This is a **preprint** version of the published article:

Qiu, S., Liu, S., Kaisar, E., Yuan, R., Yang, X., Zhang, S., ⋯ Hu, J. (2025). Balance Bird: an intergenerational balance exergame for motivating older adults' participation. *Behaviour & Information Technology*, 1–24. https://doi.org/10.1080/0144929X.2025.2587736

Balance Bird: An Intergenerational Balance Exergame for Motivating Older Adults' Participation

Shi Qiu[a], Susana Liu[a], Emiran Kaisar[a], Rui Yuan[a], Xiongjie Yang[a], Sunxiya Zhang[a], Ting Han[a,c,d], Jun Hu[b]

[a] Department of Design, Shanghai Jiao Tong University, Shanghai Post code 200240, China: qiushi11@sjtu.edu.cn; susanaliu@sjtu.edu.cn; imarsemiran@sjtu.edu.cn; YuanRui-design@sjtu.edu.cn; justin_yang@sjtu.edu.cn; zhangsunxiya@sjtu.edu.cn; hanting@sjtu.edu.cn.

[b] Department of Industrial Design, Eindhoven University of Technology, Eindhoven Post code 5600MB, The Netherlands: J.Hu@tue.nl.

[c] USC-SJTU Institute of Cultural and Creative Industry

[d] Innovation Center of Yangtze River Delta, Zhejiang University

Shi Qiu, and Susana Liu contributed equally to this work as co-first authors.

Corresponding Author 1

Ting Han, Department of Design, Shanghai Jiao Tong University

Address: 800 Dongchuan RD. Minhang District, Shanghai, China

USC-SJTU Institute of Cultural and Creative Industry

Innovation Center of Yangtze River Delta, Zhejiang University

E-mail address: hanting@sjtu.edu.cn

Corresponding Author 2

Shi Qiu, Department of Design, Shanghai Jiao Tong University

Address: 800 Dongchuan RD. Minhang District, Shanghai, China

E-mail address: qiushi11@sjtu.edu.cn

Acknowledgments

This work received support from the Youth Foundation of Humanities and Social Sciences, Ministry of Education of China [grant number 22YJC760073]; National Natural Science Foundation of China (Grant No. 72471140); the International Association of Cultural and Creative Industry Research Fund.

Abstract

Incorporating social interaction and cognitive stimulation into balance training can motivate older adults to engage more actively and consistently. We developed Balance Bird, an exergame designed to enhance physical, social, and cognitive activity for older adults. Thirty-six participants (18 older adults and 18 university students) were tested in pairs across three game modes featuring different types of bodily interaction and cognitive tasks: (1) the Single-Balance Ball Game (SBBG), where older players complete cognitive tasks independently; (2) the Dual-Balance Ball Game (DBBG), where older and younger players use lower limbs as input and complete cognitive tasks together; and (3) the Gesture-Balance Ball Game (GBBG), where older participants use gestures while younger ones use lower limbs to complete tasks. Our main findings revealed that all participants scored significantly higher in game competence in the GBBG compared to the SBBG. However, they experienced significantly less flow in the GBBG than in the DBBG. Female participants reported higher user experience and lower negative affect compared to males. Older participants also reported a better user experience overall, despite perceiving more challenges. In this study, we contribute empirical evidence that Balance Bird positively impacts physical, social and cognitive health for older adults.

Keywords

Intergenerational exergames, cognitive functions, bodily interplay, balance training.

1 Introduction

1.1 Challenges of Population Aging and Healthy Aging

Society is experiencing rapid population aging. According to the World Health Organization (2022), by 2030, one in six people worldwide will be over 60. The number of individuals aged 60 and above will increase from 1 billion in 2020 to 1.4 billion in 2030 and reach 2.1 billion by 2050. Aging is associated with a range of interconnected cognitive, physical, and social challenges. Memory decline is a major cognitive concern for older adults (Molden and Maxfield, 2017). Physically, aging leads to reduced mobility and balance, increasing the risk of falls and placing a substantial burden on healthcare systems, families, and society as a whole (Robertson et al., 2002).

Social engagement also tends to decrease with age, often due to reduced opportunities for interaction. One typical example is retirement: the departure from long-standing workplace routines and professional identities often requires substantial psychological adjustment. As daily work-based social structures dissolve, individuals may experience a decline in self-worth and a weakening of social connections, which frequently culminates in feelings of loss and loneliness. This decline in social connection can significantly impact psychological well-being (World Health Organization, 2023).

These cognitive, physiological, and social factors are closely interrelated and collectively influence quality of life in later years. Research shows that higher levels of social interaction are positively associated with cognitive functioning in older adults (Krueger et al., 2009). In contrast, social isolation can accelerate cognitive decline and impair the ability to perform daily activities (Ren et al., 2023).

1.2 Exergames as a Promising Intervention Strategy

Exergames, defined as interactive technologies that integrate physical exercise with video gaming, have increasingly been utilized in exergame-enhanced interventions aimed at health promotion, providing a multifaceted pathway for enhancing physical and cognitive performance (Manser et al., 2025). The term “exergame” was first introduced in 2008 as an

interactive video game combining physical movement, entertainment, real-time feedback, and social interaction (Smith et al., 2011). Since then, its definition has evolved. Pirovano et al. (2016) further refined the concept by emphasizing its gamified structure and distinguishing between "simple" and "therapeutic" types. Šlosar et al. (2022) identified three key elements: exergames must be based on digital devices, involve human-computer interaction, and require energy expenditure exceeding 1.5 metabolic equivalents. In recent work, Manser et al. (2025) proposed a more comprehensive definition, describing exergames as "a digital health technology that requires muscle-induced bodily movements to control the performance of, and interact with, gamified exercises or tasks that are presented via an immersive or non-immersive, two- or three-dimensional traditional, virtual, augmented, or mixed reality graphical user interface." Exergames can be categorized into two broad types: "recreational exergames", which are primarily entertainment-focused, and "serious exergames", which are goal-oriented and designed to achieve specific outcomes.

Particularly in serious exergames, exergame-enhanced interventions typically integrate elements such as strength training, balance exercises, aerobic activity, and flexibility training into highly interactive and engaging task structures (Hai et al., 2022; Oh and Yang, 2010). This level of interactivity has been shown to enhance user motivation and long-term adherence, often outperforming traditional physical or cognitive training programs in improving overall outcomes (Fujiyama et al., 2022). These advantages have drawn increasing attention to their use in aging populations. Studies have shown that exergames are both safe and beneficial for older adults. Reported benefits include enhanced motivation, improved balance and cognitive function, reduced symptoms of depression, and broader contributions to the promotion of healthy aging (Chao et al., 2015; Drazich et al., 2020; Ismail et al., 2022; Kappen et al., 2019; Rytterström et al., 2024).

1.3 Integrating Social and Cognitive Dimensions in Intergenerational Exergames

Emerging research highlights the value of exergames designed to promote social interaction, especially those facilitating intergenerational play. Recent empirical and systematic reviews reveal that when older adults engage in exergames with younger participants, they experience increased enjoyment, improved attitudes toward younger co-

players, and deeper social connectedness (Qiu et al., 2025; Qiu, Kaisar, et al., 2023) . Previous research on socially assistive systems for older adults and individuals with disabilities (Qiu, 2019; Qiu, An, et al., 2020, 2023; Qiu et al., 2022; Qiu, Hu, et al., 2020a, 2023; Qiu, Kaisar, et al., 2023) provides a foundation for exploring how social features in exergames can enhance intergenerational engagement.

Moreover, cognitive training is widely recognized as a key component of exergame-based interventions for older adults due to its proven ability to delay or reverse cognitive decline (Lövdén et al., 2010). Yet, most intergenerational exergames do not embed cognitive tasks within gameplay. Prior studies have demonstrated cognitive benefits through separate interventions, such as VR-based exergames improving executive function Huang (2020), cognitive motor games enhancing processing speed and working memory Moret et al. (2022), and adapted Ping Pong games supporting cognitive inhibition (Zhang et al., 2021). However, cognitive effects are often evaluated after gameplay rather than achieved through integrated mechanics (Santen et al., 2020).

These studies highlight the promising potential of integrating cognitive training tasks into intergenerational exergames. Exergames that combine physical activity, cognitive challenges, and social interaction represent a promising avenue for promoting healthy aging and should be further investigated.

1.4 Identified Research Gaps

Based on the existing literature on exergames and their social features, several important research gaps can be identified. First, there is limited understanding of how exergames affect the overall user experience of older adults, particularly in intergenerational contexts where social, cognitive, and physical dimensions are closely interconnected. Second, the impact of specific design elements, such as different modes of bodily interaction and the structure of cognitive tasks, on user experience across generations has not been thoroughly explored. Third, findings on gender differences in exergame engagement remain inconclusive, particularly regarding preferences for gameplay styles and collaborative experiences. These gaps highlight the need for further empirical research to support the development of inclusive, engaging, and socially meaningful intergenerational exergames.

1.5 Research Objectives and Hypotheses

To address existing research gaps, we developed *Balance Bird*, a social exergame designed to promote intergenerational interaction and encourage physical activity among older adults by integrating cognitive stimulation and gamification. The game enables older and younger players to control the same character together and offers three modes: a single-player mode (SBBG), a dual-player collaborative mode using lower-limb movements (DBBG), and a collaborative mode using gesture-based input (GBBG). These modes vary in bodily engagement and task collaboration, contrasting independent versus joint actions and lower limb versus gesture inputs. A within-subjects experiment with 18 intergenerational pairs was conducted to examine the effects of interaction mode, gender, and age on user experience and collaboration. We adhere to the Mixed Methods Reporting in Rehabilitation & Health Sciences (MMR-RHS) checklist (Tovin and Wormley, 2023) to ensure the quality and rigor of our reporting.

The following research questions are proposed:

- **RQ1:** How do older and younger players experience Balance Bird across the three different game modes?
- **RQ2:** Are there any differences in user experience based on gender?
- **RQ3:** What differences in user experience exist between older and younger players?
- **RQ4:** What design insights can be gained to inform the development of future intergenerational exergames?

The contributions of this work to the field of Human-Computer Interaction (HCI), which is generally defined as the study of how computer technology impacts human work and activities (Dix, 2009), are outlined as follows:

- We designed and implemented an exergame that combines social, cognitive, and physical exercise tasks to enhance the balance training experience for older adults.
- Empirical evidence is provided that participants have a more positive user experience, especially feeling a greater sense of competence, in the game mode with interdependent bodily interactions and collaborative cognitive tasks compared to the

independent single-player mode lacking these interactions and cognitive collaborations.

- Empirical evidence also shows that, in intergenerational exergames, females and older participants perceive a better user experience than males and younger participants.

2 Balance Bird Exergame

The main goal of the Balance Bird exergame is to offer a fun and engaging way to encourage intergenerational communication through balance training. Our game not only incorporates dual tasking (combining cognitive and physical activities simultaneously)(Zhang et al., 2021), but also enhances social interaction. Thus, the Balance Bird exergame integrates physical, cognitive, and social activities within a triple-task framework.

2.1 Tangible Interaction

The hardware of Balance Bird was developed and refined based on earlier versions of the balance ball system (Kaisar et al., 2021, 2022; Qiu, Kaisar, et al., 2023). Briefly, when a player stands on the balance ball, she can change the tilt direction and angle of the acrylic board to interact with Balance Bird. In this study, we added the Gesture-Balance Ball Game (GBBG) mode. In this mode, the older player wears a wristband equipped with an accelerometer, allowing the Balance Bird to fly vertically through up-and-down arm movements, while the younger player uses the balance ball to collaboratively control the bird's horizontal flight. The Balance Bird system consists of a round acrylic board (5mm thick and 460mm in diameter) for the user to stand on, four balance balls (50mm high and 80mm in diameter) affixed to the back of the acrylic board, two accelerometers (one attached to the back of the acrylic board and the other to the wristband), a portable battery, and a 65-inch 4K LCD display. We used a multi-functional integrated accelerometer (WT901BLE5.0C) to enhance interaction accuracy, replacing the previous version that included an Arduino Uno, Bluetooth HC05, and a gyroscope. Figure 1 shows the system overview in the DBBG mode. In prior work (Kaisar et al., 2022), the Barracuda machine learning engine was utilized for webcam-based motion capture. Unity connects with the

webcam and converts the video captured into a texture. This texture serves as input for a module created by Barracuda, where the inference engine processes the ONNX model to predict the pose of the texture and generates output prediction data. An algorithm was developed to translate this output data into the movement of the Balance Bird. The workflow is shown in Figure 2.

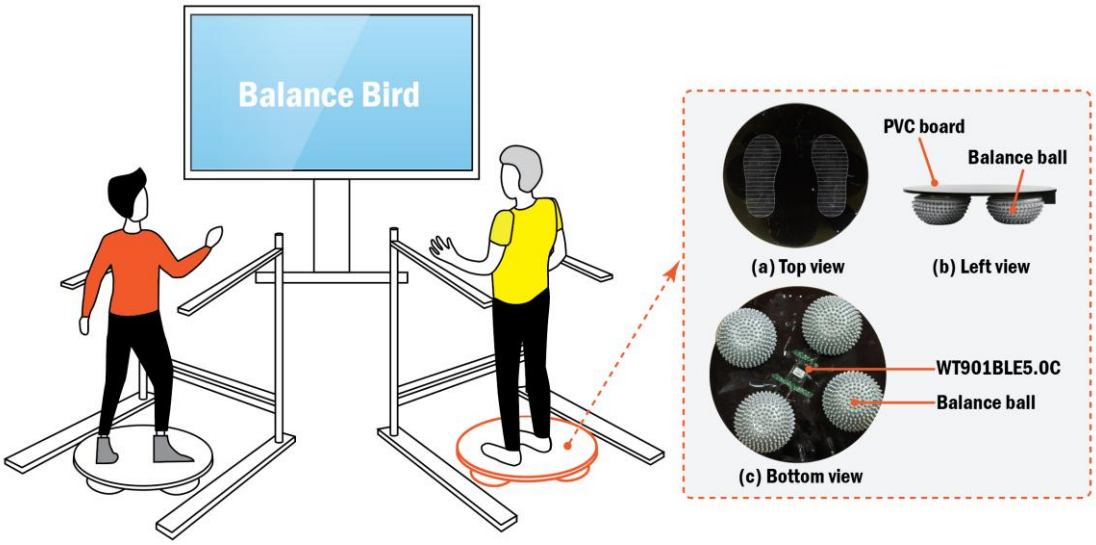

Figure 1 The system overview of the Balance Bird in the DBBG mode.

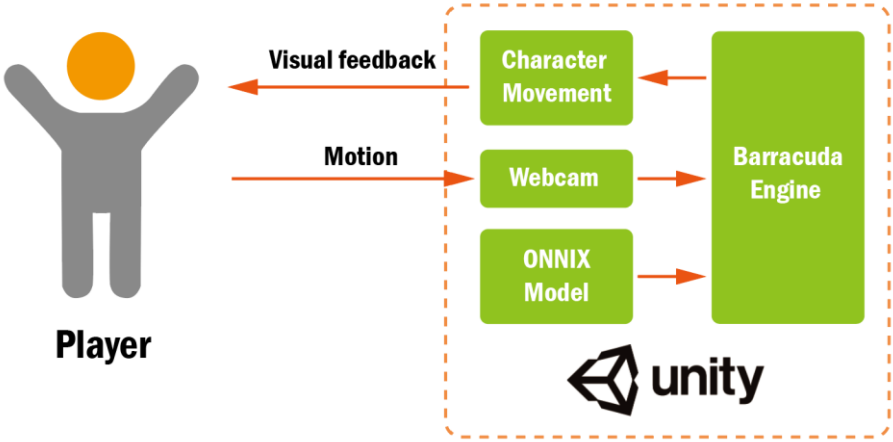

Figure 2 The workflow of motion capture using Barracuda, Unity, and webcam, with the figure redrawn based on the work of Kaisar et al. (2022b).

2.2 Game Modes: Bodily Interactions

Balance Bird has three game modes, using bodily interactions as inputs: SBBG, DBBG and GBBG. The SBBG mode is an independent play mode without bodily interplay (Mueller

et al., 2017). In this mode, a younger or older player independently uses their lower limb to control the direction and angle of the balance ball in four directions (back, forward, left, and right), thereby managing the horizontal and vertical movement of the bird in the game to hit obstacles. Both DBBG and GBBG modes are collaborative, with younger and older players working in pairs to control the same bird. According to Mueller et al. (2017), both DBBG and GBBG modes fall into the category of "interdependent exertion" within the bodily interplay dimensions. This means that players can actively use their bodies in interdependent interactions, helping each other and influencing each other's progress. In the DBBG mode, the younger player uses their lower limbs to swing the balance ball left and right, which controls the bird's horizontal movement. Meanwhile, the older player uses their lower limbs to swing the balance ball back and forth, managing the bird's vertical movement (Figure 3). In the GBBG mode, the younger player continues with the same action, while the older player uses their dominant arm, wearing a wristband to control the bird's vertical movement.

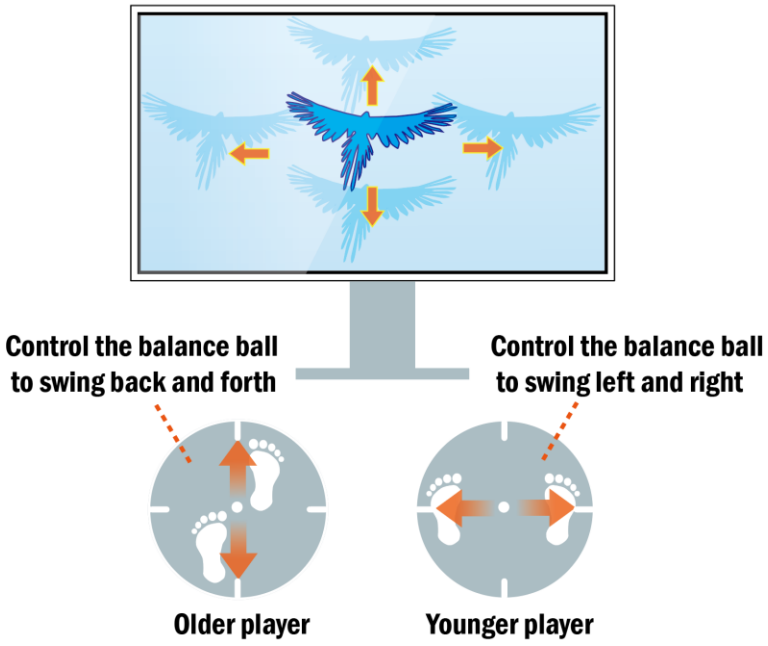

Figure 3 DBBG Mode: older player controls vertical movement of the Balance Bird, meanwhile younger player controls horizontal movement.

2.3 Game Modes: Cognitive Tasks

In the SBBG mode, the older player completes the cognitive tasks independently. In the DBBG and GBBG modes, an older player and a younger player form a team to

collaboratively complete the cognitive tasks during the game. Together, they control the bird to fly and hit the correct answer to the arithmetic problem displayed on the golden egg. While flying, the bird also needs to hit obstacles such as bubbles, stones, and snowflakes. The game lasts six minutes and is divided into three levels, each with a different setting: green forest, fiery desert, and polar ice, with each level lasting two minutes. During flying, arithmetic golden eggs appear with a 50% probability every 4-5 seconds, and obstacles appear with the same frequency and probability. The difficulty of arithmetic problems in the three game levels increases from easy to difficult. The first level focuses on single-digit addition, the second level focuses on two-digit addition and subtraction, and the third level combines two-digit addition and subtraction with single-digit multiplication (Figure 4). The tasks in the three levels increase in complexity from simple to more complex. From the first to the second level, the numerical range increases (from single-digit to two-digit numbers), while the operation complexity also increases (from addition to combined addition and subtraction). From the second to the third level, the operation complexity further increases (from two-digit addition and subtraction to combined single-digit multiplication). The design of the cognitive tasks increases the demand for information processing and cognitive switching, ensuring an increase in cognitive load across levels. In the subsequent experiment, all participants, regardless of age (older or younger), were required to complete the three game levels in a progressively increasing order of difficulty, from level 1 to level 3.

During prototype development, informal pre-tests with peers were conducted to adjust the bird's flight speed and task duration, ensuring participants could complete calculations within gameplay. The current version of the "Balance Bird" task does not include personalized difficulty adaptation; all users receive the same task configuration. However, individual differences in baseline cognition and arithmetic ability may affect performance and outcomes. Future versions will incorporate adaptive mechanisms, adjusting task difficulty (e.g., number range, operation complexity) based on real-time assessments of accuracy and response time in earlier levels.

After the game is over, a results page will appear, showing the team or individual performance, including training time, total score (five points for each correct answer to an

arithmetic question and three points for hitting an obstacle) and the percentage of arithmetic questions answered correctly (Figure 5).

Figure 4 The Polar Ice game scene in which the bird flies over the Number Maze and hits the golden egg with the answer.

Figure 5 The game results page, displaying (1) training duration, (2) total score, and (3) accuracy on numerical problems.

3 Experiment

3.1 Experimental Design

The purpose of the user experiment is to evaluate how the Balance Bird system can impact the user experience for older and younger players. Based on the four research questions outlined in Section 1, we identify three independent variables in our experimental design: game modes, gender, and generation. The fourth research question, which concerns design implications, will be addressed using qualitative data collected from post-test interviews.

The *first independent variable* is the game mode, which is treated as a within-subjects factor with three conditions: Single-Balance Ball Game (SBBG), Dual-Balance Ball Game (DBBG), and Gesture-Balance Ball Game (GBBG).

These three modes are designed based on varying levels and forms of bodily interaction, as well as the type of cognitive tasks (i.e., whether they are independent or collaborative). Whole-body interactions can enhance the gaming experience and contribute to positive outcomes (Mueller et al., 2017). SBBG does not contain interdependent interaction, serving as a baseline or control condition. DBBG and GBBG are collaborative modes, which players engage in interdependent interactions. Given that older adults have reduced lower limb flexibility due to aging, some may prefer to use upper limb actions as input in the game. In DBBG, both older and younger players use their lower limbs as input, while in GBBG, older players use upper limb gestures as input.

The *second independent variable* is gender. It is a between-subjects factor with two conditions: (1) male and (2) female.

The *third independent variable* is generation. It is a between-subjects factor, with two conditions: (1) younger players and (2) older players.

In this study, we aim to explore not only the overall differences in user experience among participants across the three game modes (the main effect), but also the interaction effects between game mode and gender, as well as between game mode and generation. Specifically, we seek to investigate whether there are significant differences in user experience between

female and male groups, as well as between older and younger groups, across the three game modes.

3.2 Participants

We used GPower software (version 3.1.9.7) (Erdfelder et al., 1996) to calculate the sample size. We set the effect size to 0.5 (a medium value according to Cohen's criteria). The same effect size of 0.5 was used in similar social exergaming experiments (Qiu, Kaisar, et al., 2023), and an effect size of 0.4 was used in behavioral research on older adults and social robots (Feng et al., 2022). This is acceptable in studies on social and cognitive behaviors. The significance level was set at 0.05, which is standard practice in behavioral and social sciences. Additionally, we chose a power of 0.80, which is a widely accepted standard in behavioral and social science research according to Cohen (2013), aiming to reduce the risk of Type II errors while providing sufficient statistical power. The final sample size for the between-group interaction comparison in this study was 31. We mainly referred to similar studies in the field of social exergaming (18 pairs, N = 36) (e.g., (Qiu, Kaisar, et al., 2023)) and dyadic interaction studies (20 pairs, N = 40) (e.g., (Qiu, Hu, et al., 2020b, 2020a)) to guide the sample size calculation. These studies typically involved medium effect sizes (around 0.5) for testing group differences and interaction effects. By adopting these assumptions, we believe that the sample size (N = 36) in this study is sufficient to detect meaningful differences across groups and provide reliable results for the objectives of the experiment.

Finally, we recruited 18 young-old pairs, totaling 36 participants. The basic information of all participants is shown in Table 1. Before conducting the experiment, we held a user experiment briefing at Yinchun Community in Minhang District, Shanghai, to introduce the Balance Bird system and outline the experimental procedures, including product features, training processes, and compensation. During this session, we confirmed the willingness of community elders to participate in the experiment and registered their availability upon expressing consent. Older and younger participants did not know each other before the experiment. In each test, an older and a younger participant of the same gender were paired together to avoid any heterosexual effects that might influence their social interactions during

the game. Each participant received 100 CNY for taking part in the one-and-a-half-hour experiment. The inclusion and exclusion criteria for the participants were as follows:

Inclusion Criteria: Older participants aged approximately 55 to 75 years were included. While existing literature generally suggests that individuals over the age of 60 tend to experience declines in cognitive function, social engagement, and physical capacity, such declines do not commence precisely at age 60. In our study, the older adults we recruited typically began retirement at age 55. The official retirement age for women in China is 55. Exiting the workforce often brings about substantial changes in social networks and identity, which can increase the risk of social isolation. Therefore, using 55 as the lower age boundary for participant recruitment is, in practice, more appropriate for the objectives of our research than the commonly referenced threshold of 60.

Moreover, social exergames require a relatively high level of motor ability to facilitate effective collaboration during gameplay. As such, our participants primarily fall within the "young-old" category (<75) (Xu et al., 2016). To ensure they met the physical and coordination demands of the experiment, we conducted pre-screening interviews and observational assessments. This approach allowed us to confirm that, despite minor differences in age, all participants met the necessary criteria for motor function and coordination. Therefore, the age of 75 was not a strict cutoff; in fact, we are interested in exploring how this social exergame impacts older-old adults (aged 75 and above) (Xu et al., 2016). However, at this stage, our primary focus is still on the experience of younger-old adults.

Older adults were recruited with a targeted gender ratio of 1:1. Individuals with medically diagnosed cognitive impairments, motor disabilities, or visual impairments were excluded from participation. They were required to have at least a junior high school education, be literate, and capable of responding to written questionnaires. They should not have a strong accent and be able to communicate fluently in spoken Chinese. The rationale behind these recruitment criteria was to ensure that older participants could accurately read, comprehend, and respond to the questionnaires, thereby maintaining the validity of the data collected. However, to achieve an adequate sample size for the experiment, the eligibility criteria were

applied with some flexibility in practice. For instance, two participants with only elementary school education (Table 1) were included in the study. Both demonstrated sufficient literacy to understand and complete the questionnaires and were able to successfully participate in the exergame tasks. Based on their demonstrated capabilities, researchers decided to accept these individuals into the final sample. Preference was given to those who regularly do morning exercises, square dancing, and those who speak Chinese Mandarin. Younger participants aged 18 to 30 were eligible. We recruited younger participants by posting information in the WeChat groups of [University name anonymized for review] dormitories, which included students from various majors. The notice stated that the test would last about 90 minutes and involve playing the exergame with an older adult. Participants were expected to demonstrate patience and empathy; however, no quantitative assessment was made, relying on the students' self-assessment.

Exclusion Criteria: Considering that the social exergame experiment required older participants to control a balance ball with their lower limbs, ensure safety, and collaborate effectively with younger participants during gameplay, certain requirements were placed on their physical, cognitive, and social abilities. Thus, older adults with conditions affecting lower limb mobility (e.g., hip fractures, hemiplegia, lower limb disabilities), those taking medications that could impair balance (e.g., psychiatric drugs), those with cognitive impairments due to neurological diseases or mental disorders, or those with language communication or cognitive function impairments were excluded.

Table 1 Demographic information of the participants.

Generation		**Older**	**Younger**
Age	Minimum	57	20
	Maximum	78	26
	Mean (Std. Deviation)	66.89(5.17)	23.78(1.63)
Gender	Male	9	9
	Female	9	9
	Total	18	18
Education	High school	7	0
	Junior high school	5	0
	Elementary school	2	0

	Master student	0	14
	Correspondence junior college	1	0
	Junior college	1	0
	Secondary normal school	1	0
	Bachelor student	1	2
	PhD student	0	2
Weekly exercise frequency (%)	Many times	11.11%	0
	7	16.67%	0
	6	11.11%	0
	5	16.67%	5.56%
	4	0	11.11%
	3	27.78%	5.56%
	2	5.56%	44.44%
	1	5.56%	16.67%
	0	0	5.56%
	Not mentioned	5.56%	11.11%

3.3 Setup

User experiments were conducted in Room 4-208 of the East Campus at [University name anonymized for review]. The experimental setup primarily consisted of an LCD display (65-inch screen size, 4K resolution), a laptop, hardware devices (including the balance ball system and the wristband device), two cameras, and two sets of desks and chairs for filling out the questionnaire. The seating was spaced apart to prevent mutual interference. During all games, an experimenter stood near the older player's position to ensure his/her safety. The Balance Bird software ran on the MAC system of a laptop, with the user interface projected onto a large LCD display. Figures 6 and Figure 7 illustrate the settings for the three game modes.

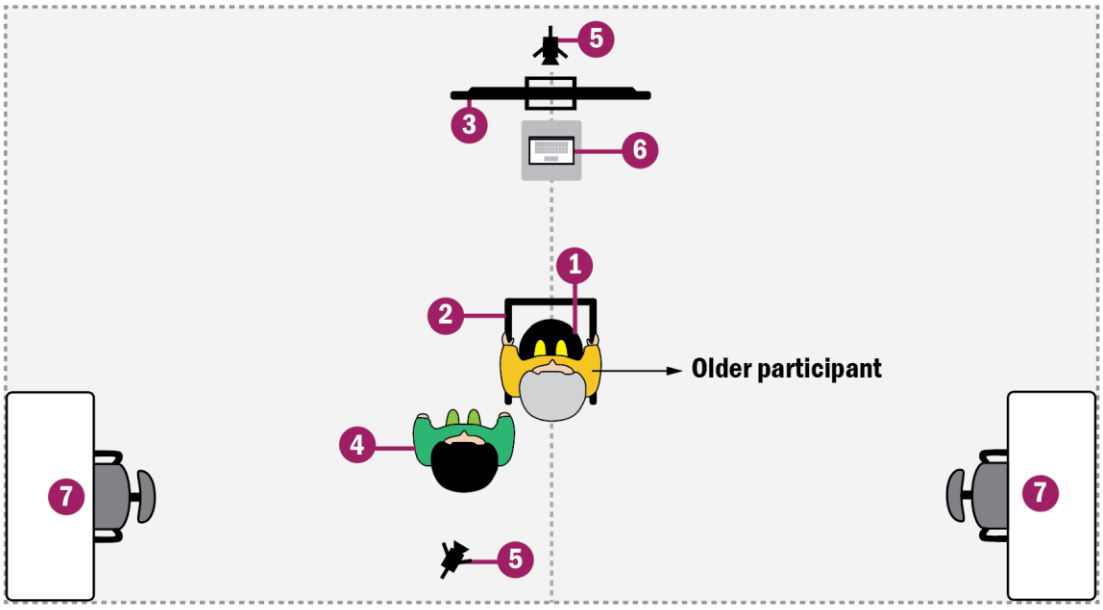

Figure 6 Experimental setup for the SBBG mode: (1) balance ball; (2) bracket with two armrests; (3) LCD display; (4) an experimenter standing next to the older participant to ensure the safety; (5) camera; (6) laptop; (7) desk and chair for completing the questionnaires.

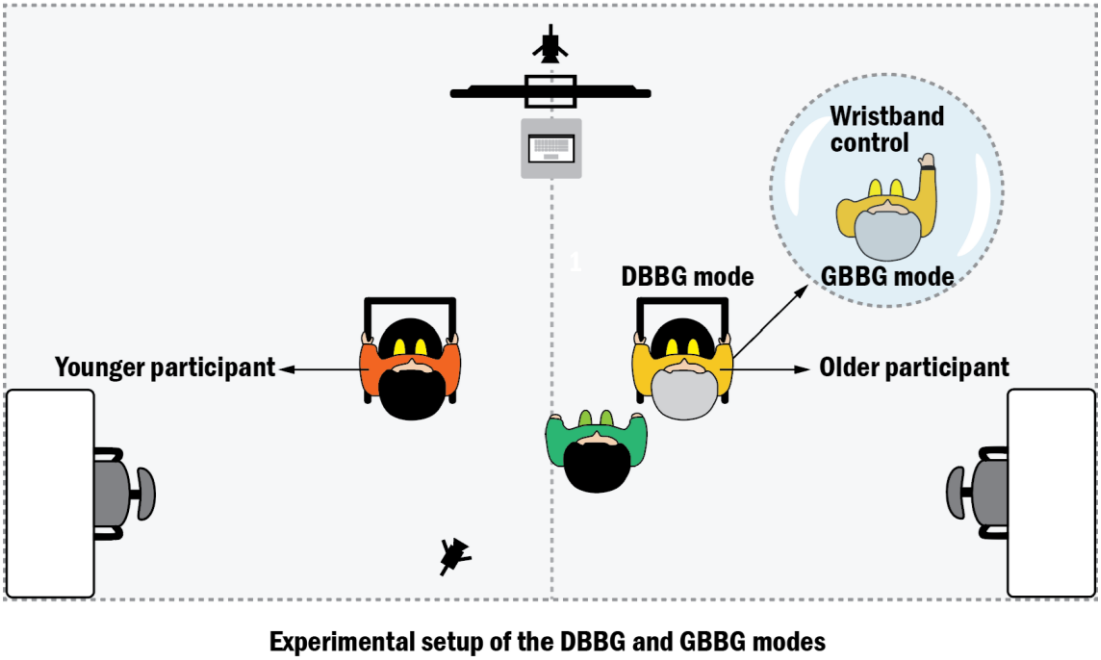

Figure 7 Experimental setup of the DBBG and GBBG modes. In the GBBG mode, the bracket was removed, and the older participant stood in the same position as in the DBBG mode.

3.4 Procedure

This study was approved by the Human Science and Technology Research Ethics Review

Committee of [University name anonymized for review], with the ethics review approval number H20240016I. At beginning of the test, all participants completed informed consent forms, which included an introduction to the experimental purpose, procedures, data collection, and other relevant information. If participants had any questions about the study, the researcher explained them clearly to ensure that all participants fully understood. To protect privacy, the participant's real name was removed and replaced with identification numbers. Three game modes (i.e., SBBG, DBBG, and GBBG) were performed in a counterbalanced order to avoid carryover effects.

At the start of the game, a user guide with simple diagrams was provided to teach players how to control the Balance Bird using a balance ball or wristband. The researcher explained how to use the system, especially for older participants. Each mode lasted six minutes, consistent with the previous study (Qiu, Kaisar, et al., 2023), ensuring that both older and younger players could participate without becoming fatigued. To maintain engagement, each game session consisted of three scenarios (i.e., green forest, fiery desert, and polar ice), each lasting two minutes. In the game, participants needed to control the Balance Bird either independently or collaboratively to hit obstacles and golden eggs that displayed the correct answer to the arithmetic problems while flying. After completing each game mode, participants filled out a post-test questionnaire. The younger participants responded to open-ended questions in written form, while the older participants were interviewed verbally by the researcher using the same questions. All interviews were audio-recorded. The entire experiment lasted approximately 90 minutes. Figure 8 illustrates the overall experimental procedure.

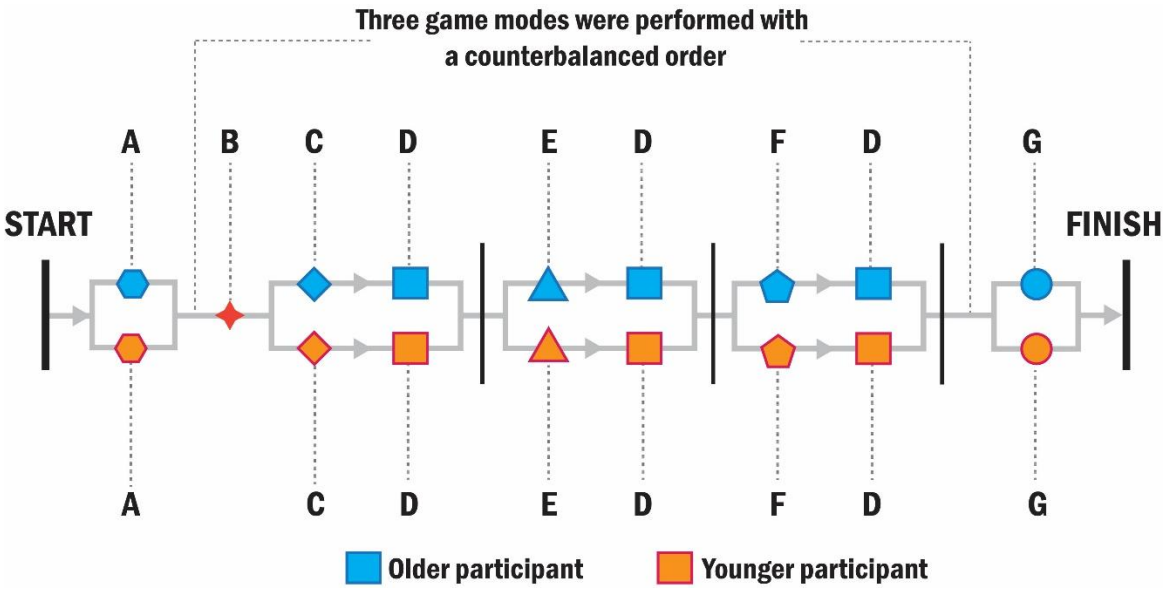

Figure 8 Overall procedure of the experiment: (A) complete consent forms; (B) (experimenters) explain game instructions; (C) test for the SBBG mode; (D) complete the post-experimental questionnaire; (E) test for the DBBG mode; (F) test for the GBBG mode; (G) open questions.

3.5 Measurements

According to Thüring and Mahlke (2007), the evaluation of user experience should go beyond a narrow focus on usefulness and usability, including a broader spectrum of emotional experiences. A comprehensive understanding of user experience is crucial, particularly when designing intergenerational exergames for older adults, as it involves more than just physical health benefits. In our study, we evaluate participants' user experience with the Balance Bird from four perspectives using standardized questionnaires: (1) game experience, (2) motivation, (3) perceived workload, and (4) usability. In addition to the standardized questionnaires, we also utilized system logs to record game scores and conducted semi-structured interviews to gain a deeper understanding of participants' perceptions of the game system and their design suggestions.

- **Game Experience Questionnaires (GEQ)** used to measure participants' game experience in the Balance Bird. The core section of the GEQ (IJsselsteijn et al., 2013) have been applied by game researchers and practitioners to assess experiences across a wide range of game types, user groups, game environments, and purposes. It evaluates game experience based on scores in seven dimensions: (1) competence, (2) sensory and

imaginative immersion, (3) flow, (4) tension/annoyance, (5) challenge, (6) negative affect, as well as (7) positive affect, with a total of 33 items.

- **Intrinsic Motivation Inventory (IMI)** was used to measure the motivation of the participants. The complete IMI scale consists of 45 items across seven scales. These areas are based on Self-Determination Theory (SDT), a psychological research framework that focuses on human motivation and personality (Deci and Ryan, 2008; Ryan and Deci, 2000). We selected three subscales, namely (1) interest/enjoyment, (2) effort/importance, and (3) value/usefulness, to examine participants' motivation and effort in the Balance Bird game.
- **NASA Task Load Index (NASA-TLX)** was used to assess participants' perceived workload during the game. The TLX is a commonly used measure of subjective mental workload. It relies on a multidimensional structure and scores based on six subscales: (1) mental demand, (2) physical demand, (3) temporal demand, (4) performance, (5) effort, and (6) frustration (Cao et al., 2009). The TLX is a user-friendly and adaptable scale for measuring an individual's self-assessed effort in gaming.
- **System Usability Scale (SUS)** was used to evaluate the usability of the game system. SUS is a widely used standardized questionnaire for assessing perceived usability (Lewis, 2018). John Brooke developed the SUS in 1996 (Brooke, 2013), which consists of 10 basic and straightforward questions regarding system usability. SUS is a useful tool for understanding the problems users face when interacting with a system, and stands out due to its wide applicability, simplicity, and rapid adoption (Bangor et al., 2008). When participants completed the game tasks, they could quickly use the SUS for scoring.
- **System logs** are JSON (JavaScript Object Notation) files generated in the Unity game that record users' total scores and accuracy rates for their answers to the arithmetic problems after completing each game mode. We can perform statistical analysis to compare the system logs across different game modes.
- **Semi-structured interviews** were conducted for older players after completing all game sessions. Older participants verbally answered two questions: (1) "*How was your experience in the three game modes?*" and (2) "*What suggestions do you have for*

improving the game?" The interviews of older participants were audio recorded and transcribed verbatim. Younger participants answered the same questions in written form. Both sets of responses were combined for analysis.

3.6 Statistical Analysis

To analyze the quantitative data from standard questionnaires and system logs, we used a mixed ANOVA, using game mode ((1) SBBG, (2) DBBG, (3) GBBG) as the within-subjects factor and generation ((1) older players, (2) younger players) and gender ((1) male, (2) female) as between-subjects factors. The data were analyzed using SPSS 26.

To analyze the interview data, we used a method named qualitative content analysis (Hsieh and Shannon, 2005), which interprets the content of textual datasets through systematic coding and categorization. All interview data were transcribed verbatim. Due to the clear pronunciation of younger participants, we used the speech-to-text software (www.iflytek.com) for direct transcription and proofreading. However, older participants often had accents that could not be directly transcribed by software, so two researchers (R.Y. and S.Z.) manually transcribed their speech and cross-checked their work. The data analysis process consisted of two steps: (1) quote selection, and (2) collaborative coding (Hsieh and Shannon, 2005). In the *quote selection* phase, two researchers selected quotes from the interview transcripts that reflected participants' perceptions of game modes and design advice. During the *collaborative coding* phase, two researchers read the selected quotes, discussed and debated the coding results, and defined and named categories. Discrepancies were resolved through discussion between two researchers (S.Q. and S.L).

4 Results

4.1 Quantitative Results

Descriptive and inferential data results are presented in Appendix Tables A1–A5.

4.1.1 Game Experience

Competence: Exergame mode significantly affected competence F (1.42, 45.28) = 5.37, p = 0.02, η_p^2 = 0.14, 95% *CI* [0.02, 0.30], with higher competence reported in the GBBG (M = 3.06) compared to the SBBG (M = 2.71).

Sensory and Imaginative Immersion: Gender significantly influenced immersion F (1, 32) = 6.16, p = 0.02, η_p^2 = 0.16, 95% *CI* [0.02, 0.36], with females (M = 2.88) feeling more immersed than males (M = 2.34). Generation also had a significant effect F (1, 32) = 35.73, $p \leq 0.001$, η_p^2 = 0.53, 95% *CI* [0.32, 0.66], with older participants (M = 3.26) experiencing more immersion than younger participants (M = 1.95).

Flow: Exergame mode significantly affected flow F (1.96, 62.77) = 3.44, $p \leq 0.05$, η_p^2 = 0.1, 95% *CI* [0.01, 0.23], with more flow in the DBBG (M = 2.77) than in the GBBG (M = 2.43). Gender also had a significant effect F (1, 32) = 8.6, p = 0.01, η_p^2 = 0.21, 95% *CI* [0.04, 0.40], with females (M = 2.93) experiencing more flow than males (M = 2.30). Additionally, generation influenced flow F (1, 32) = 20.65, $p \leq 0.001$, η_p^2 = 0.39, 95% *CI* [0.17, 0.56], with older participants (M = 3.11) reporting more flow than younger ones (M = 2.13).

Challenge: Generation significantly influenced perceived challenge F (1, 32) = 21.46, $p \leq 0.001$, η_p^2 = 0.4, 95% *CI* [0.18, 0.58], with older participants (M = 1.90) feeling more challenged than younger ones (M = 0.82). A significant interaction between exergame modes and gender was observed F (2, 64) = 4.55, p = 0.01, η_p^2 = 0.12, 95% *CI* [0.02, 0.26], although no specific pairwise differences were found. Females generally perceived higher challenge than males in the SBBG but lower challenge in the DBBG and GBBG.

Negative Affect: Gender significantly influenced negative affect F (1, 32) = 4.01, p = 0.05, η_p^2 = 0.11, 95% *CI* [0.002, 0.31], with females (M = 0.37) experiencing less negative affect than males (M = 0.79). A significant interaction between exergame modes and gender was observed F (2, 64) = 7.18, $p \leq 0.001$, η_p^2 = 0.18, 95% *CI* [0.05, 0.32], with females showing less negative affect than males in the DBBG and GBBG, but similar levels in the SBBG.

Positive Affect: Gender significantly influenced positive affect F (1, 32) = 7.04, p = 0.01, η_p^2 = 0.18, 95% *CI* [0.03, 0.38], with females (M = 3.30) experiencing more positive affect than males (M = 2.72). Generation also had a significant effect F (1, 32) = 21.59, $p \leq 0.001$, η_p^2 = 0.40, 95% *CI* [0.18, 0.58], with older participants (M = 3.52) reporting more positive affect than younger ones (M = 2.50). A significant interaction between exergame modes and gender was observed F (1.87, 59.73) = 4.18, p = 0.02, η_p^2 = 0.12, 95% *CI* [0.01, 0.26], with females showing higher positive affect in the SBBG, DBBG, and GBBG compared to males.

4.1.2 Intrinsic Motivation

Interest: Gender significantly affected interest F (1, 32) = 4.78, p = 0.04, η_p^2 = 0.13, 95% *CI* [0.01, 0.33], with females (M = 5.74) showing more interest than males (M = 5.01). Generation also had a significant effect F (1, 32) = 5.35, p = 0.03, η_p^2 = 0.14, 95% *CI* [0.02, 0.34], with older participants (M = 5.76) perceiving more interest than younger ones (M = 4.98). A significant interaction between exergame modes and gender was observed F (1.93, 61.77) = 5, p = 0.01, η_p^2 = 0.14, 95% *CI* [0.02, 0.27], with females reporting higher interest than males in the DBBG and GBBG, but similar levels in the SBBG.

Value: Gender also significantly influenced value F (1, 32) = 4.93, p = 0.03, η_p^2 = 0.13, 95% *CI* [0.01, 0.33], with females (M = 5.97) perceiving higher value than males (M = 5.41). Generation had a significant effect as well F (1, 32) = 10.05, $p \leq 0.001$, η_p^2 = 0.24, 95% *CI* [0.07, 0.43], with older participants (M = 6.10) perceiving more value than younger ones (M = 5.28). A significant interaction between exergame modes and gender was found F (1.81, 58.04) = 3.54, p = 0.04, η_p^2 = 0.10, 95% *CI* [0.01, 0.24], with females perceiving slightly higher value in the SBBG and DBBG, and significantly higher value in the GBBG.

4.1.3 Perceived Workload

A significant interaction between exergame modes and generation was observed F (1.91, 61.27) = 3.35, p = 0.04, η_p^2 = 0.09, 95% *CI* [0.002, 0.22]. Older participants perceived similar workload to younger ones in the SBBG but felt a significantly heavier workload in the DBBG and GBBG.

4.1.4 System Usability

Gender significantly influenced system usability F (1, 32) = 5.95, p = 0.02, η_p^2 = 0.16, 95% *CI* [0.02, 0.36], with females (M = 76.81) experiencing better usability than males (M = 68.61).

4.1.5 System Logs – Total Score

Gender significantly affected total scores F (1, 32) = 5.39, p = 0.03, η_p^2 = 0.14, 95% *CI* [0.01, 0.34], with females (M = 156.35) scoring higher than males (M = 135.83).

4.1.6 Summary

In response to **RQ1-3**, we analyzed the main effects of game modes, gender, and generation, as well as the interaction effects between exergame mode and gender, and between exergame mode and generation.

We observed a significant main effect of the exergame modes (Figure 9). For game experience, participants felt more competent in the GBBG compared to the SBBG. However, they experienced less flow in the GBBG than in the DBBG. We also observed a significant main effect of gender (Figure 10). In the exergames, female participants reported significantly higher levels of sensory and imaginative immersion, flow, positive affect, interest, and perceived value, as well as lower levels of negative affect compared to male participants. Additionally, they experienced better system usability and achieved higher total scores than male participants. A significant main effect of generations was also noted (Figure 11). Older participants in the exergames reported significantly higher levels of sensory and imaginative immersion, flow, positive affect, interest, and value, but they perceived a greater level of challenge compared to younger participants.

A significant interaction effect was observed between the exergame modes and gender (Figure 12). In the SBBG, female and male participants reported similar levels of negative affect and interest, or slightly higher positive affect and value. However, in the DBBG, female participants reported significantly less negative affect and more interest than male participants. Similarly, in the GBBG, female participants reported significantly less negative affect, more positive affect, more interest, and higher value compared to male participants. Furthermore, a significant interaction effect between the exergame modes and generation was

also noted (Figure 13). In the SBBG, older participants perceived the same workload as younger participants, while in the DBBG and GBBG, they perceived significantly higher workload than younger participants.

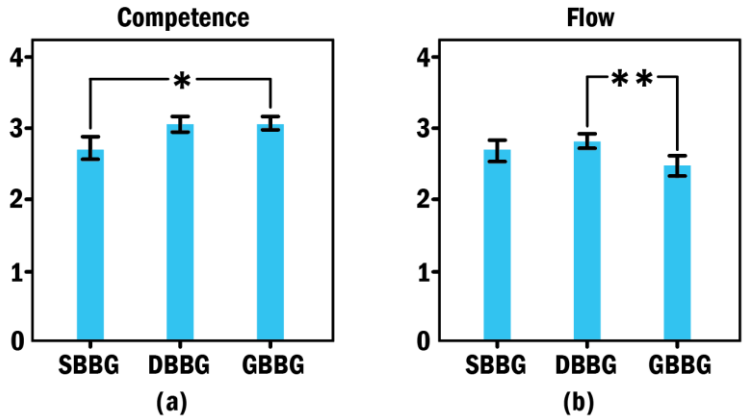

Figure 9 Main effects of the exergame modes on competence (a) and flow (b). Significant group difference: *p< .05, **p< .01.

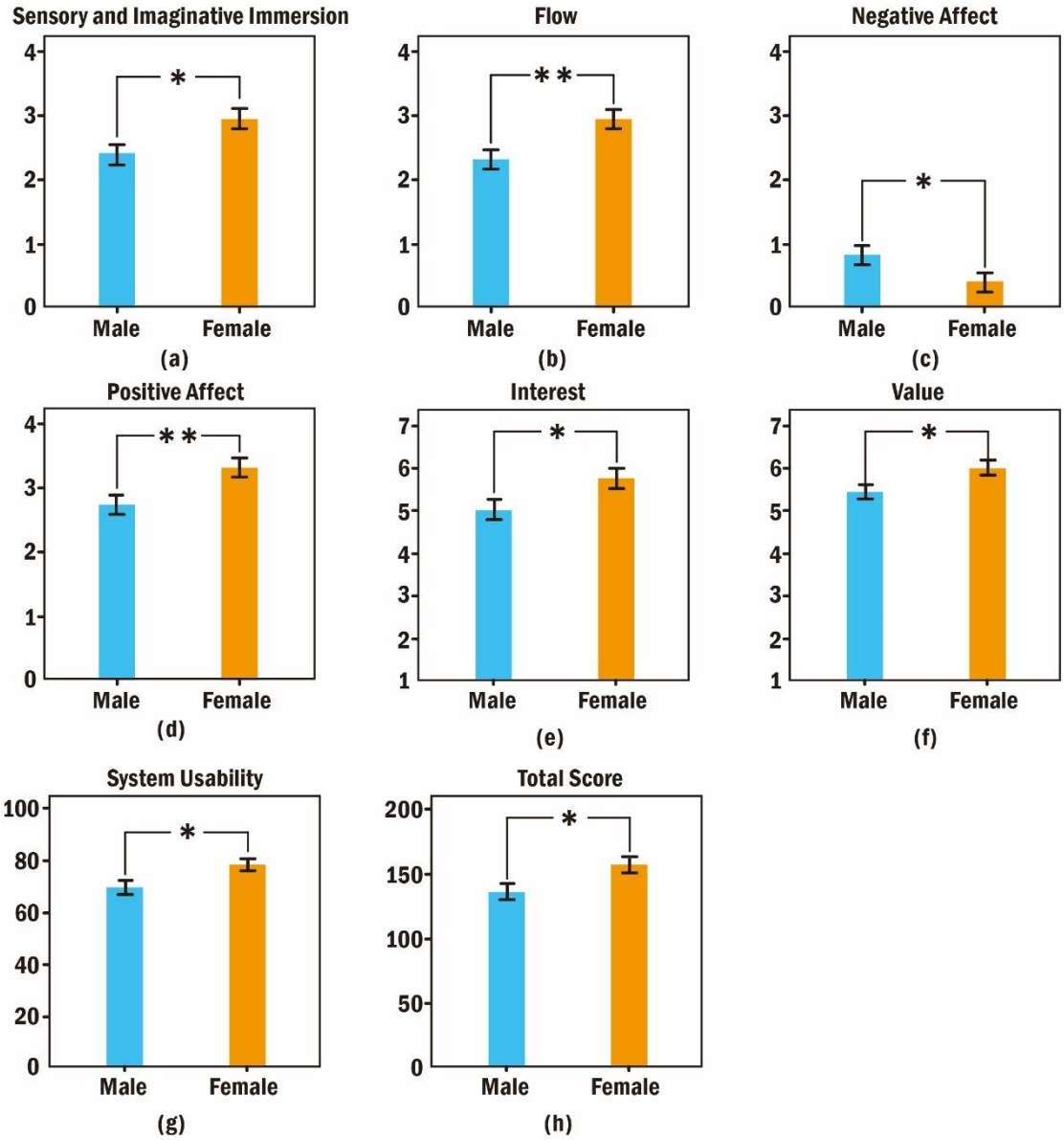

Figure 10 Main effects of gender on sensory and imaginative immersion (a), flow (b), negative affect (c), positive affect (d), interest (e), value (f), system usability (g), and total score (h). Significant group difference: **p*< .05, ***p*< .01.

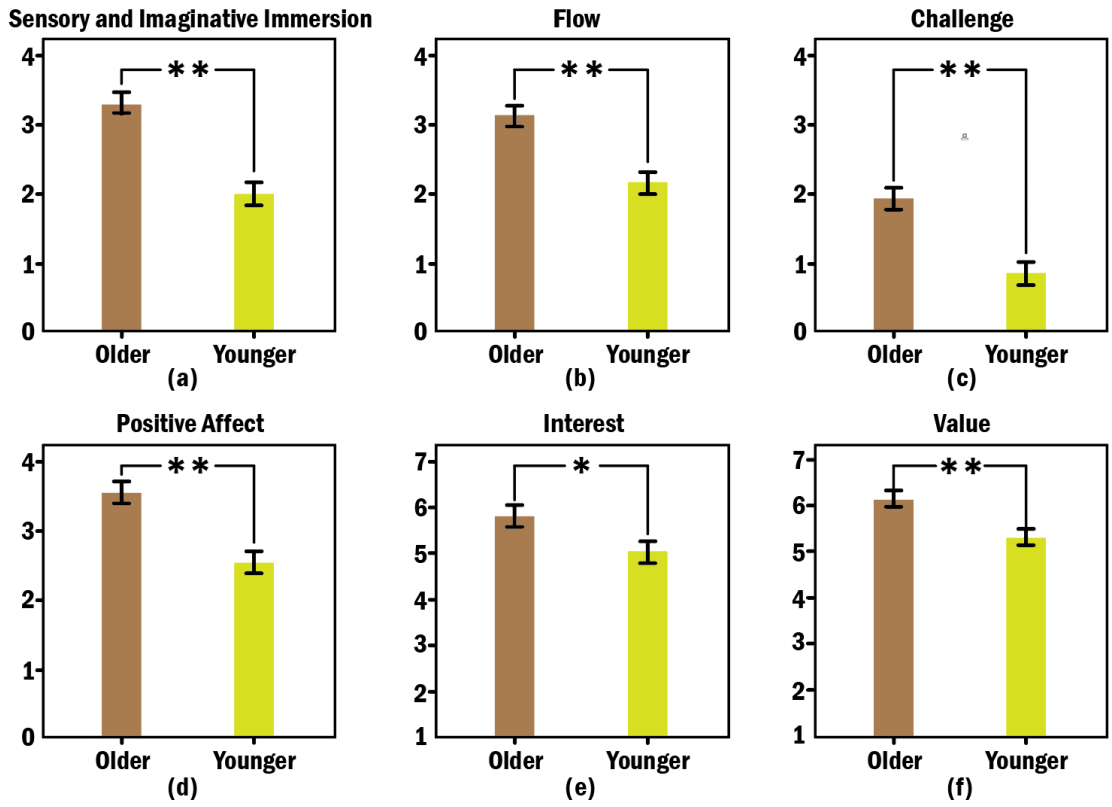

Figure 11 Main effects of generation on sensory and imaginative immersion (a), flow (b), challenge (c), positive effect (d), interest (e), and value (f). Significant group difference: **p*<.05, ***p*<.01.

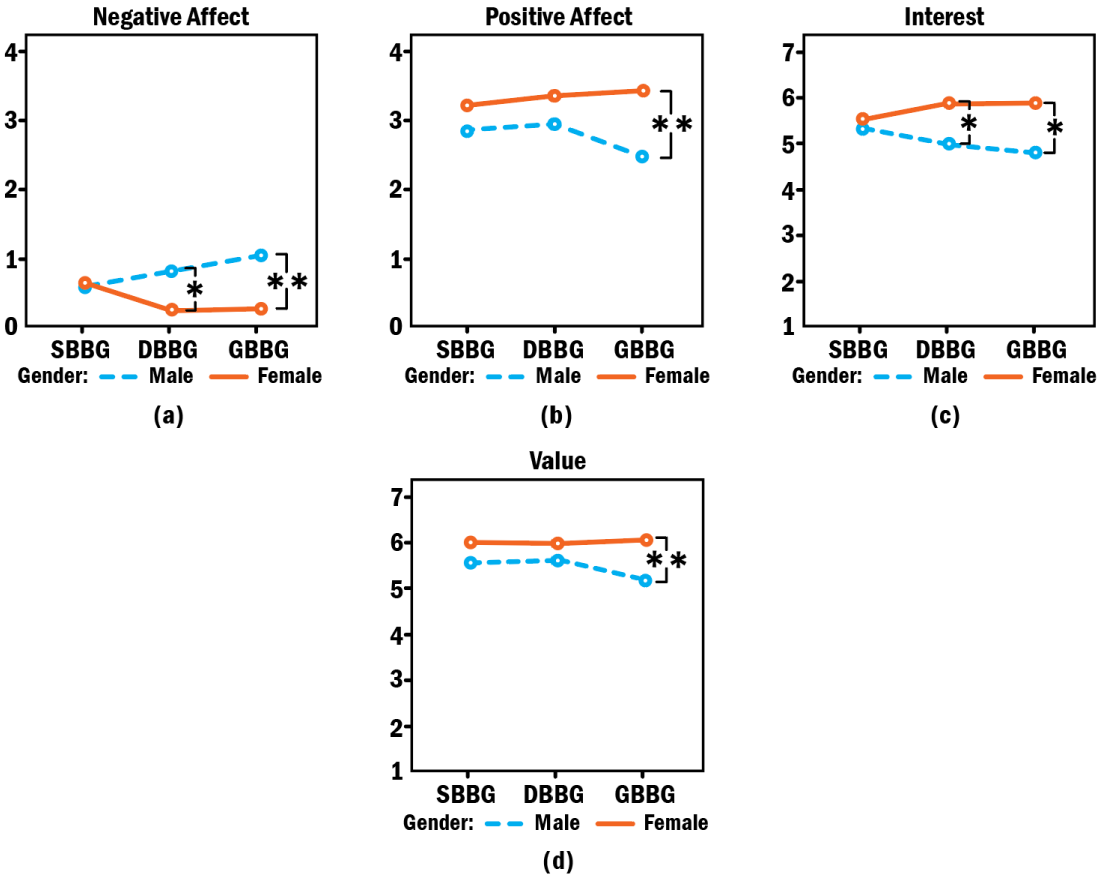

Figure 12 Interaction effects between the exergame modes and gender on negative affect (a) and positive affect (b), interest (c), and value (d). Significant group difference: **p*<.05, ***p*<.01.

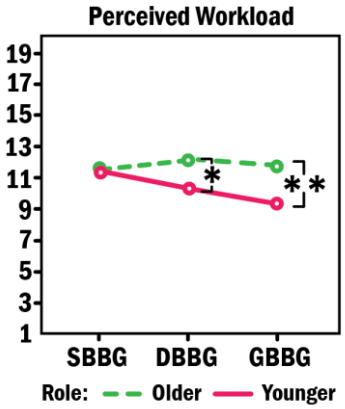

Figure 13 Interaction effects between the exergame modes and generation on perceived workload. Significant group difference: **p*<.05, ***p*<.01.

4.2 Qualitative Results

In this study, we identified categories and subcategories based on these quotes regarding the two aspects: (1) perceptions of game modes, and (2) design suggestions.

4.2.1 Perceptions of Game Modes

One hundred and twenty-three quotes reflect participants' perceptions of the three game modes. Corresponding to the quantitative data across the three game modes (Tables A1-A3), in the qualitative data analysis, we not only summarized quotes reflecting participants' perceptions of the game modes (Table 2) but also quantified quotes related to the game modes across different user groups, including gender (Table 3) and generational differences (Table 4). To indicate the source of each quote, we assigned a unique ID to every participant. "O" denotes older participants, while "Y" represents younger ones. "F" stands for female participants, and "M" stands for male participants. In reporting the qualitative results of participants' perceptions across the three game modes, we organized the positive and negative feedback for each mode into distinct themes and illustrated these themes with representative participant quotes. This approach followed the methodology of Stephan et al. (2024).

Regarding the game modes, all participants provided a total of 80 positive quotes and 43 negative quotes. In the SBBG mode, there were 20 positive quotes and 20 negative quotes. Positive feedback centered on cognitive and hand-eye coordination improvement ("I think this is a great exercise for coordination between the brain and hands, especially for older people like us (OF15)") and immersive engagement ("The immersive experience is fascinating, and exploring different scenarios is highly engaging (YF12)"). Negative feedback focused on solo-play workload ("When I play alone, I find it difficult to manage everything at once (OF15)") and time pressure with cognitive load ("There is a sense of urgency. Sometimes I feel my brain cannot keep up when I see arithmetic problems (OF27)").

In the DBBG mode, there were 29 positive quotes and only nine negative quotes. Positive feedback focused on reduced individual burden ("It feels good because I do not have to take on so much responsibility at once (OF11)") and enjoyment of partner interaction ("The emotional feedback from my partner during the game (such as cheering and praising) is one

of the most interesting aspects of the game (YM10)"). Negative feedback centered on coordination challenges ("Performance in a cooperative mode is not as good as when playing alone; good coordination is quite challenging, and calculation time is short, so it is more difficult than playing alone (YF18)") and reduced immersion due to partner dependency ("The outcome of the game depends not only on the individual but also on the performance of the partner, which slightly reduces the immersion compared to playing the game alone (YM32)").

In the GBBG mode, there were 22 positive and 14 negative quotes. Positive feedback highlighted preference and comfort with gesture control ("For me, it would be more convenient to use gestures; I prefer gestures (OF25)") and perceived efficiency from partner's gesture use ("It felt easier because my partner seemed more efficient using gestures to control the interface (YF26)"). Negative feedback primarily reflected gesture-control confusion ("The wristband device sometimes gives me the illusion of controlling left and right, as if I want to control both directions at the same time (OF15)").

Table 2 Number of positive and negative quotes of three game modes.

Modes	Number of Positive Quotes	Number of Negative Quotes
SBBG	29	20
DBBG	29	9
GBBG	22	14
Total	80	43

Table 3 Number of positive and negative quotes on game modes: male vs. female participants.

Modes	Gender	Number of Positive Quotes	Number of Negative Quotes
SBBG	Male	16	9
	Female	13	11
DBBG	Male	12	6
	Female	17	3
GBBG	Male	7	8
	Female	15	6
Total	Male	35	23
	Female	45	20

Table 4 Number of positive and negative quotes on game modes: older vs. younger participants.

Modes	Generations	Number of Positive Quotes	Number of Negative Quotes
SBBG	Older	18	10
	Younger	11	10
DBBG	Older	18	4
	Younger	11	5
GBBG	Older	14	8
	Younger	8	6
Total	Older	50	22
	Younger	30	21

4.2.2 Design Advice

Based on the qualitative results, we summarized the following recommendations that could contribute to future improvements. We did not quantify the design suggestions; instead, we selected and categorized relevant quotes based on their quality. High-quality design suggestions provide more valuable guidance for future designs. The design suggestions can be primarily categorized into three categories: (1) gameplay mechanics, (2) multisensory feedback interface, as well as (3) game settings and adjustments.

"Gameplay mechanics" consists of four subcategories: (1) cognitive tasks in games, (2) social interaction, (3) game challenge difficulty, and (4) beginner mode.

(1) Cognitive Tasks in Games. In addition to arithmetic problems, cognitive tasks in games can also include other forms, such as shape and color recognition tasks, which can enhance the variety and enjoyment of these tasks (YM14, YF18, YF36, YF40). YF18 provided some specific examples, such as allowing users to identify animal names through images, like asking, "*Which picture shows a crocodile?*" or "*What day is today?*" Additionally, OM19 emphasized that he hoped the current cognitive tasks in the game could be made even simpler.

(2) Social Interaction. Although the current game mode includes two-person collaboration, the number of participants is still relatively small. Expanding it to a multiplayer format could create a "Social + Sports" model that encourages healthy competition among users, fosters community engagement, and enhances user retention. For instance, YM20 suggested introducing a multiplayer cooperation mode, allowing not just two players but potentially expanding to larger groups or becoming a

team-building activity. YF36 noted that while the current game mode allowed for collaboration with older adults, there was limited communication; therefore, more interactive gameplay options could be added to enhance engagement.

(3) **Game Challenge Difficulty.** The game difficulty can be increased, especially for younger players. OF25 stated, "*The difficulty could be raised a bit more. It is too easy for young people. For older adults, it is still manageable*." Similarly, OM29 also expressed that the game difficulty could be increased a little. YF14 believed the game pace should be faster, similar to the challenges in *Overcooked*. YM20 suggested that incorporating both hands and feet to increase the operational difficulty would make the game more engaging.

(4) **Beginner Mode.** Beginner mode is essential, especially for older adults, as it allows them to familiarize themselves with and experience the game process in advance. OF15 stated, "*You need to have the opportunity to practice and become familiar with the game before you begin, understand what it entails, and try it out for yourself.*" YM38 also stressed the importance of establishing a learning phase, especially for older adults.

"Multisensory feedback interface" includes two subcategories: (1) background music and sound effects, and (2) prompt icons.

(1) **Background Music and Sound Effects.** YF6 provided specific suggestions for the game background music: "*Popular music can be selected, as it should be acceptable to both older and younger people from various decades.*" YF8 also believed that background music and sound effects needed to be included. YF26 provided specific suggestions for sound effects, such as adding sounds for correct and incorrect responses.

(2) **Prompt Icons.** Optimize the visual design of the UI interface, such as adding prompt icons during game interactions. As YF26 mentioned, "*Some elements in the game can be misleading. For instance, in the second level, the rock element led me to instinctively think I should avoid it during my first playthrough, [when in fact I needed to hit the rock to score]. Therefore, I hope a prompt can be added.*" Similarly, YF34

wished to add some icons in the game scene as reminders to help players react more easily.

"Game settings and adjustments" includes two subcategories: (1) game duration settings and (2) hardware sensitivity.

(1) **Game Duration Settings.** Interestingly, older players expressed different views from younger players regarding game duration settings. OF17 stated, "*The time is quite long, and it feels a bit tiring, plus the numbers later on are a bit larger [indicating a higher difficulty]*." In contrast, YM32 felt that the game required minimal physical activity and was mainly a way for young people to stretch themselves.

(2) **Hardware Sensitivity.** Game controls need to increase sensitivity. For instance, OM11 mentioned that older adults might react more slowly, and if the sensitivity was low, they had to exert extra effort to reach their goals. YF36 pointed out that when controlling the game with two players, the precision was not as accurate as when one person was in control.

5 Discussion

In this section, we discuss our research findings based on four research questions, address the limitations of this study, and propose directions for future work.

5.1 RQ1: "How do older and younger players experience Balance Bird across the three different game modes?"

For RQ1, the results clearly showed that participants in the GBBG felt more competent and capable compared to those in the SBBG. This finding supports the expectation that the social interaction embedded in the GBBG setting contributed positively to participants' sense of ability and self-efficacy. This interpretation is consistent with recent systematic reviews and meta-analyses, which have identified social interaction as a key moderating factor influencing the effectiveness of exergame-based training on cognitive functions such as attention, particularly in group-based training contexts (Rieker et al., 2022). Specifically, Rieker et al. (2022) found that combined motor-cognitive training conducted in a group

setting significantly improved cognitive performance in older adults compared to unsupervised individual training, the authors attributed this improvement, at least in part, to the motivational impact of social interaction during training. These observations also align with findings from conventional motor-cognitive training (Manser et al., 2024) and have been interpreted through the lens of social interaction as a motivational factor, which further contributes to long-term adherence and sustained cognitive performance.

Interestingly, participants reported lower flow in the GBBG than in the DBBG, likely due to mismatched input methods. In the DBBG, similar lower-limb movements fostered bodily synchrony and mutual engagement, whereas in the GBBG, different gestures reduced shared immersion. This supports the "Chameleon Effect" (Chartrand and Bargh, 1999), where physical mimicry enhances connection. These results suggest that synchronized physical input can enhance intergenerational engagement and should be considered in future exergame design.

5.2 RQ2: "Are there any differences in user experience based on gender?"

Contrary to previous findings that males are generally more engaged in digital games (Chen, 2010; Chou and Tsai, 2007; Hainey et al., 2011), our study found that female participants reported significantly better user experiences in exergaming, with higher scores in SII, flow, affect, interest, and value, as well as better performance in system logs. This contrasts with Qiu et al. (2023), who found no gender differences in social exergaming. A likely reason for the discrepancy lies in methodological differences: Qiu et al. lacked gender balance and focused only on social interaction, whereas our study ensured a 1:1 gender ratio and used a multidimensional evaluation framework covering motivation, usability, and workload, making the results more robust.

Further, gender differences were more pronounced in collaborative modes (DBBG and GBBG), where females reported higher engagement than males. This aligns with prior literature indicating that women are more motivated by social interaction in games (Choi et al., 2012; Guadagno et al., 2011; Hou, 2012; Wohn et al., 2020). Hartmann and Klimmt (2006) also noted that women tend to disengage from games lacking meaningful social context. Our findings suggest that collaborative, socially rich exergame designs may be

especially effective for engaging female players. This challenges gender stereotypes in gaming and offers actionable insights for designing inclusive, intergenerational exergame experiences.

5.3 RQ3: "What differences in user experience exist between older and younger players?"

For RQ3, consistent with (Qiu, et al., 2023), older participants reported significantly better user experiences than younger participants during exergaming, with higher ratings in SII, flow, positive affect, interest, and value, supported by more positive qualitative feedback (N = 50 vs. N = 30; Table 4). Qiu, et al. (2023) attributed this to the contrast between the monotony of retired life and the novelty of social exergames for older adults, while younger participants, exposed to more diverse gaming experiences, found such games less engaging.

Our study extends these findings by revealing that older participants also perceived greater physical and cognitive challenges. The difficulty of the numerical problems in the three levels was based on the cognitive abilities of older participants, using two-digit addition, subtraction, and single-digit multiplication, similar to content in Chinese elementary grades 2 or 3. While older participants are familiar with these concepts in daily life, some still struggled with the speed of the tasks (e.g., OF27), whereas younger participants found the tasks too simple, leading to reduced engagement. These generational differences were particularly evident in collaborative modes (DBBG and GBBG), where older adults reported significantly higher workload than younger peers, despite similar ratings in the solo mode (SBBG). This interaction effect may be explained by age-related declines in motor and cognitive abilities (Goodpaster et al., 2006; Salthouse, 2010), requiring older participants to exert more effort to match the rhythm and demands of younger teammates. These results highlight the dual nature of older adults' experience: enhanced enjoyment alongside increased exertion. Unlike prior studies that primarily emphasized emotional or social benefits, our multi-dimensional approach provides a more nuanced understanding.

5.4 RQ4: "What design insights can be gained to inform the development of future intergenerational exergames?"

Based on the interviews, we identified several key design insights:

Cognitive Task Design. In this study, the game's cognitive tasks involved solving arithmetic problems of increasing difficulty. Although the calculations were simple, the dynamic game scenarios and required coordination overwhelmed older players, making task completion challenging. Future designs could integrate alternative cognitive elements such as color- or shape-based tasks (for example, "Which picture shows a crocodile?") or flexible numerical tasks including selecting matching numbers or varying number frequency, thereby diversifying challenges and enhancing older players' engagement.

Social Interaction Expansion. We examined young-old dyads since two-person groups are ideal for observing social behavior (Charon, 1996). Intergenerational exergames involving pairs provide a foundation for studying such interactions (Hagad et al., 2011). However, future game designs should go beyond the typical two-person format by incorporating more participants, thereby enhancing family cohesion and expanding community-level engagement.

At the implementation level, current research mainly focuses on inviting community older adults to community activity centers or nursing homes to engage in face-to-face interactions with nearby young people. To further expand the interaction scenario, this study suggests incorporating the home environment, particularly for older adults who are unable to regularly participate in group activities due to mobility issues or remote living. By integrating social functions into home-based exergames, they can achieve high-quality social participation comparable to offline group activities without having to visit an activity center (Manser et al., 2025; Manser and de Bruin, 2024). Furthermore, to enhance sustained participation and engagement, social elements such as collaborative challenges and real-time leaderboards can be integrated into the training games (Manser et al., 2025; Manser and de Bruin, 2024). A progressive strategy, transitioning from community guided training to home-based training, should be implemented to promote long-term adherence among participants (Manser et al., 2025; Manser and de Bruin, 2024).

Onboarding Support for Older Adults. A beginner mode for older players can help them learn game rules and operations before joining younger players. Tutorials or training sessions (Yuan et al., 2023) could reduce reliance on younger guides and improve older players' confidence. Building on the concept of adaptive difficulty (Cantwell et al., 2012), where younger players use lower-body inputs and older players use upper-body inputs, adding a beginner mode could further ease learning, reduce stress, and enhance overall user experience.

Feedback and Reward Mechanisms. Manser et al. (2025) categorize the feedback mechanisms in exergames into "immediate" and "delayed" types, guiding users from passive participation to long-term adherence. For instance, in "Balance Bird" training, immediate feedback may progress from basic sensory cues (e.g., cheers, visual effects) to real exercise data (e.g., calories burned) and rewards (e.g., points, new levels). Subsequently, feedback includes heart rate monitoring and motion optimization through visual/audio cues. Delayed feedback reinforces long-term adherence through participation badges and virtual trophies. This design transforms external incentives into intrinsic motivation for sustaining healthy behavior.

5.5 Implications for Research and Practice

First, social interaction in the GBBG mode enhanced participants' competence and self-efficacy, highlighting its role as a motivational mechanism in group-based exergame cognitive training. This aligns with Fang et al. (2024), who showed that socially supported exercise programs improved older adults' self-efficacy and sustained engagement. Similarly, providing external motivational support, such as peer support, competition, and feedback, can also further enhance motivation and adherence (Palombi et al., 2025). Thus, future research could examine how different forms of social interaction (e.g., cooperation, competition, collaborative feedback) influence motivation and cognitive outcomes in intergenerational exergames and use longitudinal designs to assess whether socially driven motivation supports long-term adherence and cognitive benefits.

Second, this study challenges long-standing gender stereotypes by showing that women were more engaged and had better experiences in collaborative, socially interactive

exergames. This highlights social interaction as a key design element for engaging female players. For instance, incorporating dance-based movements and blending light competition with collaboration may enhance enjoyment while reducing pressure. Future research should test these mechanisms in one to two months interventions to assess their effects on motivation, cognitive gains, and gender differences in long-term adherence.

Third, despite greater physical and cognitive demands, older participants reported significantly better exergaming experiences, reflecting a dual pattern of heightened enjoyment and increased exertion. Future research could examine adaptive systems that adjust task speed, complexity, or intensity to match age-related abilities, reducing frustration for older players while sustaining younger players' engagement. Additionally, pairing older adults with peers of similar ability or using role-based collaboration with complementary tasks could ease intergenerational strain while maintaining social benefits.

5.6 Limitations and Future Work

We identify several limitations of this study and outline opportunities for future research.

First, we primarily used four subjective questionnaires to measure participants' comprehensive user experience. For older adults, the number of questions may be excessive; For instance, the GEQ questionnaire contains 33 items. Despite the participants were healthy older adults without cognitive impairments, some provided incomplete answers, requiring reminders from the experimenters. Therefore, future research involving older adults should consider using more concise questionnaires to minimize the number of subjective questions.

Second, the GBBG mode, which involves upper body movements, was originally intended to reduce the operational burden on older adults. However, it did not provide a better user experience compared to the DBBG mode. This outcome may be attributed to the less consistent execution of upper body movements by older adults, leading to system misjudgments that negatively impact the user experience. Future efforts should prioritize optimizing the GBBG mode to enhance the sensitivity and adaptability of the system. This could include improving algorithms to better recognize and accommodate irregular movements.

Third, although the arithmetic tasks were relatively simple, they still posed cognitive challenges for some older participants, indicating that simplification alone is not sufficient. As exergames should balance user experience with exercise principles, particularly progressive overload, cognitive task design must combine accessibility with suitable challenge. Rather than merely reducing complexity, emphasis should be placed on adaptive difficulty and gradual progression. For instance, dynamic difficulty adjustment can increase task complexity and timing based on performance, while incorporating language comprehension or context-based questions such as date recognition or daily scenarios can provide multimodal cognitive stimulation. Such designs enrich task variety, enhance real-world relevance, align with exercise principles, and reflect the dual nature of older adults' experiences by combining heightened enjoyment with greater effort.

Fourth, during participant recruitment, in addition to age and general health status, we included physical activity level as a screening criterion. Specifically, priority was given to older adults with regular exercise habits, such as morning workouts or square dancing. It should be noted that while our findings are applicable to older adults who engage in routine physical activity, the results may not be generalizable to "younger older adults" (aged 55–75) with irregular or no exercise habits, due to the limited sample representation of low-activity individuals.

Fifth, this study primarily focused on healthy older adults with an average age of 66.89 years, most of whom are classified as young-old (under 75) (Xu et al., 2016). Future research should consider recruiting older adults in the old-old category (over 75) (Xu et al., 2016), as their participation in intergenerational exergaming studies would present a more significant research challenge.

Lastly, while continuing to study intergenerational communication between older and younger adults, the experimental design could be extended to other types of intergenerational play. For instance, this approach could also be applied to the user experiments including children with autism and their parents, using exergames as a social intervention to evaluate their impact on children's stress levels and social adaptability.

6 Conclusion

In this study, we introduce the Balance Bird exergame, which integrates social, cognitive, and physical exercises to enhance balance training for older adults in a game-based environment. This exergame features interdependent bodily interactions between younger and older players, enabling both groups to use their upper or lower bodies to jointly control the Balance Bird's horizontal and vertical movements as it navigates obstacles. A user experiment involving 18 pairs of unfamiliar young and older adults assessed their experience with the Balance Bird. The results indicated that participants felt more competent and had a stronger sense of ability in the game mode with interdependent bodily interactions and collaborative cognitive tasks compared to the independent single-player mode without such interactions. This suggests that the Balance Bird can effectively enhance the balance training experience for older adults through intergenerational exergames. Participant feedback also provided valuable design insights for future improvements and the development of intergenerational games.

Declarations

Conflicts of interest/Competing interests

The authors declare that they have no conflict of interest/competing interests.

Declaration of Generative AI and AI assisted technologies in the writing process

The authors declare that they do not use AI tools to analyze and draw insights from data as part of the research process.

Authors' contributions

Conceptualization, S.Q.; experimental design, S.Q.; software, E.K, X.Y, R.Y.; experiment implementation, R.Y., S.Z., and X.Y.; data input: R.Y. and S.Z.; formal analysis, S.Q., S.L.; writing—original draft preparation, S.Q.; writing—review and editing, S.Q. and J.H.; revise the manuscript; S.Q.; supervision, S.Q., T.H; sponsor: S.Q.

Consent for publication

All authors read and approved the manuscript.

Reference

Aarsand, P. A. (2007). Computer and Video Games in Family Life: The digital divide as a resource in intergenerational interactions. *Childhood*, *14*(2), 235–256. https://doi.org/10.1177/0907568207078330

Bangor, A., Kortum, P. T., & Miller, J. T. (2008). An Empirical Evaluation of the System Usability Scale. *International Journal of Human-Computer Interaction*, *24*(6), 574–594. https://doi.org/10.1080/10447310802205776

Brooke, J. (2013). *SUS: A Retrospective* (Vol. 8).

Cantwell, D., Broin, D. O., Palmer, R., & Doyle, G. (2012). Motivating elderly people to exercise using a social collaborative exergame with adaptive difficulty. *European Conference on Games Based Learning*, 615. https://www.helmholtz-muenchen.de/fileadmin/JOIN/Motivating_Elderly_People_To_Exercise_Using_a_Social_Collaborative_Exergame_with_Adaptive_Difficulty.docx

Cao, A., Chintamani, K. K., Pandya, A. K., & Ellis, R. D. (2009). NASA TLX: Software for assessing subjective mental workload. *Behavior Research Methods*, *41*(1), 113–117. https://doi.org/10.3758/BRM.41.1.113

Chao, Y. Y., Scherer, Y. K., & Montgomery, C. A. (2015). Effects of using Nintendo WiiTM exergames in older adults: a review of the literature. *Journals.Sagepub.ComYY Chao, YK Scherer, CA MontgomeryJournal of Aging and Health, 2015•journals.Sagepub.Com*, *27*(3), 379–402. https://doi.org/10.1177/0898264314551171

Charon, J. M. (1996). *The Meaning of Sociology: a reader*. Prentice Hall. https://doi.org/10.2307/1318008

Chartrand, T. L., & Bargh, J. A. (1999). The Chameleon Effect: The Perception-Behavior Link and Social Interaction. In *Journal of Personality and Social Psychology* (Vol. 76, Issue 6).

Chen, L. S. L. (2010). The impact of perceived risk, intangibility and consumer characteristics on online game playing. *Computers in Human Behavior*, *26*(6), 1607–1613. https://doi.org/10.1016/J.CHB.2010.06.008

Choi, G., Chung, H., & Kim, Y. (2012). Are Stereotypes Relative to Gender Usage Applicable to Virtual Worlds? *International Journal of Human-Computer Interaction*, *28*(6), 399–405. https://doi.org/10.1080/10447318.2011.601973

Chou, C., & Tsai, M. J. (2007). Gender differences in Taiwan high school students' computer game playing. *Computers in Human Behavior*, *23*(1), 812–824. https://doi.org/10.1016/j.chb.2004.11.011

Chua, P. H., Jung, Y., Lwin, M. O., & Theng, Y. L. (2013a). Let's play together: Effects of video-game play on intergenerational perceptions among youth and elderly participants. *Computers in Human Behavior*, *29*(6), 2303–2311. https://doi.org/10.1016/j.chb.2013.04.037

Chua, P. H., Jung, Y., Lwin, M. O., & Theng, Y. L. (2013b). Let's play together: Effects of video-game play on intergenerational perceptions among youth and elderly participants. *Computers in Human Behavior*, *29*(6), 2303–2311. https://doi.org/10.1016/j.chb.2013.04.037

Cohen, J. (2013). *Statistical power analysis for the behavioral sciences*. routledge.

Cook, S. W. (1984). Cooperative interaction in multiethnic contexts. *Groups in Contact: The Psychology of Desegregation*.

De la Hera, T., Loos, E., Simons, M., & Blom, J. (2017). Benefits and Factors Influencing the Design of Intergenerational Digital Games: A Systematic Literature Review. *Societies*, *7*(3), 18. https://doi.org/10.3390/soc7030018

Deci, E. L., & Ryan, R. M. (2008). Facilitating optimal motivation and psychological well-being across life's domains. *Canadian Psychology / Psychologie Canadienne*, *49*(1), 14–23. https://doi.org/10.1037/0708-5591.49.1.14

Dix, A. (2009). Human-computer interaction. In *Encyclopedia of database systems* (pp. 1327–1331). Springer.

Drazich, B. F., Lafave, S., Crane, B. M., Szanton, S. L., Carlson, M. C., Budhathoki, C., & Taylor, J. L. (2020). Exergames and depressive symptoms in older adults: a systematic review. *Liebertpub.ComBF Drazich, S LaFave, BM Crane, SL Szanton, MC Carlson, C Budhathoki, JL TaylorGames for Health Journal, 2020•liebertpub.Com*, *9*(5), 339–345. https://doi.org/10.1089/G4H.2019.0165

Erdfelder, E., Faul, F., & Buchner, A. (1996). GPOWER: A general power analysis program. *Behavior Research Methods, Instruments, and Computers*, *28*(1), 1–11. https://doi.org/10.3758/BF03203630

Fang, J., Ren, J., Wang, J., Qiu, X., Zhang, S., Yuan, S., Wu, L., & Xie, L. (2024). Combining motivational and exercise intervention components to reverse pre-frailty and promote self-efficacy among community-dwelling pre-frail older adults: a randomized controlled trial. *BMC Geriatrics*, *24*(1). https://doi.org/10.1186/s12877-024-05464-6

Feng, Y., Perugia, G., Yu, S., Barakova, E. I., Hu, J., & Rauterberg, G. W. M. (2022). Context-Enhanced Human-Robot Interaction: Exploring the Role of System Interactivity and Multimodal Stimuli on the Engagement of People with Dementia. *International Journal of Social Robotics*, *14*(3), 807–826. https://doi.org/10.1007/s12369-021-00823-4

Fuchsberger, V., Sellner, W., Moser, C., & Tscheligi, M. (2012). Benefits and hurdles for older adults in intergenerational online interactions. *International Conference on Computers for Handicapped Persons*, 697–704. https://doi.org/10.1007/978-3-642-31522-0_104

Fujiyama, H., Song, C.-Y., Sciences, H., Sandra Bond Chapman, T., Maria Torre, M., & Temprado, J.-J. (2022). A review of combined training studies in older adults according to a new categorization of conventional interventions. *Frontiersin.Org*, *13*. https://doi.org/10.3389/FNAGI.2021.808539/FULL

Giles, H., & Gasiorek, J. (2011). Intergenerational Communication Practices. In *Handbook of the Psychology of Aging* (pp. 233–247). Elsevier. https://doi.org/10.1016/B978-0-12-380882-0.00015-2

Goodpaster, B. H., Park, S. W., Harris, T. B., Kritchevsky, S. B., Nevitt, M., Schwartz, A. V., Simonsick, E. M., Tylavsky, F. A., Visser, M., & Newman, A. B. (2006). The Loss of Skeletal Muscle Strength, Mass, and Quality in Older Adults: The Health, Aging and Body Composition Study. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, *61*(10), 1059–1064. https://doi.org/10.1093/gerona/61.10.1059

Guadagno, R. E., Muscanell, N. L., Okdie, B. M., Burk, N. M., & Ward, T. B. (2011). Even in virtual environments women shop and men build: A social role perspective on Second Life. *Computers in Human Behavior*, *27*(1), 304–308. https://doi.org/10.1016/j.chb.2010.08.008

Hagad, J. L., Legaspi, R., Numao, M., & Suarez, M. (2011). Predicting Levels of Rapport in Dyadic Interactions through Automatic Detection of Posture and Posture Congruence. *2011 IEEE Third Int'l Conference on Privacy, Security, Risk and Trust and 2011 IEEE Third Int'l Conference on Social Computing*, 613–616.

https://doi.org/10.1109/PASSAT/SocialCom.2011.143

Hai, L., Hou, H. Y., Zhou, C., & Li, H. J. (2022). The effect of exergame training on physical functioning of healthy older adults: a meta-analysis. *Games for Health Journal*, *11*(4), 207–224. https://doi.org/10.1089/G4H.2021.0173

Hainey, T., Connolly, T., Stansfield, M., & Boyle, E. (2011). The differences in motivations of online game players and offline game players: A combined analysis of three studies at higher education level. *Computers and Education*, *57*(4), 2197–2211. https://doi.org/10.1016/j.compedu.2011.06.001

Hanson, V. L. (2010). Influencing technology adoption by older adults. *Interacting with Computers*, *22*(6), 502–509. https://doi.org/10.1016/j.intcom.2010.09.001

Hargittai, E., Piper, A. M., & Morris, M. R. (2019). From internet access to internet skills: digital inequality among older adults. *Universal Access in the Information Society*, *18*(4), 881–890. https://doi.org/10.1007/s10209-018-0617-5

Hartmann, T., & Klimmt, C. (2006). Gender and Computer Games: Exploring Females' Dislikes. *Journal of Computer-Mediated Communication*, *11*(4), 910–931. https://doi.org/10.1111/j.1083-6101.2006.00301.x

Hou, H.-T. (2012). Exploring the behavioral patterns of learners in an educational massively multiple online role-playing game (MMORPG). *Computers & Education*, *58*(4), 1225–1233. https://doi.org/10.1016/j.compedu.2011.11.015

Hsieh, H. F., & Shannon, S. E. (2005). Three approaches to qualitative content analysis. *Qualitative Health Research*, *15*(9), 1277–1288. https://doi.org/10.1177/1049732305276687

Huang, K. T. (2020). Exergaming Executive Functions: An Immersive Virtual Reality-Based Cognitive Training for Adults Aged 50 and Older. *Cyberpsychology, Behavior, and Social Networking*, *23*(3), 143–149. https://doi.org/10.1089/cyber.2019.0269

IJsselsteijn, W. A., Kort, Y. A. W. de, & Poels, K. (2013). *The Game Experience Questionnaire*. Technische Universiteit Eindhoven. https://research.tue.nl/en/publications/the-game-experience-questionnaire

Ismail, N. A., Hashim, H. A., & Yusof, H. A. (2022). Physical activity and exergames among older adults: a scoping review. *Liebertpub.ComNA Ismail, HA Hashim, H Ahmad YusofGames for Health Journal, 2022•liebertpub.Com*, *11*(1), 1–17. https://doi.org/10.1089/G4H.2021.0104

Kaisar, E., Ding, R. B., Han, T., & Qiu, S. (2021). NEONEO Balance Ball: Designing an Intergenerational Interaction Exergame for In-home Balance Training. *International Conference on Human-Computer Interaction*, 78–89. https://doi.org/10.1007/978-3-030-78111-8_5

Kaisar, E., Qiu, S., Yuan, R., & Han, T. (2022). Designing Social Exergame to Enhance Intergenerational Interaction and Exercise. *International Conference on Human-Computer Interaction*, 530–541. https://doi.org/10.1007/978-3-031-05637-6_34

Kappen, D. L., Mirza-Babaei, P., & Nacke, L. E. (2019). Older adults' physical activity and exergames: a systematic review. *Taylor & FrancisDL Kappen, P Mirza-Babaei, LE NackeInternational Journal of Human–Computer Interaction, 2019•Taylor & Francis*, *35*(2), 140–167. https://doi.org/10.1080/10447318.2018.1441253

Krueger, K. R., Wilson, R. S., Kamenetsky, J. M., Barnes, L. L., Bienias, J. L., & Bennett, D. A. (2009). Social Engagement and Cognitive Function in Old Age. *Experimental Aging Research*, *35*(1), 45–60. https://doi.org/10.1080/03610730802545028

Levy, S. R., & Macdonald, J. L. (2016). Progress on Understanding Ageism. *Journal of Social Issues*, *72*(1), 5–25. https://doi.org/10.1111/josi.12153

Lewis, J. R. (2018). The System Usability Scale: Past, Present, and Future. *International Journal of Human–Computer Interaction*, *34*(7), 577–590. https://doi.org/10.1080/10447318.2018.1455307

Li, J., Erdt, M., Chen, L., Cao, Y., Lee, S. Q., & Theng, Y. L. (2018). The social effects of exergames on older adults: Systematic review and metric analysis. *Journal of Medical Internet Research*, *20*(6), 1–12. https://doi.org/10.2196/10486

Lövdén, M., Bäckman, L., Lindenberger, U., Schaefer, S., & Schmiedek, F. (2010). A theoretical framework for the study of adult cognitive plasticity. *Psychological Bulletin*, *136*(4), 659–676. https://doi.org/10.1037/a0020080

Manser, P., & de Bruin, E. D. (2024). "Brain-IT": Exergame training with biofeedback breathing in neurocognitive disorders. *Alzheimer's & Dementia*, *20*(7), 4747–4764.

Manser, P., de Bruin, E. D., Temprado, J.-J., Bherer, L., & Herold, F. (2025). Beyond "just" fun: The role of exergames in advancing health promotion and disease prevention. *Neuroscience & Biobehavioral Reviews*, 106260.

Manser, P., Herold, F., & de Bruin, E. D. (2024). Components of effective exergame-based training to improve cognitive functioning in middle-aged to older adults – A systematic review and meta-analysis. In *Ageing Research Reviews* (Vol. 99). Elsevier Ireland Ltd. https://doi.org/10.1016/j.arr.2024.102385

Mary Lee Hummert. (2010). *Age Group Identity, Age Stereotypes, and Communication in a Life Span Context*. www.peterlang.com

Molden, J., & Maxfield, M. (2017). The impact of aging stereotypes on dementia worry. *European Journal of Ageing*, *14*(1), 29–37. https://doi.org/10.1007/s10433-016-0378-z

Moret, B., Nucci, M., & Campana, G. (2022). Effects of exergames on mood and cognition in healthy older adults: A randomized pilot study. *Frontiers in Psychology*, *13*. https://doi.org/10.3389/fpsyg.2022.1018601

Mueller, F., Gibbs, M. R., Vetere, F., & Edge, D. (2017). Designing for bodily interplay in social exertion games. *ACM Transactions on Computer-Human Interaction*, *24*(3). https://doi.org/10.1145/3064938

Oh, Y., & Yang, S. (2010). Defining exergames & exergaming. *Proceedings of Meaningful Play*, *2010*, 21–23.

Palombi, T., Chirico, A., Cazzolli, B., Zacchilli, M., Alessandri, G., Filosa, L., Borghi, A., Fini, C., Antoniucci, C., Pistella, J., Alivernini, F., Baiocco, R., & Lucidi, F. (2025). Motivation, psychological needs and physical activity in older adults: a qualitative review. In *Age and Ageing* (Vol. 54, Issue 7). Oxford University Press. https://doi.org/10.1093/ageing/afaf180

Petrovčič, A., Slavec, A., & Dolničar, V. (2018). The Ten Shades of Silver: Segmentation of Older Adults in the Mobile Phone Market. *International Journal of Human–Computer Interaction*, *34*(9), 845–860. https://doi.org/10.1080/10447318.2017.1399328

Pettigrew, T. F. (1998). Intergroup contact theory. *Annualreviews.Org*, *49*, 65–85. https://doi.org/10.1146/annurev.psych.49.1.65

Pirovano, M., Surer, E., Mainetti, R., Lanzi, P. L., & Borghese, N. A. (2016). Exergaming and rehabilitation: A methodology for the design of effective and safe therapeutic exergames. *Entertainment Computing*, *14*, 55–65.

Qiu, S. (2019). *Social glasses: Designing gaze behaviors for visually impaired people*. Technische Universiteit Eindhoven.

Qiu, S., An, P., Hu, J., Han, T., & Rauterberg, M. (2020). Understanding visually impaired people's experiences of social signal perception in face-to-face communication. *Universal Access in the Information Society*, *19*(4), 873–890. https://doi.org/10.1007/s10209-019-00698-3

Qiu, S., An, P., Kang, K., Hu, J., Han, T., & Rauterberg, M. (2022). A review of data gathering methods for evaluating socially assistive systems. *Sensors*, *22*(1), 82. https://doi.org/10.3390/s22010082

Qiu, S., An, P., Kang, K., Hu, J., Han, T., & Rauterberg, M. (2023). Investigating socially assistive systems from system design and evaluation: a systematic review. *Universal Access in the Information Society*, *22*(2), 609–633. https://doi.org/10.1007/s10209-021-00852-w

Qiu, S., Hu, J., Han, T., Osawa, H., & Rauterberg, M. (2020a). An evaluation of a wearable assistive device for augmenting social interactions. *IEEE Access*, *8*, 164661–164677. https://doi.org/10.1109/ACCESS.2020.3022425

Qiu, S., Hu, J., Han, T., Osawa, H., & Rauterberg, M. (2020b). Social Glasses: Simulating Interactive Gaze for Visually Impaired People in Face-to-Face Communication. *International Journal of Human-Computer Interaction*, *36*(9), 839–855. https://doi.org/10.1080/10447318.2019.1696513

Qiu, S., Hu, J., Han, T., & Rauterberg, M. (2023). Can blindfolded users replace blind ones in product testing? an empirical study. *Behaviour and Information Technology*, 1–19. https://doi.org/10.1080/0144929X.2023.2226768

Qiu, S., Kaisar, E., Ding, R., Han, T., Hu, J., & Rauterberg, M. (2023). Social Balance Ball: Designing and Evaluating an Exergame That Promotes Social Interaction between Older and Younger Players. *International Journal of Human-Computer Interaction*, 1–24. https://doi.org/10.1080/10447318.2023.2175157

Qiu, S., Yang, X., Kang, K., Zhang, S., Han, T., Hu, J., & Rauterberg, M. (2025). Exploring the design and evaluation of social exergames for older adults: a systematic review. *Universal Access in the Information Society*. https://doi.org/10.1007/S10209-025-01237-Z

Ren, Y., Savadlou, A., Park, S., Siska, P., Epp, J. R., & Sargin, D. (2023). The impact of loneliness and social isolation on the development of cognitive decline and Alzheimer's Disease. *Frontiers in Neuroendocrinology*, *69*, 101061. https://doi.org/https://doi.org/10.1016/j.yfrne.2023.101061

Rieker, J. A., Reales, J. M., Muiños, M., & Ballesteros, S. (2022). The Effects of Combined Cognitive-Physical Interventions on Cognitive Functioning in Healthy Older Adults: A Systematic Review and Multilevel Meta-Analysis. In *Frontiers in Human Neuroscience* (Vol. 16). Frontiers Media S.A. https://doi.org/10.3389/fnhum.2022.838968

Robertson, M. C., Campbell, A. J., Gardner, M. M., & Devlin, N. (2002). Preventing Injuries in Older People by Preventing Falls: A Meta-Analysis of Individual-Level Data. *Journal of the American Geriatrics Society*, *50*(5), 905–911. https://doi.org/10.1046/j.1532-5415.2002.50218.x

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, *55*(1), 68–78. https://doi.org/10.1037/0003-066X.55.1.68

Rytterström, P., Strömberg, A., Jaarsma, T., & Klompstra, L. (2024). Exergaming to increase physical activity in older adults: feasibility and practical implications. *SpringerP Rytterström, A Strömberg, T Jaarsma, L KlompstraCurrent Heart Failure Reports, 2024•Springer*, *21*(4), 439–459. https://doi.org/10.1007/S11897-024-00675-9

Salthouse, T. A. (2010). Selective review of cognitive aging. *Journal of the International Neuropsychological Society*, *16*(5), 754–760. https://doi.org/10.1017/S1355617710000706

Sarkisian, N., & Gerstel, N. (2012). *Nuclear Family Values, Extended Family Lives*. Routledge. https://doi.org/10.4324/9780203141977

Šlosar, L., Voelcker-Rehage, C., Paravlić, A. H., Abazovic, E., de Bruin, E. D., & Marusic, U. (2022). Combining physical and virtual worlds for motor-cognitive training interventions: position paper with guidelines on technology classification in movement-related research. *Frontiers in Psychology*, *13*, 1009052.

Smith, S. T., Sherrington, C., Studenski, S., Schoene, D., & Lord, S. R. (2011). A novel Dance Dance Revolution (DDR) system for in-home training of stepping ability: basic parameters of system use by older adults. *British Journal of Sports Medicine*, *45*(5), 441–445.

Stephan, A. T., McVey, A., Steele, T., Cooper, K., Phillips, C. B., & Ross, L. A. (2024). Older Adults' Perceptions of Cognitive, Physical, and Exergame Training: A Mixed Methods Investigation of a Four-Armed Randomized Controlled Pilot Trial. *Games for Health Journal*, *13*(6), 459–465. https://doi.org/10.1089/G4H.2023.0208

Thüring, M., & Mahlke, S. (2007). Usability, aesthetics and emotions in human–technology interaction. *International Journal of Psychology*, *42*(4), 253–264. https://doi.org/10.1080/00207590701396674

Tovin, M. M., & Wormley, M. E. (2023). Systematic development of standards for mixed methods reporting in rehabilitation health sciences research. *Physical Therapy*,

103(11), pzad084.

van Santen, J., Dröes, R.-M., Twisk, J. W. R., Blanson Henkemans, O. A., van Straten, A., & Meiland, F. J. M. (2020). Effects of Exergaming on Cognitive and Social Functioning of People with Dementia: A Randomized Controlled Trial. *Journal of the American Medical Directors Association*, *21*(12), 1958-1967.e5. https://doi.org/10.1016/j.jamda.2020.04.018

Voida, A., & Greenberg, S. (2012). Console gaming across generations: exploring intergenerational interactions in collocated console gaming. *Universal Access in the Information Society*, *11*(1), 45–56. https://doi.org/10.1007/s10209-011-0232-1

Williams, A., & Giles, H. (1996). Intergenerational Conversations Young Adults' Retrospective Accounts. *Human Communication Research*, *23*(2), 220–250. https://doi.org/10.1111/j.1468-2958.1996.tb00393.x

Wohn, D. Y., Ratan, R., & Cherchiglia, L. (2020). *Gender and Genre Differences in Multiplayer Gaming Motivations* (pp. 233–248). https://doi.org/10.1007/978-3-030-50164-8_16

World Health Organization. (2022, October 1). *Ageing and health*. https://www.who.int/news-room/fact-sheets/detail/ageing-and-health#:~:text=By%202030%2C%201%20in%206,will%20double%20(2.1%20billion).

World Health Organization. (2023). *Reducing social isolation and loneliness among older people*. https://www.who.int/activities/reducing-social-isolation-and-loneliness-among-older-people

Xu, X., Theng, Y. L., Li, J., & Phat, P. T. (2016). Investigating effects of exergames on exercise intentions among young-old and old-old. *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems*, 2961–2968. https://doi.org/10.1145/2851581.2892296

Yuan, R., Zhang, S., Yang, X., Qiu, S., Han, T., & Hu, J. (2023). A Review of Prototyping and Evaluating Approaches for Intergenerational Social Exergames Targeting Older Adults. *ACM International Conference Proceeding Series*, 146–157. https://doi.org/10.1145/3629606.3629621

Zhang, H., Shen, Z., Liu, S., Yuan, D., & Miao, C. (2021). Ping Pong: An Exergame for Cognitive Inhibition Training. *International Journal of Human–Computer Interaction*, *37*(12), 1104–1115. https://doi.org/10.1080/10447318.2020.1870826

Appendix

Table A1 Items of game experience, intrinsic motivation, perceived workload, system usability and logs in the SBBG, DBBG, and GBBG.

	Experimental conditions mean (Std. Deviation)		
	SBBG	**DBBG**	**GBBG**
Game Experience			
-Com	2.71(0.94)	3.04(0.63)	3.06(0.63)
-SII	2.61(1.00)	2.63(1.00)	2.58(1.00)
-Flo	2.64(0.99)	2.77(0.89)	2.43(1.02)
-Ten	0.41(0.81)	0.28(0.61)	0.29(0.56)
-Cha	1.52(0.98)	1.28(1.01)	1.27(0.94)
-Neg	0.60(0.69)	0.51(0.85)	0.63(0.78)
-Pos	3.00(0.93)	3.12(0.88)	2.91(1.07)
Intrinsic Motivation			
-Int	5.40(1.03)	5.41(1.26)	5.30(1.34)
-Eff	4.97(0.80)	4.78(0.77)	4.79(0.82)
-Val	5.75(0.93)	5.75(0.89)	5.57(1.07)
Perceived Workload			
-PWL	11.36(3.44)	11.07(2.42)	10.42(2.69)
System Usability			
-SU	72.08(12.6)	73.33(12.38)	72.71(12.15)
System Logs			
-Tot	134.67(46.8)	155.33(39.47)	148.28(34.33)
-ARNQ	47.89(33.14)	53.67(22.63)	56.5(22.55)

Abbreviations: Com: competence; SII: sensory and imaginative immersion; Flo: flow; Ten: tension; Cha: Challenge; Neg: negative affect; Pos: positive affect; Int: interest; Eff: effort; Val: value; PWL: perceived workload; SU: system usability; Tot: total score; ARNQ: accuracy rate for numerical questions.

Table A2 Items of game experience, intrinsic motivation, perceived workload, system usability and logs in male and female participants in three game modes.

	Experimental conditions mean (Std. Deviation)					
	SBBG		DBBG		GBBG	
	Male	Female	Male	Female	Male	Female
Game Experience						
-Com	2.77(0.86)	2.66(1.04)	3.02(0.49)	3.07(0.76)	2.84(0.70)	3.28(0.48)
-SII	2.34(1.09)	2.88(0.86)	2.41(1.12)	2.85(0.84)	2.26(1.04)	2.91(0.86)
-Flo	2.43(1.06)	2.86(0.89)	2.41(0.97)	3.13(0.64)	2.06(1.17)	2.81(0.68)
-Ten	0.46(0.93)	0.35(0.70)	0.48(0.81)	0.07(0.18)	0.50(0.72)	0.07(0.18)
-Cha	1.34(0.84)	1.70(1.10)	1.44(1.16)	1.12(0.85)	1.41(1.07)	1.13(0.79)
-Neg	0.57(0.65)	0.63(0.75)	0.79(1.10)	0.22(0.32)	1.01(0.89)	0.25(0.39)
-Pos	2.81(0.99)	3.19(0.85)	2.91(1.13)	3.32(0.50)	2.43(1.23)	3.39(0.59)
Intrinsic Motivation						
-Int	5.29(1.21)	5.50(0.84)	4.95(1.54)	5.87(0.66)	4.77(1.61)	5.83(0.72)
-Eff	4.71(0.95)	5.23(0.51)	4.56(0.93)	5.01(0.51)	4.68(0.88)	4.91(0.77)
-Val	5.52(0.95)	5.98(0.88)	5.57(0.96)	5.94(0.79)	5.13(1.11)	6.01(0.86)
Perceived Workload						
-PWL	10.77(3.20)	11.95(3.67)	11.18(2.93)	10.96(1.85)	10.02(3.16)	10.83(2.13)
System Usability						
-SU	69.31(11.94)	74.86(12.96)	68.19(11.30)	78.47(11.48)	68.33(11.88)	77.08(11.06)
System Logs						
-Tot	130.39(50.77)	138.94(43.51)	140.56(46.48)	170.11(24.16)	136.56(36.12)	160(28.81)
-ARNQ	43.11(35.52)	52.67(30.85)	48.67(22.91)	58.67(21.82)	49.56(19.48)	63.44(23.77)

Abbreviations: Com: competence; SII: sensory and imaginative immersion; Flo: flow; Ten: tension; Cha: Challenge; Neg: negative affect; Pos: positive affect; Int: interest; Eff: effort; Val: value; PWL: perceived workload; SU: system usability; Tot: total score; ARNQ: accuracy rate for numerical questions.

Table A3 Items of game experience, intrinsic motivation, perceived workload, system usability and logs in older and younger participants in three game modes.

	Experimental conditions mean (Std. Deviation)					
	SBBG		DBBG		GBBG	
	Older	Younger	Older	Younger	Older	Younger
Game Experience						
-Com	2.92(0.68)	2.50(1.13)	3.08(0.76)	3.01(0.50)	3.21(0.62)	2.91(0.63)
-SII	3.27(0.56)	1.95(0.92)	3.22(0.72)	2.04(0.90)	3.30(0.51)	1.87(0.84)
-Flo	3.08(0.91)	2.21(0.90)	3.30(0.60)	2.24(0.82)	2.94(0.89)	1.92(0.89)
-Ten	0.33(0.64)	0.48(0.97)	0.35(0.80)	0.20(0.35)	0.20(0.38)	0.37(0.69)
-Cha	2.00(0.88)	1.04(0.86)	1.90(0.95)	0.67(0.63)	1.79(0.93)	0.76(0.62)
-Neg	0.58(0.86)	0.61(0.49)	0.61(1.06)	0.40(0.56)	0.56(0.87)	0.71(0.70)
-Pos	3.56(0.67)	2.44(0.82)	3.60(0.50)	2.63(0.93)	3.40(0.84)	2.42(1.06)
Intrinsic Motivation						
-Int	5.86(0.62)	4.94(1.17)	5.81(1.07)	5.02(1.34)	5.60(1.24)	5.00(1.40)
-Eff	5.31(0.60)	4.63(0.84)	4.98(0.73)	4.59(0.79)	4.84(0.90)	4.74(0.76)
-Val	6.21(0.84)	5.28(0.79)	6.22(0.71)	5.29(0.81)	5.85(1.16)	5.29(0.92)
Perceived Workload						
-PWL	11.48(1.91)	11.24(4.55)	11.95(2.04)	10.18(2.50)	11.57(2.49)	9.28(2.43)
System Usability						
-SU	70.42(12.28)	73.75(13.04)	69.44(13.87)	77.22(9.54)	70.28(13.47)	75.14(10.48)
System Logs						
-Tot	133.06(45.27)	136.28(49.54)	155.33(40.04)	155.33(40.04)	148.28(34.83)	148.28(34.83)
-ARNQ	45.11(29.86)	50.67(36.79)	53.67(22.96)	53.67(22.96)	56.5(22.88)	56.5(22.88)

Abbreviations: Com: competence; SII: sensory and imaginative immersion; Flo: flow; Ten: tension; Cha: Challenge; Neg: negative affect; Pos: positive affect; Int: interest; Eff: effort; Val: value; PWL: perceived workload; SU: system usability; Tot: total score; ARNQ: accuracy rate for numerical questions.

Table A4 Results of main and interaction effects of game experience, intrinsic motivation, perceived workload, system usability, and system logs for the exergame modes.

Source		*df*	*SS*	*MS*	*F*	*p*	η_p^2
Exergame Mode	Com	1.42	2.81	1.98	5.37	**0.02***	0.14
	SII	2	0.04	0.02	0.18	0.84	0.01
	Flo	1.96	2.11	1.08	3.44	**0.04***	0.1
	Ten	2	0.38	0.19	0.7	0.5	0.02
	Cha	2	1.44	0.72	2.53	0.09	0.07
	Neg	2	0.30	0.15	0.65	0.52	0.02
	Pos	1.87	0.77	0.41	1.69	0.2	0.05
	Int	1.93	0.26	0.13	0.34	0.7	0.01
	Eff	2	0.81	0.40	1.5	0.23	0.04
	Val	1.81	0.80	0.44	2.09	0.14	0.06
	PWL	1.91	16.50	8.62	2.72	0.08	0.08
	SU	2	28.13	14.06	0.29	0.75	0.01
	Tot	2	7945.85	3972.93	2.94	0.06	0.08
	ARNQ	1.38	1386.74	1002.24	1.4	0.25	0.04
Exergame Mode * Gender	Com	1.42	1.42	1	2.71	0.09	0.08
	SII	2	0.19	0.09	0.85	0.43	0.03
	Flo	1.96	0.61	0.31	0.99	0.38	0.03
	Ten	2	0.56	0.28	1.04	0.36	0.03
	Cha	2	2.59	1.29	4.55	**0.01****	0.12
	Neg	2	3.30	1.65	7.18	**0.00****	0.18
	Pos	1.87	1.89	1.01	4.18	**0.02***	0.12
	Int	1.93	3.80	1.97	5	**0.01****	0.14
	Eff	2	0.41	0.21	0.76	0.47	0.02
	Val	1.81	1.36	0.75	3.54	**0.04***	0.1
	PWL	1.91	9.46	4.94	1.56	0.22	0.05
	SU	2	104.51	52.26	1.06	0.35	0.03
	Tot	2	2100.07	1050.04	0.78	0.46	0.02
	ARNQ	1.38	102.30	73.93	0.1	0.83	0
Exergame Mode * Generation	Com	1.42	0.59	0.42	1.12	0.32	0.03
	SII	2	0.26	0.13	1.19	0.31	0.04
	Flo	1.96	0.18	0.09	0.3	0.74	0.01
	Ten	2	0.56	0.28	1.04	0.36	0.03
	Cha	2	0.37	0.18	0.65	0.53	0.02
	Neg	2	0.61	0.30	1.32	0.27	0.04
	Pos	1.87	0.12	0.06	0.26	0.76	0.01
	Int	1.93	0.46	0.24	0.61	0.54	0.02
	Eff	2	1.50	0.75	2.78	0.07	0.08
	Val	1.81	0.83	0.46	2.18	0.13	0.06
	PWL	1.91	20.29	10.60	3.35	**0.04***	0.09
	SU	2	91.78	45.89	0.93	0.4	0.03
	Tot	2	62.30	31.15	0.02	0.98	0
	ARNQ	1.38	185.19	133.84	0.19	0.75	0.01
	Com	45.28	16.72	0.37			
	SII	64	7.03	0.11			

Error (Exergame Mode)	Flo	62.77	19.63	0.31
	Ten	64	17.33	0.27
	Cha	64	18.18	0.28
	Neg	64	14.70	0.23
	Pos	59.73	14.49	0.24
	Int	61.77	24.30	0.39
	Eff	64	17.27	0.27
	Val	58.04	12.27	0.21
	PWL	61.27	193.83	3.16
	SU	64	3143.98	49.12
	Tot	64	86527.04	1351.98
	ARNQ	44.28	31677.93	715.46

Abbreviations: Com: competence; SII: sensory and imaginative immersion; Flo: flow; Ten: tension; Cha: Challenge; Neg: negative affect; Pos: positive affect; Int: interest; Eff: effort; Val: value; PWL: perceived workload; SU: system usability; Tot: total score; ARNQ: accuracy rate for numerical questions. Significant group difference; *$p \leq 0.05$, **$p \leq 0.01$.

Table A5 Results of main and interaction effects of game experience, intrinsic motivation, perceived workload, system usability, and system logs for gender and generation.

Source		*df*	*SS*	*MS*	*F*	*p*	η_p^2
Gender	Com	1	0.13	0.13	0.37	0.55	0.01
	SII	1	2.66	2.66	6.16	**0.02***	0.16
	Flo	1	3.61	3.61	8.6	**0.01****	0.21
	Ten	1	0.89	0.89	3.35	0.08	0.09
	Cha	1	0.06	0.06	0.12	0.73	0
	Neg	1	1.63	1.63	4.01	**0.05***	0.11
	Pos	1	3.04	3.04	7.04	**0.01****	0.18
	Int	1	4.80	4.8	4.78	**0.04***	0.13
	Eff	1	1.47	1.47	3.7	0.06	0.1
	Val	1	2.91	2.91	4.93	**0.03***	0.13
	PWL	1	3.15	3.15	0.53	0.47	0.02
	SU	1	604.34	604.34	5.95	**0.02***	0.16
	Tot	1	3789.09	3789.09	5.39	**0.03***	0.14
	ARNQ	1	1118.53	1118.53	2.81	0.1	0.08
Generation	Com	1	0.62	0.62	1.69	0.2	0.05
	SII	1	15.41	15.41	35.73	**<0.001****	0.53
	Flo	1	8.67	8.67	20.65	**<0.001****	0.39
	Ten	1	0.03	0.03	0.1	0.75	0
	Cha	1	10.38	10.38	21.46	**<0.001****	0.4
	Neg	1	0	0	0	0.97	0
	Pos	1	9.34	9.34	21.59	**<0.001****	0.4
	Int	1	5.37	5.37	5.35	**0.03***	0.14
	Eff	1	1.36	1.36	3.43	0.07	0.1
	Val	1	5.94	5.94	10.05	**<0.001****	0.24
	PWL	1	18.4	18.4	3.08	0.09	0.09
	SU	1	255.11	255.11	2.51	0.12	0.07
	Tot	1	10.38	10.38	0.01	0.9	0
	ARNQ	1	30.86	30.86	0.08	0.78	0
Gender * Generation	Com	1	0.07	0.07	0.18	0.68	0.01
	SII	1	0.63	0.63	1.47	0.23	0.04
	Flo	1	0.25	0.25	0.6	0.45	0.02
	Ten	1	0.1	0.1	0.37	0.55	0.01
	Cha	1	0.47	0.47	0.98	0.33	0.03
	Neg	1	0.2	0.2	0.49	0.49	0.01
	Pos	1	0.21	0.21	0.48	0.49	0.01
	Int	1	0.03	0.03	0.03	0.88	0
	Eff	1	0.02	0.02	0.05	0.82	0
	Val	1	0.03	0.03	0.06	0.81	0
	PWL	1	2.82	2.82	0.47	0.5	0.01
	SU	1	126.56	126.56	1.25	0.27	0.04
	Tot	1	534.12	534.12	0.76	0.39	0.02
	ARNQ	1	57.09	57.09	0.14	0.71	0
Error	Com	32	11.79	0.37			
	SII	32	13.8	0.43			

Flo	32	13.44	0.42
Ten	32	8.52	0.27
Cha	32	15.49	0.48
Neg	32	13.03	0.41
Pos	32	13.84	0.43
Int	32	32.15	1
Eff	32	12.69	0.4
Val	32	18.9	0.59
PWL	32	190.95	5.97
SU	32	3250.62	101.58
Tot	32	22504.77	703.27
ARNQ	32	12745.06	398.28

Abbreviations: Com: competence; SII: sensory and imaginative immersion; Flo: flow; Ten: tension; Cha: Challenge; Neg: negative affect; Pos: positive affect; Int: interest; Eff: effort; Val: value; PWL: perceived workload; SU: system usability; Tot: total score; ARNQ: accuracy rate for numerical questions. Significant group difference; *$p \leq 0.05$, **$p \leq 0.01$.