish it from simpler digital models or digital shadows. Jones et al. [19] expand this definition by introducing the concept of twinning rate, which uses the direction and frequency of the data flow to differentiate between digital twins in the literature. Still, as previous papers show, the term is often used in contexts where real-time coupling is not required. In such cases, which have started to outnumber the ones following the strict definition, the term digital twin-inspired simulation should be used to avoid confusing the field with the broader educational implementation of virtual laboratories.

The Proportional-Integral-Derivative (PID) controller is one of the most widely used algorithms due to its simplicity and applicability across different scenarios. Leveraging modern approaches such as digital twins and serious games opens new perspectives for teaching PID controllers [20,21].

This paper presents a serious game containing a digital twin-inspired simulation that has been designed to enhance the teaching of PID controllers. The game is built around a physically accurate digital model of a one-axis drone that is presented in a visually clear Unity scene. An analytics back-end records the error value at 10 Hz, which allows instructors to visualize it and provide targeted feedback to students. This paper documents the design and implementation steps taken for developing this serious game and analyzes the data obtained by testing it on a group of 21 students with diverse cultural and technical backgrounds. Both confirmatory and exploratory analyses were performed for answering the following three research questions:

- whether collected error values are useful for diagnosing students' issues with tuning;
- whether brief, unsupervised use of the serious game leads to measurable gains in student knowledge;
- whether a second attempt reduces tuning effort without degrading tune quality.

Overall, the contribution proved to be a safe, low-cost, and reliable teaching aid that complements traditional teaching methods and helps bridge the gap between theory and practice.

The paper is organized into five sections and contains twelve Figures and four Tables. Section 1 briefly presents the concepts underlying this study. Section 2 reviews related work and offers a short comparison across existing approaches. Section 3 examines the design and development of the proposed serious game. Section 4 reports the evaluation of quiz and analytics data collected during the student study. Section 5 summarizes the findings and outlines directions for future development.

2. Related Work

As discussed in the previous chapter, digital education requires approaches that are centered on the human factor, in a way that empowers human work. Industry 4.0 and 5.0 place digital education among the top priorities in terms of their approaches [22,23]. For this, the recent scientific literature highlights three major concepts: serious games, used to increase engagement through interactivity and active learning; digital twins, which offer a simulation faithful to reality on which students can experiment; and automation and controls, which are fundamental domains of industry [24,25]. These domains should not evolve separately. Instead, they should complement each other, shaping the education to be based on applications and serious games relevant to industrial concepts. This chapter examines the literature, split in three subsections:

- Serious Games in Education;
- Digital Twins in Education and Industry;
- Automation and Control in Industry 4.0/5.0.

2.1. Serious Games in Education

In recent studies, serious games play an important role in developing the competencies required by the digital industry. Pacheco-Velázquez et al. show that educational simulators can support training in logistics for Industry 4.0 by facilitating the understanding of complex processes in an interactive environment [26]. A systematic analysis performed in [27] indicates that the use of serious games in higher education has high potential within the Education 4.0 context. It also outlines key challenges related to development costs, integration in the curriculum, and lack of standardized evaluation methods. An automated evaluation system applied in serious VR games is presented in [28], where automated assessments were deemed to be comparable to human ones. Bertozzi et al. present an experimental study in industrial engineering education in which a serious game that has been integrated into a traditional course improved motivation and understanding of technical concepts [29]. Rodriguez-Calzada et al. report that the integration of serious games with active pedagogical methods increases engagement and student involvement. In a comparative study, agile learning through VR serious games outperformed the LEGO Serious Play method in teaching Scrum [30,31].

A subsection of serious games is represented by simulation-centric games, which emphasize realistic system behavior over having traditional win states. A review performed by [32] shows that this form of serious games is especially widespread in engineering education. Similar to Pacheco-Velázquez et al., a paper by Sánchez et al. presents three simulation-based educational games for teaching automatic control concepts [33].

In serious games, the main focus rests on motivation and engagement. However, realism is another element that is essential for a relevant educational environment. This

role is fulfilled by the digital twin, a technology that brings learning closer to real-life applications and industrial contexts.

2.2. Digital Twin in Education and Industry

The industry uses digital twins especially for educating or training operators so that they better understand processes. Ortiz et al. [34] developed a digital twin for the system that processes cocoa beans. When used as an educational resource, users gained a better understanding of the process, significantly reducing operating times. Another digital twin increased motivation by facilitating access to virtual laboratories and involved reduced costs [35]. Many serious games are powered by digital twins, through which the student's state and progress can be tracked [36]. Zhang et al. report a quasi-experimental study with a learning system based on digital twins used for teaching landscape architecture courses. Students who use the system report a significant improvement in critical thinking, academic performance, and learning experience [37].

Another field where digital twins are present is medicine. Rajamäki et al. provide a review of the use of digital twins in higher medical education. The study highlights the advantages offered by simulating, diagnosing, and training in a controlled environment [38]. A recent concept called Digital Twin-Assisted Surgery offers digital replicas of the patient's anatomy, thus improving situational awareness and safety in complex scenarios [39]. Likewise, in oncology, digital twins are used for personalized therapeutic modeling. A systematic review presented applications for diagnosis, treatment decisions, prognosis prediction, and surgical planning by integrating medical imaging, multi-omics data, and AI techniques [40].

Digital twins provide realistic simulations and are very valuable when applied in fundamental technical fields. A good example field is automation and control, where digital twins integrated in serious games of virtual laboratories, paired with modern teaching methods, support adapting education to the requirements of Industry 4.0 and 5.0.

2.3. Automation and Control in Industry 4.0 and 5.0

If Industry 4.0 focused on digitalization and interconnection, Industry 5.0 emphasizes the human factor along with sustainability. Teaching these concepts requires adaptation to new technological paradigms. Visioli highlights the need to modernize control courses to include IoT, machine learning, and cyber–physical systems [41]. Azofeifa et al. explore the technological integrations present in Industry 4.0 and how they are used in the education of engineers. They propose competence visualization (KSA) and adaptive learning as ways to adapt to current challenges [42]. Ahmad et al. emphasize the benefits that Industry 4.0 technologies such as AI and robotics can have when applied to educational settings. This reinforces the link between teaching automation and new technologies [43]. While the PID controller is a fundamental part of automation, Industry 4.0 leverages it to support self-optimizing processes and predictive maintenance [44]. Dapkute and colleagues present the integration of traditional PID controllers with digital twin technologies in industrial process control. They propose a solution that can enable adaptive tuning and predictive maintenance through real-time simulations [45]. Even additive manufacturing can benefit from such an approach. There, the use of model predictive control (MPC) supported by digital twins and multivariable neural networks has been demonstrated to facilitate real-time decisions related to the manufacturing process [46].

The scientific literature shows that serious games, digital twins, and automatic control do not represent separate domains that are not linked, but are, instead, complementary parts of a modern educational ecosystem. Together, they create an educational approach in

which theory, practice, and technology are used together, thus facilitating the development of competencies with direct effects in the industry.

3. Proposed Solution

The solution takes the form of a serious game designed to be used as an auxiliary tool when teaching PID controllers. The game should be designed to contain a digital twin-like physical system in which a PID controller is used to regulate a specific parameter. By providing a visual representation of the real-world system, along with interactive elements, the game aims to increase engagement and improve conceptual understanding of PID controllers. Players should adjust controller parameters (proportional, integral, and derivative coefficients) and, through real-time feedback, offered both in the form of text (system stats) and visually (graphics rendered by the game engine), immediately observe resulting changes in system behavior. This proposed solution not only reinforces theoretical concepts but also bridges the gap between abstract control theory and its practical implementation in engineering systems.

3.1. Choosing a Physical Twin

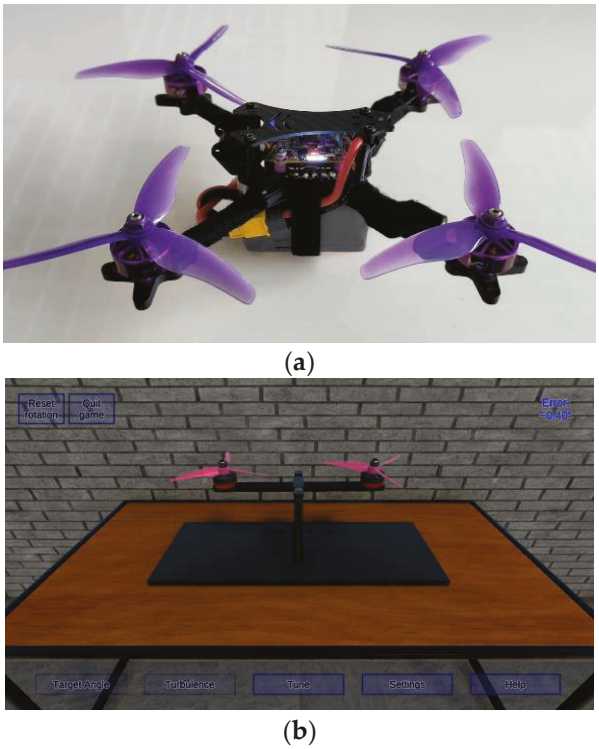
Before developing the game, a choice had to be made. Namely, which system was going to serve as the physical inspiration for the digital model used by the above-mentioned teaching process. There are a plethora of simple systems that could have been chosen, from cruise controls (traditional or adaptive) to temperature control or water level control in a tank. While the inner workings of a drone are technically harder to understand than other examples, its applicability in real-life is instantly recognizable, and the effects of tuning the controller are visually apparent. For these reasons, a drone was selected to teach students how a PID controller works. However, in order to reduce system complexity and to require a single set of tuning parameters, all but one degree of freedom was dropped, resulting in a simplified drone with a single degree of freedom. The remaining degree of freedom represents the roll axis, and can be influenced by adjusting the speed of the two remaining motors. This simplified one-axis drone serves as the real-world counterpart, or physical twin.

In Figure 1, we can observe the differences between a normal, hobby-grade drone with four motors (quadcopter—(a)) and the digital model of our simplified one-axis drone (b). While a quadcopter has four brushless motors, providing it with six degrees of freedom, our proposed model contains only two motors. On top of this, the one-axis drone is fixed on all but the roll angle, leaving it with a single degree of freedom. The biggest advantage of simplifying the physical twin is reducing the number of degrees of freedom, thus requiring PID-controller tuning for a single axis. Removing the yaw axis was especially important for two reasons: firstly, tuning the yaw axis is not as intuitive, as it is dependent on motor torque, and, secondly, implementing yaw control would require calculating the reaction torque of motors, which increases implementation complexity.

Some drawbacks to simplifying the physical twin are apparent, such as the loss of realism compared to a fully functional quadcopter and the omission of interactions between multiple degrees of freedom, but the advantages gained certainly outweigh these limitations in the context of this paper. While the reduced model cannot fully replicate the aerodynamic effects and control challenges present in its real-world counterpart, this level of fidelity is not essential for the primary objective of the project, which is to facilitate the understanding of PID-controller principles in a safe, interactive format. By limiting the system to a single degree of freedom, the cognitive burden on the students is reduced,

helping them better interpret the relationship between the PID coefficients' adjustments and the system's response.

A digital twin-inspired simulation of the simplified one-axis drone has been chosen over a physical prototype for multiple reasons. The most important being that a digital implementation eliminates the costs associated with hardware, maintenance, and repairs. Aside from the initial time investment, the project is freely replicable and accessible to as many students as needed. Another advantage comes from the removal of safety concerns that would arise from working with high-powered brushless motors and propellers. Such safety concerns range from equipment damage to personal injuries. Avoiding this means a safer and more productive learning environment. Lastly, the reduced complexity of a virtual system means that experiments will be easier to replicate consistently, without being influenced by the variability or malfunctions that physical hardware can introduce.



(**a**)

(**b**)

Figure 1. (**a**) The author's real-life racing quadcopter and (**b**) the digital model of a hypothetical one-axis drone.

3.2. Core Functionality

The serious game was designed around a limited number of core functions that the player can perform, with the goal of enabling experimentation with the PID controller in a safe environment. The feature that sits at the center of the game is the digital twin-inspired model of a simplified one-axis drone, which is simulated in real-time to respond to external disturbances based on the tuning received by the player.

Figure 2 presents the main actions undertaken by the student when playing the game. The process of tuning the digital model starts by observing the drone's physical state and is followed by tuning the parameters. After this, players can keep observing and tuning the parameters as long as they want, until a good tune is reached. After a good tune for counteracting gravity is reached, the player can impart disturbances through the use of balls falling from above the drone and tune the system accordingly. The described loop can then be performed until a satisfactory tune is reached. On top of this, the player can,

at any point, slightly de-tune the system to observe how a bad tune would manifest on a physical twin.

Another complementary feature is present, namely the ability to modify the target roll angle. It can be set to either a fixed value, or to a slowly oscillating signal. Together with the ability to apply disturbances, these features provide a learning environment in which students can iteratively refine the tune of the system, thereby increasing understanding of the PID controller.

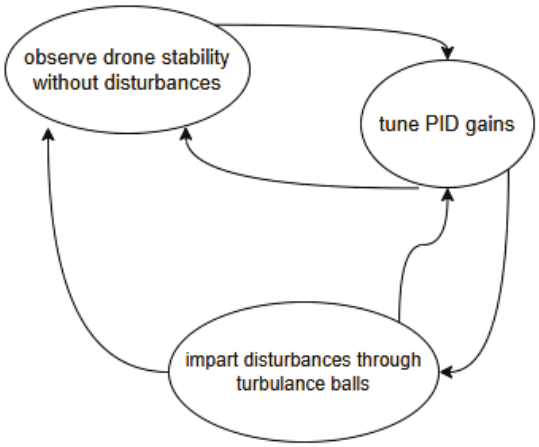


Figure 2. Core gameplay loop.

3.3. Implementation

The serious game was implemented in Unity 6000.1.10f1, ensuring a visually engaging environment. The physics simulation was carefully set up so that the drone's dynamics would closely resemble the physical twin in order to better support the intended learning outcomes.

3.3.1. Technologies Used

Most of the development was carried out in the Unity game engine as it provides high-fidelity rendering and physics simulation. The PID controller, other drone logic, and supporting functionalities were implemented in C# scripts attached to *GameObjects*. Blender 4.4.3 was used for modeling and adapting 3D models used in the scene. For version control, Git was used to manage the whole project [47].

3.3.2. Scene Setup

The main scene was designed to provide a realistic-looking and intuitive environment in which the digital model would be found. Aside from the simulated drone, the environment consists of a wooden table with metal legs situated on a tiled floor and in front of a grey brick wall. Another visually significant object is the metallic drone support. The support is not physics-simulated as it only serves to make it feel like the drone is fixed there, instead of floating.

At the core of the serious game lies the digital model of the simplified one-axis drone, represented in Unity as a parent *GameObject* called *OneAxisDrone* with multiple children. The parent contains the physics components, while the child *GameObjects* mostly ensure visual accuracy (motors, props, and drone body). Each motor is divided into two parts so that only the correct part rotates. The propeller is attached to the part that rotates.

Figure 3 illustrates the modular structure of the one-axis drone. The parent *GameObject* with its attached components handle physics and control of the drone, while the children are for purely aesthetic purposes, either static relative to the *OneAxisDrone* (*MRBody*, *Pivot*,

FrontBearing, *BackBearing*, and stators), or dynamic (rotors, along with their respective props and nuts).

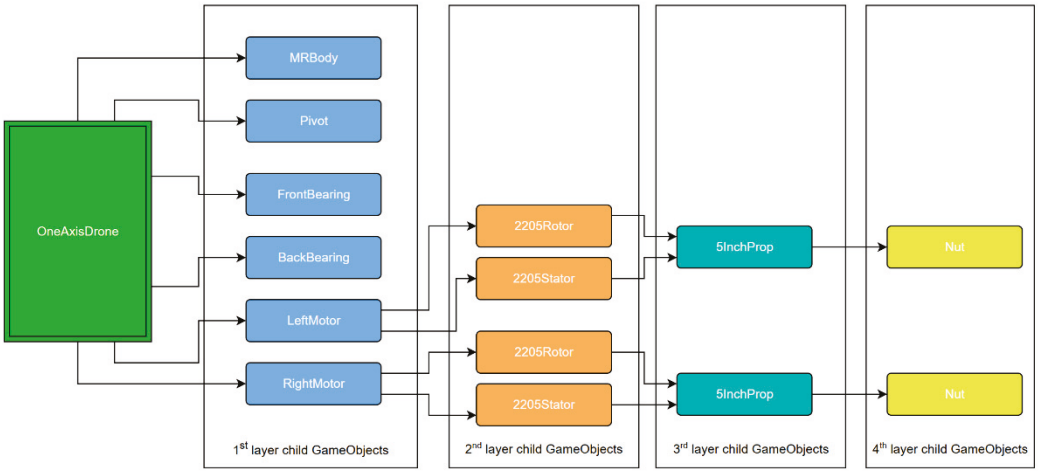


Figure 3. *OneAxisDrone GameObject* hierarchy, with colors representing hierarchical depth.

3.3.3. Physics Simulation

The physics simulation of the digital model is governed by the physics engine present in Unity. The parent *GameObject* (*OneAxisDrone*) has a *Rigidbody* component, enabling it to be simulated in the physics engine and to correctly respond to forces applied to it (gravity, disturbances, and lift). In order to ensure that the simulation reflected the expected physical structure, the *Rigidbody's* center of mass was adjusted to where the center of mass should be on a drone (higher than the center of *MRBody*) as motors and props are situated on top of the body.

Since the simulation should allow a single degree of freedom, namely on the roll axis, a configurable joint was used, constrained in all degrees of freedom except rotation around the roll axis. It is not attached to any other *GameObject*, thus fixing the drone to the world.

Control inputs were applied at the position and taking into account the orientation of the two motors, thus generating torque that imparts rotation on the roll axis to the drone. The magnitude of these forces is determined by a script called *OneAxisDroneByMotorForce*, which reads the output of the PID controller and calculates the lift forces required to maintain the drone at the correct roll angle.

Finally, external disturbances were implemented in the form of turbulence balls. They are lightweight spheres (50 g by default) that are spawned over the drone and fall, colliding with the drone, thus imparting a visually intuitive perturbation to its roll angle. This feature is vital for the serious game as it allows students to evaluate the robustness of their tuning, thus completing the core loop of the serious game.

3.3.4. One-Axis Drone Digital Model Implementation

The digital twin-inspired model is governed by a set of custom C# scripts attached to the parent *GameObject* and its motor child *GameObjects*.

Figure 4 shows the *GameObjects* to which all major scripts that make up and control the digital model are attached. In their totality, they simulate the control system, where the PID controller calculates the desired action and the *OneAxisDroneByMotorForce* reads it and actuates the two motors. In this way, the scripts regulate the drone's roll angle and maintain it to the desired value.

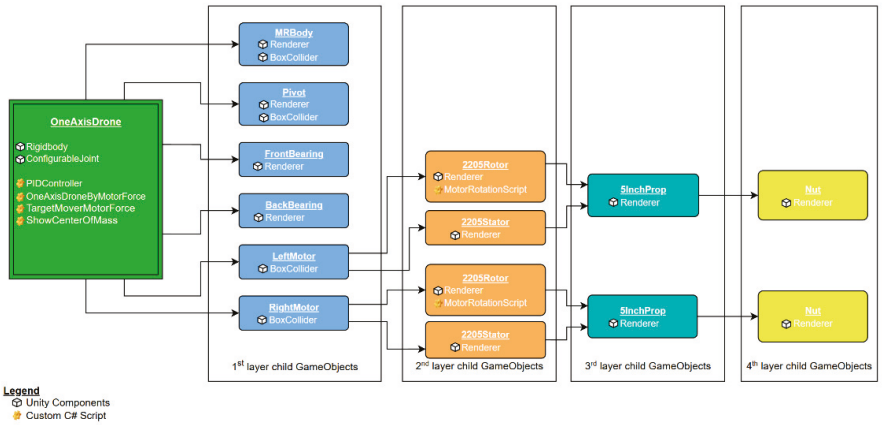


Figure 4. Major custom C# scripts and *GameObjects* holding them, with colors representing hierarchical depth.

1. PID Controller: The PID-controller logic is situated in the script called *PIDController*. The PID controller, in its continuous form, is expressed as:

$$u(t) = K_P e(t) + K_I \int_0^t e(\tau)d\tau + K_D \frac{de(t)}{dt} \tag{1}$$

where *e(t)* is the error between the target and the current angle, and K_P, K_I, and K_D are the proportional, integral, and derivative gains. The error is calculated by subtracting the current value from the target value. However, since the simulation runs in discrete time, the equation looks like this:

$$u(t) = K_P e(t) + K_I \int_0^t e(\tau)d\tau + K_D \frac{de(t)}{dt} \tag{2}$$

To ensure stability in cases when the error persists for long periods of time, the implementation includes an anti-windup mechanism, which clamps the integral term between a minimum and maximum value. Furthermore, the final output of the PID controller is clamped between configurable limits in order to prevent unrealistic values. Normally, for the derivative term, the change in error is calculated and used. However, when the target angle suffers a sudden change, this implementation can cause the derivative term to reach very high values for one frame, resulting in a derivative kick. To mitigate this effect, the derivative term was instead computed from the measured rate of change in the roll angle. By calculating the rate of change from the process variable, we ensure that the derivative term responds only to changes in the system state, and not to jumps in the target value.

2. Motor force application and visual feedback: The script called *OneAxisDroneByMotorForce* has the purpose of taking the output of the *PIDController* script and actuating the motors. At each physics step (*FixedUpdate* in Unity), it computes the required thrust for each motor by scaling the normalized controller output with a maximum thrust constant and clamping the values to ones accurate to a generic 2205 brushless motor. A minimum thrust is maintained as normal brushless motors of drones always keep a minimum RPM. The computed thrusts are applied at the motor position, taking into account the motor orientation, based on the current drone orientation. This application of thrust at motor positions generates the required torque for controlling the rotation on the roll axis while the *ConfigurableJoint* component constrains all degrees

of freedom except for the roll axis.
Since this script already reads the output of the *PIDController* and computes the forces applied to motors, it made sense to extend it to also compute the motors' rotational speed. The maximum speed is determined from parameters of real-life motors (KV value) and the battery powering them (battery cell count), while the instantaneous speed is set proportionally to the determined thrust. The speed values are passed to the *MotorRotationScript*, ensuring consistency between the visual representation and simulated physics. Even though the motors and propellers are not required for the drone to function correctly, they are integral to strengthening the student's perception of motor activity and its effect on the whole system, making the serious game more pedagogically effective.

3. Target angle control: *TargetMoverMotorForce* is the script that sets the value that the PID controller uses as the target. It allows the developer (and the player, through the UI) to set it manually to a fixed value or to set rules for it to oscillate between two values, with the desired speed. The script increases the flexibility of the serious game by enabling demonstrations of both static and dynamic targets.
4. Simulation flow: The scripts described above make up the digital model that sits at the forefront of the serious game, linking user control logic and physics-based response with user interaction. Each script is responsible for a specific function, such as computing control actions, applying forces, animating visual elements, or introducing disturbances, and together they form and control a coherent digital model that allows students to experience how parameter tuning affects system behavior.
Figure 5 shows the flow of information and actions between elements of the simulation. The player defines a static or dynamic target (or leaves it at the default value of zero), which is then compared by the drone's current roll angle, obtained from the *Rigidbody* component. Their difference constitutes the error, which is further used by the PID controller to compute the control output. The resulting value is read by *OneAxisDroneByMotorForce*, which translates it into differential thrust values for the two motors. Forces with these magnitudes are applied as if they originated from the motors themselves, generating the torque that rotates the drone on the roll axis. During the next frame, the cycle is repeated.

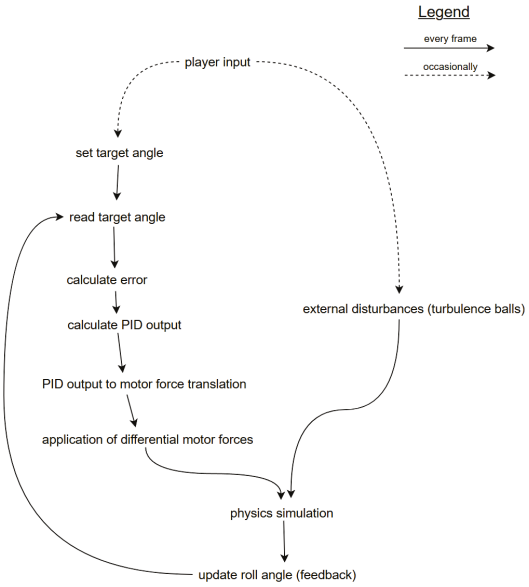


Figure 5. Interaction between components of the control loop in one frame.

External disturbances interact with the system through the physics simulation, and they influence the *Rigidbody's* state, thus eliciting a response from the PID controller, which will try to maintain their desired roll angle. This ensures players receive immediate, visually observable feedback to their tuning efforts.

5. Adjustable parameters: As described in this chapter, a modular approach was used when building the drone simulation, with each component containing several parameters that can be modified, some by the developer, inside Unity, and other by the programmer, inside the UI. Table 1 summarizes the main adjustable parameters and their functions, along with whether they are available only for developers or also for players.

Table 1. Adjustable parameters of the one-axis drone digital model.

Parameter	Description	Editable by
Proportional gain	Elicits a response that is proportional to the error	Player (UI)
Integral gain	Reacts to steady-state error	Player (UI)
Derivative gain	Dampens output based on error change	Player (UI)
Integral saturation	Clamps the integral term to prevent windup	Player (UI)
Target angle	Desired roll angle of the drone if oscillation is disabled	Player (UI)
Target movement speed	Rate at which the target value oscillates if automatic mode is enabled	Player (UI)
Turbulence ball mass	Controls disturbance strength during collisions	Player (UI)
Controller output limits	Clamps the PID-controller output	Developer
Motor KV rating	For determining the max RPM of the motor (visual only)	Developer
Battery cell count	For determining the max RPM of the motor (visual only)	Developer
Maximum motor thrust	Defines maximum force a single motor can produce	Developer
Minimum thrust percentage	Ensures propellers never stop fully, maintaining stability and visual realism	Developer

This split in configuration responsibility allows developers to tailor the simulation to reflect more than one physical twin configuration (different motors, batteries, props, and controllers), while at the same time enabling players to learn by experimenting with different tunes and disturbance scenarios in a safe and accessible manner.

3.3.5. External Disturbance Simulation

In order to fully benefit from a digital twin-inspired model that can be customized by developers and tuned by players, the game offers the option to test the tuning against external disturbances. This functionality is implemented through the *BallSpawner*, which spawns spherical objects called turbulence balls above the drone. Upon spawning, the ball falls and collides with the drone, thus imparting a force upon it, which allows the player to visually observe the behavior of their tuned drone. The mass of the ball can be configured in the user interface so players can see how lighter or heavier disturbances affect system stability and response.

For pedagogical reasons, the choice was made to use falling balls as disturbances. This is because collisions are highly visible, and its effects can help players interpret the effectiveness of their PID tuning. A potential future improvement to this style of applying disturbance to the drone would be to also scale the balls visually, according to their assigned mass, thereby strengthening the connection between the visual cue and the physical property it represents.

In Figure 6, we can observe turbulence balls falling (a) and colliding (b) with the drone. This introduces external disturbances that can be clearly and intuitively observed by vision alone.

Although the current disturbance implementation provides a clear and visually expressive method for visualizing disturbances, it does not resemble any real-world challenge faced by drones, the most common being gusts of wind. Adding player-controlled gusts of wind would improve realism, but it introduces the challenge of making wind, which is an inherently invisible phenomenon, visually observable. Additional visual cues such as dust particles, foliage movement (if outdoors), or subtle, cartoon-like visual effects usually associated with air movement would need to be implemented to convey wind presence and power to the player. The current choice of turbulence balls represents a practical compromise between pedagogical effectiveness, implementation simplicity, and realism.

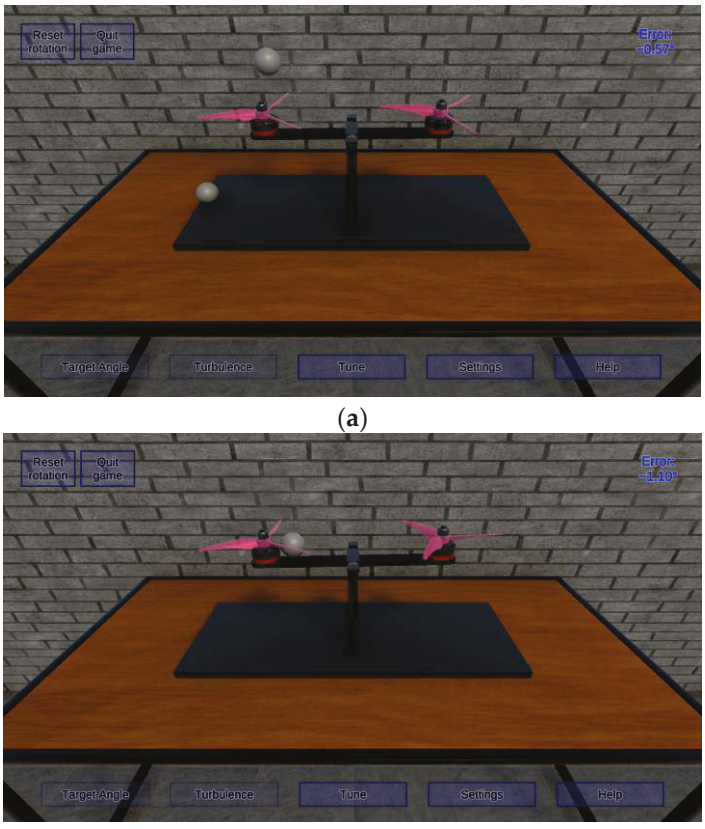


Figure 6. Turbulence balls falling and interacting with the digital model—(**a**) while ball is falling and (**b**) while the ball is interacting with drone through the physics engine.

3.3.6. User Interface

The user interface provides easy access to the simulation's main functions by offering rapid access to all the essential functions and parameters needed for experimenting with the digital model.

Figure 7 illustrates the in-game interface, which is controlled by the bottom panel containing five buttons, each opening its respective panel.

The panels seen in Figure 7 and explained in Table 2 offer comprehensive control over the simulation, allowing players to adjust the PID-controller values, to set the desired target (static or dynamic), and to also introduce external disturbances. This compact and intuitive

interface ensures that students have all the tools needed to learn through experimentation at their disposal.

Auxiliary buttons are included for quitting the game and for resetting the drone rotation. The latter is especially useful when the student inputs values into the PID controller that would cause uncontrollable movement that breaks the simulation. One way that the simulation can break is by getting the drone rotating so much that it becomes stuck in the table. While this rarely happens, an easily accessible button fixes this and any other issue that may potentially arise.



Figure 7. In-game interface, with the Tune panel open.

Table 2. Panels allowing the player to interact with the digital model.

Description	Panels
Allows players to set a static target value or generate automatic oscillations	Target angle panel
Allows players to set the mass of the turbulence balls	Turbulence panel
Allows players to modify K_P, K_I, K_D, and ISat (integral saturation)	Tune panel
Will contain settings in future releases.	Settings panel
Provides short instructions describing the main actions undertaken by the player	Help panel

4. Results

4.1. Testing Methodology

4.1.1. Goal and Research Questions

The goal of the study was to check if a serious game containing the digital twin-like model of a hypothetical one-axis drone can support instructors during PID teaching. Specifically, the serious game's usefulness in diagnosing students' tuning difficulties was evaluated using metrics collected in-game as the game was designed to provide a safe, low-risk environment for students to tinker in. On top of this, a confirmatory analysis was performed to check if unsupervised sessions can also produce measurable gains in tuning performance.

To better meet the stated goal, the following research questions were addressed:

- RQ1: Do error values provide actionable diagnostic signals that can help instructors identify PID tuning issues that students are having?
- RQ2: Does a short, unsupervised tinkering session improve quiz scores and stability-oriented metrics?
- RQ3: Does tinkering effort decrease from the first tinkering session to the second one without degrading stability-oriented performance?

Research question 1 will be answered in an exploratory manner. For this, metrics will be plotted and conclusions will be drawn. When discussing the error values, we will

show how common PID issues manifest themselves. In the confirmatory part (research questions 2 and 3), pre/post quiz scores and stability-oriented metrics will be formally tested across the two tinkering sessions.

4.1.2. Participants and Context

There were a total of 21 participants in this study, aged 18–33, all part of a lab session at an international student festival in Timisoara (ISWinT 2025). The data was collected using anonymous IDs. The students had very diverse cultural and technical backgrounds, along with different prior PID-controller knowledge. The backgrounds are as follows: computer science, multimedia, aerospace, and electrical engineering. The field with the most students was computer science, with 14 students. Multimedia and electrical engineering had 3 students each, while aerospace contained a single student.

4.1.3. Study Design

The study followed a within-subjects design, with each student serving as their own control across the quizzes and two rounds of playing the serious game under identical task conditions. As allowing students to change the target angle would make identifying the error from analytics data impossible without additional information, the target angle was set to 0 degrees. The option to customize the turbulence balls, while implemented, was disabled in the official testing rounds as the balls came already set to a relevant value, and further changing the mass could have led to incredibly strong or weak disturbances, which would lower their relevance. To reduce the complexity and to maintain focus on fundamental concepts, participants were advised not to modify the integral gain and saturation as they were not treated in-depth in the theory part due to time constraints.

The participants were instructed to tune the PID controller until the system stabilized itself at 0.4° roll angle. Optionally, they were informed of the ability to spawn turbulence balls to observe the tuned system's response and gauge its stability. Tinkering sessions were limited to 10 min, after which the game would automatically stop, while also allowing any student who was satisfied with their result to quit early via a dedicated button. Upon quitting (automatically or as a choice), the game sent the collected analytics data to the back-end.

4.1.4. Tools

Data was collected using two different means: forms and metrics telemetry, giving us both subjective, student questionnaires, along with objective metrics that could be analyzed.

For making quizzes, Google Forms have been used for their ease of setup, distribution, and because they offer the possibility to export to CSV files. Students were given the option to choose between randomly generated IDs, which were present in all quiz exports and both analytics results, enabling the merger of all data collected from a user. IDs are vital as the study relies on measuring differences between different stages of knowledge. There is no way of associating the real student to a username available as the choice was neither tracked not written anywhere, and the data provided is not enough to make any link between a real person and a username.

A custom back-end was developed for logging analytics data to ensure full control over recorded data. Data is saved locally, on the author's private server, with no reliance on third-party services, and can be easily exported to JSON files. The back-end also relied on the IDs for merging data for the same user, so no personally identifiable information was recorded. The implementation details of the custom back-end are outside the scope of this

paper. Later subsections will describe the relevant information, mainly the data captured and its usage.

4.1.5. Procedure

The study consisted of multiple steps, provided in fixed order and undertaken in two days, with the first five steps being performed on the first day, and the last two on the second day.

Step 1: Background quiz

Students were given a short survey for understanding their background and matching it to a unique ID. The survey contained questions about their technical background, country of origin, country they were currently studying in, their age, and their prior experience with PID controllers.

Step 2: 30-min theory/practical course

A concise course covering PID controllers was taught. It presented topics ranging from the need for having these controllers to the math needed to understand why they work the way they do. This was followed by two examples of PID controllers unrelated to the app and was used to build intuition about PID controllers and the way they work.

Step 3: App introduction and controller functioning quiz

The serious game was presented, showing the available features and highlighting the ability to change the parameters for tuning the digital model. Following the presentation, students completed a quiz targeting their understanding of parameters related to and controlled by the PID controller (controller's goal, error computation, and how the PID output is used). These questions align directly with the game's learning objectives.

Step 4: First tinkering phase

Students were instructed to tune K_P, K_I, and K_D to stabilize the roll angle as close to 0° as possible. While they were working, the custom back-end recoded the following session parameters:

- error value in degrees for the entire session
- last error value of the sessions
- total session duration in seconds, up to a maximum of 600 s, as the game stopped at that time
- number of parameter changes performed during the session

Step 5: Post-tinkering quiz

Using the same questions as the previous quiz, this quiz aimed to gauge the impact that tinkering with a digital reproduction of a real system had on the understanding of core concepts.

Step 6: Second tinkering phase

Students replayed the serious game and tuned it again, under the same conditions as the previous tinkering phase. The second set of data was to be compared to the first one in order to observe if playing the serious game improved the students' ability to tune a PID controller.

Step 7: Final quiz

At the end of the study, students were asked to self-assess the improvement in their understanding of the inner workings of the PID controller, and to provide feedback in the form of improvements to the app that could increase teaching effectiveness.

4.1.6. Collected and Derived Metrics

The multiple datasets were merged using the anonymous unique IDs. Quizzes provided participants' background details (age, country of origin, country of study, technical background, and prior PID exposure) and pre- and post-tinkering-session self-assessed scores for their understanding of the PID controller, along with feedback on the serious game. Figure 8 shows the format of the JSON file in which all quiz-provided information was saved. The tinkering sessions contributed with metrics taken from inside the game, which can be grouped into two sections: collected (values of game state, sent directly to the analytics back-end and derived) and determined through various means from the collected metrics.


```
{
  "username": "username",
  "Quiz1":{
    "What is your background? (programming, automation, media, art etc.)":"background",
    "Where are you from (country)?":"country",
    "Where do you study (country)?":"country",
    "How old are you?":0,
    "Do you have any experience with PID controllers?":"Yes/No",
    "By this, I give my consent for the data collected with the fictitious user to be collected and processed.  ":"Yes/No"
    },
  "Quiz2":{
    "What is the goal of the PID controller in the example app?":"Answer",
    "Which parameter is the PID controller trying to regulate?":"Answer",
    "How is the error calculated in the PID controller?":"Answer",
    "How is the PID controller output used?":"Answer"
  },
  "Quiz3":{
    "What is the goal of the PID controller in the example app?":"Answer",
    "Which parameter is the PID controller trying to regulate?":"Answer",
    "How is the error calculated in the PID controller?":"Answer",
    "How is the PID controller output used?":"Answer"
  },
  "Quiz4":{
    "Do you have a better understanding of the goal, workings, and usage of a PID controller?":"Yes/No",
    "Has your understanding increased by using the digital twin or was it enough after the first step?":"Yes/No",
    "What would you improve in the app to increase its usefulness for teaching PID controllers?":"Answer"
  },
  "collectionDate_Before":"yyyy-mm-dd-tt",
  "data_Before":{
  "ErrorValues_Before":[0,0,0],
  "LastSecondError_Before": 0,
  "PlayTimeSeconds_Before": 0,
  "ParameterChanges_Before": 0
  },
  "collectionDate_After":"yyyy-mm-dd-tt",
  "data_After":{
  "ErrorValues_After":[0,0,0],
  "LastSecondError_After": 0,
  "PlayTimeSeconds_After": 0,
  "ParameterChanges_After": 0
  }
}
```


Figure 8. JSON file structure.

Quizzes recorded the following data:

- Quiz 1 (background information): age, country of origin, country of study, technical background, and prior exposure to PID controllers.
- Quiz 2 (knowledge application on the digital model's PID controller): purpose of the PID controller, error computation, and how the PID controller is used by the system.
- Quiz 3 (post tinkering knowledge application on the digital model's PID controller): same format as Quiz 2, enabling direct pre–post comparison.
- Quiz 4 (final quiz): self-assessment of the perceived improvement in understanding PID controllers, plus open-ended feedback on improving the serious game.

The custom analytics back-end recorded the following metrics:

- Submission time: sent at the end of the tinkering session. Contains the date and time of the end of the current tinkering session.
- Play time: sent at the end of the tinkering session. Contains the time that the player spent in-game, in seconds.

- Number of parameter changes: tracked throughout the tinkering session and sent at the end. Contains the number of times the player applied a change to the values used by the PID controller.
- Error values: tracked throughout the tinkering session and sent at the end. Contains the value of the roll angle, sampled at 10 Hz for the entire tinkering session. Stored as a list of values in degrees.
- Final error: sent at the end of the tinkering session. Contains the last recorded error value in degrees.

From the collected metrics, the following derived metrics were computed:

- Stability threshold: the tinkering phase instructions stated that an error of less than 0.4° was set as the goal to be reached by the students and was considered good enough for the purpose of this study. For this, the minimum and maximum values present in the earliest 10 s section with the lowest error value were calculated, along with the mean error value on that segment and its starting point in time.
- Normalized area under absolute error: this metric was devised to provide a session-wide stability summary for easy comparison between sessions. It was calculated by taking the absolute value of all collected error values, followed by calculating the area under them. Finally, the obtained value was time-normalized (per minute) as comparisons between sessions must not be affected by session length.
- Parameter edit rate: because participants could end the tinkering session earlier than the maximum allocated time of 10 min, the parameter edit rate normalizes their effort by time spent in-game. It is calculated by dividing the number of parameter changes perform during a tinkering stage by the total play time.

While disturbance events (spawning of turbulence balls) were not logged explicitly, their influence is visually apparent in error plots as short-lived increases in error values, followed by quick re-stabilization. Re-stabilization time also tells us if a system has reached stability by waiting until oscillations were small enough, or if the system was tuned correctly, without oscillations. Since spawning turbulence balls was optional, this is something we can observe in only some of the metrics data. A drawback to the lack of disturbance event logging is that they can be observed through the error values only on a system that is at least partially stable.

Per-participant differences between the two tinkering sessions were calculated and analyzed in the data interpretation subsections to summarize within-subject change under identical conditions.

4.2. Data Interpretation

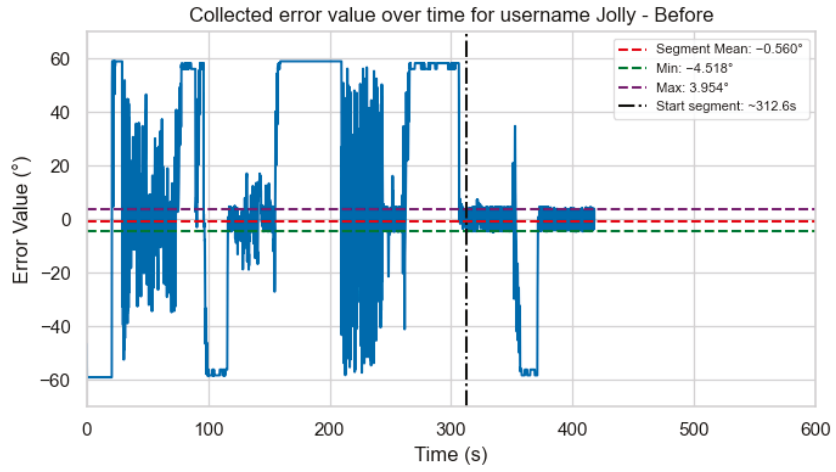
This section contains the analysis of the collected data with respect to the previously stated research questions. The diagnostic value of collected analytics metrics (RQ1), the effects of unsupervised tinkering on knowledge (RQ2), and improvements in efficiency without degradation of tuning performance (RQ3) will be explored.

4.2.1. RQ1—Do Error Values Provide Actionable Diagnostic Signals That Can Help Instructors Identify PID Tuning Issues That Students Are Having?

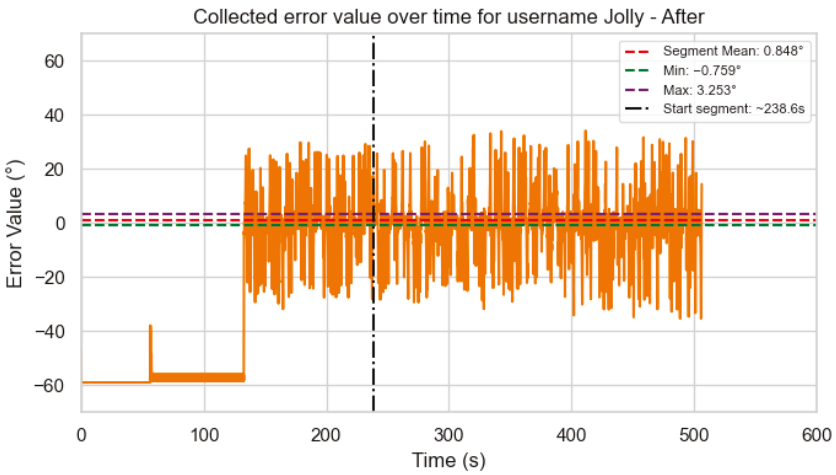
The collected error values sampled at 10 Hz can be easily plotted and then visually analyzed alongside each other for both the day 1 and day 2 sessions (before and after). This was performed for each student to see if practical hints can be prepared by an instructor to help students overcome potential PID tuning issues. In the following paragraphs we will explore the error values of three students and see what conclusions about their knowledge

can be drawn. Each plot shows the initial position of the drone. When the game starts, the drone is laying on its side, with a roll value of around $-60°$.

Figure 9a shows the error values recorded from a computer science student during the first tinkering session (before). In it, we can observe a high error value, which reaches even negative or positive 60°. This clearly indicates a very weak power applied to the digital model, likely as a result of a K_P value that is too small for the system. This implies a lack of understanding of the main concepts, including the proportional term of a PID controller. This is in contrast to the second tinkering stage (after), presented in Figure 9b, in which the one-axis drone maintains an oscillation of $\pm 30°$. From the plot shape we understand that the system has an adequate K_P value, since it has enough power to overcome gravity, but the K_D term is tuned incorrectly, as shown by the massive overshoot. Both plots contain an additional four terms related to the best (smallest) continuous 100 error values: min error value of the segment, max error value of the segment, mean of the error values present in the segment, and the start time of the segment. This was calculated to more reliably highlight potential differences in the error. In this case, the error value clearly improved, since the distance between the lines determined by the minimum and maximum values of the segment was lowered significantly. After the first session, an instructor can easily identify an inadequate P term and help the student focus on that concept first. After the second session, the instructor should encourage the student to revise the PID-controller theory with a focus on the D term in order to improvement tuning in a potential third session.



(**a**)



(**b**)

Figure 9. Computer science student's (username: Jolly) error value over time. (**a**) Before and (**b**) after.

The analysis of this student's error values shows an acceptable understanding of the PID controller from the start (Figure 10a), with errors between $\pm 20^\circ$ for the most part during the first session. This shows that the student has easily learned how to tune the P term, but still struggles with tuning the D term. The second session, however, shows a near-perfectly tuned system, with the best segment (defined by the best 100 continuous error values) being between -0.053° and 0.022°. Small, temporary, and negatively inclined deviations from stability can be observed even after the system has been tuned. They are the result of the player choosing to spawn the aforementioned turbulence balls with the purpose of testing the system's stability when disturbed by outside forces. The system handles both short, singular turbulence balls (observed as short-lived deviations) and large groups of balls (observed as longer-lasting deviations) with minimal oscillations and overshoot. Finally, the analysis reveals that the player required significantly less time to tune the system while reaching higher tuning performance.

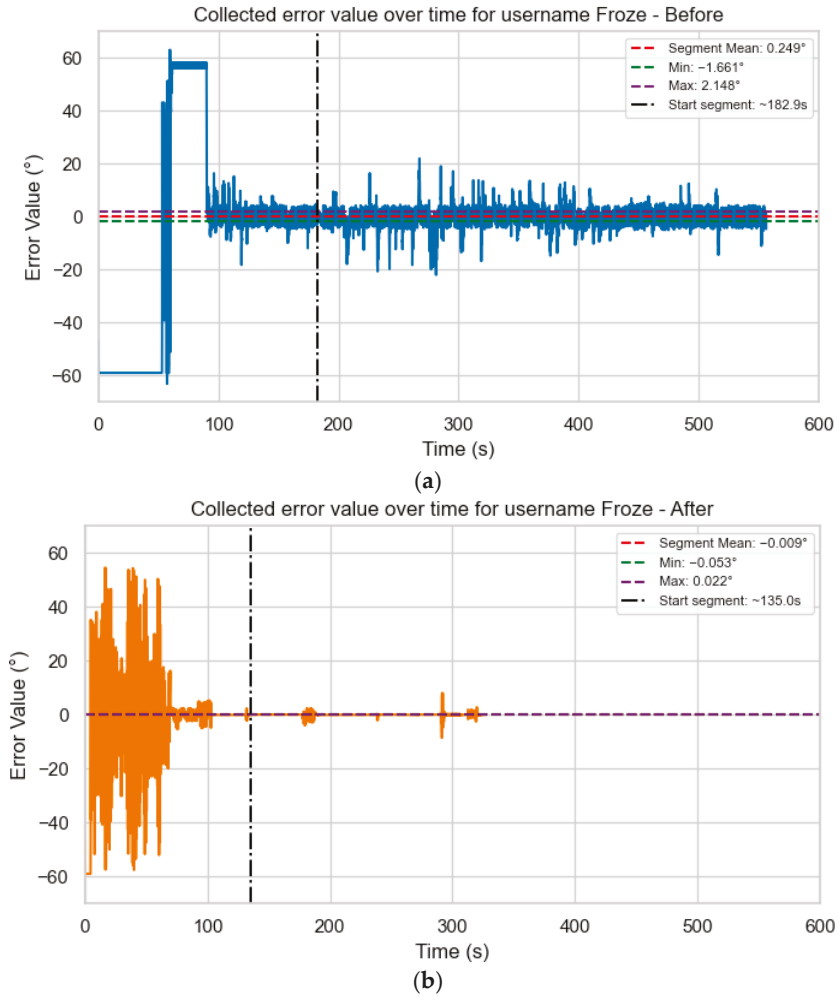


Figure 10. Computer science student's (username: Froze) error value over time. (**a**) Before and (**b**) after.

While the previous two students had a computer science background, the student whose error is plotted in Figure 11a has a multimedia background. A struggle to obtain the error to improve can easily be observed in the first half of Figure 11a, followed by a somewhat stable period 300 s after the session's start. An acceptable understanding of PID controllers is, therefore, present even after the first tinkering session, with instructors well placed to advise the student to improve K_D. It can be observed that the knowledge

is retained until the second tinkering session as the student quickly brings the error in the same interval as in the first session. Then, after about 150 s, the student starts tuning the D term, thus lowering the error to near-perfect values. After this stage, the controller is tuned correctly and responds to external disturbances well.

These example analyses show how information can be inferred from the plotted error values. Instructors can identify patterns such as drift, overshoot, and jitter and point students to their likely cause and give targeted advice after each attempt. In many cases, the second session either reduces the error or reaches the same error in significantly less time. Overall, the serious game works best as a supervised learning tool that helps focus feedback rather than as a standalone tutor. For information on all students who participated in this study, the full set of plots can be consulted in the provided Git repository [47].

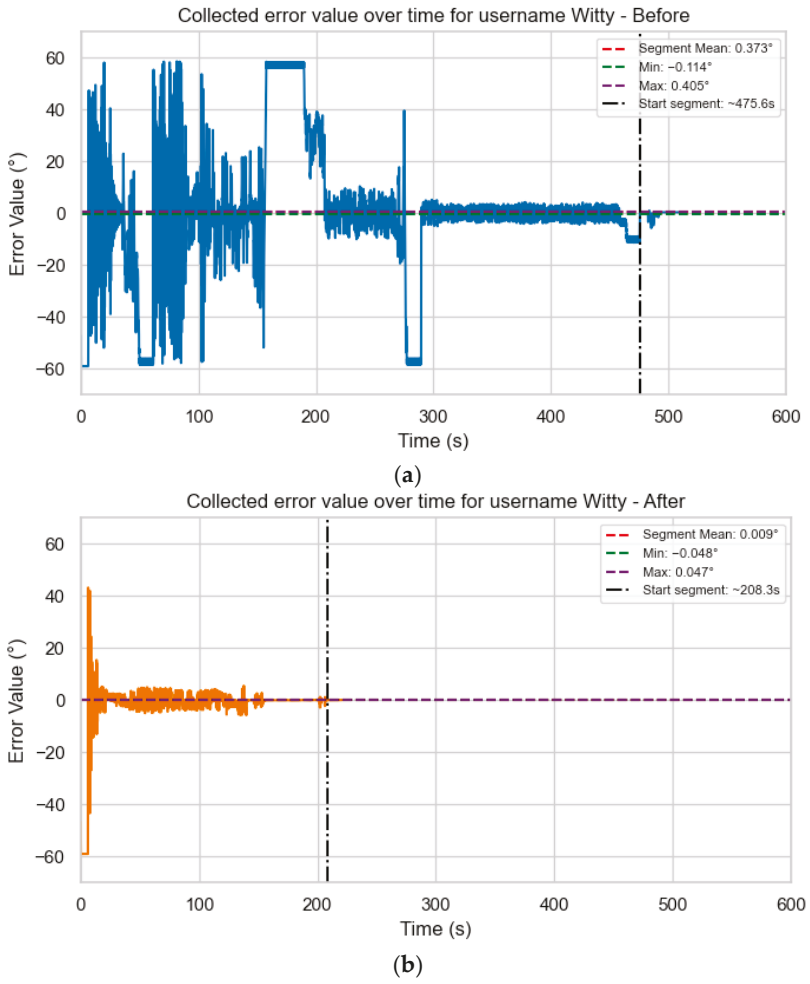


Figure 11. Multimedia student's (username: Witty) error value over time. (**a**) Before and (**b**) after.

4.2.2. RQ2—Does a Short, Unsupervised Tinkering Session Improve Quiz Scores and Stability-Oriented Metrics?

In order to see if unsupervised sessions of playing the serious game helped students improve the theoretical understanding of PID controllers, a score was calculated for each student who provided answers to Quiz 2 (administered before playing the serious game) and Quiz 3 (administered after playing the serious game for the first time). Responses were assigned scores that reflected the percentage of correct answers, which means that the only possible values were 0, 0.25, 0.5, 0.75, and 1.

Students scored 0.666 on average before the first tinkering session, and 0.619 after the first tinkering session. This leads to an insignificant-looking mean change of −0.047. To check the true relevance of this change, the Wilcoxon signed-rank test was chosen as an inferential procedure since it is suitable even when data is discrete and does not follow normal distribution [48].

Table 3 was computed while performing the Wilcoxon signed-rank test. The null hypothesis was H0: there is no systematic change in scores. The alternative hypothesis was H1: there is systematic change in scores. Since statistical analysis resulted in W = 53.5, with $p = 0.479$, we failed to reject the null hypothesis. Therefore, we can conclude that there is no systematic change in scores.

Table 3. Values obtained while performing the Wilcoxon signed-rank test on Quiz 2 and 3 scores.

Total Score Quiz 2	Total Score Quiz 3	Difference	Absolute Difference	Difference Sign	Rank	T_-	T_+	W
0.75	0.75	0	0	0	-			
1	1	0	0	0	-			
0.75	0.75	0	0	0	-			
0.75	0.75	0	0	0	-			
0.75	0.75	0	0	0	-			
0.75	0.5	−0.25	0.25	−1	6.5			
0.75	1	0.25	0.25	1	6.5			
0.75	0.5	−0.25	0.25	−1	6.5			
0.75	1	0.25	0.25	1	6.5			
0.25	0.5	0.25	0.25	1	6.5	82.5	53.5	53.5
0.75	0.5	−0.25	0.25	−1	6.5			
0.75	1	0.25	0.25	1	6.5			
1	0.75	−0.25	0.25	−1	6.5			
0.25	0.5	0.25	0.25	1	6.5			
0.5	0.25	−0.25	0.25	−1	6.5			
0.25	0	−0.25	0.25	−1	6.5			
0.25	0.5	0.25	0.25	1	6.5			
0.75	0.25	−0.5	0.5	−1	14.5			
0.75	0.25	−0.5	0.5	−1	14.5			
1	0.5	−0.5	0.5	−1	14.5			
0.5	1	0.5	0.5	1	14.5			

The effect size, expressed as a rank-biserial correlation was −0.213. This indicates a negligible difference, while the median change was 0, showing that, in fact, most students did not change their score between the two quizzes.

For checking improvements in the measured metrics, the error values were used to calculate the normalized area under absolute error. This derived metric is calculated by taking the absolute value of all collected error values, calculating the area under them, then normalizing it to be comparable regardless of time spent in-game. These metric aggregates measure deviation from 0° across the whole attempt and are comparable even when students quit the tinkering session early. Using this metric, a Wilcoxon signed-rank test was performed across the two tinkering sessions as it is appropriate for discrete data that does not follow a normal distribution.

Table 4 shows the values resulting from computing the Wilcoxon signed-rank test. The null hypothesis was H0: there is no systematic change in the normalized area under absolute error, while the alternative hypothesis was H1: there was systematic improvement. A two-tailed Wilcoxon test reported $p = 0.3737$ and W = 89. With a critical rank sum of 67, we fail to reject H0. This means that unsupervised use does not lead to a detectable reduction in normalized area under absolute error between sessions.

Taken together, the confirmatory tests show no measurable gain from short, unsupervised tinkering sessions as both quiz scores and error values did not suffer improvements between sessions. Consequently, this serious game should be viewed as an auxiliary learning tool to be used with instructor guidance, where error values are interpreted by the instructor and transformed into actionable feedback for the students.

Table 4. Values obtained while performing the Wilcoxon signed-rank test on the area under error value.

Day1_nAUE	Day2_nAUE	Difference	Absolute Difference	Difference Sign	Rank	T_-	T_+	W
243.30624	228.72069	−14.585555	14.585555	−1	1			
864.69223	901.28067	36.588443	36.588443	1	2			
894.53042	938.78995	44.25953	44.25953	1	3			
201.01026	326.56464	125.55438	125.55438	1	4			
979.56527	848.03606	−131.529215	131.52922	−1	5			
582.98658	776.97678	193.9902	193.9902	1	6			
676.55098	923.2512	246.70023	246.70023	1	7			
1063.6568	783.04613	−280.610644	280.61064	−1	8			
1025.5287	741.60723	−283.921514	283.92151	−1	9			
1189.2225	1514.816	325.59347	325.59347	1	10	142	89	89
1159.6068	776.09	−383.51679	383.51679	−1	11			
695.76433	289.4231	−406.341222	406.34122	−1	12			
962.47359	512.89044	−449.583156	449.58316	−1	13			
970.08941	497.38339	−472.706021	472.70602	−1	14			
1768.0974	1272.3261	−495.77132	495.77132	−1	15			
833.1392	196.43131	−636.707888	636.70789	−1	16			
1448.2651	666.45058	−781.81449	781.81449	−1	17			
694.37792	1500.5228	806.14484	806.14484	1	18			
796.06865	1620.2031	824.13441	824.13441	1	19			
1611.9923	2501.8061	889.81381	889.81381	1	20			
3491.1501	137.16073	−3353.989352	3353.9894	−1	21			

4.2.3. RQ3—Does Tinkering Effort Decrease from the First Tinkering Session to the Second One Without Degrading Stability-Oriented Performance?

To evaluate students' efficiency in tuning the PID controller, per-participant effort between the sessions was compared. The effort was measured directly by two metrics: the number of parameter changes per session and play time. The rate of change was derived from the previously mentioned metrics and shows the effort normalized by play time.

As shown in Figure 12, almost all students spent significantly less time during the second session. Many also performed fewer parameter changes, while keeping a somewhat stable rate of change. This implies that students worked at a similar per-minute pace but finished sooner and by performing fewer edits. This is consistent with the quicker error convergence observed visually during the analysis of the second session performed while answering RQ1.

As reported in the answer to the second research question, while there was no systematic change in quiz scores and no reliable improvement in error, there was also an absence of worsening of performance. This demonstrates that the reduction in effort did not come at the expense of stability, so students simply reached satisfactory PID tunes faster.

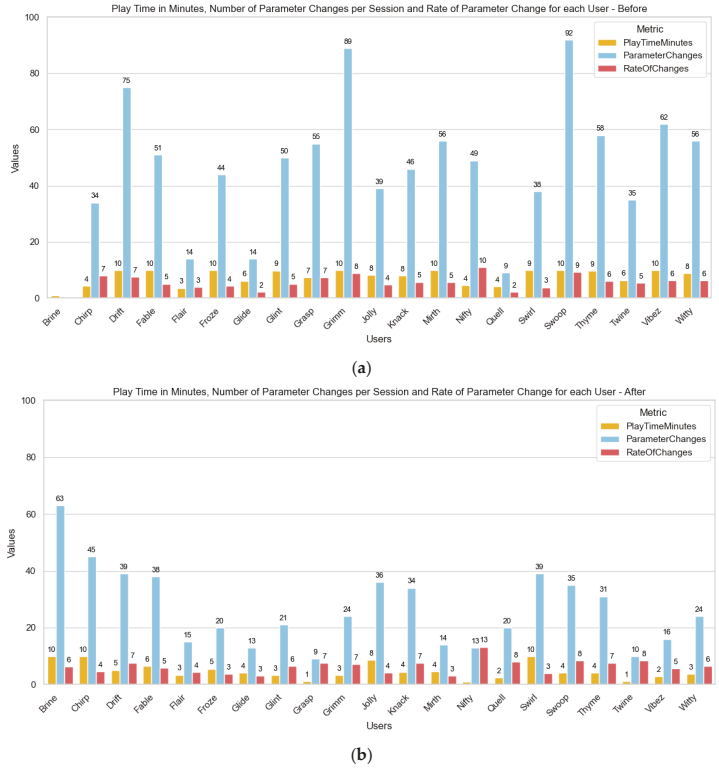


Figure 12. Bar charts containing the play time, number of parameter changes, and rate of change. (**a**) Before and (**b**) after.

5. Conclusions

This paper explored a serious game built around the digital twin-inspired model of a hypothetical one-axis drone, developed with the goal of supporting instructors when teaching PID controllers. The primary contribution is the simplified, safe, and free to deploy game, in which students can tinker with K_P, K_I, and K_D, while error values are sent to an analytics back-end for later analysis. The simplified drone with one degree of freedom keeps concepts approachable while reducing cognitive load, which makes tuning the system easier. Using a serious game comes with significant advantages over a physical system, such as eliminating safety risks and avoiding additional hardware and maintenance costs. The digital model of a simplified drone with a single degree of freedom proved effective for teaching PID tuning fundamentals as it enabled approachable and repeatable experiments while avoiding cognitive load in students. Turbulence balls spawned over the drone fulfilled their goal of imparting predictable and visually observable disturbances on the drone.

The research questions are revisited below, and the performed analyses are summarized. Collected data was explored to see what metrics can tell instructors regarding whether brief, unsupervised tinkering sessions yield significant gains and whether efficiency improves without harming stability.

RQ1: the first research question asked whether error values provide actionable diagnostic signals that help instructors identify PID tuning issues that students have. The analysis showed that error values collected at 10 Hz can be interpreted visually, from a plot, to identify error patterns and to generate clear, actionable hints for the students.

RQ2: the second research question asked whether a brief, unsupervised tinkering session shows clear improvements in quiz scores and/or stability metrics. No systematic

change was found between pre/post quiz scores and the normalized area under absolute error between the two tinkering sessions.

RQ3: the third research question asked whether tinkering effort decreases from the first to the second tinkering session without degrading stability metrics. Analytics data showed substantially fewer edits and less time spent in the second session, while stability metrics did not suffer any improvement or degradation.

The results proved that the serious game is most efficient when used as an auxiliary tool, under the supervision and guidance of instructors, as they are qualified to quickly diagnose common PID issues and give targeted hints to help consolidate understanding.

The analysis of the data revealed a couple of limitations that affected this study. The number of participants was small, sitting at 21 students, and unsupervised use of the game likely limited learning gains to undetectable levels. The lack of logging of disturbance events and very strict stability thresholds also played a small role in limiting the analysis of the data, while the quiz content provided a low-resolution scale that does not allow measurements of small changes. Further studies will include guided sessions, refined assessments, and qualitative feedback, and will be used to better evaluate learning outcomes.

Nevertheless, the analysis yielded useful insights into how digital twin-inspired models can support and enhance the teaching of automation and control concepts. Lessons learned from this study form a solid basis for future research aimed at the development of scalable teaching tools that better prepare students for the demands of Industry 4.0 and 5.0.

Data from both the quizzes and the analytics were collected anonymously, with participants giving their explicit consent to data collection with the fictitious user ID. All analytics data is available in the public repository as a JSON file. Furthermore, plots comparing analytics metrics for all students (plots and bar charts), two Python scripts (version 3.11.3) for regenerating the plots and bar charts, a CSV file that has all relevant values from the plots, and two Excel files detailing how the two Wilcoxon signed-rank tests were performed are available.

Future work will involve building an instructor dashboard that plots live error values, calculates derived metrics, and highlights students who might require help. Parameter change and disturbance events' timestamps will be logged for facilitating a more accurate study of error values. Tuning performance testing will be separated from the tinkering scene to ensure identical test conditions across students. A wind model with clear visual cues (moving leaves and dust movement effects) is planned to extend turbulence generation and enhance realism. For students, a cheat sheet with common tuning problems, the error shape associated with them, and how they can be fixed, will be provided to be used as a reference. Furthermore, the assessment will be refined, with more quizzes being more sensitive to change and the studies being performed on larger groups. Finally, this paper represents only a part of the ongoing research and development in this area as it has not yet progressed past the initial goal of developing a customizable digital model. The next steps will aim to establish a bidirectional connection between the digital model and a physical system. Achieving this link would transform the current standalone simulation into a true digital twin, fully synchronized with its physical counterpart.

Feedback provided by participants in the last quiz mostly aligned with the improvements presented in the previous paragraph. Students suggested clearer guidance during tuning, the ability to visualize error shape in plots, more realistic disturbances, and for the serious game to contain theoretical and tuning information, instead of it being presented separately, before tinkering. These suggestions will guide the next iterations of this research.

In summary, the digital model-based serious game proved to be a practical, safe, and cheaply scalable way to enhance the teaching of PID fundamentals. Even though unsupervised use yielded no performance improvements on its own, repeated exposure improved tuning efficiency without loss in tune performance. When used in tandem with guidance by instructors, the tool can help students apply the theory in a digital twin-inspired simulation, thus laying the foundation for future classrooms that bridge the gap between theory and practice.

Author Contributions: Conceptualization, R.B.; Methodology, R.B. and S.N.; Software, R.B. and S.N.; Validation, R.B. and S.N.; Formal analysis, R.B. and S.N.; Resources, R.B. and S.N.; Data Curation, R.B. and S.N.; Writing—original draft, R.B. and S.N.; Writing—review & editing, R.B. and S.N.; Visualization, R.B. and S.N.; Supervision, H.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: This work is about the design and technical evaluation of a serious game. It is not a clinical study and does not report results from a clinical trial. The activity was conducted with adult volunteers, who gave their informed consent, using anonymous IDs, and with no personally identifiable information being collected. Under our institution's policies for minimal-risk educational research with anonymous data, ethical review was not required.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study. Electronic informed consent was collected individually and stored alongside the collected data; documentation is available from the corresponding author upon reasonable request. The public repository contains only data collected anonymously. Written informed consent has been obtained from all participants for the publication of anonymous results and illustrative Figures derived from their data.

Data Availability Statement: The dataset supporting this study is openly available in our repository [47] and includes: (i) the complete set of 21 raw session logs merged and saved in JSON format, (ii) two Excel files with the both Wilcoxon signed-rank tests, and (iii) two Python files for data processing and plot generation. All generated plots are also available in the open repository.

Acknowledgments: The authors express their gratitude to the Polytechnic University of Timisoara (UPT) and the student organization Liga AC for their support in carrying out this research. The ISWinT student festival provided a good place to test the serious game with students of varied backgrounds, which contributed significantly to the relevance and quality of the results. Our gratitude also extends to the 21 students who participated in the study.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

VR	Virtual Reality
AR	Augmented Reality
PID	Proportional-Integral-Derivative
DT	Digital Twin
MPC	Model Predictive Control
UI	User Interface
ISWinT	International Student Week in Timisoara
JSON	JavaScript Object Notation
ID	Identifier
CSV	Comma-separated values

References

1. Raave, D.K.; Saks, K.; Pedaste, M.; Roldan Roa, E. How and Why Teachers Use Technology: Distinct Integration Practices in K-12 Education. *Educ. Sci.* **2024**, *14*, 1301. [CrossRef]
2. Kearney, S.; Maakrun, J. Let's Get Engaged: The Nexus between Digital Technologies, Engagement and Learning. *Educ. Sci.* **2020**, *10*, 357. [CrossRef]
3. Lampropoulos, G.; Fernández-Arias, P.; de Bosque, A.; Vergara, D. Virtual Reality in Engineering Education: A Scoping Review. *Educ. Sci.* **2025**, *15*, 1027. [CrossRef]
4. Jaramillo-Mediavilla, L.; Basantes-Andrade, A.; Cabezas-González, M.; Casillas-Martín, S. Impact of Gamification on Motivation and Academic Performance: A Systematic Review. *Educ. Sci.* **2024**, *14*, 639. [CrossRef]
5. Ramírez de Dampierre, M.; Gaya-López, M.C.; Lara-Bercial, P.J. Evaluation of the Implementation of Project-Based-Learning in Engineering Programs: A Review of the Literature. *Educ. Sci.* **2024**, *14*, 1107. [CrossRef]
6. Suleman, F.; Videira, P.; Araújo, E. Higher Education and Employability Skills: Barriers and Facilitators of Employer Engagement at Local Level. *Educ. Sci.* **2021**, *11*, 51. [CrossRef]
7. Christiansen, L.; Hvidsten, T.E.; Kristensen, J.H.; Gebhardt, J.; Mahmood, K.; Otto, T.; Lassen, A.H.; Brunoe, T.D.; Schou, C.; Laursen, E.S. A Framework for Developing Educational Industry 4.0 Activities and Study Materials. *Educ. Sci.* **2022**, *12*, 659. [CrossRef]
8. Reis, V.; Santos Baptista, J.; Duarte, J. Immersive Tools in Engineering Education—A Systematic Review. *Appl. Sci.* **2025**, *15*, 6339. [CrossRef]
9. Tobarra, L.; Utrilla, A.; Robles-Gómez, A.; Pastor-Vargas, R.; Hernández, R. A Cloud Game-Based Educative Platform Architecture: The CyberScratch Project. *Appl. Sci.* **2021**, *11*, 807. [CrossRef]
10. Di Nardo, V.; Fino, R.; Fiore, M.; Mignogna, G.; Mongiello, M.; Simeone, G. Usage of Gamification Techniques in Software Engineering Education and Training: A Systematic Review. *Computers* **2024**, *13*, 196. [CrossRef]
11. Abhadiomhen, S.E.; Nzeakor, E.O.; Oyibo, K. Health Risk Assessment Using Machine Learning: Systematic Review. *Electronics* **2024**, *13*, 4405. [CrossRef]
12. Laine, T.H.; Lindberg, R.S.N. Designing Engaging Games for Education: A Systematic Review on Game Motivators and Design Principles. *IEEE Trans. Learn. Technol.* **2020**, *13*, 804–821. [CrossRef]
13. Alshiha, M.B.; Al-Abdullatif, A.M. Gamification in Flipped Classrooms for Sustainable Digital Education: The Influence of Competitive and Cooperative Gamification on Learning Outcomes. *Sustainability* **2024**, *16*, 10734. [CrossRef]
14. Kim, J.; Castelli, D.M. Effects of Gamification on Behavioral Change in Education: A Meta-Analysis. *Int. J. Environ. Res. Public Health* **2021**, *18*, 3550. [CrossRef] [PubMed]
15. Barricelli, B.R.; Casiraghi, E.; Fogli, D. A Survey on Digital Twin: Definitions, Characteristics, Applications, and Design Implications. *IEEE Access* **2019**, *7*, 167653–167671. [CrossRef]
16. Ruiu, P.; Nitti, M.; Pilloni, V.; Cadoni, M.; Grosso, E.; Fadda, M. Metaverse & Human Digital Twin: Digital Identity, Biometrics, and Privacy in the Future Virtual Worlds. *Multimodal Technol. Interact.* **2024**, *8*, 48. [CrossRef]
17. Lu, M.; Hu, Z. Digital Twin-Enhanced Programming Education: An Empirical Study on Learning Engagement and Skill Acquisition. *Computers* **2025**, *14*, 322. [CrossRef]
18. Kritzinger, W.; Karner, M.; Traar, G.; Henjes, J.; Sihn, W. Digital Twin in manufacturing: A categorical literature review and classification. *IFAC-Pap.* **2018**, *51*, 1016–1022. [CrossRef]
19. Jones, D.; Snider, C.; Nassehi, A.; Yon, J.; Hicks, B. Characterising the Digital Twin: A systematic literature review. *CIRP J. Manuf. Sci. Technol.* **2020**, *29*, 36–52. [CrossRef]
20. Guajardo-Cuéllar, A.; Corona-Echauri, R.; Meza-Flores, R.A.; Vázquez, C.R.; Rodríguez-Arreola, A.; Navarro-Gutiérrez, M. Mixed Reality Laboratory for Teaching Control Concepts: Design, Validation, and Implementation. *Educ. Sci.* **2025**, *15*, 883. [CrossRef]
21. Hu, W.; Lei, Z.; Zhou, H.; Liu, G.-P.; Deng, Q.; Zhou, D.; Liu, Z.-W. Plug-in Free Web-Based 3-D Interactive Laboratory for Control Engineering Education. *IEEE Trans. Ind. Electron.* **2017**, *64*, 3808–3818. [CrossRef]
22. Zalozhnev, A.Y.; Ginz, V.N. Industry 4.0: Underlying technologies. Industry 5.0: Human–computer interaction as a tech bridge from Industry 4.0 to Industry 5.0. In Proceedings of the 2023 9th International Conference on Web Research (ICWR), Tehran, Iran, 3–4 May 2023; pp. 232–236.
23. Islam, M.T.; Sepanloo, K.; Woo, S.; Woo, S.H.; Son, Y.-J. A Review of the Industry 4.0 to 5.0 Transition: Exploring the Intersection, Challenges, and Opportunities of Technology and Human–Machine Collaboration. *Machines* **2025**, *13*, 267. [CrossRef]
24. Domínguez, L.G.I. Digital Twins in Industry 5.0—Systematic Literature Review. *Espacios (EPSIR)* **2024**, *9*, 641.
25. Folgado, F.J.; Calderón, D.; González, I.; Calderón, A.J. Review of Industry 4.0 from the Perspective of Automation and Supervision Systems: Definitions, Architectures and Recent Trends. *Electronics* **2024**, *13*, 782. [CrossRef]

26. Pacheco-Velázquez, A.; Ramírez-Noriega, A.; Ruiz-Ramírez, J.; Ramírez-Morales, I. Educational Simulators for Logistics Training in Industry 4.0 Contexts. *Front. Educ.* **2024**, *9*, 1331911.
27. Brandl, L.C.; Schrader, A. Realizing Ambient Serious Games in Higher Education—Concept and Heuristic Evaluation. *Trends High. Educ.* **2025**, *4*, 52. [CrossRef]
28. Žilak, M.; Car, Ž. A Framework for Improving Accessibility of Serious Games in Handheld Augmented Reality Based on User Interaction Data. *Appl. Sci.* **2025**, *15*, 2161. [CrossRef]
29. Bertozzi, G.; Paciarotti, C.; Schiraldi, M. Implementing Serious Games through a Pedagogical Lens in Engineering Education: An Experimental Study. *Eur. J. Eng. Educ.* **2024**, *49*, 1131–1157. [CrossRef]
30. Rodríguez-Calzada, L.; Paredes-Velasco, M.; Urquiza-Fuentes, J. The Educational Impact of a Comprehensive Serious Game within the University Setting: Improving Learning and Fostering Motivation. *Heliyon* **2024**, *10*, e35608. [CrossRef] [PubMed]
31. Gordillo, A.; López-Fernández, D.; Mayor, J. Examining and Comparing the Effectiveness of Virtual Reality Serious Games and LEGO Serious Play for Learning Scrum. *Appl. Sci.* **2024**, *14*, 830. [CrossRef]
32. Núñez Pacheco, R.; Espinoza Montoya, C.; Yucra Quispe, L.M.; Turpo Gebera, O.; Aguaded, I. Serious video games in engineering education: A scoping review. *J. Technol. Sci. Educ.* **2023**, *13*, 446–460. [CrossRef]
33. Sánchez, J.; Dormido-Canto, S.; Farías, G.; Godoy, F.; Dormido, S. Understanding automatic control concepts by playing games. *Int. J. Eng. Educ.* **2011**, *27*, 558–573.
34. Ortiz, J.S.; Quishpe, E.K.; Sailema, G.X.; Guamán, N.S. Digital Twin-Based Active Learning for Industrial Process Control and Supervision in Industry 4.0. *Sensors* **2025**, *25*, 2076. [CrossRef] [PubMed]
35. Guc, F.; Viola, J.; Chen, Y. Digital twins enabled remote laboratory learning experience for mechatronics education. In Proceedings of the 2021 IEEE 1st International Conference on Digital Twins and Parallel Intelligence (DTPI), Beijing, China, 15–20 July 2021; pp. 242–245.
36. Loizou, S.; Andreou, A.S. A Framework for Standardizing the Development of Serious Games with Real-Time Self-Adaptation Capabilities Using Digital Twins. *Technologies* **2025**, *13*, 369. [CrossRef]
37. Zhang, J.; Zhu, J.; Tu, W.; Wang, M.; Yang, Y.; Qian, F.; Xu, Y. The Effectiveness of a Digital Twin Learning System in Assisting Engineering Education Courses: A Case of Landscape Architecture. *Appl. Sci.* **2024**, *14*, 6484. [CrossRef]
38. Rajamäki, J. Digital Twin Technology Training and Research in Health Higher Education: A Review. *Explor. Digit. Health Technol.* **2024**, *2*, 188–201. [CrossRef]
39. Asciak, L.; Kyeremeh, J.; Luo, X.; Kazakidi, A.; Connolly, P.; Picard, F.; O'Neill, K.; Tsaftaris, S.A.; Stewart, G.D.; Shu, W.; et al. Digital twin-assisted surgery: Concept, opportunities, and challenges. *NPJ Digit. Med.* **2025**, *8*, 32. [CrossRef]
40. Cellina, M.; Cè, M.; Alì, M.; Irmici, G.; Ibba, S.; Caloro, E.; Fazzini, D.; Oliva, G.; Papa, S. Digital Twins: The New Frontier for Personalized Medicine? *Appl. Sci.* **2023**, *13*, 7940. [CrossRef]
41. Visioli, A. Teaching Control in the Era of Industry 4.0. *Front. Control Eng.* **2023**, *4*, 1228462.
42. Azofeifa, J.D.; Rueda-Castro, V.; Camacho-Zuñiga, C.; Chans, G.M.; Membrillo-Hernández, J.; Caratozzolo, P. Future Skills for Industry 4.0 Integration and Innovative Learning for Continuing Engineering Education. *Front. Educ.* **2024**, *9*, 1412018. [CrossRef]
43. Ahmad, I.; Sharma, S.; Singh, R.; Gehlot, A.; Gupta, L.R.; Thakur, A.K.; Priyadarshi, N.; Twala, B. Inclusive Learning Using Industry 4.0 Technologies: Addressing Student Diversity in Modern Education. *Cogent Educ.* **2024**, *11*, 2330235. [CrossRef]
44. Global PID Controllers Market Report. PRNewswire 2023. Available online: https://www.prnewswire.com/news-releases/global-pid-controllers-market-report-2023-industry-4-0-to-give-market-impetus{-}{-}forecast-to-2030{-}{-}302002422.html (accessed on 18 August 2025).
45. Dapkute, A.; Siozinys, V.; Jonaitis, M.; Kaminickas, M.; Siozinys, M. Enhancing Industrial Process Control: Integrating Intelligent Digital Twin Technology with Proportional-Integral-Derivative Regulators. *Machines* **2024**, *12*, 319. [CrossRef]
46. Chen, Y.-P.; Karkaria, V.; Tsai, Y.-K.; Rolark, F.; Quispe, D.; Gao, R.X.; Cao, J.; Chen, W. Real-time decision-making for Digital Twin in additive manufacturing with Model Predictive Control using time-series deep neural networks. *J. Manuf. Syst.* **2025**, *80*, 412–424. [CrossRef]
47. Brumar, R. OneAxisDrone-OpenRepository; GitHub Repository. Available online: https://github.com/AtthosTheGreat/OneAxisDrone-OpenRepository (accessed on 3 September 2025).
48. Yeo, I.-K. An algorithm for computing the exact distribution of the Wilcoxon signed-rank statistic. *J. Korean Stat. Soc.* **2017**, *46*, 328–338. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.

MDPI AG
Grosspeteranlage 5
4052 Basel
Switzerland
Tel.: +41 61 683 77 34

Multimodal Technologies and Interaction Editorial Office
E-mail: mti@mdpi.com
www.mdpi.com/journal/mti



Disclaimer/Publisher's Note: The title and front matter of this reprint are at the discretion of the Guest Editors. The publisher is not responsible for their content or any associated concerns. The statements, opinions and data contained in all individual articles are solely those of the individual Editors and contributors and not of MDPI. MDPI disclaims responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.





Academic Open
Access Publishing

mdpi.com

ISBN 978-3-7258-7920-5