Original Paper

Game-Based Cognitive Health Training and Monitoring for Older Adults: Prospective, Single-Arm, Longitudinal, Convergent, Mixed Methods Usability and Acceptability Study

Jessica Demarchi[1], MSc; Alessandro Puiatti[2], MEng; Massimo Bortolamei[2], MEng; Matteo Metaldi[2], MEng; Masiar Babazadeh[2], PhD; Sara Levati[1], PhD

[1]Department of Business Economics, Health and Social Care (DEASS), University of Applied Sciences and Arts of Southern Switzerland, Manno, Ticino, Switzerland

[2]Department of Innovative Technologies (DTI), University of Applied Sciences and Arts of Southern Switzerland, Viganello, Ticino, Switzerland

Corresponding Author:

Sara Levati, PhD
Department of Business Economics, Health and Social Care (DEASS)
University of Applied Sciences and Arts of Southern Switzerland
Via Violino 11
Manno, Ticino 6928
Switzerland
Phone: 41 586666740
Email: sara.levati@supsi.ch

Abstract

Background: Serious games are increasingly investigated as engaging digital interventions to support cognitive health in older adults. In parallel, underdiagnosis and insufficient continuous monitoring of cognitive decline remain major challenges in routine care, highlighting the need for digital solutions that combine cognitive training with longitudinal monitoring functionalities. The potential of serious games to serve not only as cognitive training tools but also as indirect sources of behavioral data for longitudinal monitoring remains underexplored, particularly for the home care setting.

Objective: This study aimed to evaluate the usability and acceptability of a set of serious games designed to train cognitive functions in older adults, as well as the preliminary feedback on the perceived usefulness of a monitoring platform for home care nurses.

Methods: A prospective, single-arm, longitudinal, convergent mixed methods usability and acceptability study in 2 consecutive rounds involving distinct cohorts of older adults receiving home care was conducted. Participants used the beSerious gaming platform over approximately 3 months. Qualitative data were obtained through focus groups and interviews with older adults and home care nurses, while quantitative data were collected through usability questionnaires, a study-specific game evaluation questionnaire, and gameplay analytics. No evaluation of the training effects was conducted at this stage.

Results: A total of 18 older adults (mean age 79 years, SD 6.6; 12/18 women) and 3 geriatric nurses participated across 2 study rounds. A total of 17 older adults completed the intervention, corresponding to a 94% retention rate. Among questionnaire respondents, game usability was rated highly, with favorable responses exceeding 90% (61/64) across usability items. Acceptability was similarly positive, with more than 85% (59/64) favorable responses across the corresponding items. These findings were echoed in the interviews and focus groups, which highlighted the accessibility, enjoyment, and perceived cognitive stimulation offered by the games. Gameplay analytics further enabled the characterization of participants' engagement, revealing heterogeneous usage profiles in terms of both temporal patterns of play and game preferences. Nurses perceived the monitoring platform as a valuable complement to routine home care by providing longitudinal information on gameplay and engagement. Minor usability issues were identified, informing future refinements of the platform.

Conclusions: The beSerious gaming platform demonstrated high usability and acceptability among older adults receiving home care. Its monitoring component was favorably evaluated by home care nurses as a potential complement to routine practice. By combining cognitive training with longitudinal gameplay monitoring, the platform offers a foundation for sustained use of serious games and could enable nurses to derive clinically meaningful information about patients' health.

These findings provide preliminary evidence of feasibility for a combined approach in home care settings. Further studies with larger and more diverse populations are needed to optimize the games for integration into older adults' daily routines, develop gameplay-derived indicators, and determine their clinical relevance.

JMIR Serious Games 2026;14:e91956; doi: 10.2196/91956

Keywords: serious games; cognitive training; cognitive monitoring; older adults; nurses; digital nursing; digital health

Introduction

Over the past decades, advances in medicine, public health, and living conditions have increased life expectancy worldwide. According to the latest United Nations report, the average longevity has risen by more than 8 years since the mid-1990s and is projected to continue increasing, reaching approximately 77 years by 2054. By the late 2070s, the number of people aged 65 years and older is expected to exceed 2.2 billion worldwide [1]. Aging is frequently associated with gradual declines in both physical and cognitive functions, with a direct impact on independence, daily functioning, and quality of life [2]. Therefore, supporting cognitive health has become a key priority in the field of geriatric care and public health [3].

In this context, serious games have emerged as a promising nonpharmacological approach to support cognitive training and mental stimulation in older adults [4,5]. These digital games combine entertainment with specific learning, therapeutic, or training objectives, and may offer an engaging alternative or complement to conventional paper-and-pencil cognitive activities. In the field of cognitive health in older adults, recent evidence shows growing interest in serious games for both healthy individuals and people affected by cognitive impairment, highlighting their potential to stimulate multiple cognitive domains [6]. Memory, attention, executive functions, and visuospatial abilities are among the domains most frequently targeted in previous serious game-based cognitive training studies with older adults [6]. Tablet-based serious games may be particularly suitable for this population because tablets are portable, relatively accessible, and rely on touch-based interaction, which may be easier to use than traditional mouse-and-keyboard interfaces [6].

However, technology adoption in the older adults population remains low, with usability and acceptability challenges, along with ambivalent attitudes toward technology, that frequently hinder uptake [7,8]. Evidence shows that education, income, prior digital experience, cognitive skills, and willingness to learn facilitate adoption, whereas perceiving learning as burdensome affects it. Perceived usefulness (eg, clear benefits, acceptable cost, and privacy) and perceived ease of use (eg, plain language, intuitive navigation, readable text, and easy clicking/tapping) are also key, alongside emotional factors such as enjoyment and technology-related anxiety [8].

Alongside these elements, timely recognition of changes in cognitive functioning remains a major challenge in the field of cognitive health, with cognitive impairment and dementia often underdiagnosed or detected only after symptoms have begun to affect daily functioning [9]. More broadly, neurologists and other physicians assess older adults only during specific clinical visits using standardized and validated tests, whereas home care nurses may have more frequent contact with older adults, but their assessments are typically based on observations made during scheduled care visits. As a result, subtle changes in cognition, motivation, daily routines, or engagement that occur between encounters may go unnoticed. Digital tools used repeatedly in everyday life have the potential to provide complementary longitudinal information on these changes. As such, a recent work has explored the feasibility of repeated game-based digital approaches to collect longitudinal data on daily cognitive functioning in older adults [10]. Similar approaches based on game-play data may help identify within-person changes that warrant further attention, always without replacing formal clinical assessment.

In light of the above, we developed beSerious, a tablet-based application featuring a set of serious games designed for older adults and a web-based monitoring platform for home care nurses. The novelty of beSerious lies in its integrated design: the games were codeveloped from scratch to target multiple cognitive domains relevant to aging, including those commonly affected by age-related cognitive decline, while also generating gameplay data that could be displayed through the monitoring platform. The tablet-based format and simple game mechanics were intended to facilitate usability and adoption among older adults. The platform was designed to provide nurses with complementary information on gameplay patterns and possible within-person changes over time, without serving as a clinical assessment or diagnostic tool. To our knowledge, no previously published system combines a purpose-built suite of serious games targeting multiple cognitive domains with a nurse-facing platform for longitudinal behavioral monitoring, with both evaluated together in a real-world home care setting. Related work has instead addressed these components (training and monitoring) separately, rather than within a single integrated platform. Several platforms evaluated purely for cognitive training lack any clinician-facing monitoring component. Project EVO targets a single cognitive domain and was evaluated in a short-term randomized controlled trial (RCT) for late-life depression [11]; Mind Frontiers offers a multidomain game suite comparable in scope to beSerious but was assessed only through pre/post testing over a single month [12]; and InSight, benchmarked against the off-the-shelf game Crazy Taxi (Hitmaker/Sega; PlayStation 2 version by Acclaim Studios Cheltenham), was evaluated through a pre/post/3-month follow-up battery in a home-based RCT, again without continuous monitoring or home care integration [13].

On the other hand, studies extracting gameplay-derived behavioral or biomarker indicators have not been embedded within an integrated training-plus-monitoring platform. The MassGeneral Institute for Neurodegenerative Disease (MIND) GamePack relies on 4 preexisting casual games and collected four months of semipassive daily gameplay for research analysis only, without a caregiver-facing interface or home care integration [10]. Gielis et al [14] extracted digital biomarkers from a single exploratory session with an off-the-shelf card game in a clinical sample for mild cognitive impairment, rather than through longitudinal use. Finally, a serious game has been developed to train nursing students in geriatric assessment skills rather than as a patient-facing tool paired with a monitoring platform, and was validated in a classroom setting with students and experts rather than older adults or practicing home care nurses [15].

As such, beSerious was designed to partially support 2 complementary functions: behavioral monitoring and cognitive training rather than formal cognitive assessment. Behavioral monitoring refers to the longitudinal collection of gameplay-derived indicators that may suggest changes in engagement and functioning over time [10], whereas cognitive training refers to the use of games to stimulate cognitive function through repeated practice [5]. In contrast, cognitive assessment involves the formal evaluation of cognitive status using standardized instruments. Regarding the training component specifically, whether gains from cognitive training generalize beyond the trained tasks to broader cognitive domains or everyday functioning—so-called far transfer—remains debated in the literature: large-scale studies and reviews report robust near-transfer effects but limited and inconsistent evidence of far transfer, including in older adult populations [16-18]. In accordance, beSerious's training component is not intended as a validated means of enhancing broader cognitive functioning, but rather as repeated, engaging practice that sustains user engagement and generates gameplay data for the monitoring platform, with any transfer effects lying beyond the scope of the current evaluation. This study therefore focuses on evaluating the usability and acceptability of the cognitive training games and the behavioral monitoring platform, without aiming to replace or replicate standardized cognitive assessment tools.

Following the initial development of the games [19], this study evaluated: (1) the usability and acceptability of the serious games among older adults, and (2) home care nurses' preliminary feedback on the perceived usefulness of the monitoring platform.

Methods

Study Design

Overview

The Serious Games per l'allenamento e il monitoraggio delle funzioni cognitive delle persone anziane (SIGMA project) was conceived and developed following the Medical Research Council framework for complex interventions [20], which identifies sequential and iterative phases including intervention development, feasibility assessment, evaluation, and implementation. The project comprised 2 main phases: (1) the development and refinement of a set of serious games for cognitive training to be used by older adults and a platform for cognitive monitoring to be used by home care nurses, and (2) the evaluation of their usability, acceptability, and user engagement. The development phase of the SIGMA project followed a participatory co-design approach, actively involving end users and relevant stakeholders throughout the iterative design and refinement process to ensure that the intervention was user-centered, contextually appropriate, and aligned with users' needs and preferences. Findings from the first phase are described in detail elsewhere [19]. This study reports findings from the second phase, a usability and acceptability evaluation of the gaming platform through a convergent mixed methods design integrating qualitative and quantitative feedback with gameplay analytics. This design was deemed methodologically appropriate given that quantitative data could characterize user-reported evaluations and gameplay-derived behavioral patterns, whereas qualitative data could yield an in-depth, contextualized understanding of older adults' and home care nurses' lived experiences, identifying dimensions that are not captured by numerical measures. The quantitative strand followed a prospective single-arm, longitudinal design, with baseline characterization of participants and outcome data collected at defined points throughout the study period, without a comparison group (see Procedure and Assessment Tools section for details). The qualitative strand followed a qualitative descriptive approach using reflexive thematic analysis, based on semistructured interviews with older adults and focus groups with both home care nurses and older adults. The 2 strands were implemented concurrently within the same rounds, analyzed independently in accordance with a convergent parallel mixed methods design, and subsequently integrated at the interpretation stage through methodological triangulation to identify convergence, complementarity, and potential discrepancies across findings.

This study was designed to assess how easily users could engage with the intervention and how appropriate, satisfactory, and feasible it was perceived to be by the intended users. Unlike an effectiveness evaluation, which examines whether an intervention achieves its intended health outcomes, this study focused on user experience, usability, and perceived suitability to inform refinement and potential future implementation. Subjective usability and acceptability were evaluated using questionnaires and semistructured interviews with end users. Objective user engagement was assessed through system-generated usage data (gaming data), capturing real-world interaction with the digital platform. Integrating self-reported and qualitative data with behavioral usage data enabled triangulation of findings and provided a comprehensive assessment of both actual engagement and perceived experience of the intervention.

Specifically, usability was defined as the extent to which users were able to effectively, efficiently, and satisfactorily interact with the digital gaming platform to achieve their intended tasks, including ease of use, learnability, and clarity of system functions. Acceptability was defined as the

extent to which the intervention was perceived as appropriate, satisfactory, and suitable for use in routine, including users' favorite games, users' perceptions of relevance, comfort, perceived benefit, and willingness to use the system. These constructs were used to guide both data collection and analysis, ensuring that questionnaire items, focus groups, interview questions, and thematic analysis were aligned with established conceptual domains of usability and acceptability [21,22].

Setting

The study was conducted with the support of the home care service Servizio di Assistenza e Cura a Domicilio del Mendrisiotto e Basso Ceresio (ACD Mendrisiotto) in Southern Switzerland, providing nursing care to older adults in their homes. The service delivers community-based geriatric care aimed at supporting aging in place and managing complex health needs in the home setting.

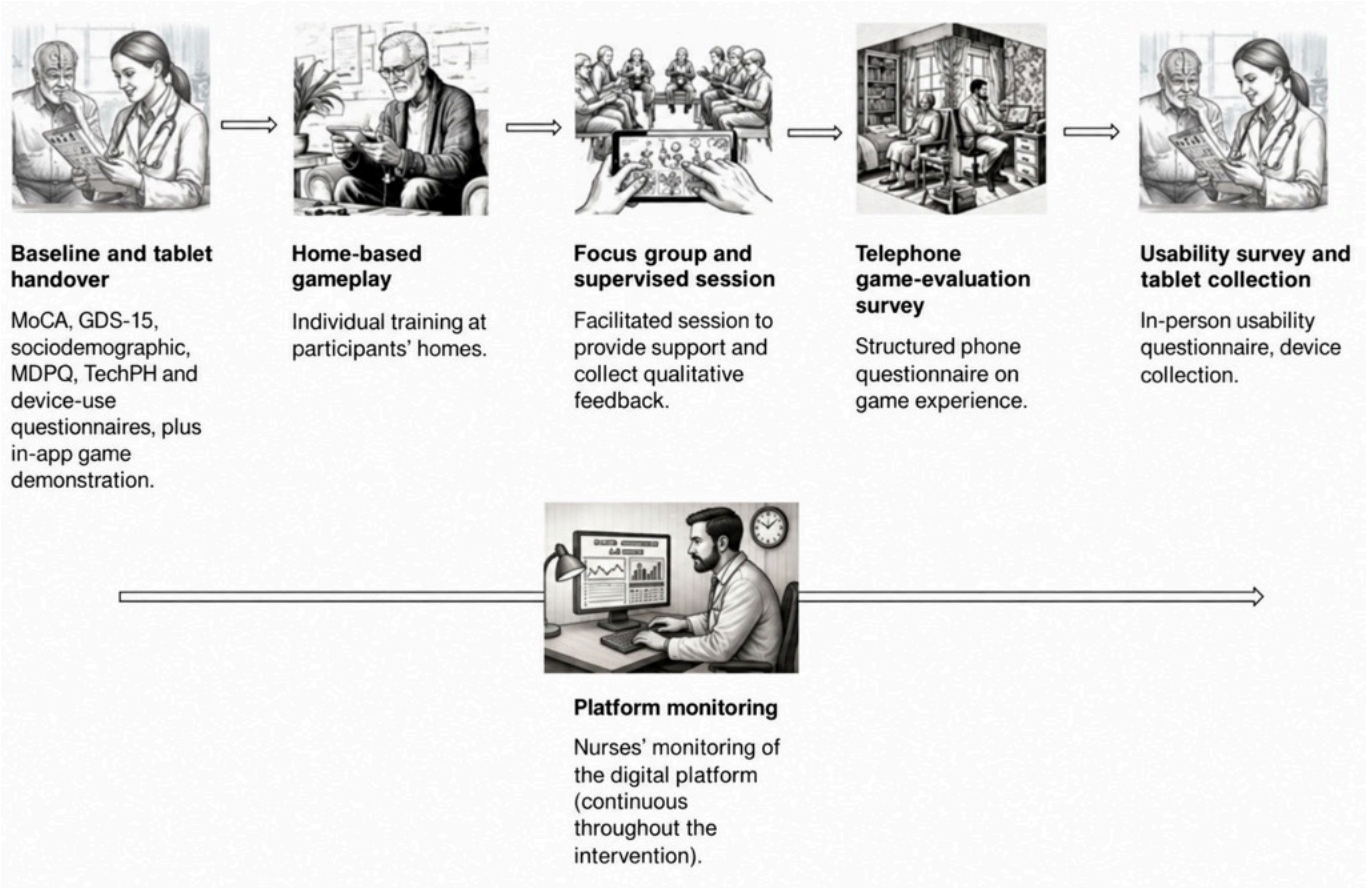
Participants

Community-dwelling older adults and geriatric nurses were invited to take part in this usability and acceptability study. Older adults were recruited from the home care service ACD Mendrisiotto in Switzerland, among their home care patients or among older people attending community social events organized by the service. A convenience sampling approach was used. Given the exploratory nature of the study, broad inclusion criteria were adopted to reflect the real-world heterogeneity of older people receiving care at home. The criteria applied were: age of 60 years and older; adequate visual and motor function to interact with a tablet; and capacity to provide informed consent. No additional exclusion criteria were applied; participants not meeting the inclusion criteria listed above were excluded from the study. As this was a usability and acceptability study, the sample size of older adults was determined pragmatically based on feasibility and the aim of obtaining in-depth feedback rather than statistical power. All geriatric nurses (ie, nurses with a specialization in geriatric care) working within the service were recruited and invited to participate, for their advanced expertise in aging-related health issues and their support to colleagues in the care of older adults with cognitive issues. All eligible nurses within the service were invited to participate, and the study size was determined by the total number of staff meeting the inclusion criteria.

Procedure and Assessment Tools

The overall study procedure is illustrated in Figure 1. The usability and acceptability study was carried out in 2 rounds, each involving a distinct cohort of participants and lasting approximately 3 months (June 2024-September 2024, October 2024-January 2025). At baseline, participants completed a sociodemographic questionnaire; a device-use frequency checklist (adapted from [23]) together with the 6-item Technophilia scale [24]; the Mobile Device Proficiency Questionnaire (MDPQ) [25]; the Montreal Cognitive Assessment (MoCA), with the standard education correction applied; and the 15-item Geriatric Depression Scale (GDS-15). All instruments were administered in Italian. At intervention onset, after the baseline assessment, participants were provided with a tablet preinstalled with the beSerious app. They received step-by-step instructions on using the application and on how to play the games. Participants were then encouraged to use it at their own pace throughout the study period, recommending a minimum usage of once a week. No changes were made to the beSerious app between the 2 rounds. Approximately one month after receiving the tablet, a focus group was conducted with participants to gather feedback on the overall gaming experience. Members of the project team were available over the phone to provide assistance with the games if needed. A study-specific game evaluation questionnaire, developed by the research team and not previously validated, was administered via telephone at the midpoint of the intervention. At the end of each round, a usability and acceptability questionnaire was administered. Specifically, an adapted version of the System Usability Scale, a widely used and well-established questionnaire-based tool in user experience research, was used [26].

Figure 1. Study procedure. GDS-15: Geriatric Depression Scale-15; MDPQ: Mobile Device Proficiency Questionnaire; MoCA: Montreal Cognitive Assessment; TechPH: Technophilia.



Data from the first round were used to finalize the development of the monitoring platform to be used by home care nurses. Therefore, in the second round of the study, the set of serious games and the monitoring platform were explored simultaneously by a group of older people and geriatric nurses.

Games

Overview

The 6 games are embedded within a progressive travel storyline in which participants "visit" 47 European countries by completing game sessions. Progression through the countries is unlocked after completing a set of games, which introduces a sense of progression and curiosity and is intended to enhance motivation and sustained engagement with the intervention. In each country, seniors could play all 6 games. Upon successfully completing at least 3 of 6 games, participants unlocked a "curiosity pill" related to the country virtually visited and could move to the next one. A demonstration video showcasing the set of games is available in Multimedia Appendix 1. The targeted cognitive functions and performance metrics of the games are described below. A synoptic summary of metrics is provided in Table 1.

Table 1. Canonical versus implemented gameplay metrics across games.

Game	Reference paradigm/test	Canonical metrics	Metrics implemented in this study
Find the word	Mastermind/Wordle	Attempts; information gain/entropy; time per attempt; lexical difficulty/frequency	Success/failure and number of attempts; hint use
Guess the sequence	Corsi Block-Tapping Test (Spatial Span)	Span length (forward/backward); total correct per length; positional/order errors	Success/failure and number of attempts (fixed 5-card sequence)
Jigsaw	Object Assembly/Hooper Visual Organization Test	Accuracy; completion time; qualitative placement/orientation errors	Completion time
Whack-a-mole	CPT/SART	Hit rate; false alarms/omissions; d′; reaction-time mean/variability	Percentage of correct hits (RT recorded but excluded from criteria)
Three towers	Tower of London	Excess moves over optimal; rule violations; initiation latency; completion time; problems solved	Successful trials; mean attempts per level; initiation latency
Labyrinth	Porteus Maze	Completion time; errors (dead ends/violations); path optimality; perseverations	Completion time; valid-step proportion; mean time per valid step

Find the Word

In this game, players are required to identify a 5-letter target word within 6 attempts. After each attempt, the interface delivers immediate, color-coded feedback that indicates the accuracy of each letter. If the player fails to identify the correct word within 6 attempts, the game records an unsuccessful attempt and displays a pop-up encouraging the player to try a different word. Selecting "Continue" returns the player to the game selection screen.

This game was inspired by deductive reasoning paradigms similar to Mastermind, in which common performance metrics include the number of attempts to solve the puzzle, information gain or efficiency across guesses, and time per attempt. In our adaptation, however, the game was implemented in a verbal, Wordle-like format (original game developed by Josh Wardle) [27,28].

We aligned with Mastermind/Wordle-like paradigms by tracking attempts to solve (successful completion and number of attempts) and the use of hints. The Italian word list was obtained from an open-source GitHub repository (available at [29]), and then manually curated to remove archaic, highly technical, or otherwise uncommon entries, thereby reducing the risk of participant frustration and ceiling/floor effects.

Guess the Sequence

In Guess the Sequence, 5 cards light up on the screen in a random order. The player is asked to reproduce the sequence by tapping the cards in the correct order, with a maximum of 7 attempts allowed. If the player uses all 7 attempts without completing the sequence correctly, the game records the attempt as unsuccessful and displays a pop-up message inviting the player to try again with a new sequence. Selecting "Continue" returns the player to the game selection screen.

This task was inspired by the Corsi Block-Tapping Test (Spatial Span), which probes visuospatial short-term memory, working memory, and selective attention. Canonical metrics in Corsi include forward/backward span (maximum sequence length reproduced), total correct sequences at each length, and order/positional errors [30].

Echoing Corsi/Spatial Span, we indexed sequence-reproduction accuracy/efficiency with successful completion of the game and number of attempts, prioritizing repeatability and brief administration over staircase designs. We excluded span estimation and fine-grained error coding, opting instead for a fixed 5-card sequence as a mid-range difficulty.

Jigsaw

This game is a digital adaptation of the classic jigsaw puzzle. An image is divided into multiple tiles (16, 36, or 64) that must be repositioned within a frame to reconstruct the original picture. Only if the player leaves the puzzle and returns to the game selection screen without completing it, a loss is recorded.

This game was conceptually aligned with the visuoconstructive and visual-organization tasks, such as Wechsler Adult Intelligence Scale (WAIS), Object Assembly, and the Hooper Visual Organization Test, which gauge visuospatial reasoning, perceptual organization, and constructive abilities. Typical metrics include accuracy, time to completion, and qualitative placement/orientation errors [31].

In line with Object-Assembly/Hooper tasks, we used completion time as a continuous indicator of visuoconstructive solution efficiency.

Whack-a-Mole

In this game, 40 moles appear briefly and at random from holes during a fixed 60-second round. Players are required to tap the moles as they appear to hit them. When the time limit expires, the player is shown the number of moles successfully hit. No unsuccessful attempts are recorded for this game.

This game was developed following the Continuous Performance Test (CPT) [32] and Sustained Attention to Response Task (SART) paradigms [33], commonly used to evaluate sustained and selective attention, vigilance, and, in certain variants, response inhibition. Canonical metrics include hit rate and false alarms or omissions, signal-detection indices, and reaction-time (RT) mean/variability [34].

Consistent with CPT/SART logic, we summarized attentional accuracy under time pressure via the percentage of correct hits. RTs were recorded (mean and variability) but excluded from the criteria to keep the performance index straightforward to normalize.

Three Towers

This game is a modified version of an already existing digital adaptation of the Tower of London (ToL) test [35]. It features 3 pegs and 3 colored disks displayed in a predefined start and goal configuration. The scope of the game is to reproduce the goal configuration by moving the disks between the pegs. Each game round consists of 3 attempts, each with a predefined maximum number of moves. An attempt is considered unsuccessful if the player fails to reproduce the target configuration within the allowed moves. If all three attempts end without completing the task successfully, the game records an unsuccessful attempt and displays a pop-up message inviting the player to try again. Selecting "Continue" returns the player to the game selection screen.

The original neuropsychological ToL test is a planning paradigm used to assess multistep problem solving, working memory for intermediate states, and rule maintenance [36]. In ToL, performance is typically quantified via optimal (minimum) moves versus excess moves, rule violations, and timing measures such as initiation latency (time before the first move) and execution time (time to completion), as well as the number of problems solved within constraints [37,38].

Borrowing from ToL conventions, we focused on planning efficiency and strategy cost using successful trials, mean attempts per level, and initiation latency. Completion time was recorded but excluded from the criteria to keep the index streamlined in a small-N pilot and to avoid redundancy with efficient components (successful trials/attempts). Given

the fixed optimal-move design, completion time largely reflects execution latency rather than planning; we therefore prioritized initiation latency as a cleaner planning proxy.

Labyrinth

In this game, players guide a small car icon onscreen through a top-down maze from the starting point to the exit using stepwise movements, with three selectable difficulty levels of increasing complexity. A loss is recorded only if the player leaves the game and returns to the game selection screen without having completed it. This task was adapted from paper-and-pencil maze tasks such as the Porteus Maze [39], which assesses visuospatial navigation, route planning, error monitoring, and flexibility during path recalculation. Typical metrics include completion time, error counts (dead ends and rule violations), path optimality (deviation from the shortest path), and perseverations.

Following maze-navigation conventions, we combined completion time (global efficiency) with the proportion of valid steps (path quality) and the mean time per valid step (local decision speed). In our mazes, path optimality (deviation from the shortest path) is not applicable, as the optimal path is fixed by design. Errors are instead summarized implicitly via the valid-step proportion and their impact on completion time.

Monitoring Platform

The web-based monitoring platform was developed using Node.js (version 20.19.5; OpenJS Foundation), enabling geriatric nurses to remotely track patients' gaming performance over time. Data collected from each game session was automatically processed within the platform. The interface displays multiple types of data (Figure 2): (1) a general dashboard of gameplay, including comparisons with the previous week or 2 weeks, (2) the playing time spent on each game, (3) the distribution of games played across different time slots in the 24 hours, (4) the number of games won and lost, and (5) the distribution of rounds played per game. By clicking on a bar in panel (5), a game-specific performance plot appears (6), allowing nurses to visually examine performance trends over time.

Gameplay data were collected during each session, securely stored on a dedicated server, processed, and subsequently made available to the nurses on the monitoring platform. Figure 3 shows the schematic diagram of the system architecture with full data-acquisition and processing flow.

Figure 2. Monitoring platform dashboard. (A) Summary statistics from the start of the study: total rounds played, total gameplay time, and week-over-week percentage changes in rounds played, gameplay time, proportion of games won, and proportion of games lost. (B) Time spent playing each game. (C) Total rounds played, broken down by 4-hour time slots across the 24-hour day. (D) Total rounds won versus lost. (E) Total rounds played per game. (F) Example of a game-specific performance plot (Jigsaw): the y-axis shows completion time (min), the x-axis shows the dates of play, and colors indicate difficulty level (light blue: easy; purple: medium; red: hard).

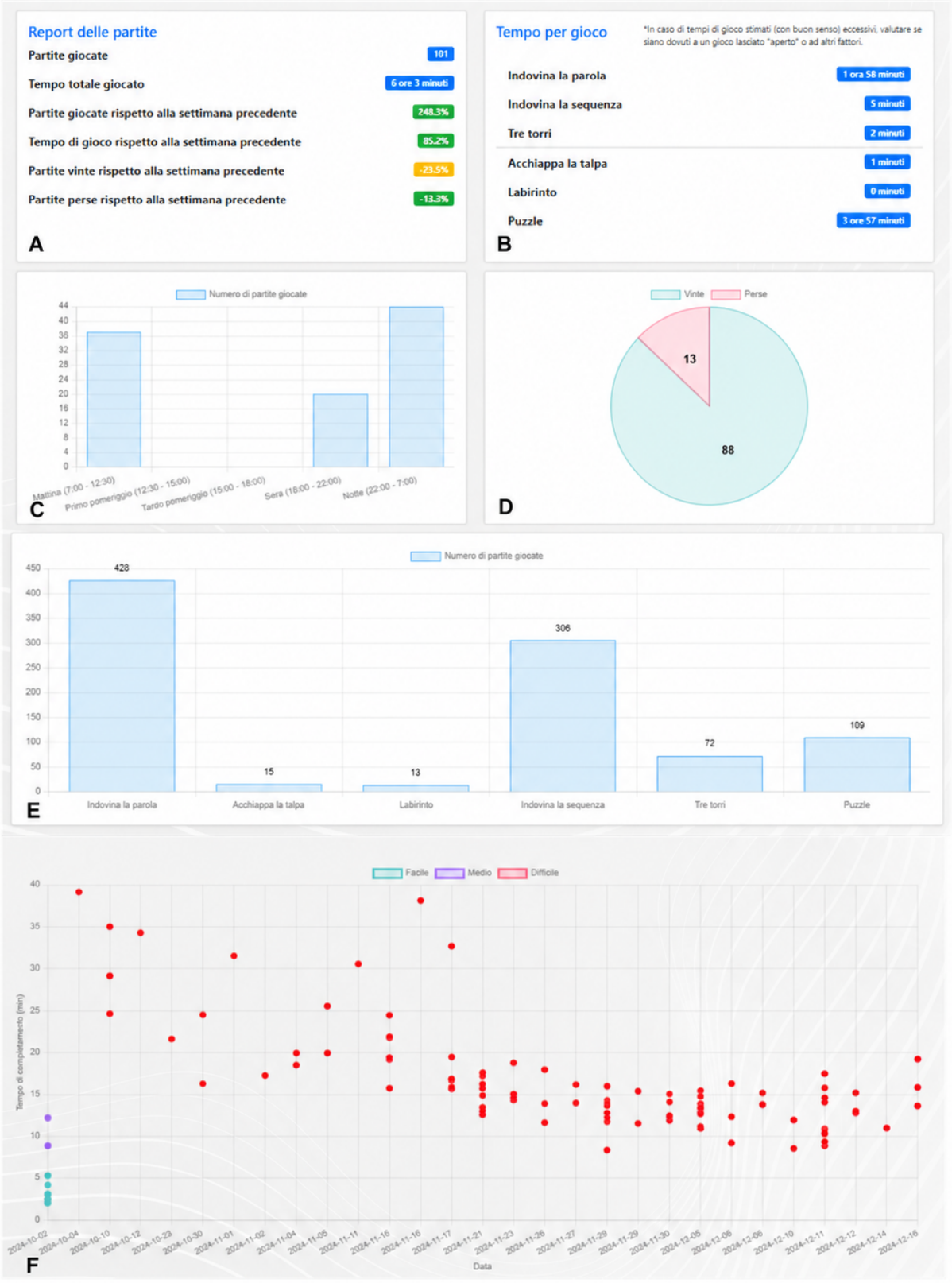
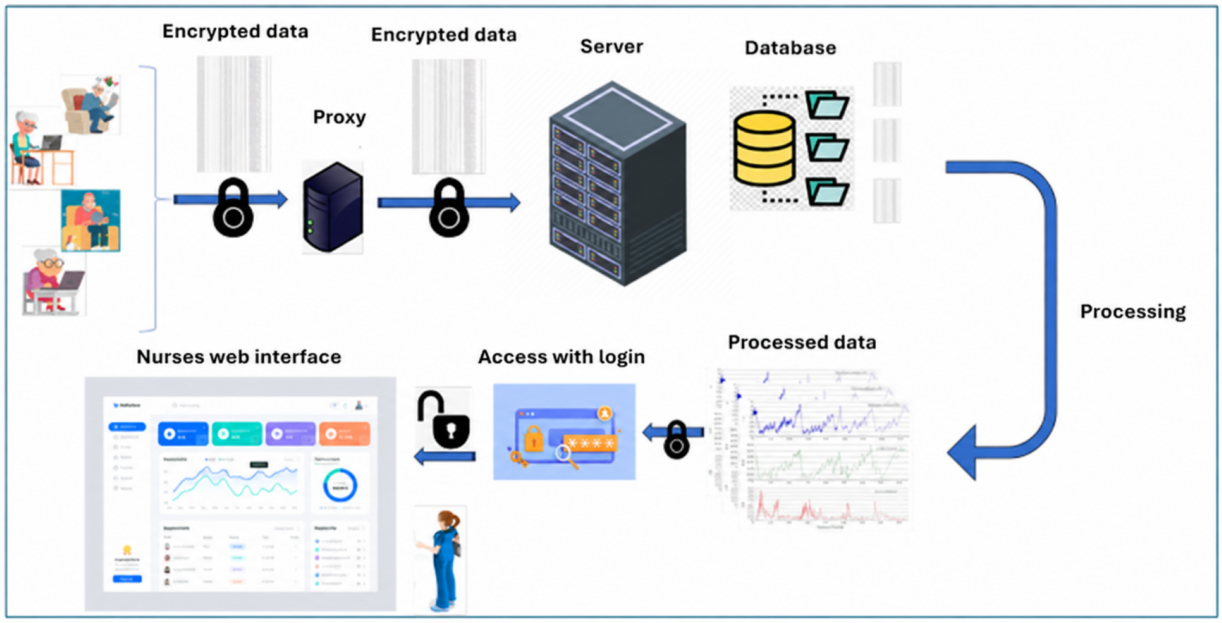


Figure 3. Data acquisition system architecture and web platform.



Data Collection and Analysis

Both qualitative and quantitative data were collected and analyzed in this study. Qualitative data were collected through semistructured interviews and focus groups, while quantitative data included system usage metrics and questionnaire responses.

Qualitative Data Collection and Analysis

During each round, qualitative data were collected through open-ended questions, interviews, and focus groups. All older adults enrolled in the study were invited to take part in the qualitative strand; participants had first-hand experience with the games through their involvement in the study and had voluntarily agreed to participate, and the sample was heterogeneous in terms of age, gender, and digital literacy, supporting the capture of diverse perspectives relevant to the research goals. The same applied to the nurses who took part in the study. Audio recordings were transcribed verbatim using the open-source Whisper automatic speech recognition system (version 1.5; OpenAI), executed locally to ensure data privacy. Files were manually reviewed for accuracy by a researcher (JD). Collected data were analyzed thematically by 2 independent researchers (JD and SL) to explore seniors' feedback on the use and perception of the games, and nurses' opinions on the platform [40]. A total of 2 focus groups with older adults were conducted, involving 7 participants in the first round and 5 participants in the second round. In addition, 2 individual interviews were carried out with older adults and 3 with nurses (interviewers: JD and SL). The individual interviews with older adults were conducted to accommodate participants who were unable to take part in the focus groups.

Interview data were analyzed using reflexive thematic analysis following Braun and Clarke [40]. An inductive, data-driven approach was adopted to allow themes to emerge from the dataset without imposing pre-existing categories. An initial coding framework was developed and iteratively refined into a structured codebook through discussion and consensus among the research team. Using an inductive approach, several subthemes were identified from the collected data and subsequently organized into overarching themes. Throughout the analytic process, the researchers engaged in regular discussions to compare interpretations and resolve discrepancies, supporting analytical rigor. Rather than calculating formal intercoder reliability, consistency was ensured through iterative consensus-building, in line with the reflexive thematic analysis approach. Reflexivity was considered throughout the analysis, with the researchers reflecting on their professional backgrounds and potential influence on interpretation during team meetings and analysis discussions [41].

Quantitative Data Collection and Analysis

Quantitative data were derived from participants' gameplay sessions (Table 1: Metrics implemented in this study) and answers to the questionnaires. Analyses were made using R software, version 4.3.3 (R Foundation for Statistical Computing).

Descriptive statistics were computed for all relevant variables, including mean (SD) and medians. Frequency distributions were analyzed to describe device usage and characterize participants' gameplay frequency across the different games.

Given the limited number of participants in the study and the nonnormal distribution of Technophilia data, attitudes toward technology differences between anxiety and enthusiasm were explored descriptively. Participants were classified into 4 profiles based on their median score on the 3-item Technology Enthusiasm and 3-item Technology Anxiety subscales, using the scale midpoint (3) as a threshold: enthusiasts (enthusiasm>3, anxiety≤3), competent (enthusiasm>3, anxiety>3), anxious (enthusiasm≤3, anxiety>3), and indifferent (enthusiasm ≤3, anxiety ≤3).

Engagement profiles were similarly explored descriptively: individual gameplay frequency data are presented. Objective gameplay analytics—including play frequency, session duration, and game preference—were used as complementary descriptive indicators of engagement, providing additional context for interpreting the usability and acceptability findings. Kruskal-Wallis tests were used to compare age and years of education across usage profiles, while Fisher exact tests were applied to examine differences in living situations and MoCA-based cognitive status.

In addition, combined performance indices were constructed to summarize gameplay performance. Performance indices were derived by aggregating the gameplay metrics for each game, described in the "Game" section and summarized in Table 1. When necessary, metrics were directionally aligned so that higher values reflected better performance. Each metric was then normalized using winsorized min-max scaling based on the game-specific global minimum and maximum values computed across all participants and sessions. Winsorized min-max normalization was preferred over z-score standardization because it preserves absolute differences across games, yields interpretable values on a common 0-1 scale, and is more appropriate for small and nonnormal distributions. Standard min-max was further refined by replacing the scale anchors (minimum and maximum) with the 5th and 95th percentiles computed within each game, with values outside this interval clamped to 0 or 1. This winsorization approach prevents single extreme values from distorting the scale and is consistent with recommendations for min-max normalization with nonnormal distributions [42]. For each game, the final performance index was computed as the unweighted mean of the normalized metrics. Differential weighting was not applied, as the heterogeneous nature of the measures did not justify assuming that any single metric represented performance more strongly than the others.

These performance indices were subsequently used to explore associations between gameplay performance and participant characteristics (eg, gender and education level) and to compare average performance across games, with the aim of identifying games that were more feasible versus those that were more challenging.

Given the minimal extent of missingness and the exploratory nature of the study, complete-case analysis was applied. Formal MCAR testing was not performed, as standard tests lack adequate statistical power with samples of this size. No imputation was conducted, as it would introduce artificial variability in an already small and heterogeneous sample.

The game evaluation questionnaire included items focusing on key domains (ie, usability and acceptability). Specifically, usability was assessed by participants' answers on whether they understood the game rules and the overall aim of the games, while acceptability was assessed on whether they perceived any learning from gameplay, their level of enjoyment, the appreciation of the visual design, and their willingness to continue playing the games. For each domain, the overall percentage of favorable responses was calculated. Favorable responses were defined as the 2 most positive response options (on a 5-item Likert scale). At the item-level, the number of participants providing a favorable response out of the total number of respondents was reported.

The final questionnaire assessed mainly the domains of usability and acceptability. Usability-related items addressed the clarity of the game instructions, how easy it was for participants to learn to use the tablet and games, how easy it was to use tablet controls such as swiping and tapping, and whether technical problems occurred during use. Acceptability-related items addressed participants' appreciation of the games' graphic design, enjoyment of the games, perceived usefulness in helping them keep their minds active, and willingness to recommend the games to other older adults. This questionnaire was analyzed using the same descriptive approach, with favorable responses defined as the 2 most positive response options for each item. In addition, the self-reported weekly playing time was analyzed descriptively as an engagement indicator, reporting the number of participants providing each response out of the total number of respondents.

Data Integration

Qualitative data, questionnaire responses, and gameplay analytics were analyzed independently. Integration occurred at the interpretation stage through triangulation, where findings from each dataset were synthesized and compared to examine convergence and divergence across results through a joint display (Multimedia Appendix 2). This process allowed quantitative measures of usability and engagement to be contextualized and enriched by qualitative insights into user experiences, thereby providing a comprehensive evaluation of the serious games and the platform developed.

Ethical Considerations

The study was conducted in accordance with the Declaration of Helsinki and received approval from the regional ethics committee (Ticino, ID Req-2024-01211) in accordance with Swiss national regulations. All participants provided informed written consent prior to participation, after receiving clear information about the study purpose, procedures, and their right to withdraw at any time without consequences. Data were handled in compliance with applicable privacy and data protection regulations, with all personal information anonymized and stored securely on password-protected systems accessible only to the research team. Participants (both older adults and nurses) did not receive compensation for their time. Separate consent was obtained for the use of any images or identifiable material, ensuring participants' explicit agreement for such data to be included in study outputs. No images of participants or staff are included in this manuscript or its supplementary materials; all figures depict the study procedure, the software interface, the system architecture, or aggregated study data.

Results

Baseline Assessment

A total of 18 older adults (10 in the first round and 8 in the second round, 12 of whom were women and 6 were men; mean age 79 years, SD 6.6, 95% CI 75.72-82.28) were included in the study. Participants' sociodemographic characteristics are summarized in Table 2. At baseline, cognitive performance (MoCA) indicated that 10 participants scored within the normal range (≥26), whereas 8 fell within the mild cognitive impairment range (18-24), with no participants reporting moderate to severe impairment. Regarding depressive symptoms (GDS-15), 16 participants scored below the clinical threshold (<5), while 2 presented scores in the moderate range (9-15).

Table 2. Sociodemographic information (n=18).

Variables	Values (n)
Sex	
Male	6
Female	12
Age (years), mean (SD)	79.2 (6.6)
Education	
Elementary school	2
Middle school	6
High school/vocational education	8
University	2
Family type	
Having both children and grandchildren	17
Having just children	1
Living situation	
Alone	8
With somebody	10
Time receiving home-care assistance	
<1 year	8
Between 1 and 3 years	2
>3 years	8
MoCA[a]	
Normal (≥26)	10
Mild impairment (18-24)	8
GDS-15[b]	
No depression (0-5)	16
Mild depression (6-8)	0
At least moderate depression (9-15)	2

[a]MoCA: Montreal Cognitive Assessment.
[b]GDS-15: Geriatric Depression Scale-15.

Of the 18 participants recruited, one withdrew before completing the study, resulting in an analytic sample of 17 participants for gameplay and most questionnaire analyses. Two participants did not play Jigsaw or Labyrinth, reducing the analytic sample to 16 for those games. The end-of-study usability and acceptability questionnaire was completed by 16 participants. Sample sizes are reported for each analysis where they differ from the full sample.

Regarding digital competence, Table 3 presents the results from the MDPQ. These results yielded an overall mean score of 3.02 (SD 1.83; median 4, IQR 1-5), with the highest values observed in the basic skills domain (mean 3.80, SD 1.68; median 5, IQR 2-5), which includes abilities such as turning the device on and off, using the touchscreen keyboard, and adjusting volume or brightness. The entertainment domain (mean 2.99, SD 1.87; median 4, IQR 1-5), covering activities such as watching videos, listening to music, and taking photos or videos, also indicated moderate proficiency, while more complex domains such as data, file storage, troubleshooting, and software management showed comparatively lower scores (Table 3).

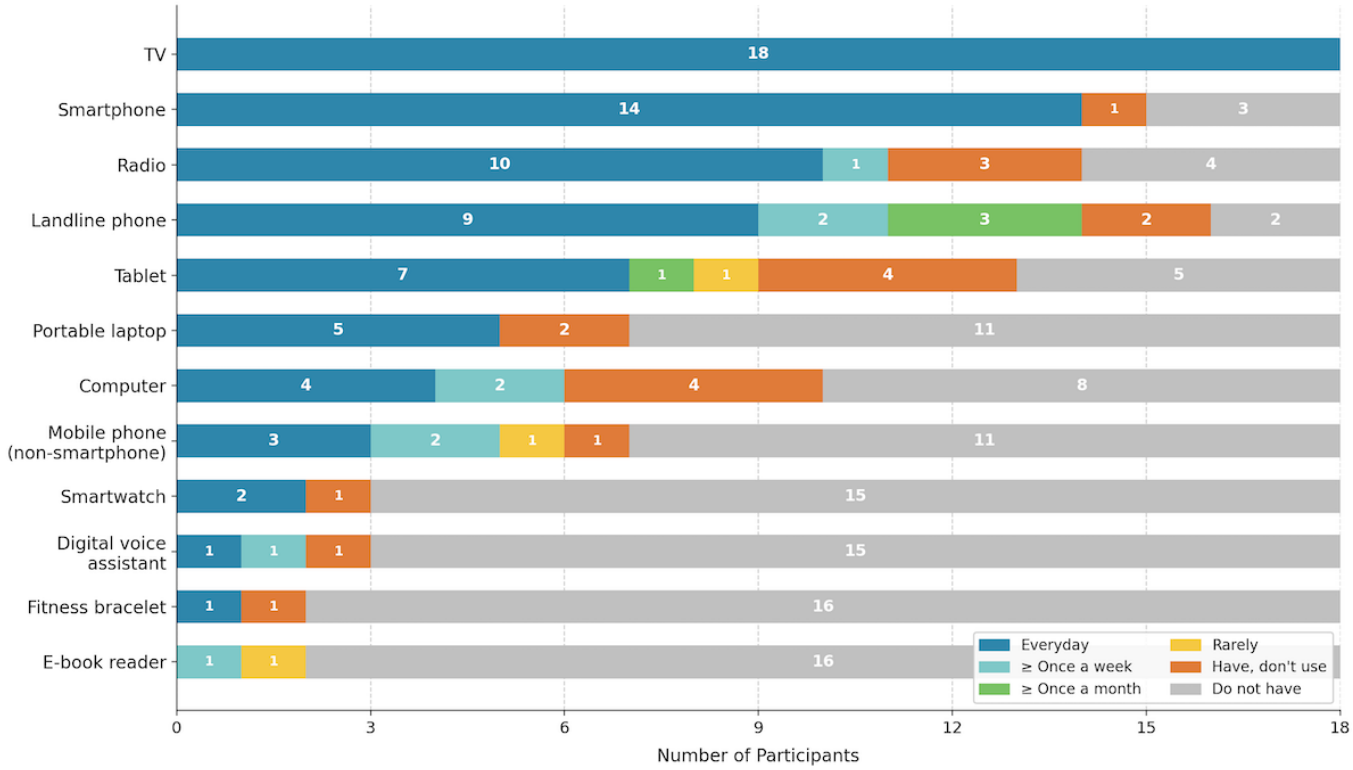
Table 3. Results of the MDPQ[a] questionnaire (n=17).

Subscale	Mean (SD)	95% CI[b]
MDPQ overall	3.02 (1.83)	(2.08-3.96)
Basic skills	3.80 (1.68)	(2.94-4.66)
Communication	2.94 (1.87)	(1.98-3.90)
Data and file storage	2.06 (1.53)	(1.27-2.85)
Internet	3.16 (1.79)	(2.24-4.08)
Calendar	2.90 (1.73)	(2.01-3.79)
Entertainment	2.99 (1.87)	(2.03-3.95)
Privacy	2.79 (1.87)	(1.83-3.75)
Troubleshooting	2.57 (1.80)	(1.64-3.50)

[a]MDPQ: Mobile Device Proficiency Questionnaire.
[b]Scores range from 1 to 5. CIs computed using t-distribution (df=16).

Figure 4 shows the frequency of device use in our sample. The first thing that emerged is that television was used daily by all participants. Smartphones (14/18), tablets (7/18), radios (10/18), and landline phones (9/18) were also part of the participants' daily routines, suggesting that both digital and traditional technologies are integrated into their everyday lives. In contrast, most participants reported not owning or rarely using devices such as smartwatches (15/18), laptops (11/18), basic mobile phones without internet access (11/18), fitness trackers (16/18), ebook readers (16/18), digital voice assistants (15/18), or desktop computers (8/18).

Figure 4. Frequency of technology device usage among study participants (n=18). Devices are ordered by the number of participants reporting everyday use. Response categories range from daily use (Everyday) to complete non-ownership (Do not have).



Prestudy profiling using the 6-item Technophilia scale revealed heterogeneous technological profiles among participants. Based on median subscale scores and using the scale midpoint as the classification threshold, 8 of 18 (44.4%) participants were classified as competent (high enthusiasm combined with high anxiety), 6 (33.3%) as anxious (low enthusiasm combined with high anxiety), and 2 (11.1%) each as enthusiasts (high enthusiasm and low anxiety) or indifferent (low enthusiasm and low anxiety).

Findings from the qualitative data, questionnaires, and gameplay analytics are presented below and summarized in a joint display table, which also provides a comparison of the different data sources (Multimedia Appendix 2).

Games Results

Qualitative Data

Qualitative findings are based on thematic analysis of 2 focus groups and semistructured interviews conducted with study participants. Four overarching themes emerged from the analysis: (1) usability, (2) acceptability, (3) cognitive demand and mental stimulation, and (4) practical considerations and suggestions. The themes are presented below and illustrated with representative participant quotations.

Usability

Overall, older adults taking part in the 2 rounds indicated that the interaction with the games appeared intuitive. The instructions were concise and clear, and participants appreciated the easy-to-follow tutorial, which facilitated rapid onboarding and smooth tablet handling. A few participants mentioned that some elements were less intuitive, such as the navigation across countries, with the pop-up button not always visible (the button was becoming clickable only after the player had successfully completed 3 of the 6 games). A few participants also described that some on-screen hints were infrequently used because they were not sufficiently visible in the interface for some users (eg, hints in the Find the word game). Minor technical issues were mentioned, including reduced brightness when the battery was low and occasional difficulties with tiles dragging in the Jigsaw game:

> *I didn't have real technical problems, sometimes the light made the screen less clear, but it never stopped me from playing.* [Senior]

It also emerged that using a tablet pen improved precision and reduced friction during interaction.

Acceptability

Participants reported high overall appreciation of the games, describing them as enjoyable and visually pleasing. Moreover, the travel narrative sustained curiosity and interest across sessions. Many participants described strong engagement, including extended play and perception that time passed quickly, supported by intrinsic motivation that combined enjoyment, a sense of competition, cognitive challenge, and reduced feelings of loneliness:

> *Because time flies when you're playing.* [Senior]

> *At least, you don't feel as lonely.* [Senior]

This favorable experience was reflected in a unanimous recommendation to use the games and in many participants' willingness to continue playing after the study:

> *You need the tablet back, right? I'll try to install it on mine if possible… are those games accessible?* [Senior]

Participants reported enjoying all of the games, although Find the word, Jigsaw, and Labyrinth emerged as the most frequently preferred ones across focus groups and interviews.

In relation to engagement, participants reported considerable variability in gameplay duration, ranging from occasional short sessions of around 30 minutes to more frequent and sustained engagement. Some participants described a particularly consistent pattern of use, with one stating:

> *I spend up to 45 minutes playing, sometimes even twice a day.*

Reported play times also ranged from morning to late evening, indicating flexible and person-dependent use.

Overall, the qualitative findings suggest a high degree of acceptability, characterized by flexible usage patterns that are shaped by users' living circumstances and daily contexts. Playing frequency appeared to vary according to individual circumstances such as social commitments, personal routines, and motivation to play. With respect to social routines, descriptions of usage and engagement with the games suggested a link between digital use and social connectedness. Participants with richer social networks tended to engage less frequently with the digital intervention, while seniors with fewer social ties played the games more regularly.

Cognitive Demand and Mental Stimulation

Overall, the cognitive challenge posed by the games was positively perceived. Participants reported that gameplay promoted concentration, strategic reflection, and perseverance rather than task abandonment, with one stating:

> *[.] if it becomes too easy, what's the point?* [Senior]

Although difficulty was usually considered appropriate, some games were perceived as more demanding, particularly Three towers and Find the word, with additional mentions of Guess the sequence and Jigsaw (difficult level). Cognitively demanding tasks were also the ones most frequently described as mentally stimulating. Several participants reported that games requiring reasoning, planning, and problem-solving fostered a sense of mental engagement, allowing some participants to perceive cognitive benefits beyond enjoyment, including a general feeling of maintaining mental activity.

In general, participants reported developing spontaneous strategies to approach the games. In the Jigsaw game, for example, some participants described starting from the border and leveraging the snap-in confirmation of correct placements to complete the image; while in the Find the word game, one of the participants kept note of every word identified in the game in a notebook, in order to expand her vocabulary. With respect to the hints provided in the games, the "first letter" hint in the Find the word game was appreciated by some, whereas others preferred not to use it to preserve the full challenge.

Practical Considerations and Suggestions

Considerations reported by participants were limited and primarily of a practical nature. One participant, in particular, noted mild neck discomfort related to sustained posture during game sessions. A mild form of technophobia was mentioned as a potential barrier towards an initial engagement with the tablet and the games; however, participants in our study generally demonstrated good digital competence and, in some cases, explicit interest in technology.

Suggested improvements included varying word length in Find the word game, favoring more familiar terms, and increasing as well as diversifying images in the Jigsaw game, which was perceived as limited and somewhat repetitive.

Quantitative Data: Gameplay Analytics

Game Patterns

To provide an initial descriptive overview, violin plots were used to visualize the distribution of total rounds played per game (Figure 5), highlighting the substantial heterogeneity in engagement across participants and games. Descriptive analysis of total rounds played (Table 4) further characterizes this engagement variability. In both Guess the sequence and Three towers, a small number of participants (n=2 and n=1, respectively) displayed markedly higher gameplay frequencies than the rest of the sample. Across most games, the distribution was right-skewed, with a subset of participants playing infrequently (below 10 rounds) and others engaging at moderate levels, as reflected by the interquartile ranges: Guess the sequence (median 45, IQR 10-77), Three towers (median 55, IQR 6-77), Jigsaw (median 13, IQR 3-62), Labyrinth (median 14, IQR 5-61), and Whack-a-mole (median 17, IQR 8-60). In contrast, Find the word showed a distinct profile with a substantially wider range of engagement (median 30, IQR 10-138, range 1-2027) and a small group of participants with markedly higher play volume than observed in any other game.

Figure 5. Distribution of total rounds played per participant for each game, displayed on a $\log_{10}$ scale. Each violin plot represents the kernel density estimate across participants. Internal box plots show the interquartile range (IQR), median (horizontal line), and Tukey whiskers (Q1 – 1.5·IQR to Q3+1.5·IQR). Individual points above the upper whisker represent outliers.

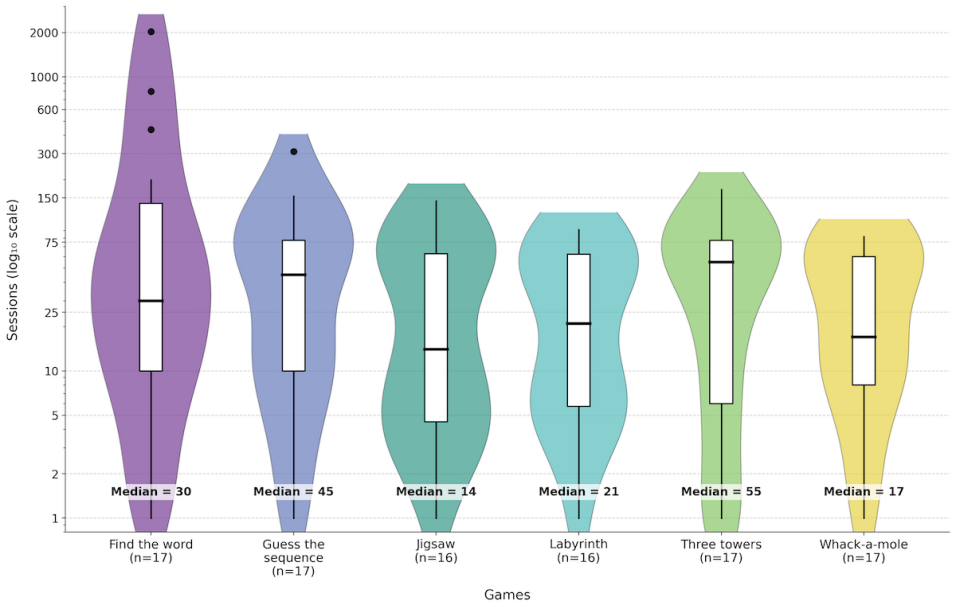


Table 4. Descriptive statistics of total rounds played per participant, by game[a].

Game	N[b]	Median (IQR)	Range (min-max)	Mean (SD)
Find the word	17	30 (10-138)	1-2027	229.2 (507.0)
Guess the sequence	17	45 (10-77)	1-310	61.8 (77.8)
Three towers	17	55 (6-77)	1-171	50.0 (46.6)
Jigsaw	16	14 (4-62)	1-143	37.9 (43.0)
Labyrinth	16	21 (6-62)	1-91	33.2 (31.4)
Whack-a-mole	17	17 (8-60)	1-82	35.9 (30.7)

[a]Median (IQR), range, and mean (SD) refer to total rounds played per participant. Two participants did not play Jigsaw or Labyrinth (N=16).
[b]Number of participants who played at least one round.

Play frequency varied markedly across participants, reflecting flexible and context-dependent usage patterns. Across all active user-week pairs, participants completed a median of 33 rounds per week (IQR 13-66), corresponding to a median of

3 active days per week (IQR 1-5). More than half of active weeks (70/133, 52.6%) included at least 3 distinct play days, and 29.3% (39/133) included 5 or more days. The median intersession gap was 1 day (IQR 1-2 days; range: 1-61 days).

Also, round duration varied considerably across games, reflecting inherent differences in game mechanics and design (Table 5). Whack-a-mole has a fixed play duration of 60 seconds per round, as rounds are time-controlled. Among open-ended games, Guess the sequence had the shortest median per-participant round duration (median 32, IQR 17-61 seconds; range: 15-122 seconds), followed by Three towers (median 106, IQR 84-121 seconds) and Labyrinth (median 110, IQR 81-157 seconds). Find the word required the longest engagement (median 215, IQR 176-302 seconds), while Jigsaw had the longest median round duration (median 370, IQR 225-835 seconds; approximately 6 minutes), consistent with the greater spatial and motor demands of puzzle completion. Substantial interindividual variability was observed for Jigsaw (range: 44-1,164 seconds) and Find the word (range: 84-552 seconds).

Table 5. Round duration statistics per game (per-participant medians).

Game	n[a]	N rounds[b]	Median (IQR), seconds	Mean (SD), seconds	Range (s)
Find the word	17	3897	215 (176-302)	236 (113)	84-552
Guess the sequence	17	1050	32 (17-61)	43 (34)	15-122
Three towers	17	850	106 (84-121)	115 (46)	57-248
Jigsaw	16	606	370 (225-835)	521 (364)	44-1164
Labyrinth	16	532	110 (81-157)	126 (67)	41-296
Whack-a-mole[c]	17	611	60 (60-60)	60 (<1)	60-61

[a]n: participants.
[b]N rounds: total rounds across all study periods.
[c]Whack-a-mole rounds have a fixed duration of 60 seconds by game design.

Gameplay was distributed across all hours of the day, from morning to late night. The 3 most active 4-hour bands were early afternoon 12 PM to 3:59 PM (1944/7546, 25.8% of all rounds), late afternoon/evening 4:00 PM to 7:59 PM (1897/7546, 25.1%), and morning 8 AM to 11:59 AM (1732/7546, 23.0%), together accounting for 73.9% (5573/7546) of all gameplay. Early morning 04:00 AM 07:59 AM represented 14.6% (1101/7546) of rounds, suggesting that a minority of participants played before conventional waking hours. Night play (12 AM to 03:59 AM and 8 PM to 11:59 PM combined) accounted for 11.6% (872/7546) of rounds. At the individual level, 6 of 17 participants (35.3%) showed a morning preference, 8 (47.1%) preferred afternoon or late afternoon, and 3 (17.6%) played predominantly in the evening or early morning, confirming the highly personalized nature of usage schedules. The peak activity hour across the full sample was 4 PM (668 rounds), followed by 12 PM (558 rounds), and 3 PM (555 rounds).

Inspection of usage patterns at the individual level revealed a generally consistent engagement across games within participants; however, a few exceptions were observed. These less homogeneous profiles may reflect individual game preferences rather than overall engagement with the games.

Engagement levels were explored descriptively in relation to the cognitive status (MoCA: normal vs impaired), age, years of education, living situation (living alone vs with somebody), digital competence (MDPQ), and attitudes toward technology (Technophilia). No clear differences emerged for any variables.

In the Guess the sequence game, the proportion of participants living alone appears to increase with higher engagement level (14% in the low-engagement, n=7; 54% in the moderate-engagement, n=7; and 100% in the high-engagement, n=2). This observation should be interpreted with caution, given the very small sample size. Similarly, in the Jigsaw game, participants in the lower-engagement group (n=9) had a lower proportion of normal MoCA scores (≥26) compared to those in the higher-engagement group (n=7): 33% versus 85%, respectively.

Overall Performance and Participant Characteristics

Analyses explored whether individual characteristics influenced performance levels. No appreciable differences were observed according to age (≤79 years, n=7; >79 years, n=10), sex (female, n=11; male, n=6), digital competence (low, n=8; high, n=8), or attitude toward technology (technophilic, n=2; technophobic, n=6; mixed, n=10). However, a statistically significant effect of education emerged: participants with a high school or vocational degree performed significantly better than those with a middle school education. This latter result should be interpreted cautiously and warrants further investigation, as the overall sample was small and participants were distributed across four education categories, resulting in very few individuals per group.

Game-Specific Performance

Performance indices were then analyzed separately for each of the six games to identify those that were more accessible to the participants versus the more demanding. All games were overall feasible, with mean performance levels above 0.5. Among them, Whack-a-mole (mean 0.83, SD 0.27, 95% CI 0.81-0.85) and Guess the sequence (mean 0.80, SD 0.28, 95% CI 0.79-0.82) showed the highest performance levels. Labyrinth (mean 0.66, SD 0.17, 95% CI 0.65-0.68), Three Towers (mean 0.65, SD 0.23, 95% CI 0.64-0.67), and Jigsaw

(mean 0.62, SD 0.30, 95% CI 0.59-0.64) showed intermediate levels. In contrast, Find the word emerged as the most demanding game, with the lowest mean performance score (mean 0.57, SD 0.21, 95% CI 0.56-0.58).

Game Evaluation Questionnaire

Usability-related responses were favorable (33/36, 91.7%), with most of the participants reporting that they understood the goals of the games (17/18) and that they understood the game rules (16/18). Acceptability-related responses were also positive (63/72, 87.5%). All participants appreciated the visual design of the games, particularly the images (18/18). Most participants indicated that the games were enjoyable (15/18), that they perceived some learning benefits from the games (16/18), and that they were willing to continue playing the games (14/18).

Usability and Acceptability Questionnaire

Questionnaire responses were grouped into usability, acceptability, and self-reported weekly playing time domains. Overall, usability ratings were highly favorable across the 4 usability items (61/64, 95.3%). All participants found the game instructions clear or very clear (16/16) and reported never or rarely encountering technical issues (16/16). Most participants also rated learning to use the tablet and games as easy or very easy (15/16) and the tablet controls as easy or very easy (14/16).

Acceptability ratings were also positive, with 59 of 64 responses across the four acceptability items classified as favorable (59/64, 92.2%). All participants appreciated the graphic design of the games (16/16) and indicated that they would recommend the games to other older adults (16/16). Most participants also reported that the games were enjoyable (13/16) and perceived them as helpful for keeping their minds active (14/16).

Answers on the self-reported weekly playing time showed that half of the participants reported playing for more than 3 hours per week (8/16), while some participants reported playing for a maximum of 30 minutes per week (5/16), and just a few reported playing between 30 minutes and 1 hour (2/16) or between 2 and 3 hours (1/16).

Qualitative Data: Monitoring Platform

Overview

Qualitative findings are based on thematic analysis of semistructured interviews conducted with geriatric nurses testing the platform. Data are merged into 4 overarching themes reflecting their perceptions and experiences with the monitoring platform: (1) clarity and ease of use, (2) clinical value of adherence, performance, and behavioral indicators, (3) organizational feasibility, and (4) improvements for clinical use.

Clarity and Ease of Use

The platform was perceived as clear, intuitive, and easy to navigate. The visual representation through graphs and tables allowed nurses to quickly grasp the clinical meaning of the information without needing additional explanations.

> *It's very clear, I think it's really well made, simple but super clear. The plots are also very easy to understand.* [Nurse]

Nurses described a smooth and barrier-free user experience, emphasizing the simplicity and accessibility of the system. Overall, the platform was evaluated as immediately understandable and efficient for routine consultation.

Clinical Value of Adherence, Performance, and Behavioral Indicators

Nurses viewed the graphical and tabular feedback as highly informative for monitoring patients' progress over time, both in terms of adherence and performance in games. They noted that the absence of activity itself could be meaningful, indicating possible motivational or health-related changes.

> *[…] He temporarily stopped playing; when I followed up with him, he reported that this was partly due to the holiday period and partly to health-related reasons.* [Nurse]

The platform also offered insights into sleep-wake cycles (eg, playing during the night). Information such as frequency, duration, and difficulty levels were perceived as measures of user engagement, allowing nurses to detect variations and, when necessary, encourage adherence. Overall, participants agreed that these metrics provided valuable behavioral data for patient follow-up and contextual understanding.

> *We also need to focus on prevention and early diagnosis of dementia. These games could be used from midlife onwards to identify early warning signs, so I'm absolutely in favor.* [Nurse]

Organizational Feasibility

While some nurses reported being able to review data within short sessions (about 10-20 minutes per week), others indicated limited access due to schedule and time constraints. Occasional peaks in workload during holidays were mentioned as a barrier to regular consultation. Nevertheless, the platform was described as easy to access, and the dashboard was easy to grasp. In light of this, nurses favored a potential integration into existing workflows.

Improvements for Clinical Use

Overall, the platform was considered comprehensive and fit for purpose, with most nurses reporting that all relevant data were already included. However, few targeted suggestions were proposed to enhance its implementation potential. These included clearer time labeling in the plots to distinguish between early-morning and late-night gaming sessions, and a longer observation window to capture longer temporal trends.

> *[…] knowing when participants engage with the games is crucial to understand their sleep-wake patterns and daily rhythm, as well as the potential effects of medications such as sleep inducers or sleep-maintenance drugs*. [Nurse]

Despite these minor considerations, the digital health system was consistently described as valuable and well-suited for monitoring patients' gaming activity.

Discussion

Principal Findings

This study provides an initial evaluation of the serious games included in the beSerious app, focusing on usability and acceptability among older adults, and exploring nurses' preliminary perceptions of the monitoring platform's usefulness.

Overall, the beSerious gaming platform was perceived as usable, acceptable, and meaningful by older adults. The convergence of qualitative data, questionnaire responses, and gameplay analytics indicated that participants generally found the tablet and games easy to use and understood both game rules and training objectives. Variability in prior subjects' technology-related attitudes did not translate into substantial usability barriers. Although some participants reported low confidence with digital tools, these concerns were not reflected in indicators of task completion or game performance. Gameplay data further indicated rapid interface familiarization and consistently high performance across sessions, suggesting that participants were able to adapt quickly to the system. Reported usability issues were limited and included occasional navigation difficulties between countries, low visibility of help functions, and reduced tablet screen brightness when battery levels were low. However, these issues did not substantially affect overall interaction with the system.

In terms of acceptability, participants generally reported positive experiences, describing the games as enjoyable, visually appealing, and cognitively stimulating. They perceived benefits in attention, memory, vocabulary, and strategic thinking, and frequently reported that the games helped them keep their minds active. Some also described broader psychosocial benefits, such as reduced sense of loneliness and a positive impact on daily life. Gameplay data supported these findings, showing sustained engagement, voluntary reengagement after inactivity, and progression to higher difficulty levels, indicating intrinsic motivation to continue playing. Nevertheless, usage data revealed variability in long-term engagement, with only a subset of participants remaining active throughout the study. This potentially suggests that overall acceptability does not necessarily translate into sustained use for all subjects. Reported reasons for reduced engagement included personal issues and competing social or family commitments. Differences in game preference and perceived difficulty also emerged as adherence-contributing factors, suggesting that future iterations could benefit from adaptive difficulty mechanisms to better balance accessibility and challenge. Minor refinements to content, such as increasing lexical diversity in Find the word and expanding image sets in Jigsaw, may also help sustain engagement. Overall, participants' willingness to recommend the games and their expressed intention to continue playing them further support acceptability.

Taken together, these findings suggest that the beSerious platform is feasible for repeated use in older adults and is perceived as both usable and meaningful. Ease of use, perceived cognitive stimulation, and positive engagement experiences are likely to facilitate technology adoption in this population [8]. These characteristics also support the feasibility of repeated gameplay, which is central to the broader aim of our study: combining cognitive stimulation with the gathering of longitudinal gameplay-derived information. Future work should refine game difficulty, identify user profiles associated with sustained engagement, and explore integration within home care services.

These findings are consistent with previous research showing that older adults are more likely to engage with serious games and digital cognitive interventions when systems are intuitive, enjoyable, and adapted to their abilities [43,44]. Systematic reviews have highlighted that simple interfaces, clear instructions, and well-designed game mechanics support usability, engagement, and sustained use, including in populations with cognitive impairment [43,44].

Beyond cognitive training, the findings also align with emerging research on serious games as potential sources of digital behavioral indicators [14,45]. In-game metrics and active digital measures have been proposed as tools for capturing longitudinal changes in cognition and behavior that may not be observable during episodic clinical assessments, although these approaches still require further validation and standardization [14,45]. While the development and validation of such indicators was beyond the scope of this study, the perceptions of both older adults and nurses suggest that gameplay-derived data could be integrated into home care services to complement traditional cognitive assessment. In particular, longitudinal gameplay information may support home care nurses by providing additional insight into changes in behavior and engagement between routine visits, potentially aiding the early identification of changes that warrant further clinical evaluation. This is consistent with broader literature indicating that digital health technologies can enhance home care by providing continuous, ecologically valid information between scheduled encounters, supporting more personalized monitoring and clinical decision-making [46].

A complex relationship between digital engagement and social context also emerged from qualitative findings. Participants reporting stronger social networks tended to engage less frequently with the games, prioritizing face-to-face interactions, whereas more socially isolated participants reported more frequent use, sometimes describing the games as a source of stimulation and companionship. Given the protective role of social relationships [47], exclusive reliance

on digital engagement may risk reinforcing social withdrawal in some individuals. This indicates a need to tailor digital interventions not only to the individual's abilities but also to their social context. Future implementations could explore complementary strategies, such as gamification, group-based formats, or guided sessions to promote shared engagement and reduce isolation [48,49].

Interestingly, games associated with lower performance were also perceived as the most mentally stimulating. This suggests that lower performance does not necessarily reflect poor usability or acceptability. Rather, this pattern aligns with Flow Theory, which proposes that engagement is optimized when task demands are appropriately balanced and matched with users' abilities [50]. Similarly, adaptive difficulty frameworks emphasize the importance of maintaining an optimal challenge level to sustain motivation and engagement [51,52]. Although adaptive difficulty was not implemented in this study, participants appeared to value cognitively demanding games as long as they remained understandable and achievable. Future iterations could therefore explore adaptive mechanisms that tailor difficulty to individual cognitive abilities, potentially enhancing engagement and long-term adherence.

Previous research has examined serious games either as tools for cognitive stimulation or as instruments for cognitive assessment, with fewer studies addressing their integration into clinical workflows [5]. More recently, serious games have been explored as potential sources of digital behavioral indicators, but these applications remain in early stages and require further validation [14]. This study extends this literature by evaluating a dual-purpose ecosystem in which the same serious games simultaneously support cognitive stimulation in older adults, while generating longitudinal gameplay-derived information for home care nurses. Rather than replacing formal neuropsychological assessment, these data are intended as complementary information to support the identification of changes in behavior and engagement between routine visits. This approach moves beyond the traditional separation between intervention and monitoring by integrating both functions within a single ad hoc-developed system and explicitly considering the perspectives of both older adults and health care professionals. Existing systems have addressed cognitive training and gameplay-derived monitoring in isolation, or have applied serious games for professional education rather than nurse-facing use [10-15]. No previously evaluated system has combined both functions, purpose-built cognitive training games, and longitudinal gameplay-derived monitoring within a single platform assessed for usability and acceptability with both older adults and home care nurses.

Strengths and Limitations

A major strength of this study is the evaluation of a serious game–based digital system within a real-world home care setting, providing ecologically valid evidence on its usability and acceptability among older adults. Another key strength lies in the convergent mixed methods design, which integrated gameplay analytics, questionnaires, and qualitative feedback from older adults and home care nurses. This approach enabled methodological triangulation and complementarity, allowing us to corroborate findings across different data sources while explaining objective gameplay behaviors through participants' experiences and perceptions. Together, these complementary data sources provided a comprehensive understanding of usability, acceptability, and engagement with the beSerious platform that would not have been achievable using either quantitative or qualitative methods alone.

Several limitations should be considered when interpreting the findings. First, the exploratory nature of the study and the relatively small, pragmatically determined sample limit the generalizability of the results and preclude statistical confirmation of the observed patterns. While this sample size is broadly consistent with established conventions for usability testing [53], it was not derived from a power calculation and was too small to reliably detect all but very large effects in the exploratory subgroup comparisons; consequently, the absence of statistically significant differences in engagement across cognitive status, age, education, living situation, digital competence, and attitudes toward technology should not be interpreted as evidence that no such differences exist, but rather reflects the limited power of these exploratory analyses. Recruitment from a single home care service may have further limited representativeness, although participants with heterogeneous demographic and cognitive characteristics were included to evaluate the accessibility and usability of the platform across a range of potential users. In addition, all participants presented normal cognition or only mild cognitive impairment, and therefore the findings cannot be generalized to older adults with more advanced cognitive decline.

Second, although the study included both subjective and objective measures, self-reported questionnaires may be affected by recall and social desirability bias. Conversely, the quantitative analyses, including descriptive comparisons across participant subgroups (eg, by cognitive status, age, or living situation), should be interpreted cautiously given the limited sample size and their exploratory purpose. These analyses were intended to identify preliminary patterns and generate hypotheses rather than provide confirmatory evidence.

Regarding the qualitative strand, transcripts and findings were not returned to participants for member checking; while consensus-based coding and researcher reflexivity were used to support analytical rigor, incorporating participant feedback on the interpretations could have further strengthened the credibility of the findings in future work. In addition, given the exploratory nature and limited sample size, thematic saturation was not achieved; findings should therefore be considered indicative rather than exhaustive.

Finally, the beSerious platform should be considered a proof of concept at this stage. Although the findings demonstrate promising usability and acceptability, they do not provide evidence of clinical effectiveness or validate gameplay-derived measures as indicators of cognitive decline.

In addition, the value of the training component itself rests on assumptions that remain actively debated in the cognitive training literature, particularly regarding the extent to which task-specific gains transfer to broader cognitive domains or everyday functioning [16-18]; this study did not test training efficacy and should not be read as evidence that gameplay produces such transfer.

Reaction time was intentionally excluded from the composite performance index because the primary aim of this study was to evaluate usability and acceptability rather than cognitive performance. Moreover, reaction time in older adults may be substantially influenced by non-cognitive factors, such as motor slowing and familiarity with touchscreen devices. Nevertheless, reaction time is continuously recorded by the platform and may be incorporated into future validation studies. Similarly, because each game's performance index reflects different cognitive demands, comparisons across games should be interpreted cautiously. In addition, improvements in gameplay performance over time may partly reflect practice and learning effects rather than genuine cognitive change. Longitudinal studies involving larger and more clinically diverse populations, together with appropriate comparison groups, will be essential to distinguish learning effects from meaningful cognitive trajectories, and to determine the potential role of gameplay-derived metrics as complementary tools for cognitive monitoring in home care.

Several additional considerations bear on the transferability of these findings. The study was conducted within a single home care service and reflects a specific model of care organized around periodic nurse home visits; the perceived value of gameplay-derived monitoring for filling gaps between visits may not extend directly to settings with different care structures, such as residential facilities with on-site staff, community-based day programs, or health systems with different visit frequencies. Similarly, nurses' preliminary perceptions of the monitoring platform were drawn from a small group within this same service, and their views on the usefulness and interpretability of gameplay data may not be representative of nursing teams with different staffing models, caseloads, or levels of digital literacy. The games themselves include language- and culture-specific content (eg, lexical tasks, country-based navigation), which may limit direct transferability to other linguistic or cultural contexts without adaptation. Findings are also specific to the tablet device used in this study; usability issues such as reduced screen brightness under low battery reflect this particular hardware and interface, and results may not generalize to other device form factors. Finally, the absence of a comparison group means these findings cannot establish how the beSerious platform's usability or acceptability compares to alternative digital cognitive tools, and the observation window, while sufficient to detect variability in engagement, was too short to support conclusions about adoption or engagement over longer time horizons. Readers applying these findings to other populations, care models, languages, devices, or timeframes should therefore treat them as preliminary and context-bound rather than broadly generalizable.

Future Directions

The project findings identify several opportunities to further refine the beSerious platform. Future iterations should improve the visibility of secondary interface elements, including country navigation controls and in-game hints, to further support independent use among older adults with limited digital experience. The monitoring dashboard could also be refined to improve the visualization and interpretation of longitudinal gameplay data for home care nurses. Beyond home care services, the flexibility of the platform offers opportunities for implementation in residential care facilities and community-based programs, where more structured, guided, or group-based gameplay could foster both cognitive engagement and social interaction [48,49].

Conclusion

This study provides a preliminary evaluation of an integrated system combining serious games for older people with a monitoring platform for home care nurses. Overall, the serious games were perceived as usable, acceptable, enjoyable, and meaningful, suggesting that they may support repeated use for cognitive stimulation in older adults. Importantly, repeated use also represents a prerequisite for generating game-derived information over time. Preliminary feedback from nurses indicated that the monitoring platform may provide useful complementary information on adherence, gameplay patterns, and possible within-person changes. Further research with larger samples is needed to refine the games, support sustained engagement, develop digital behavioral indicators, and evaluate the integration of the system into real-world home-care practice.

Acknowledgments

The authors would like to thank the Swiss home-care service Servizio di assistenza e cura a domicilio del Mendrisiotto e basso Ceresio (ACD Mendrisiotto) for their collaboration, in particular for the involvement of their nursing staff and for supporting the recruitment of participants among their service users.

Generative AI Use

The authors declare the use of generative AI (GenAI) in the research and writing process. According to the GAIDeT taxonomy (2025), the following tasks were delegated to GenAI tools under full human supervision: The GenAI tool used for translation was ChatGPT-5. Responsibility for the final manuscript lies entirely with the authors. GenAI tools are not listed as authors and do not bear responsibility for the final outcomes. In addition, we used ChatGPT-5 to generate the image of the study procedure.

Funding

This work was supported by Fondazione Daccò. The funder had no role in the study design, data collection, analysis, interpretation of data, or preparation of the manuscript.

Data Availability

The datasets generated or analyzed during this study are not publicly available due to ethical and legal restrictions related to the sensitive nature of cognitive and health-related data collected from older adults, but are available from the corresponding author on reasonable request.

Conflicts of Interest

None declared.

Multimedia Appendix 1

This video provides a brief demonstration of the 6 serious games included in the intervention. Licensed under Creative Commons Attribution 4.0 International (CC BY 4.0).
[MP4 File (MP4 video File), 31 KB-Multimedia Appendix 1]

Multimedia Appendix 2

. Joint Display.
[XLSX File (Microsoft Excel File), 31 KB-Multimedia Appendix 2]

References

1. World Population Prospects. World Health Organization; 2024. URL: https://population.un.org/wpp/publications [Accessed 2026-11-03]
2. World report on ageing and health. World Health Organization; 2015. URL: https://www.who.int/publications/i/item/9789241565042 [Accessed 2026-10-29]
3. Global action plan on the public health response to dementia 2017 - 2025. World Health Organization. 2017. URL: https://www.who.int/publications/i/item/global-action-plan-on-the-public-health-response-to-dementia-2017---2025 [Accessed 2026-10-29]
4. Bonnechère B, Klass M, Langley C, Sahakian BJ. Brain training using cognitive apps can improve cognitive performance and processing speed in older adults. Sci Rep. Jun 10, 2021;11(1):12313. [doi: 10.1038/s41598-021-91867-z] [Medline: 34112925]
5. Abd-Alrazaq A, Alajlani M, Alhuwail D, et al. The effectiveness and safety of serious games for improving cognitive abilities among elderly people with cognitive impairment: systematic review and meta-analysis. JMIR Serious Games. Mar 10, 2022;10(1):e34592. [doi: 10.2196/34592] [Medline: 35266877]
6. Canapa G, Paternò F, Santoro C. Interactive serious games for cognitive training of older adults: a systematic review. IEEE Trans Comput Soc Syst. Oct 2025;12(5):2445-2468. [doi: 10.1109/TCSS.2025.3532307]
7. Gomez-Hernandez M, Ferre X, Moral C, Villalba-Mora E. Design guidelines of mobile apps for older adults: systematic review and thematic analysis. JMIR Mhealth Uhealth. Sep 21, 2023;11:e43186. [doi: 10.2196/43186] [Medline: 37733401]
8. Zhang M. Older people's attitudes towards emerging technologies: a systematic literature review. Public Underst Sci. Nov 2023;32(8):948-968. [doi: 10.1177/09636625231171677] [Medline: 37204075]
9. Liu Y, Jun H, Becker A, Wallick C, Mattke S. Detection rates of mild cognitive impairment in primary care for the United States Medicare population. J Prev Alzheimers Dis. 2024;11(1):7-12. [doi: 10.14283/jpad.2023.131] [Medline: 38230712]
10. Schwab N, Wu CY, Galler J, et al. Feasibility of common, enjoyable game play for assessing daily cognitive functioning in older adults. Front Neurol. 2023;14:1258216. [doi: 10.3389/fneur.2023.1258216] [Medline: 37900599]
11. Anguera JA, Gunning FM, Areán PA. Improving late life depression and cognitive control through the use of therapeutic video game technology: a proof-of-concept randomized trial. Depress Anxiety. Jun 2017;34(6):508-517. [doi: 10.1002/da.22588] [Medline: 28052513]
12. Souders DJ, Boot WR, Blocker K, Vitale T, Roque NA, Charness N. Evidence for narrow transfer after short-term cognitive training in older adults. Front Aging Neurosci. 2017;9:41. [doi: 10.3389/fnagi.2017.00041] [Medline: 28293188]
13. Belchior P, Yam A, Thomas KR, et al. Computer and videogame interventions for older adults' cognitive and everyday functioning. Games Health J. Apr 2019;8(2):129-143. [doi: 10.1089/g4h.2017.0092] [Medline: 30273002]

14. Gielis K, Vanden Abeele ME, De Croon R, et al. Dissecting digital card games to yield digital biomarkers for the assessment of mild cognitive impairment: methodological approach and exploratory study. JMIR Serious Games. Nov 4, 2021;9(4):e18359. [doi: 10.2196/18359] [Medline: 34734825]
15. Busca E, Caristia S, Bidone SP, et al. Serious game for the nursing assessment of home-dwelling older adults: development and validation study. JMIR Serious Games. Nov 26, 2024;12:e52644. [doi: 10.2196/52644] [Medline: 39591607]
16. Sala G, Aksayli ND, Tatlidil KS, Tatsumi T, Gondo Y, Gobet F. Near and far transfer in cognitive training: a second-order meta-analysis. Collabra. Jan 1, 2019;5(1):18. [doi: 10.1525/collabra.203]
17. Simons DJ, Boot WR, Charness N, et al. Do "Brain-Training" programs work? Psychol Sci Public Interest. Oct 2016;17(3):103-186. [doi: 10.1177/1529100616661983] [Medline: 27697851]
18. Nguyen L, Murphy K, Andrews G. A game a day keeps cognitive decline away? A systematic review and meta-analysis of commercially-available brain training programs in healthy and cognitively impaired older adults. Neuropsychol Rev. Sep 2022;32(3):601-630. [doi: 10.1007/s11065-021-09515-2] [Medline: 34251578]
19. Levati S, Bortolamei M, Morellini L, et al. Game-based training of cognitive functions: an exploratory study involving seniors in Switzerland. ECGBL. Oct 7, 2024;18(1):541-550. [doi: 10.34190/ecgbl.18.1.2683]
20. Skivington K, Matthews L, Simpson SA, et al. A new framework for developing and evaluating complex interventions: update of Medical Research Council guidance. BMJ. Sep 30, 2021;374:n2061. [doi: 10.1136/bmj.n2061] [Medline: 34593508]
21. Sousa VEC, Dunn Lopez K. Towards usable e-health. A systematic review of usability questionnaires. Appl Clin Inform. May 10, 2017;8(2):470-490. [doi: 10.4338/ACI-2016-10-R-0170] [Medline: 28487932]
22. Sekhon M, Cartwright M, Francis JJ. Acceptability of healthcare interventions: an overview of reviews and development of a theoretical framework. BMC Health Serv Res. Jan 26, 2017;17(1):88. [doi: 10.1186/s12913-017-2031-8] [Medline: 28126032]
23. Börsch-Supan A, Brandt M, Hunkler C, et al. Data resource profile: the Survey of Health, Ageing and Retirement in Europe (SHARE). Int J Epidemiol. Aug 2013;42(4):992-1001. [doi: 10.1093/ije/dyt088] [Medline: 23778574]
24. Anderberg P, Eivazzadeh S, Berglund J. TechPH - an instrument for measuring older people's attitudes towards technology. JMIR Preprints. Preprint posted online on Mar 7, 2019. [doi: 10.2196/preprints.13951]
25. Roque NA, Boot WR. A new tool for assessing mobile device proficiency in older adults: the mobile device proficiency questionnaire. J Appl Gerontol. Feb 2018;37(2):131-156. [doi: 10.1177/0733464816642582] [Medline: 27255686]
26. Soegaard M. System usability scale for data-driven UX. Interaction Design Foundation. 2025. URL: https://ixdf.org/literature/article/system-usability-scale [Accessed 2026-07-01]
27. Best JB. Knowledge acquisition and strategic action in "Mastermind" problems. Mem Cogn. Jan 1990;18(1):54-64. [doi: 10.3758/BF03202646]
28. Knuth D. The computer as master mind. J Recreat Math. 1977. URL: https://www.semanticscholar.org/paper/The-Computer-as-Master-Mind-Knuth/b137221e486cdf6dbfd2a028ead77db740519696 [Accessed 2025-10-20]
29. Wordle. GitHub. 2026. URL: https://github.com/Hugo0/wordle [Accessed 2026-07-01]
30. Kessels RPC, van Zandvoort MJE, Postma A, Kappelle LJ, de Haan EHF. The Corsi Block-Tapping Task: standardization and normative data. Appl Neuropsychol. 2000;7(4):252-258. [doi: 10.1207/S15324826AN0704_8] [Medline: 11296689]
31. Wechsler D. Wechsler adult intelligence scale. APA PsycTests. Preprint posted online on 2012. [doi: 10.1037/t15169-000]
32. Beck LH, Bransome ED Jr, Mirsky AF, Rosvold HE, Sarason I. A continuous performance test of brain damage. J Consult Psychol. Oct 1956;20(5):343-350. [doi: 10.1037/h0043220] [Medline: 13367264]
33. Robertson IH, Manly T, Andrade J, Baddeley BT, Yiend J. "Oops!": performance correlates of everyday attentional failures in traumatic brain injured and normal subjects. Neuropsychologia. Jun 1997;35(6):747-758. [doi: 10.1016/s0028-3932(97)00015-8] [Medline: 9204482]
34. Roebuck H, Freigang C, Barry JG. Continuous performance tasks: not just about sustaining attention. J Speech Lang Hear Res. Jun 1, 2016;59(3):501-510. [doi: 10.1044/2015_JSLHR-L-15-0068] [Medline: 27124083]
35. Marras A, Babazadeh M. A video game approach to assessing executive function in schools: a digital "Tower of London" [Article in Italian]. JIMTLT. 2023;3(3). [doi: 10.32043/jimtlt.v3i3.101]
36. Shallice T. Specific impairments of planning. Philos Trans R Soc Lond B Biol Sci. Jun 25, 1982;298(1089):199-209. [doi: 10.1098/rstb.1982.0082] [Medline: 6125971]
37. Culbertson WC, Zillmer EA. The Tower of LondonDX: a standardized approach to assessing executive functioning in children. Arch Clin Neuropsychol. Apr 1998;13(3):285-301. [doi: 10.1093/arclin/13.3.285] [Medline: 14590643]

38. Luciana M, Collins PF, Olson EA, Schissel AM. Tower of London performance in healthy adolescents: the development of planning skills and associations with self-reported inattention and impulsivity. Dev Neuropsychol. 2009;34(4):461-475. [doi: 10.1080/87565640902964540] [Medline: 20183711]
39. Porteus SD. Porteus Maze Test: Fifty Years' Application. Pacific Books; 1965.
40. Braun V, Clarke V. Reflecting on reflexive thematic analysis. Qual Res Sport Exerc Health. Aug 8, 2019;11(4):589-597. [doi: 10.1080/2159676X.2019.1628806]
41. Dodgson JE. Reflexivity in qualitative research. J Hum Lact. May 2019;35(2):220-222. [doi: 10.1177/0890334419830990] [Medline: 30849272]
42. Moeller J. Why and when you should avoid using z-scores in graphs displaying profile or group differences. J Pers Oriented Res. 2025;11(2):58-78. [doi: 10.17505/jpor.2025.28091] [Medline: 40641852]
43. Gutiérrez-Pérez BM, Martín-García AV, Murciano-Hueso A, de Oliveira Cardoso AP. Use of serious games with older adults: systematic literature review. Humanit Soc Sci Commun. 2023;10(1):939. [doi: 10.1057/s41599-023-02432-0]
44. Rienzo A, Cubillos C. Playability and player experience in digital games for elderly: a systematic literature review. Sensors (Basel). Jul 16, 2020;20(14):3958. [doi: 10.3390/s20143958] [Medline: 32708679]
45. Longley V, Wilkey J, Opdebeeck C. Outcome measurement of cognitive impairment and dementia in serious digital games: a scoping review. Disabil Rehabil Assist Technol. Apr 2025;20(3):499-509. [doi: 10.1080/17483107.2024.2405894] [Medline: 39311520]
46. Piau A, Wild K, Mattek N, Kaye J. Current state of digital biomarker technologies for real-life, home-based monitoring of cognitive function for mild cognitive impairment to mild Alzheimer disease and implications for clinical care: systematic review. J Med Internet Res. Aug 30, 2019;21(8):e12785. [doi: 10.2196/12785] [Medline: 31471958]
47. Li J, Shi L, Jiang N, Wang L, Wang Y, Zheng Z. Associations between social relationships and cognitive impairment in older adults: a systematic review of risk and protective factors. Medicine (Baltimore). 2026;105(10):e47753. [doi: 10.1097/MD.0000000000047753]
48. Bodner E, Segev A, Chernitsky R, Barak Y. Playing kioku reduces loneliness in older adults: a pilot study. Games Health J. Dec 2024;13(6):428-435. [doi: 10.1089/g4h.2023.0130] [Medline: 38808474]
49. Jamei F, Saleh Moghadam AR, Mazlom SR, Namazinia M, Asghari Nekah SM, Mohajer S. Effect of group play therapy as holistic geriatric interventional modality on feelings of loneliness among older people: a randomized control trial. Geriatrics Gerontology Int. 2024;24:1067-1076. [doi: 10.1111/ggi.14972]
50. Buchanan R, Csikszentmihalyi M. Flow: the psychology of optimal experience. Design Issues. 1991;8(1):80. [doi: 10.2307/1511458]
51. Zohaib M. Dynamic difficulty adjustment (DDA) in computer games: a review. Adv Hum-Comput Interact. Nov 1, 2018;2018:1-12. [doi: 10.1155/2018/5681652]
52. Víteková L, Eichhorn C, Pirker J, Plecher DA. Dynamic difficulty adjustment in serious games: a literature review. Information. 2026;17(1):96. [doi: 10.3390/info17010096]
53. 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]

Abbreviations

ACD Mendrisiotto: Servizio di Assistenza e Cura a Domicilio del Mendrisiotto e Basso Ceresio
CPT: Continuous Performance Test
GDS-15: 15-item Geriatric Depression Scale
MDPQ: Mobile Device Proficiency Questionnaire
MIND: MassGeneral Institute for Neurodegenerative Disease
MoCA: Montreal Cognitive Assessment
RT: reaction-time
SART: sustained attention to response task
SIGMA: Serious Games per l'allenamento e il monitoraggio delle funzioni cognitive delle persone anziane
ToL: Tower of London
WAIS: Wechsler Adult Intelligence Scale

Edited by Stefano Brini; peer-reviewed by Fabrício Noveletto, Jeong-Heon Song, Mireille Gagnon-Roy, Rhys Mantell; submitted 22.Jan.2026; final revised version received 26.Sep.2026; accepted 30.Sep.2026; published 20.Nov.2026

Please cite as:
Demarchi J, Puiatti A, Bortolamei M, Metaldi M, Babazadeh M, Levati S

Game-Based Cognitive Health Training and Monitoring for Older Adults: Prospective, Single-Arm, Longitudinal, Convergent, Mixed Methods Usability and Acceptability Study
JMIR Serious Games 2026;14:e91956
URL: https://games.jmir.org/2026/1/e91956
doi: 10.2196/91956

© Jessica Demarchi, Alessandro Puiatti, Massimo Bortolamei, Matteo Metaldi, Masiar Babazadeh, Sara Levati. Originally published in JMIR Serious Games (https://games.jmir.org), 20.Nov.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.