

Original Paper

Tonal Auditory Discrimination Training in Youths: Design, Implementation, and Evaluation of Game-Based Versus Non-Game-Based Systems in a Cross-Sectional Study

Sergio M-Cam, MSc; Daniela Acosta; Karla I Cruz Ornelas; Priscila Montoya Beltrán, MSc; Luz M Alonso Valerdi, PhD; David I Ibarra Zarate, PhD

Escuela de Ingeniería y Ciencias, Tecnológico de Monterrey, Monterrey, NL, Mexico

Corresponding Author:

Sergio M-Cam, MSc
Escuela de Ingeniería y Ciencias
Tecnológico de Monterrey
Ave Eugenio Garza Sada 2501
Monterrey, NL 64849
Mexico
Phone: 52 (81) 8358-2000
Email: A01336974@tec.mx

Abstract

Background: Auditory discrimination training is widely used to supplement aural habilitation and rehabilitation in individuals with hearing or auditory challenges. Recently, gamification has been introduced as a strategy to enhance attention, motivation, and engagement during training. However, the relationship between user experience (UX) and auditory discrimination performance remains unclear. In addition, maintaining long-term interest in auditory training is still a challenge, particularly in younger populations, highlighting the need for more engaging and user-centered approaches.

Objective: The objective of our study was to develop and compare 2 pure-tone auditory discrimination training systems: a game-based system with dual-task gamified activities and a non-game-based control system with identical auditory tasks but without gamified elements. We further aimed to evaluate differences in UX, engagement, usability, and behavioral auditory discrimination performance between the two systems.

Methods: A 3-stage process (design, implementation, and evaluation) yielded beta versions of both systems. Both platforms were developed using identical auditory stimuli, task logic, and performance metrics to isolate the effect of gamification, differing only in the interface and game elements. In the evaluation stage, a cross-sectional study with 11 young adults (aged 18-30 y) was conducted, where participants completed usability, UX, and engagement questionnaires after using each system. Behavioral performance was assessed through mean response time, proportion of correct responses, the Balanced Integration Score, and a novel Auditory Discrimination Performance Index integrating the Weber fraction and response time. UX was measured using the User Experience Questionnaire, the Post-Study System Usability Questionnaire, and the User Engagement Scale-Short Form. Statistical analyses included Shapiro-Wilk tests, paired *t* tests, and Wilcoxon signed rank tests.

Results: The game-based system produced significantly higher scores in 6 of 11 evaluated questionnaire dimensions, including focused attention, aesthetic appeal, reward, attractiveness, stimulation, and novelty, while no significant differences were found in most auditory discrimination performance metrics. Specifically, measures such as response time, accuracy, and Balanced Integration Score showed comparable results between systems. Usability-related outcomes showed a slight advantage for the non-game-based system, whereas all engagement-related domains consistently favored the game-based system, indicating improved user involvement and overall engagement without compromising performance.

Conclusions: These findings suggest that gamification can substantially improve UX and engagement without degrading short-term discrimination performance. This indicates that gamified auditory training systems may enhance user adherence without compromising task effectiveness. This study is innovative in that it directly compares game-based and non-game-based auditory training systems under controlled conditions, isolating the effect of gamification. Unlike other studies that primarily evaluate gamified systems without a control condition, this work clarifies the specific contribution of game elements to UX and performance. These findings contribute to the development of more effective, user-centered auditory training

technologies and highlight the need for future longitudinal studies to determine whether these advantages translate into sustained auditory benefits across tasks.

Trial Registration: ISRCTN ISRCTN443661994; <https://www.isrctn.com/ISRCTN43661994>

International Registered Report Identifier (IRRID): RR2-10.1371/journal.pone.0313478

JMIR Serious Games 2026;14:e92496; doi: [10.2196/92496](https://doi.org/10.2196/92496)

Keywords: game-based training; gamification; auditory discrimination; auditory rehabilitation; user experience; usability

Introduction

Aural habilitation and rehabilitation support individuals in adapting to challenging hearing situations such as the use of hearing aids [1], bone-conduction implants [2], or cochlear implants [3]. These approaches are also relevant for populations requiring improved auditory performance, such as those with auditory processing disorders, including difficulties in localization, lateralization, auditory discrimination, pattern recognition, and temporal processing. Adaptation at early stages can be a complex task, as individuals must learn to process whole sound domains in the best way [4]. Main difficulties include accurately identifying and discriminating speech from environmental sounds (including noise), locating sound sources, and particularly listening to music [5,6], which can be quite difficult. This adaptation process may extend over months to years, since it relies not only on sensory changes but also on neural plasticity [7]. Recommended strategies for this population include auditory-verbal therapy, mainly in children [8], and auditory training [9], which contain core exercises that enhance perceptual refinement [10]. As these activities extend over long aural rehabilitation periods, individuals often engage in lengthy sessions that must be integrated into their daily routines [8]. In this way, user-centered and engaging approaches may have better reception among young people and children [11], enhancing motivation and interest in longitudinal activities [12].

In recent decades, a series of systems and tools have been developed to improve accessibility and personalization [1,13], enhancing the population's attention and rehabilitation performance. New technologies and approaches have been followed by a user-centered design and have been validated through user experience (UX) tests [14]. Yet not all of them have followed processes of user-centered design and UX evaluation that support engagement in attractive and modern gameplay and genres, fitting auditory discrimination tasks and ensuring usability—critical elements for achieving game goals. Thus, studying the real impact and analyzing the efficiency of gamification in auditory discrimination training could be quite complex [15], especially when identifying key improvement elements when no control groups are considered.

Recent studies have evaluated game-based auditory training through questionnaires assessing usability, gameplay, and attractiveness [16,17], building design and evaluation methodologies ranging from game-user research to beta testing and assessment. Other recent studies have compared gamified auditory training performance with control groups

performing different tasks [18]. However, very few studies to date have compared game-based auditory training with equivalent non-game-based auditory discrimination activities to measure both differences in general UX (focused on usability and engagement) and behavioral performance [19, 20].

A key consideration of auditory training is the selected stimuli. Most speech-based training tools are mainly available in English, which makes accessibility difficult for youths and children whose first language is not English. As an alternative, some studies have suggested a transfer effect of other auditory stimuli that may improve complex auditory perception, such as speech [21,22]. Some analyses about the effects of nonverbal training, such as music and other types of auditory temporal training, have demonstrated improvement in language skills [23], while some others have reported no differences [24]. Still, this approach addresses the challenges of language-dependent systems development.

Addressing these gaps is crucial for developing inclusive, evidence-based tools that support diverse populations and, importantly, link enhanced UX to long-term performance improvements [25]. The aim of this study was to develop and directly compare 2 pure-tone auditory discrimination training systems—a game-based system and a non-game-based system—while controlling for auditory stimuli and task structure. Specifically, this study sought to evaluate differences in UX, engagement, usability, and auditory discrimination performance between both approaches. We hypothesized that the game-based system would improve engagement and UX without negatively affecting auditory discrimination performance compared to the non-game-based system. In this manner, this work makes 3 main contributions. First, it presents the design and beta evaluation of 2 matched auditory discrimination systems that isolate the effect of gamification by keeping auditory stimuli and task logic identical. Second, it introduces an Auditory Discrimination Performance Index (ADPI) that combines just-noticeable differences (JNDs) and reaction times into a single performance metric. Third, it provides empirical evidence on how specific gamification elements influence UX, engagement, and behavioral performance in pure-tone training tasks.

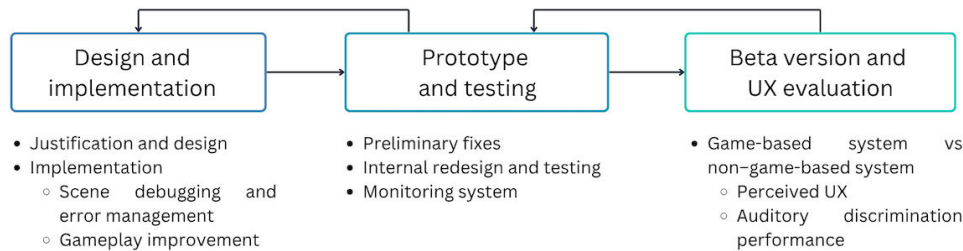
Methods

Development Stages

The methodology for both systems' design, implementation, and evaluation was adapted from [16,19] and consisted of

(1) problem definition and initial design, (2) creation of a prototype and testing, and (3) UX, usability, and engagement evaluation of the beta version, as described in Figure 1.

Figure 1. Overview of the methodology for design, implementation, testing, and evaluation of both game-based and non-game-based auditory training systems. UX: user experience.



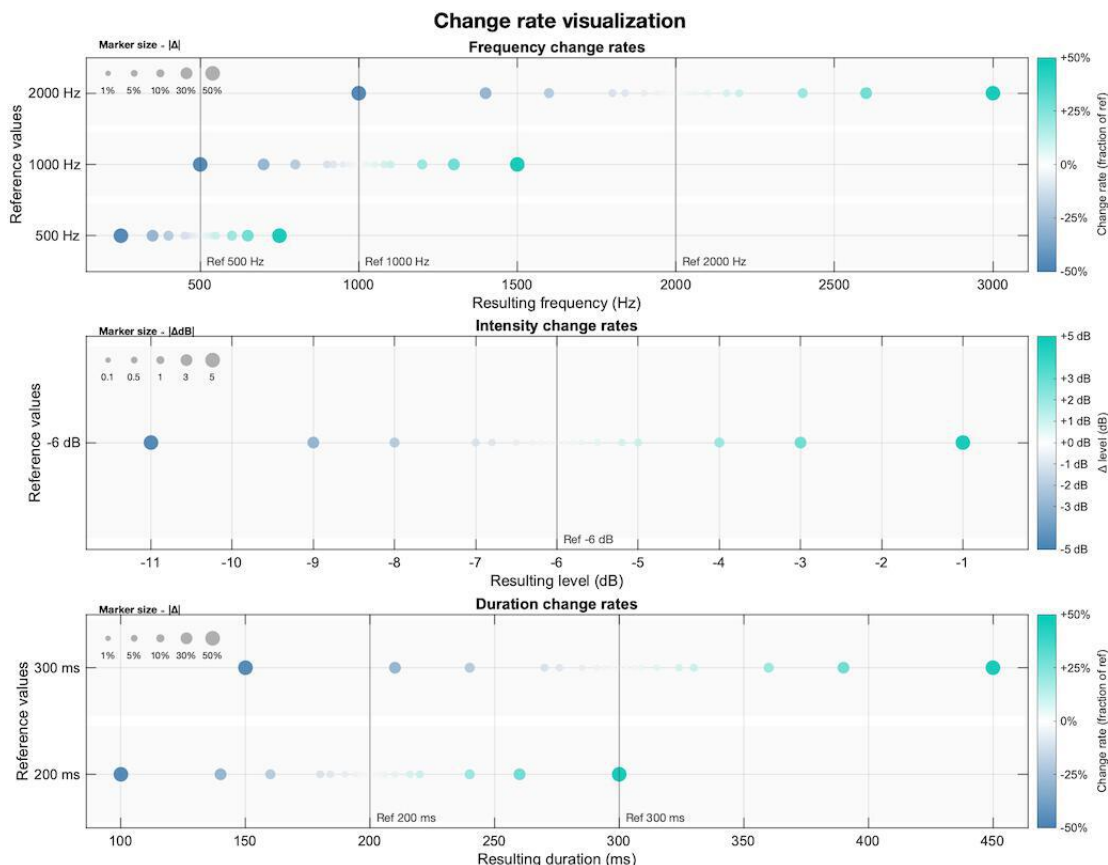
Systems Design and Implementation

Auditory Discrimination Tasks

Two different auditory training systems were developed: a game-based training system (GBS) and a non-game-based training system (NGBS). Each of them delivers sound stimuli for auditory discrimination of pitch, duration, and intensity, as the basis of the auditory discrimination hierarchy proposed in Sindrey [26]. Pure-tone pitch and intensity references and targets were selected based on the main speech features across languages [27], as shown in Figure 2. Duration

references and targets were selected based on consonant-vowel timing [28], as shown in Figure 2. In the GBS, however, these tasks are presented in a gamified format based on an infinite runner genre (Licensed Infinite Runner Engine in Unity), as used by other similar game-based training systems [16,22], while the NGBS offers the same pure-tone auditory discrimination tasks in a nongame platform. To guarantee comparability, both systems used identical auditory stimuli, task logic, and performance metrics. Only the interface and gamification elements differed.

Figure 2. Pitch, duration, and frequency references and targets. Top: pure-tone pitch references of 0.5, 1, and 2 kHz and targets from 1% to 50% change rates. Middle: pure-tone intensity reference of -6 dB FS and targets from 0.1 to 5 dB changes. Bottom: pure-tone duration references of 200 and 300 ms and targets from 1% to 50% change rates. Reference tones mainly have a pitch of 0.5 kHz, an intensity of -6 dB FS, a duration of 200 ms, and an envelope with fade-in and fade-out of 5 ms.



Non-Game-Based Training Platform

The NGBS delivered the same auditory stimuli and discrimination tasks using a static, minimalist graphical interface. Visual feedback was restricted to the participant's response to each individual trial, with no indication of score, animation, time pressure, or level of engagement. This methodology required participants to focus only on sequences of pure-tone discrimination tasks, providing a low-engagement control condition.

Game-Based Training Platform

During GBS training, participants played in an infinite-runner format in which they guided an animated character through an obstacle course as they completed different auditory discrimination tasks. This dual-task structure was intended to increase cognitive load and promote sustained engagement by combining behavioral training with interactive gameplay, following a classical dual-task paradigm [25], where the primary task is a pure-tone discrimination task and the secondary task is the infinite-runner genre game. Correct responses on the primary discrimination task allowed them to continue playing and triggered positive audiovisual feedback, whereas errors reduced lives and elicited corrective audiovisual feedback. Correct responses on the secondary task allowed them to continue increasing their score but did not directly affect the primary task's performance metrics.

Implementation Process

Both platforms were implemented in Unity (C#) using the FMOD audio engine and following an iterative, multistage development cycle.

Prototype and Alpha Phase

The design and genre were selected based on popular infinite-runner mobile games and previous implementations in auditory tasks [16,22]. Core mechanics (jumping and maneuvering) were selected based on their connection with the main discrimination goal, as described in [29], and were implemented and internally tested to ensure timing precision and data integrity. Enhancements included improved instructions, sound-level adjustment, and basic visual feedback.

Beta Phase

Dynamic scoring and refined user interface elements were introduced. Feedback from pilot testing guided the correction of perceptual calibration issues, menu structure, and visual clarity. The stable versions used for validation and experimentation incorporated all refinements and enabled the automated export of performance metrics.

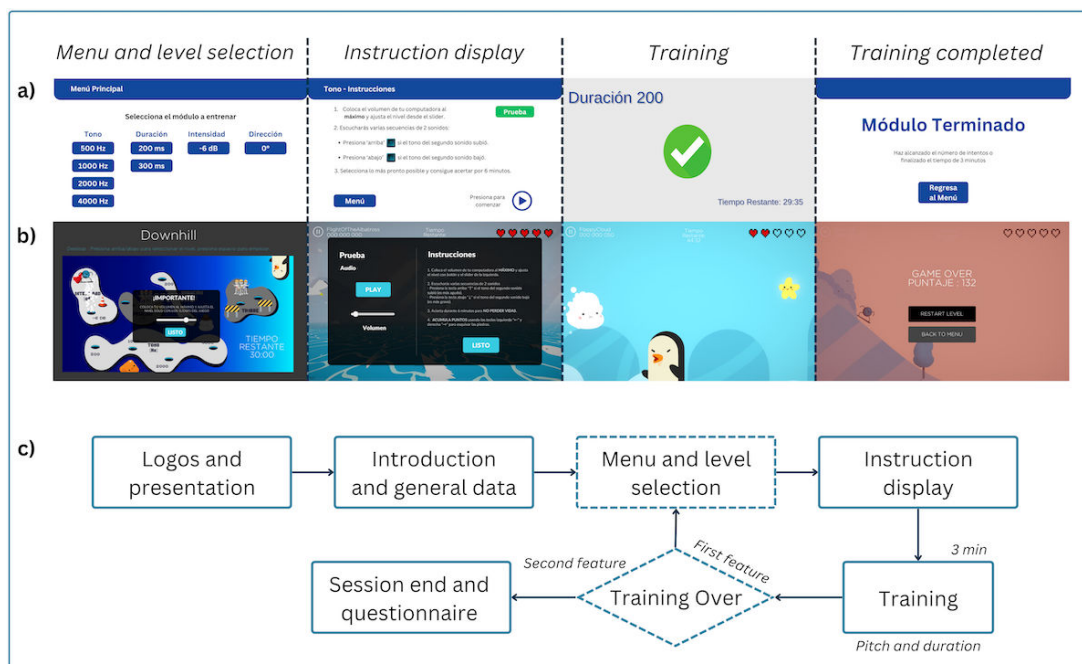
A comparison of platform features is presented in [Table 1](#), summarizing differences in task presentation, feedback, scoring, and data logging between the GBS and NGBS. This table highlights how gamification elements were deliberately implemented in the GBS and omitted from the NGBS to isolate the effect of game mechanics. A more detailed game design sheet is shown in [Table S1 \(Multimedia Appendix 1\)](#).

A version of each auditory discrimination system was exported for Windows OS for controlled testing. Each version, as well as the system user flow, is shown in [Figure 3](#).

Table 1. Comparison of game-based and non-game-based systems features.

| Feature | Non-game-based system | Game-based system |
|---------------------------|-------------------------------------|--|
| Task presentation | Static auditory stimuli | Embedded within a dynamic obstacle-navigation game |
| Feedback | Text-based success or error | Animated success or error, points, lives, and rewards |
| Engagement cues | Minimal | Visual animations, background music, rewards, and challenges |
| Data logging | Responses, intervals, training time | Responses, intervals, scores, lives, and session data |
| Questionnaire integration | Pop-up reminders | Integrated pop-ups within the game flow |
| Instructions | Text instructions | Text and interactive in-game guidance |

Figure 3. Systems interfaces and user flow. (A) Non–game-based system and (B) game-based system interfaces of 4 stages (menu and level selection, instructions display, training, and training over). (C) User flow that participants followed when testing each system for both sound features (duration and pitch).



Systems Testing and Evaluation

Stage 1: Prototype Testing

The first validation stage focused on technical reliability and preliminary user feedback. Six adult volunteers completed 5- to 10-minute sessions with each platform. Participants performed sample tasks, reported interface issues, and provided qualitative feedback on clarity, responsiveness, and sound calibration. Observations from this stage informed adjustments to auditory cue presentation, error handling, and menu navigation.

Stage 2: Beta Testing

The goal of the second stage was to evaluate usability, engagement, and performance in the near-final versions. A total of 11 participants completed structured training blocks on both platforms in counterbalanced order. After each session, they completed 3 standardized questionnaires: (1) the User Experience Questionnaire (UEQ), (2) the Post-Study System Usability Questionnaire (PSSUQ), and (3) the User Engagement Scale-Short Form (UES-SF). Feedback guided final refinements, including clearer instruction prompts, improved sound-level calibration, and enhanced data-logging reliability.

The current work focuses on findings from the beta testing stage prior to the full-scale deployment of the final version.

Ethical Considerations

This study was conducted in accordance with the principles outlined in the Declaration of Helsinki. Ethical approval was obtained from the Ethical Committee of the Instituto Tecnológico y de Estudios Superiores de Monterrey Engineering School with ID EHE-2023-11 and was registered

in the International Standard Randomized Controlled Trial Number registry (ISRCTN43661994).

Participants were given informed consent prior to study enrollment. The consent stated the procedure and described the materials, ensuring voluntary participation. Participants' data were anonymized prior to analysis to protect participants' privacy and confidentiality. No identification of individual participants or users is included in this study. No monetary compensation was given to the participants.

Participants

Sample Size, Power, and Precision

The sample for the beta testing consisted of 11 participants. The sample size was determined based on usability testing guidelines. In general, a cohort between 5 and 10 participants is often considered sufficient to identify major usability problems [19,30]. The achieved sample size met these recommendations and corresponded to the total number of participants available during the study period.

Sampling Procedures and Participant Characteristics

Participants were recruited via online flyers and personal invitations through a voluntary response sampling process from the beginning of September to the end of October 2024. All participants were adults aged 18 to 30 years (mean 21.82, SD 3.13 y) and were enrolled in either a bachelor's or master's program at Tecnológico de Monterrey. The sample consisted of 11 participants with a male-to-female ratio of 1.75:1. All participants recruited participated in the study.

Sessions were done at the NeuroTechs laboratory within Tecnológico de Monterrey. At the beginning of the session,

participants were informed about the overall purpose, procedure, and relevance of the study. Following written informed consent, participants completed an initial questionnaire comprising demographic data and questions aimed at evaluating their overall health condition and gaming hours per week.

Inclusion and Exclusion Criteria

Inclusion criteria required the participants to be aged between 18 and 30 years, report normal hearing abilities, and have no diagnosis of neurological, auditory, or attentional disorders. In addition, no participant was undergoing any medical treatment at the time of the experimental procedure. If any of these requirements were not met, participants would not be enrolled in the study.

Study Design and Procedure

Conditions and Design

This study followed a within-subjects cross-sectional experimental design in which participants interacted with both systems. The order of exposure to the systems (NGBS or GBS) was randomized for each participant to avoid potential order effects. This corresponds to an experimental manipulation with a randomized condition order.

Instrumentation

A Class-1 sound level meter (B&K Type 2270) was set to register the background noise levels during the use of each system to ensure that the experimental phase was conducted in a quiet environment and that noise was not a relevant factor affecting performance. Participants were seated comfortably in front of a laptop with the required software installed.

Steren AUD 230-NE headphones were provided to ensure uniform quality of the delivered audio across all sessions with participants.

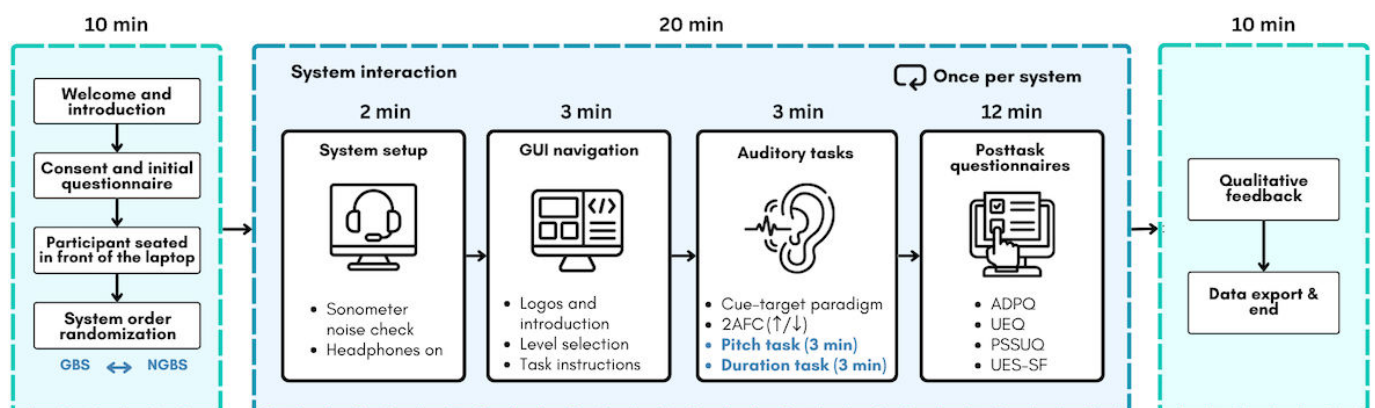
Procedure and Data Collection

Participants were asked, via a written and established set of instructions, to proceed with the interaction with the first system (randomly assigned). Participants were required to navigate through the *Logos*, *Introduction*, and *Level Selection* scenes before starting the first task to ensure the graphical user interface was understandable.

Once the first feature was selected (pitch or duration), instructions were displayed with the auditory discrimination tasks they would perform. Participants had to listen to continuous auditory sequences of 2 sounds, following a cue-target paradigm with a 2-alternative forced-choice response, and determine whether the target sound had increased (pressing the up button) or decreased (pressing the down button) compared to the cue in the sound feature that was being tested (pitch or duration). Subsequently, they could select the following sound feature (pitch or duration) in the same way. Auditory discrimination tasks were run for 3 minutes per feature. Upon completion, participants completed the UEQ, the PSSUQ, the UES-SF, and the Auditory Discrimination Performance Test for Discrimination Tasks.

The procedure was then repeated for the second system, from setting up the sonometer to the completion of the questionnaires mentioned above. All data generated throughout the session were automatically saved and exported to a compressed ZIP file available for download from the systems for further analysis. An overview of the session and experimental procedure times is illustrated in Figure 4.

Figure 4. Overview of the experimental procedure for both systems. Each system (randomly assigned) received its own interaction phase with behavioral performance and user experience assessment. 2AFC: 2-alternative forced-choice; APDQ: Auditory Discrimination Performance Questionnaire; GBS: game-based training system; GUI: graphical user interface; NGBS: non-game-based training system; PSSUQ: Post-Study System Usability Questionnaire; UEQ: User Experience Questionnaire; UES-SF: User Engagement Scale-Short Form.



Measures and Covariates

Balanced Integration Score

The Balanced Integration Score (BIS) is a measure that combines speed and accuracy from task responses while attenuating the speed-accuracy trade-off by giving equal

weights to both components [31]. BIS is defined as Equation 1.

$$\text{BIS} = z_{PC} - z_{MRT} \quad (1)$$

where z_{PC} and z_{MRT} correspond to the standardized values (z scores) across the observed proportion of correct answers

(PC) and mean reaction time (MRT) in seconds from all subjects and conditions.

Auditory Discrimination Performance Index

The ADPI is a proposed integrated version of the inverse efficiency score [32] for auditory discrimination tasks, considering the Weber fraction K obtained by dividing the JND by the reference value of the corresponding attempt, and the MRT of JND, defined as the average time in seconds that each participant took to correctly answer the attempts made only at the selected JND. JND and the Weber fraction have been used to measure discrimination thresholds in challenging hearing conditions [33] or increased cognitive load tasks [34]. ADPI was computed as shown in Equation 2.

$$\text{ADPI} = \frac{\text{MRT}_{\text{JND}}}{K \times 100} \quad (2)$$

A JND was considered valid only if the participant achieved at least 75% accuracy for that specific stimulus. If the threshold was not yielded, the next smallest JND was evaluated until a valid one was identified. For instance, suppose a participant selected the pitch modality with a reference value of 500 Hz and was tested with pitch differences of 5, 10, and 20 Hz relative to the reference value. If they did not achieve 75% accuracy at 5 Hz, but they did at 10 Hz, then the valid JND would be set at 10 Hz. Subsequently, the corresponding K value would be $10/500=0.02$. In addition, if the participant correctly answered 4 times for that JND stimulus, the mean response time of those 4 attempts is the MRT of the JND.

Questionnaires

To compare user engagement, usability, and experiential perceptions between systems, 3 validated questionnaires were presented to participants: UEQ, PSSUQ, and UES-SF. UEQ evaluates attractiveness, perspicuity, efficiency, dependability, stimulation, and novelty on a differential scale of 26 bipolar adjective pairs rated on a 7-point Likert scale transformed into a numerical scale of -3 to $+3$ [35]. PSSUQ assesses system usefulness, information quality, and interface quality through a 1 to 7 scale, where lower scores indicate better usability [36]. UES-SF measures focused attention, perceived usability, aesthetic appeal, and reward on a 1 to 7 Likert scale, where higher scores indicate stronger engagement [37].

Covariates

No covariates were included in the analyses.

Quality of Measurements and Data Diagnostics

Measurement quality was supported using a controlled laboratory environment, standardized hardware (headphones and computer setup), and consistent task instructions across participants. All sessions and procedures were supervised from informed consent until the last task was completed by each participant. No missing data were reported. Prior to performing statistical tests, data were first inspected for normality (Shapiro-Wilk tests) to select an appropriate analysis strategy.

Analytic Strategy

Analyses were conducted to evaluate differences between systems across behavioral and questionnaire-based outcomes. For the in-game key performance indicators (KPIs), Python 3.12.12 was used to perform all statistical computations with the SciPy Stats library. The tests were carried out by comparing NGBS and GBS conditions paired by training feature (pitch or duration). Data were first inspected for normality using Shapiro-Wilk tests. The parametric paired t test was used when normality conditions were met. Otherwise, the nonparametric alternative, Wilcoxon signed rank test, was applied. Statistical significance was set at $P<.05$ for all tests conducted.

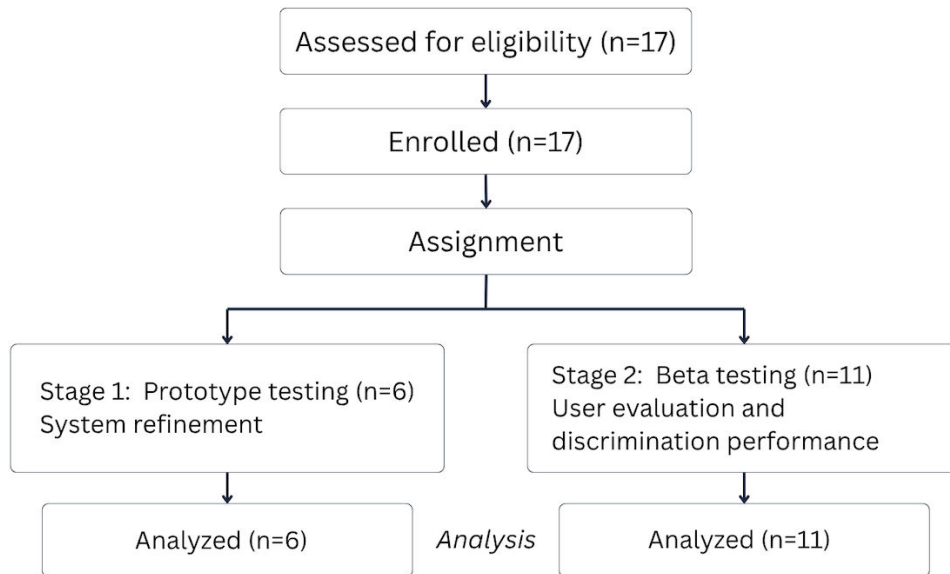
For questionnaire assessment, the Wilcoxon signed rank test using MATLAB 2025a and the Statistics and Machine Learning Toolbox was used to determine statistically significant differences between both systems across all questionnaire dimensions, considering the ordinal nature of the data. Mean scores, SD, P values, and effect sizes (r) values were calculated for each dimension per user, per questionnaire, and per system.

Results

Preliminary Data

All participants reported no hearing or visual impairment, no attentional or neurological disorders, and no current use of medication at the time of the experiment. As general information for the study, participants self-reported information regarding their sleep, eating, and hydration habits, as well as their average weekly time spent in hours playing video games. A participant flow diagram adapted from JARS guidelines is presented in Figure 5.

Figure 5. Flow of participants through each stage describing enrollment, assignment, and analysis samples.



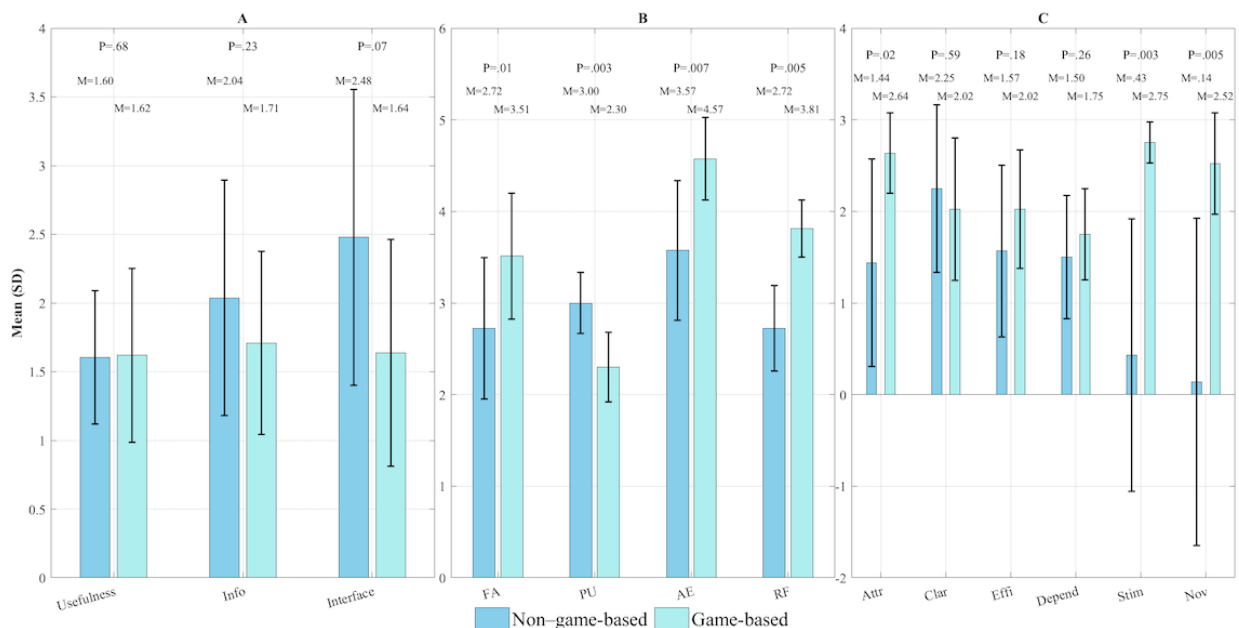
Regarding gaming exposure, 9 participants reported playing video games less than 4 hours per week, 1 participant reported between 4 and 8 hours, and 1 participant reported between 8 and 12 hours. This variable was included as a proxy for prior familiarity with interactive and gamified environments, which can influence both usability and engagement outcomes.

Due to limited variability in gaming exposure, no clear exploratory relationship was found between demographics and average weekly hours spent playing video games.

Questionnaire Outcomes

For all 3 questionnaires, participants recorded their answers across every dimension and system. For comparison between the GBS and NGBS, bar plots were generated to represent results and highlight differences in UX and overall performance patterns, as shown in Figure 6. Questionnaire responses showed significant differences and greater results in focused attention, aesthetic appeal, reward factor, attractiveness, stimulation, and novelty domains when using the gamified version.

Figure 6. Questionnaires’ results for both systems. (A) Bar graphs showing mean, SD, and *P* value results in the Post-Study System Usability Questionnaire per dimension. (B) Bar graphs showing mean, SD, and *P* value results in the User Engagement Scale Short-Form per dimension. (C) Bar graphs showing mean, SD, and *P* value results in the User Experience Questionnaire per dimension. AE: aesthetic appeal; Attr: attractiveness; Clar: clarity; Depend: dependability; Effi: efficiency; FA: focused attention; Nov: novelty; PU: perceived usability; RF: reward factor; Stim: stimulation.



Although the PSSUQ domains and the perceived usability domain of the UES-SF scored slightly higher in the NGBS

than in the GBS, this resulted in a significant difference in only 1 of the 4 domains.

Afterward, a comprehensive statistical analysis was performed, which included the calculation of metrics such as mean score, SD, P value, and r coefficient, as presented in Tables 2-4.

Table 2. Statistical results of Post-Study System Usability Questionnaire domain for both systems.

| Statistic | Dimensions | | | | | |
|-----------|-------------------------|-------------------|-------------------|-------------|------------------------|--------------|
| | Usefulness ^a | | Info ^b | | Interface ^c | |
| | GBS ^d | NGBS ^e | GBS | NGBS | GBS | NGBS |
| Mean (SD) | 1.619 (0.63) | 1.60 (0.48) | 1.70 (0.66) | 2.03 (0.85) | 1.63 (0.82) | 2.47 (1.07) |
| 95% CI | 1.2 to 2.04 | 1.28 to 1.92 | 1.26 to 2.14 | 1.46 to 2.6 | 1.08 to 2.18 | 1.75 to 3.19 |

^a $P=.68$; $r=0.02$.

^b $P=.22$; $r=-0.32$.

^c $P=.07$; $r=-0.76$.

^dGBS: game-based training system.

^eNGBS: non-game-based training system.

Table 3. Statistical results of User Engagement Scale-Short Form questionnaire domain for both systems. Significance was defined as $P<.05$ and indicated with an asterisk.

| Statistic | Dimensions | | | | | | | |
|-----------|--------------------------------|-------------------|----------------------------------|--------------|-------------------------------|--------------|---------------------|--------------|
| | Focused attention ^a | | Perceived usability ^b | | Aesthetic appeal ^c | | Reward ^d | |
| | GBS ^e | NGBS ^f | GBS | NGBS | GBS | NGBS | GBS | NGBS |
| Mean (SD) | 3.51 (0.68) | 2.72 (0.77) | 2.29 (0.37) | 2.99 (0.33) | 4.57 (0.45) | 3.57 (0.76) | 3.81 (0.31) | 2.72 (0.46) |
| 95% CI | 3.05 to 3.97 | 2.2 to 3.24 | 2.04 to 2.54 | 2.77 to 3.21 | 4.27 to 4.87 | 3.06 to 4.08 | 3.6 to 4.02 | 2.41 to 3.03 |

^a $P=.01^*$; $r=1.12$.

^b $P=.003^*$; $r=-1.7$.

^c $P=.007^*$; $r=1.26$.

^d $P=.005^*$; $r=2.01$.

^eGBS: game-based training system.

^fNGBS: non-game-based training system.

Table 4. Statistical results of User Experience Questionnaire domain for both systems. Significance was defined as $P < .05$ and indicated with an asterisk.

| Statistic | Dimensions | | Clarity ^b | | Efficiency ^c | | Dependability ^d | | Stimulation ^e | | Novelty ^f | |
|-----------|------------------|-------------------|----------------------|--------------|-------------------------|--------------|----------------------------|--------------|--------------------------|---------------|----------------------|---------------|
| | GBS ^g | NGBS ^h | GBS | NGBS | GBS | NGBS | GBS | NGBS | GBS | NGBS | GBS | NGBS |
| Mean (SD) | 2.63 (0.43) | 1.43 (1.13) | 2.02 (0.77) | 2.25 (0.91) | 2.02 (0.64) | 1.56 (0.93) | 1.75 (0.5) | 1.5 (0.67) | 2.75 (0.22) | 0.43 (1.48) | 2.52 (0.55) | 0.13 (1.78) |
| 95% CI | 2.34 to 2.92 | 0.67 to 2.19 | 1.5 to 2.54 | 1.64 to 2.86 | 1.59 to 2.45 | 0.94 to 2.18 | 1.41 to 2.09 | 1.05 to 1.95 | 2.6 to 2.9 | -0.56 to 1.42 | 2.15 to 2.89 | -1.07 to 1.33 |

^a $P = .02^*$; $r = 0.95$.
^b $P = .59$; $r = -0.18$.
^c $P = .18$; $r = 0.45$.
^d $P = .26$; $r = 0.36$.
^e $P = .003^*$; $r = 1.66$.
^f $P = .005^*$; $r = 1.44$.
^gGBS: game-based training system.
^hNGBS: non-game-based training system.

In-Game KPIs

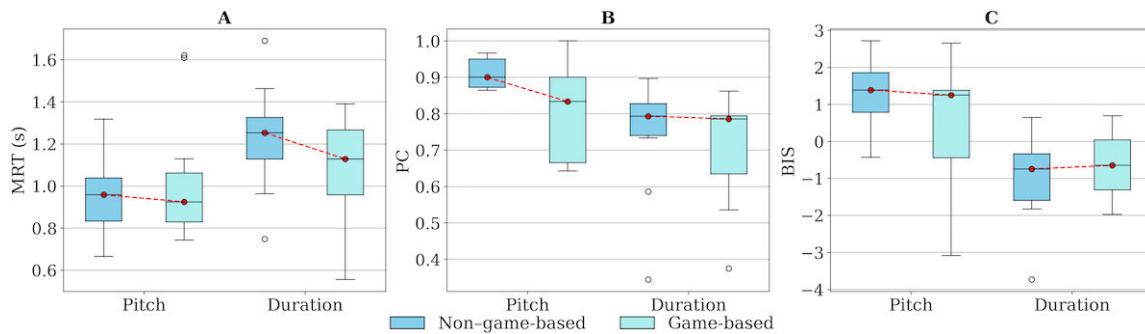
During auditory discrimination tasks, background noise was recorded when both systems were used. Background noise for the GBS had a mean equivalent continuous A-weighted sound pressure level of 43.21 (SD 3.39), and for the NGBS, it had a mean equivalent continuous A-weighted sound pressure level of 37.79 (SD 2.08), which means that for both systems, the background noise was very low during training and may not affect discrimination performance.

Balanced Integration Score

Figure 7 presents the results of MRT, PC, and BIS grouped by system and training feature. To facilitate visual comparisons across systems, a red dotted line connects the medians corresponding to the same training feature.

As can be seen for MRT and PC, the GBS medians tend to be slightly lower than those of the NGBS. However, this is not clearly observed for BIS.

Figure 7. Performance results for both systems. (A) Boxplots of mean reaction time (MRT) by system and training feature. (B) Boxplots of proportion of correct answers (PC) by system and training feature. (C) Boxplots of Balanced Integration Score (BIS) by system and training feature.



Auditory Discrimination Performance Index

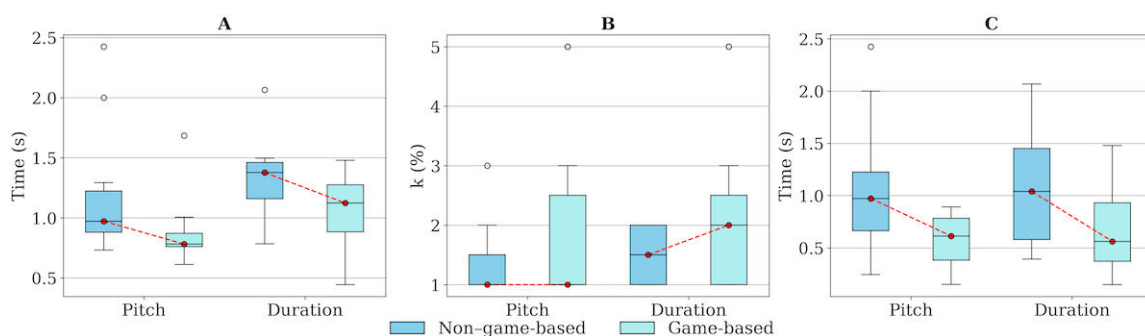
Data used to compute ADPI, along with the resulting ADPI values, are illustrated as boxplots grouped by system and training feature in Figure 8. As in the previous figure, medians are indicated by red dots, and red dotted lines connect the medians of each training modality across systems.

From the MRT of JND and ADPI subfigures, a clear trend can be observed, with the GBS showing lower values than its

NGBS counterpart. The trend is not present in the percentual Weber fraction graph, which shows no variation in pitch medians and a higher value for the GBS compared to the NGBS in the duration training feature.

To determine whether there were significant differences between systems, a statistical analysis composed of a Shapiro-Wilk normality test, followed by a paired *t* test or a Wilcoxon signed rank test, was performed.

Figure 8. Performance results for both systems. (A) Boxplots of mean reaction time (MRT) by system and training feature. (B) Boxplots of Weber fraction (*K*) by system and training feature. (C) Boxplots of Auditory Discrimination Performance Index (ADPI) grouped by system and training feature.



As shown in Tables 5-8, 9 out of 12 groups could not reject the null hypothesis of normality and therefore used a paired *t* test to analyze the relationship between the 2 systems.

Where normality could not be determined, the Wilcoxon signed rank test was performed instead.

According to the results, only 5 of the tests showed statistically significant differences ($P < .05$): MRT duration, PC pitch, MRT of JND duration, MRT of JND pitch, and ADPI pitch.

Table 5. Statistical results for the in-game key performance indicators.^a

| Tests | In-game KPIs ^b | | |
|---------------------------|--|------------------|------------------|
| | MRT ^c of JND ^d pitch | K pitch | K duration |
| Shapiro-Wilk test | | | |
| W value | 0.804 | 0.816 | 0.807 |
| P value | .01 ^e | .02 ^e | .02 ^e |
| Wilcoxon signed rank test | | | |
| W value | 8 | 1.5 | 6 |
| P value | .02 ^e | .19 | .53 |

^aThe top row of the in-game key performance indicators does not meet the normality assumption.

^bKPI: key performance indicator.

^cMRT: mean reaction time.

^dJND: just-noticeable difference.

^eSignificance was defined as $P < .05$.

Table 6. Statistical results for normally distributed in-game key performance indicators (part 1).^a

| Tests | In-game KPIs ^b | | |
|------------------------------|---------------------------|------------------|-----------------------|
| | MRT ^c pitch | MRT duration | PC ^d pitch |
| Shapiro-Wilk test | | | |
| W value | 0.964 | 0.917 | 0.904 |
| P value | .82 | .34 | .21 |
| Paired <i>t</i> test | | | |
| <i>t</i> value (<i>df</i>) | -0.832 (10) | 2.692 (9) | 2.515 (10) |
| P value | .43 | .03 ^e | .03 ^e |

^aThe KPIs in this table met the normality assumption.

^bKPI: key performance indicator.

^cMRT: mean reaction time.

^dPC: proportion of correct answers.

^eSignificance was defined as $P < .05$.

Table 7. Statistical results for normally distributed in-game key performance indicators (part 2).^a

| Tests | In-game KPIs ^b | | |
|------------------------------|---------------------------|---|-------------------------|
| | PC ^c duration | MRT ^d of JND ^e duration | ADPI ^f pitch |
| Shapiro-Wilk test | | | |
| W value | 0.972 | 0.909 | 0.926 |
| P value | .91 | .27 | .37 |
| Paired <i>t</i> test | | | |
| <i>t</i> value (<i>df</i>) | 0.113 (9) | 3.501 (9) | 2.730 (10) |
| P value | .91 | .007 ^g | .02 ^g |

^aThe KPIs in this table met the normality assumption.

^bKPI: key performance indicator.

^cPC: proportion of correct answers.

^dMRT: mean reaction time.

^eJND: just-noticeable difference.

^fADPI: Auditory Discrimination Performance Index.

^gSignificance was defined as $P < .05$.

Table 8. Statistical results for normally distributed in-game key performance indicators (part 3).^a

| Tests | In-game KPIs ^b | | |
|------------------------------|----------------------------|------------------------|--------------|
| | ADPI ^c duration | BIS ^d pitch | BIS duration |
| Shapiro-Wilk test | | | |
| W value | 0.968 | 0.964 | 0.854 |
| P value | .88 | .83 | .07 |
| Paired <i>t</i> test | | | |
| <i>t</i> value (<i>df</i>) | 2.111 (9) | 1.768 (10) | -1.294 (9) |

| Tests | In-game KPIs ^b | BIS ^d pitch | BIS duration |
|----------------|----------------------------|------------------------|--------------|
| | ADPI ^c duration | | |
| <i>P</i> value | .06 | .11 | .23 |

^aThe KPIs in this table met the normality assumption.

^bKPI: key performance indicator.

^cADPI: Auditory Discrimination Performance Index.

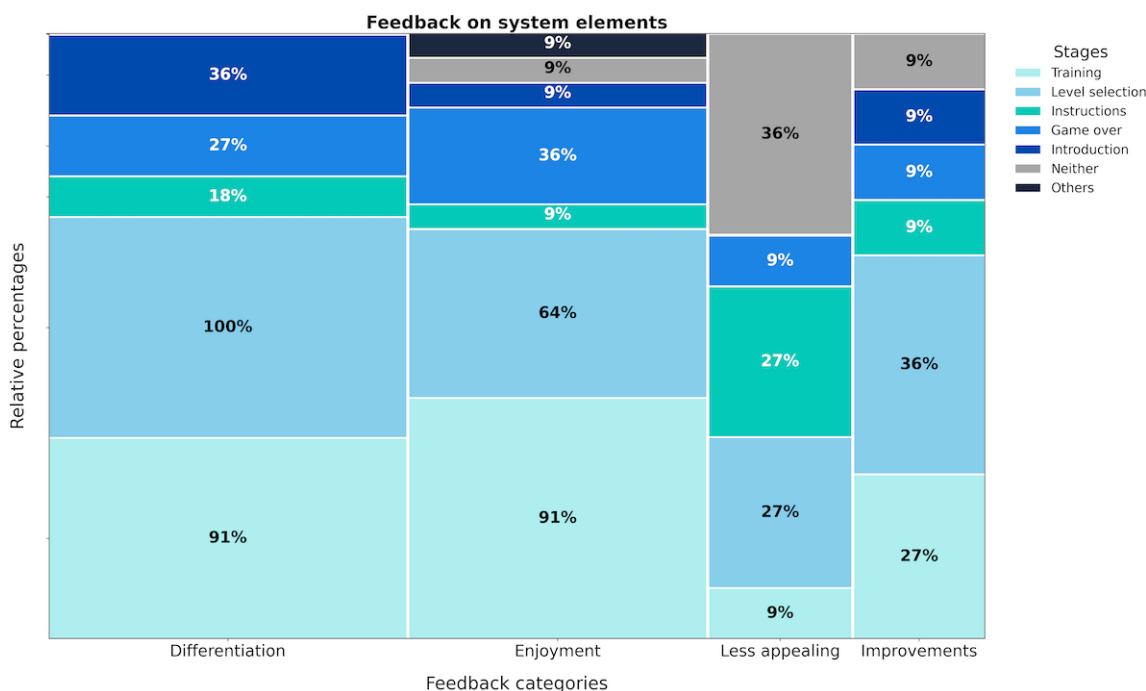
^dBIS: Balanced Integration Score.

System Elements Comparison

At the end of the experiment, participants were asked to compare both systems and answer 4 questions, mainly about the elements of the GBS that were enhanced: differentiation, enjoyment, less appealing, and improvements compared to the NGBS, as can be observed in Figure 9. They could select as many options as they considered appropriate for each question. It was observed that *Training* and *Level Selection*

were the stages of the GBS that were majorly differentiated (10/11, 91% and 11/11, 100%) and enjoyed (10/11, 91% and 7/11, 64%) by the participants compared to the NGBS. Some participants suggested smooth improvements on *Training* (3/11, 27%) for having a practical tutorial and enhancement of the reward system. *Level Selection* improvements (4/11, 36%) were suggested for clearer elements and level labels.

Figure 9. Participants’ percentual responses regarding game-based training system elements compared with the non–game-based training system in differentiation, enjoyment, less appealing, and improvements within system stages. Percentages for each individual section were calculated from the total number of participants.



Discussion

Principal Findings

This study aimed to design a game-based auditory discrimination training system and compare it with an equivalent non–game-based version through UEQs and behavioral performance. Main findings of this work suggest that in terms of UX, the game-based system significantly improved engagement-related and hedonic dimensions while maintaining comparable auditory discrimination performance across most metrics. In contrast, the non–game-based system showed a slight advantage in perceived usability, suggesting a trade-off between simplicity and engaging elements [38]. Behavioral results showed that the non–game-based system tended

toward slower but slightly more accurate responses. Despite added cognitive load, the game-based system maintained similar efficiency and no degradation of auditory discrimination, suggesting that gamification enhanced engagement without impairing performance [39].

Our study has revealed that UEQ dimensions such as attractiveness, stimulation, and novelty were consistently higher in the GBS, indicating a more engaging and appealing UX. In contrast, usability dimensions such as clarity, efficiency, and dependability showed no relevant differences between systems. These results imply that the design of game-based auditory training did not compromise usability aspects as recommended [40]. This is important since if the system is insufficiently easy to use, the user may choose to avoid the system or use it improperly even if other aspects,

such as functionality, are well thought out [41], and this can drastically impact the UX and training outcomes [42]. PSSUQ results indicated no significant differences between systems in perceived usability, suggesting that both interfaces were functionally comparable. These findings suggest that the NGBS design may be more straightforward in terms of usability, which may not be the case in more interactive systems [40]. Similarly, engagement-related dimensions in the UES-SF consistently favored the GBS, particularly in terms of focused attention, aesthetic appeal, and perceived reward. Overall, these findings indicate that gamification may enhance user engagement while not differing in levels of usability.

A factor that could be relevant and influential to interpreting UX outcomes is the rate of exposure to video games. Most participants reported “low” gaming activity (<4 h/wk), indicating limited familiarity with gamified environments. Repetitive gaming experience (replay value) can influence how users perceive and interact with gamified systems [43]. Users who regularly play video games are accustomed to dynamic interfaces, reward structures, and multitasking demands, which may reduce cognitive load, in contrast with users with limited gaming experience, who may initially perceive gamified environments as more complex and less intuitive, potentially contributing to the slightly higher usability ratings observed for the non-game-based system [44]. Still, out of the 11 participants, only 2 reported higher use of video games (4-8 h and 8-12 h), and no exploratory trend was found.

Interestingly, most of the KPIs showed nonsignificant differences between systems, with the exception of MRT (duration), MRT of JND (pitch and duration), PC (pitch), and ADPI (pitch), suggesting comparable auditory performance across systems. KPIs related to BIS showed slightly higher MRT values for the NGBS when compared to the GBS across both training features, indicating an overall slower response that could be attributed to the absence of dynamic elements in the NGBS, as supported by studies comparing tasks with and without game elements [45]. NGBS also exhibited slightly higher PC results, supported by the traditional speed-accuracy trade-off, where reaction time correlates with accuracy [46, 47]. However, when comparing training features, BIS shows a clear trend: both systems exhibit positive values in the pitch feature, indicating fast and accurate performance, whereas the negative values in duration training reflect slower and inaccurate performance. Regarding MRT of JND, related to ADPI, NGBS showed longer reaction times than GBS, while Weber fraction had no variation in pitch but a larger value for GBS duration, possibly indicating lower sensibility. However, significant differences were only found for MRT of JND in both pitch and duration, as well as for ADPI pitch. Additionally, since participants in the NGBS achieved higher MRTs but lower Weber fraction values, slightly higher ADPI results were obtained. This enhances the efficiency compensation of the speed-accuracy trade-off [31].

The dual-task element introduced in the game-based system did not lead to a decline in performance, as shown by BIS and ADPI results. Conversely, participants in the GBS

maintained similar BIS values to the NGBS and even showed an improvement in ADPI. Similar results were found in a study involving an auditory discrimination gamified system [48], where the enhancement of learning was attributed to the reward signals playing a reinforcement role. In the present study, this may suggest that the different aspects of the game-based system (eg, animated success/error, points, lives, visual stimuli, background music) served to mitigate the monotony of the primary task—repetitive auditory discrimination—rather than acting as a source of distraction, thereby enhancing engagement and maintaining performance despite the additional cognitive load [49]. This finding also highlights the potential benefits of the infinite runner game genre in these tasks, suggesting recommended levels of cognitive load to mediate task difficulty and user performance [50].

Studying the relationships between engagement and outcomes has been suggested in game-assisted learning, as motivation may play a crucial role in effectiveness [51]. This is supported by other studies that suggest that user-perceived experience and cognitive load may be good predictors of cognitive training performance [52]. Thus, players’ interactions with the systems’ mechanics (dynamics) and aesthetics (feelings of players during those interactions) [53] may enhance user-perceived experience while balancing cognitive load to optimize performance.

This study has several limitations that should be considered. First, the sample size was modestly small and limited to young adults. This may bias behavioral performance significance and restrict the generalizability of the findings to other populations but may serve as a pilot study for treating auditory processing disorders [54] or hearing device users [55]. The sample size is considered enough for addressing usability-related measurements; however, the significance of differences in behavioral performance should be interpreted with caution. Second, the sample presented a slightly unbalanced gender distribution with a male-to-female ratio of 1.75:1. Although acceptable for initial studies, it can introduce bias in UX perceptions. Future studies should aim for a more balanced sample in gender distribution to ensure broader generalizability of the findings. Moreover, the evaluation focused on short-term performance and UX and therefore does not allow conclusions about long-term training effects, as auditory processing is subject to different time courses of plasticity [56]. Although the systems were carefully controlled to isolate the effect of gamification, other factors such as individual familiarity with video games and personal preferences may have influenced engagement outcomes [57].

The usability-engagement tradeoff seen in the results between both systems could be addressed through adaptive, user-centered interface design. Interface complexity may adapt to user performance and familiarity, offering simpler layouts for novices and more dynamic environments for experienced users [58]. Additionally, progressive disclosure can introduce game mechanics, scoring, and feedback incrementally, maintaining clarity while leveraging motivational benefits [59]. Over time, neuroplasticity-related adaptation may further enhance usability, as repeated

exposure to gamified environments improves users' efficiency and perception of the interface [49]. Consequently, although non-game-based systems may provide immediate usability advantages, game-based systems could match or surpass them as familiarity increases and cognitive load decreases [59].

Taken together, the present findings highlight the value of incorporating gamification into auditory discrimination training, particularly as a strategy to enhance engagement and sustained attention without negatively affecting performance [52]. A key contribution of this study lies in directly comparing game-based and non-game-based systems under controlled conditions, which offers a more nuanced perspective on how specific game elements may influence UX and behavioral outcomes, although game elements may vary across conditions and other outcomes [60]. From an applied standpoint, the observed improvements in engagement-related domains suggest that such approaches may support greater adherence to auditory training protocols, which is critical for compliance in the real-world aural rehabilitation process [61]. Future research should further explore these effects over longer periods to determine the systems' efficiency in longitudinal studies, as it is suggested that gamification may perform more efficiently compared to control groups [62].

Conclusion

This study compared the use of gamification in auditory discrimination tasks. The introduction of gamified dual tasks favors a better qualitative perception of UX, particularly those related to engagement- and attention-related domains but may complicate interface understanding in a first approach, since more elements are displayed that require higher attention and cognitive workload when starting with system interaction. This is supported by qualitative feedback items suggesting

longer tutorials. The introduction of a dual task did not significantly affect performance in auditory discrimination tasks for duration and pitch features in the GBS compared to the NGBS, as shown by BIS and ADPI, even presenting slightly better scores for the GBS, although not significant except for ADPI in pitch. This also suggests that gamified discrimination tasks tend to decrease reaction time as part of the game mechanics, yielding focused attention and faster responses.

The current research also analyzes some key gamification elements that impact user perception. Findings suggest that the infinite runner game genre is flexible with auditory tasks and enhances engagement and attentional resources, supported by both questionnaire and discrimination performance results, providing important insights into genre selection for training compatibility. Additionally, longer exposure time is something to consider, as only one training session was evaluated and could yield slower responses due to first impressions. This may guide further research on the long-term impact of gamification on user evaluation and performance in auditory tasks, where improvements can be measured to compare both systems based on users' responses. More research is still needed in game-based training and user-centered design for auditory technology development to deliver better experiences to youths and children in aural habilitation and rehabilitation processes that enhance attention, fun, and lead to better auditory discrimination. Overall, our findings support the use of gamified, dual-task paradigms as a promising strategy to enhance UX and engagement in auditory discrimination training without compromising short-term performance. Future longitudinal studies should examine whether these experiential benefits translate into sustained improvements in auditory performance and clinical outcomes.

Acknowledgments

This study was supported by the Monterrey Institute of Technology and Higher Education (ITESM) and the National Council of Humanities, Sciences, and Technologies (CONAHCYT).

Funding

This study was funded by the Monterrey Institute of Technology and Higher Education (ITESM).

Data Availability

The datasets generated or analyzed during this study are available from the corresponding author upon reasonable request.

Authors' Contributions

Conceptualization: SM-C

Methodology: SM-C, DA, PMB, KICO, LMAV, DIIZ

Writing – original draft: SM-C, DA, PMB, KICO

Writing – review & editing: SM-C, DA, PMB, KICO, LMAV, DIIZ

Conflicts of Interest

None declared.

Multimedia Appendix 1

Differences in game-based training and non-game-based training features.

[\[DOCX File \(Microsoft Word File\), 18 KB-Multimedia Appendix 1\]](#)

References

1. Van Wilderode M, Vermaete E, Francart T, Wouters J, van Wieringen A. Effectiveness of auditory training in experienced hearing-aid users, and an exploration of their health-related quality of life and coping strategies. *Trends Hear.* 2023;27:23312165231198380. [doi: [10.1177/23312165231198380](https://doi.org/10.1177/23312165231198380)] [Medline: [37709273](https://pubmed.ncbi.nlm.nih.gov/37709273/)]
2. Caspers CJI, Janssen AM, Agterberg MJH, Cremers CWRJ, Hol MKS, Bosman AJ. Sound localization with bilateral bone conduction devices. *Eur Arch Otorhinolaryngol.* Apr 2022;279(4):1751-1764. [doi: [10.1007/s00405-021-06842-1](https://doi.org/10.1007/s00405-021-06842-1)] [Medline: [33956208](https://pubmed.ncbi.nlm.nih.gov/33956208/)]
3. Astefanei O, Martu C, Cozma S, Radulescu L. Cochlear and bone conduction implants in asymmetric hearing loss and single-sided deafness: effects on localization, speech in noise, and quality of life. *Audiol Res.* Apr 27, 2025;15(3):49. [doi: [10.3390/audiolres15030049](https://doi.org/10.3390/audiolres15030049)] [Medline: [40407663](https://pubmed.ncbi.nlm.nih.gov/40407663/)]
4. Alkhloof AM, Che Omar MB, Kazakzeh M. The role of early intervention in developing auditory perception skills in first-stage children. *Int J Acad Res Bus Soc Sci.* 2025;15(3). [doi: [10.6007/IJARBS/v15-i3/25040](https://doi.org/10.6007/IJARBS/v15-i3/25040)]
5. Mustafa M, Krishnamurthy A. A comprehensive review on real-world challenges faced by hearing aid users and innovative solutions for background noise—from lab to life. *Egypt J Otolaryngol.* 2025;41(1):62. [doi: [10.1186/s43163-025-00814-6](https://doi.org/10.1186/s43163-025-00814-6)]
6. Moore BCJ. Listening to music through hearing aids: potential lessons for cochlear implants. *Trends Hear.* 2022;26:23312165211072969. [doi: [10.1177/23312165211072969](https://doi.org/10.1177/23312165211072969)] [Medline: [35179052](https://pubmed.ncbi.nlm.nih.gov/35179052/)]
7. Maor I, Shwartz-Ziv R, Feigin L, Elyada Y, Sompolinsky H, Mizrahi A. Neural correlates of learning pure tones or natural sounds in the auditory cortex. *Front Neural Circuits.* 2020;13:82. [doi: [10.3389/fncir.2019.00082](https://doi.org/10.3389/fncir.2019.00082)] [Medline: [32047424](https://pubmed.ncbi.nlm.nih.gov/32047424/)]
8. Aaberg K, Friis IJ, Slynge CB, Britze A, Devantier L. Differences in language outcomes after cochlear implantation in children: exploring the impact of auditory verbal therapy duration (0, 1, or 3 years). *Int J Pediatr Otorhinolaryngol.* Jul 2025;194:112371. [doi: [10.1016/j.ijporl.2025.112371](https://doi.org/10.1016/j.ijporl.2025.112371)] [Medline: [40334637](https://pubmed.ncbi.nlm.nih.gov/40334637/)]
9. Lai CYY, Ng PS, Chan AHD, Wong FCK. Effects of auditory training in older adults. *J Speech Lang Hear Res.* Oct 4, 2023;66(10):4137-4149. [doi: [10.1044/2023_JSLHR-22-00621](https://doi.org/10.1044/2023_JSLHR-22-00621)] [Medline: [37656601](https://pubmed.ncbi.nlm.nih.gov/37656601/)]
10. Karah H, Karawani H. Auditory perceptual exercises in adults adapting to the use of hearing aids. *Front Psychol.* 2022;13:832100. [doi: [10.3389/fpsyg.2022.832100](https://doi.org/10.3389/fpsyg.2022.832100)] [Medline: [35664209](https://pubmed.ncbi.nlm.nih.gov/35664209/)]
11. Chandak P, Chandak A. User-centered design and rapid prototyping in healthcare education technologies for youth: a systematic review. *Prog Med Sci.* 2025;9(1):1-6. [doi: [10.47363/PMS/2025\(9\)E174](https://doi.org/10.47363/PMS/2025(9)E174)]
12. Bülow A, Janssen LHC, Dietvorst E, et al. From burden to enjoyment: a user-centered approach to engage adolescents in intensive longitudinal research. *J Adolesc.* Jun 2025;97(4):886-900. [doi: [10.1002/jad.12478](https://doi.org/10.1002/jad.12478)] [Medline: [39956787](https://pubmed.ncbi.nlm.nih.gov/39956787/)]
13. Pinzón-Díaz MC, Martínez-Moreno O, Castellanos-Gómez NM, et al. Technologies and auditory rehabilitation beyond hearing aids: an exploratory systematic review. *Audiol Res.* Jul 3, 2025;15(4):80. [doi: [10.3390/audiolres15040080](https://doi.org/10.3390/audiolres15040080)] [Medline: [40700223](https://pubmed.ncbi.nlm.nih.gov/40700223/)]
14. Katsikerou M, Ioannou D, Katsanos C. CiApplication: a user-centered web app for postoperative cochlear implant rehabilitation. Presented at: Proceedings of the 3rd International Conference of the ACM Greek SIGCHI Chapter, CHI Greece 2025; Sep 24-26, 2025:33-38; Syros, Greece. [doi: [10.1145/3749012.3749051](https://doi.org/10.1145/3749012.3749051)]
15. M-Cam S, Alonso-Valerdi LM, Ibarra-Zarate DI. Auditory discrimination improvement in young adults by a digital auditory game-based training protocol. *PLoS One.* 2025;20(10):e0313478. [doi: [10.1371/journal.pone.0313478](https://doi.org/10.1371/journal.pone.0313478)] [Medline: [41129514](https://pubmed.ncbi.nlm.nih.gov/41129514/)]
16. Garadat SN. Development and beta testing of serious game-based auditory training application to enhance perceptual learning of speech in cochlear implant recipients. *Am J Audiol.* Jun 2023;32(2):261-273. [doi: [10.1044/2023_AJA-22-00194](https://doi.org/10.1044/2023_AJA-22-00194)] [Medline: [37040306](https://pubmed.ncbi.nlm.nih.gov/37040306/)]
17. Tuz D, Isikhan SY, Yücel E. Developing the computer-based auditory training program for adults with hearing impairment. *Med Biol Eng Comput.* Jan 2021;59(1):175-186. [doi: [10.1007/s11517-020-02298-3](https://doi.org/10.1007/s11517-020-02298-3)] [Medline: [33400137](https://pubmed.ncbi.nlm.nih.gov/33400137/)]
18. Ockelmann J, Scherpiet S, Stropahl M, Giroud N. Personalized and gamified auditory-cognitive training improves naturalistic speech-in-noise comprehension in older adults with hearing loss. *NPJ Sci Learn.* Nov 19, 2025;10(1):80. [doi: [10.1038/s41539-025-00369-4](https://doi.org/10.1038/s41539-025-00369-4)] [Medline: [41258026](https://pubmed.ncbi.nlm.nih.gov/41258026/)]
19. Meral Çetinkaya M, Konukseven Ö, İralı AE. World of sounds (Seslerin dünyası): a mobile auditory training game for children with cochlear implants. *Int J Pediatr Otorhinolaryngol.* Apr 2024;179:111908. [doi: [10.1016/j.ijporl.2024.111908](https://doi.org/10.1016/j.ijporl.2024.111908)] [Medline: [38461681](https://pubmed.ncbi.nlm.nih.gov/38461681/)]
20. Berber G, Meral Çetinkaya M. Mobile game-based auditory training as an adjunct to aural rehabilitation in preschool children with cochlear implants: a randomised feasibility trial. *Int J Audiol.* Jan 5, 2026:1-11. [doi: [10.1080/14992027.2025.2607535](https://doi.org/10.1080/14992027.2025.2607535)] [Medline: [41487059](https://pubmed.ncbi.nlm.nih.gov/41487059/)]

21. Besson M, Chobert J, Marie C. Transfer of training between music and speech: common processing, attention, and memory. *Front Psychol*. 2011;2:94. [doi: [10.3389/fpsyg.2011.00094](https://doi.org/10.3389/fpsyg.2011.00094)] [Medline: [21738519](https://pubmed.ncbi.nlm.nih.gov/21738519/)]
22. de Larrea-Mancera ESL, Philipp MA, Stavropoulos T, et al. Training with an auditory perceptual learning game transfers to speech in competition. *J Cogn Enhanc*. 2022;6(1):47-66. [doi: [10.1007/s41465-021-00224-5](https://doi.org/10.1007/s41465-021-00224-5)] [Medline: [34568741](https://pubmed.ncbi.nlm.nih.gov/34568741/)]
23. Murphy CFB, Schochat E. Effects of different types of auditory temporal training on language skills: a systematic review. *Clinics (Sao Paulo)*. Oct 2013;68(10):1364-1370. [doi: [10.6061/clinics/2013\(10\)12](https://doi.org/10.6061/clinics/2013(10)12)] [Medline: [24212845](https://pubmed.ncbi.nlm.nih.gov/24212845/)]
24. Hu W, Zhang J, Chen T, Sun Y. Effects of auditory training on children with developmental language disorder: a systematic review. *Front Hum Neurosci*. 2025;19:1606860. [doi: [10.3389/fnhum.2025.1606860](https://doi.org/10.3389/fnhum.2025.1606860)] [Medline: [40606495](https://pubmed.ncbi.nlm.nih.gov/40606495/)]
25. Gagné JP, Besser J, Lemke U. Behavioral assessment of listening effort using a dual-task paradigm: a review. *Trends Hear*. Jan 2017;21:2331216516687287. [doi: [10.1177/2331216516687287](https://doi.org/10.1177/2331216516687287)] [Medline: [28091178](https://pubmed.ncbi.nlm.nih.gov/28091178/)]
26. Sindrey D. *Cochlear Implant: Auditory Training Guidebook*. Word Play Publications; 1997. URL: <https://books.google.co.in/books?id=8gcjtwAACAAJ> [Accessed 2026-05-21]
27. Roy A, Bradlow A, Souza P. Comparing long-term average speech spectra (LTAS) across languages using natural and AI speech samples. Presented at: 186th Meeting of the Acoustical Society of America and the Canadian Acoustical Association; May 13-17, 2024; Ottawa, Ontario, Canada. [doi: [10.1121/2.0002045](https://doi.org/10.1121/2.0002045)]
28. Wang C, Xu Y, Zhang J. Functional timing or rhythmical timing, or both? A corpus study of English and Mandarin duration. *Front Psychol*. 2023;13:869049. [doi: [10.3389/fpsyg.2022.869049](https://doi.org/10.3389/fpsyg.2022.869049)] [Medline: [36743611](https://pubmed.ncbi.nlm.nih.gov/36743611/)]
29. Östblad PA, Engström H, Brusk J. Game design and player experience evaluation in games for cochlear implant rehabilitation: a literature review. Presented at: 9th International GamiFIN 2025 (GamiFIN 2025); Apr 1-4, 2025; Ylläs, Finland. URL: <https://ceur-ws.org/Vol-4012/paper5.pdf> [Accessed 2026-05-21]
30. Faulkner L. Beyond the five-user assumption: benefits of increased sample sizes in usability testing. *Behav Res Methods Instrum Comput*. Aug 2003;35(3):379-383. [doi: [10.3758/bf03195514](https://doi.org/10.3758/bf03195514)] [Medline: [14587545](https://pubmed.ncbi.nlm.nih.gov/14587545/)]
31. Liesefeld HR, Janczyk M. Combining speed and accuracy to control for speed-accuracy trade-offs(?). *Behav Res Methods*. Feb 2019;51(1):40-60. [doi: [10.3758/s13428-018-1076-x](https://doi.org/10.3758/s13428-018-1076-x)] [Medline: [30022459](https://pubmed.ncbi.nlm.nih.gov/30022459/)]
32. Bruyer R, Brysbaert M. Combining speed and accuracy in cognitive psychology: is the inverse efficiency score (IES) a better dependent variable than the mean reaction time (RT) and the percentage of errors (PE)? *Psychol Belg*. 2011;51(1):5-13. [doi: [10.5334/pb-51-1-5](https://doi.org/10.5334/pb-51-1-5)]
33. Zeng FG, Tang Q, Lu T. Abnormal pitch perception produced by cochlear implant stimulation. *PLoS One*. 2014;9(2):e88662. [doi: [10.1371/journal.pone.0088662](https://doi.org/10.1371/journal.pone.0088662)] [Medline: [24551131](https://pubmed.ncbi.nlm.nih.gov/24551131/)]
34. Chiu F, Rakusen LL, Mattys SL. Cognitive load elevates discrimination thresholds of duration, intensity, and f_0 for a synthesized vowel. *J Acoust Soc Am*. Aug 2019;146(2):1077. [doi: [10.1121/1.5120404](https://doi.org/10.1121/1.5120404)] [Medline: [31472597](https://pubmed.ncbi.nlm.nih.gov/31472597/)]
35. Schrepp M, Hinderks A, Thomaschewski J. Applying the User Experience Questionnaire (UEQ) in different evaluation scenarios. In: Marcus A, editor. *Design, User Experience, and Usability Theories, Methods, and Tools for Designing the User Experience DUXU 2014 Lecture Notes in Computer Science, Vol 8517*. Springer; 2014. [doi: [10.1007/978-3-319-07668-3_37](https://doi.org/10.1007/978-3-319-07668-3_37)]
36. Lewis JR. Psychometric evaluation of the Post-Study System Usability Questionnaire: the PSSUQ. *Proc Hum Factors Ergon Soc Annu Meet*. Oct 1992;36(16):1259-1260. [doi: [10.1177/154193129203601617](https://doi.org/10.1177/154193129203601617)]
37. O'Brien HL, Cairns P, Hall M. A practical approach to measuring user engagement with the refined User Engagement Scale (UES) and new UES short form. *Int J Hum Comput Stud*. Apr 2018;112:28-39. [doi: [10.1016/j.ijhcs.2018.01.004](https://doi.org/10.1016/j.ijhcs.2018.01.004)]
38. Ntoa S, Ntagianta A, Flores F, Kolek L, Petrova A, Apostolakis KC, et al. Serious games beyond entertainment and learning: an evaluation methodology for assessing awareness raising, empathy, and social change. In: Marcus A, Rosenzweig E, Soares MM, Rau PLP, Moallem A, editors. *HCI International 2024 – Late Breaking Papers: 26th International Conference on Human-Computer Interaction, HCII 2024, Washington, DC, USA, June 29 – July 4, 2024, Proceedings, Part VII*. Springer; 2024:141-164. [doi: [10.1007/978-3-031-76821-7_11](https://doi.org/10.1007/978-3-031-76821-7_11)]
39. Turkstra LM, Johnson BA, Kartha A, Dagnelie G, Beyeler M. Gamification Enhances User Engagement and Task Performance in Prosthetic Vision Testing. *Transl Vis Sci Technol*. May 1, 2026;15(5):16. [doi: [10.1167/tvst.15.5.16](https://doi.org/10.1167/tvst.15.5.16)] [Medline: [42165642](https://pubmed.ncbi.nlm.nih.gov/42165642/)]
40. Moreno-Ger P, Torrente J, Hsieh YG, Lester WT. Usability testing for serious games: making informed design decisions with user data. *Adv Hum-Comput Interact*. 2012;2012(1):3696371. [doi: [10.1155/2012/3696371](https://doi.org/10.1155/2012/3696371)]
41. Magylaitė K, Kapočius K, Butleris R, Čeponienė L. Towards high usability in gamified systems: a systematic review of key concepts and approaches. *Appl Sci*. 2022;12(16):8188. [doi: [10.3390/app12168188](https://doi.org/10.3390/app12168188)]
42. Olsen T, Procci K, Bowers C. Serious games usability testing: how to ensure proper usability, playability, and effectiveness. In: Marcus A, editor. *Design, User Experience, and Usability Theory, Methods, Tools and Practice: First International Conference, DUXU 2011, Held as Part of HCI International 2011, Orlando, FL, USA, July 9-14, 2011, Proceedings, Part II*. Springer; 2011:625-634. [doi: [10.1007/978-3-642-21708-1_70](https://doi.org/10.1007/978-3-642-21708-1_70)]

43. Roth C, Vermeulen I, Vorderer P, Klimmt C. Exploring replay value: shifts and continuities in user experiences between first and second exposure to an interactive story. *Cyberpsychol Behav Soc Netw*. Jul 2012;15(7):378-381. [doi: [10.1089/cyber.2011.0437](https://doi.org/10.1089/cyber.2011.0437)] [Medline: [22780996](https://pubmed.ncbi.nlm.nih.gov/22780996/)]
44. Hilla Y, Stasch SM, Mack W. Habitual video gaming predicts multitasking performance while the role of cognitive capacity remains inconclusive. *Sci Rep*. Oct 27, 2025;15(1):37362. [doi: [10.1038/s41598-025-25260-5](https://doi.org/10.1038/s41598-025-25260-5)] [Medline: [41145652](https://pubmed.ncbi.nlm.nih.gov/41145652/)]
45. Wiley K, Berger P, Friehs MA, Mandryk RL. Measuring the reliability of a gamified Stroop task: quantitative experiment. *JMIR Serious Games*. Apr 10, 2024;12:e50315. [doi: [10.2196/50315](https://doi.org/10.2196/50315)] [Medline: [38598265](https://pubmed.ncbi.nlm.nih.gov/38598265/)]
46. Luce RD. *Response Times: Their Role in Inferring Elementary Mental Organization*. Oxford University Press; 1991. [doi: [10.1093/acprof:oso/9780195070019.001.0001](https://doi.org/10.1093/acprof:oso/9780195070019.001.0001)]
47. Liang Z. More haste, less speed? Relationship between response time and response accuracy in gamified online quizzes in an undergraduate engineering course. *Front Educ*. 2024;9:1412954. [doi: [10.3389/feduc.2024.1412954](https://doi.org/10.3389/feduc.2024.1412954)]
48. Zhang YX, Tang DL, Moore DR, Amitay S. Supramodal enhancement of auditory perceptual and cognitive learning by video game playing. *Front Psychol*. 2017;8:1086. [doi: [10.3389/fpsyg.2017.01086](https://doi.org/10.3389/fpsyg.2017.01086)] [Medline: [28701989](https://pubmed.ncbi.nlm.nih.gov/28701989/)]
49. Redlinger E, Glas B, Rong Y. Impact of visual game-like features on cognitive performance in a virtual reality working memory task: within-subjects experiment. *JMIR Serious Games*. Apr 28, 2022;10(2):e35295. [doi: [10.2196/35295](https://doi.org/10.2196/35295)] [Medline: [35482373](https://pubmed.ncbi.nlm.nih.gov/35482373/)]
50. Jafar A, Shabnam S, Khan A. The interplay of cognitive load and user experience: a systematic literature review across domains. *Theor Issues Ergon Sci*. 2026:1-21. [doi: [10.1080/1463922X.2026.2644284](https://doi.org/10.1080/1463922X.2026.2644284)]
51. Yu Z, Gao M, Wang L. The effect of educational games on learning outcomes, student motivation, engagement and satisfaction. *J Educ Comput Res*. 2021;59(3):522-546. [doi: [10.1177/0735633120969214](https://doi.org/10.1177/0735633120969214)]
52. Shaban A, Pearson E, Chang V. Evaluation of user experience, cognitive load, and training performance of a gamified cognitive training application for children with learning disabilities. *Front Comput Sci*. 2021;3:617056. [doi: [10.3389/fcomp.2021.617056](https://doi.org/10.3389/fcomp.2021.617056)]
53. Strmečki D, Bernik A, Radošević D. Gamification in e-learning: introducing gamified design elements into e-learning systems. *J Comput Sci*. Dec 2015;11(12):1108-1117. [doi: [10.3844/jcssp.2015.1108.1117](https://doi.org/10.3844/jcssp.2015.1108.1117)]
54. Weihing J, Chermak GD, Musiek FE. Auditory training for central auditory processing disorder. *Semin Hear*. Nov 2015;36(4):199-215. [doi: [10.1055/s-0035-1564458](https://doi.org/10.1055/s-0035-1564458)] [Medline: [27587909](https://pubmed.ncbi.nlm.nih.gov/27587909/)]
55. Nanjundaswamy M, Prabhu P, Rajanna RK, Ningegowda RG, Sharma M. Computer-based auditory training programs for children with hearing impairment – a scoping review. *Int Arch Otorhinolaryngol*. Jan 2018;22(1):88-93. [doi: [10.1055/s-0037-1602797](https://doi.org/10.1055/s-0037-1602797)] [Medline: [29371904](https://pubmed.ncbi.nlm.nih.gov/29371904/)]
56. MacLean J, Stirn J, Sisson A, Bidelman GM. Short- and long-term neuroplasticity interact during the perceptual learning of concurrent speech. *Cereb Cortex*. Jan 31, 2024;34(2):bhad543. [doi: [10.1093/cercor/bhad543](https://doi.org/10.1093/cercor/bhad543)] [Medline: [38212291](https://pubmed.ncbi.nlm.nih.gov/38212291/)]
57. Bai H, Shakya D, Wilhelmsson U, et al. The influence of visual recognition and preference in serious game development: a mixed-method study in Nepal. *Int J Serious Games*. 2025;12(3):47-68. [doi: [10.17083/229hsa32](https://doi.org/10.17083/229hsa32)]
58. Bunt A, Conati C, McGrenere J. What role can adaptive support play in an adaptable system? Presented at: The 9th International Conference on Intelligent User Interfaces (IUI '04); Jan 13-16, 2004; Funchal, Madeira, Portugal. URL: https://www.cs.ubc.ca/~joanna/papers/IUI2004_Bunt.pdf [Accessed 2026-05-21]
59. Vidaković M, Lara M, Duchi L, Whitcomb A, Paas F. Two-part onboarding for game-based learning environments. *Front Educ*. 2023;8:980881. [doi: [10.3389/feduc.2023.980881](https://doi.org/10.3389/feduc.2023.980881)]
60. Schrader C. Serious games and game-based learning. In: *Handbook of Open, Distance and Digital Education*. Springer Nature Singapore; 2023:1255-1268. [doi: [10.1007/978-981-19-2080-6_74](https://doi.org/10.1007/978-981-19-2080-6_74)]
61. Sweetow RW, Sabes JH. Auditory training and challenges associated with participation and compliance. *J Am Acad Audiol*. Oct 2010;21(9):586-593. [doi: [10.3766/jaaa.21.9.4](https://doi.org/10.3766/jaaa.21.9.4)] [Medline: [21241646](https://pubmed.ncbi.nlm.nih.gov/21241646/)]
62. Rodriguez-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*. Aug 6, 2024;10(16):e35608. [doi: [10.1016/j.heliyon.2024.e35608](https://doi.org/10.1016/j.heliyon.2024.e35608)] [Medline: [39220894](https://pubmed.ncbi.nlm.nih.gov/39220894/)]

Abbreviations

- ADPI:** Auditory Discrimination Performance Index
- BIS:** Balanced Integration Score
- GBS:** game-based training system
- JND:** just-noticeable difference
- KPI:** key performance indicator
- MRT:** mean reaction time
- NGBS:** non-game-based training system

PC: proportion of correct answers
PSSUQ: Post-Study System Usability Questionnaire
UEQ: User Experience Questionnaire
UES-SF: User Engagement Scale-Short Form
UX: user experience

Edited by Stefano Brini; peer-reviewed by Edgard Benitez-Guerrero, Raymond Sutjiadi; submitted 03.Feb.2026; final revised version received 01.May.2026; accepted 02.May.2026; published 22.Jun.2026

Please cite as:

M-Cam S, Acosta D, Cruz Ornelas KI, Montoya Beltrán P, Alonso Valerdi LM, Ibarra Zarate DI

Tonal Auditory Discrimination Training in Youths: Design, Implementation, and Evaluation of Game-Based Versus Non-Game-Based Systems in a Cross-Sectional Study

JMIR Serious Games 2026;14:e92496

URL: <https://games.jmir.org/2026/1/e92496>

doi: [10.2196/92496](https://doi.org/10.2196/92496)

© Sergio M-Cam, Daniela Acosta, Karla I Cruz Ornelas, Priscila Montoya Beltrán, Luz M Alonso Valerdi, David I Ibarra Zarate. Originally published in JMIR Serious Games (<https://games.jmir.org>), 22.Jun.2026. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Serious Games, is properly cited. The complete bibliographic information, a link to the original publication on <https://games.jmir.org>, as well as this copyright and license information must be included.