Fitness Applications for Healthy Older Adults Using Large Projection Displays
Methodology, Design, Assessment and Field Validation

DOCTORAL THESIS

Afonso Rodrigues Gonçalves

DOCTORATE IN INFORMATICS ENGINEERING
SPECIALTY IN HUMAN-COMPUTER INTERACTION

UNIVERSIDADE da MADEIRA
A Nossa Universidade
www.uma.pt

June | 2021

Fitness Applications for Healthy Older Adults Using Large Projection Displays

Methodology, design, assessment, and field validation

DOCTORAL THESIS

Afonso Rodrigues Gonçalves
DOCTORATE IN INFORMATICS ENGINEERING
SPECIALTY IN HUMAN-COMPUTER INTERACTION

ORIENTATION
Sergi Bermúdez i Badia

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

Due to low birth rates and rising life expectancy, the population of developed countries is aging. Concurrently, physical inactivity is an identified major health risk and significantly more prevalent in older adults, who experience the consequences related to inactivity more frequently. Exergames for the elderly are an affordable option to prevent sedentarism and complement traditional exercise training, which can otherwise suffer from low adherence and personalization. They facilitate moderate-intensity physical activity levels and positively impact fitness, health, balance, postural control, mobility, and motivation. However, due to the lack of knowledge of seniors' game preferences and technology literacy, there are challenges in designing exergames that match the users' needs and motivators with game elements. While there is an extensive body of research in this field, there are critical gaps: most of the research is done in laboratory environments, is focused on balance and ignore other motor performance domains, use commercial games which are not designed for older adults, and fail to explore the longitudinal effects of exergames.

In this thesis, there are three sequential contributions:

1) Develop a technology to facilitate exergaming in the elderly population by integrating easy-to-use full-body interaction with large projection displays. Resulting in software for low-cost virtual reality surround-screen projection systems, validated through user studies and compared with the conventional alternatives.
2) Leverage the technology and design customized exergames to promote fitness in older adults by:
   a) Evaluate the capability to automate fitness assessment using gesture detectors by testing their performance in the field with 22 elderly end-users and compare it to traditional methods administered by an expert. Resulting in a high accuracy system, consistent with the traditional fitness assessment method.
   b) Study older adults' interaction preferences with floor projection displays by developing and testing two natural user interfaces with 19 elderly participants. The participants' preference for a feet-controlled interface was identified when usability, perceived workload, and performance indicators were assessed.
   c) Apply human-centered design methodologies in the gamification of fitness training routines by focusing on insights from inquiries to improve game elements and game iterations based on playtesting sessions to produce exergames. Resulting in a set of four exergames created to train the critical functional fitness areas of older adults.
3) Measure the benefits in older adults' motor performance, quality of life, and physical activity levels during a longitudinal multidimensional training combining custom-made exergames and traditional exercise in a complementary fashion. Achieved through a 12-week long randomized controlled trial of bi-weekly exercise sessions with 31 elderly participants. Outcome measures on fitness, balance, and health-related quality of life were measured at the start, during, and after the intervention, and physical activity levels were measured at each session. This resulted in exergame players having a significant increase in strength compared to control, and both conditions improving balance and the mental component of health-related quality of life, with improvements in the latter being greater for exergame players. Additionally, during exergames' sessions, participants spent less energy but maintained the recommended physical activity levels for more extended periods.

Our results show that integrating personalized exergames designed for multidimensional fitness training in traditional settings can effectively enhance older adults' motor performance and mental well-being. This technology is a viable low-cost option to be deployed in the context of elderly fitness programs.

RESUMO

Devido às baixas taxas de natalidade e ao aumento da esperança média de vida a população dos países desenvolvidos está a tornar-se envelhecida. Ao mesmo tempo, a falta de atividade física está identificada como um importante fator de risco para a saúde, com grande prevalência em idosos, que também experienciam as suas consequências com maior frequência. Os exergames para idosos são uma opção económica para a prevenção de sedentarismo e complemento do treino físico tradicional, que pode sofrer de baixa aderência e personalização. Estes jogos promovem níveis moderados de intensidade da atividade física e benefícios ao nível do fitness, saúde, equilíbrio, postura, mobilidade e motivação. No entanto, devido ao desconhecimento sobre as preferências de jogos e literacia tecnológica dos idosos, há desafios no desenho de exergames que se adequem às suas necessidades e motivações. Apesar de haver extensa investigação nesta área, existem lacunas críticas: a maior parte da investigação é feita em ambiente laboratorial, foca-se no equilíbrio e ignora os outros domínios motores, usa jogos comerciais que não foram desenhados para esta população e não explora os efeitos longitudinais dos exergames.

Nesta tese, fazemos três contribuições, apresentadas de forma sequencial:

1) Desenvolvimento de tecnologia para a prática de exergaming pela população idosa, integrando interação corporal fácil de usar com projeções de grandes dimensões. Resultando num software para uso em sistemas de realidade virtual através de projeção de baixo custo, validado através de estudos com utilizadores e comparado com alternativas convencionais.
2) Uso da tecnologia e design de exergames feitos à medida para a promoção do fitness em idosos:
    a) Apreciação da capacidade de automação da avaliação do fitness através do uso de detetores de gestos, testando o seu desempenho no terreno com 22 utilizadores idosos. Resultando num sistema de elevada exatidão, consistente com os métodos tradicionais.
    b) Estudo das preferências de idosos na interação com projeções no solo, através do desenvolvimento e teste de dois interfaces naturais com 19 participantes idosos. Identificada uma preferência pelo interface controlado pelos pés através da avaliação de usabilidade, carga de trabalho sentida e indicadores de desempenho.
    c) Aplicação de metodologias de desenho centrado em humanos à gamificação de rotinas de treino físico através do foco na compreensão de inquéritos, baseados em sessões de jogo, aplicados para a melhoria e iteração dos mesmos. Resultando num conjunto de quatro exergames criados para treinar áreas críticas de fitness funcional em idosos.
3) Medição do desempenho motor, qualidade de vida e intensidade da atividade física de idosos durante treino longitudinal multidimensional combinando exergames feitos à medida com exercício tradicional de forma complementar. Alcançada através de um estudo randomizado controlado de 12 semanas, com sessões de exercício físico bissemanais por parte de 31 participantes idosos. Com medições de fitness, equilíbrio e qualidade de vida relacionada com a saúde medidas no início, durante e após a intervenção, e níveis de atividade física medidos em todas as sessões. Resultando num aumento significativo de força nos jogadores de exergames relativamente ao controlo e numa melhoria em ambos os grupos no equilíbrio e na componente mental da qualidade de vida, com uma melhoria maior por parte dos jogadores nesta última. Adicionalmente, registado um menor dispêndio energético, mas uma manutenção mais prolongada da intensidade da atividade física recomendada nas sessões de exergames.

Os nossos resultados mostram que a integração de exergames personalizados, desenhados para o treino multidimensional de fitness, em ambientes tradicionais de treino podem efetivamente melhorar os ganhos de desempenho motor e bem-estar mental. Assim, esta tecnologia é uma opção económica viável para ser usada no contexto de programas de fitness para idosos.

INTRODUCTION

Most developed countries are undergoing a demographic shift towards a more aged population due to low birth rates and rising life expectancy [1]. In 2018, the Portuguese population aged over 60 was 28.2% [2], and according to projections, nearly one-third of European citizens will be 65 or over by 2060 [3]. In addition to this demographic change, physical inactivity is an identified health risk negatively associated with several health outcomes. While physical activity (PA) is recognized as a global public health priority, more than 30% of the world population is not meeting the PA's minimum recommended levels. Sedentarism is the 4th main risk factor in worldwide mortality, associated with 6% of deaths [4], and 6-10% of all deaths from non-communicable diseases may be attributed to it [5]. Additionally, it is known that sedentary behaviors are more prevalent in older adults than any other age group, with 65-80% of their awake time spent sitting [6]. The combination of aging with sedentary behaviors is a growing concern and puts a high strain on modern societies and their health systems.

Evidence shows that regular PA produces significant and extensive health benefits, of particular importance in older adults, as they experience more frequently the outcomes related to inactivity [4]. The American College of Sports and Medicine has made recommendations and guidelines for physical activity in old age [7]. These guidelines target multidimensional training in aerobic fitness, musculoskeletal function, flexibility, and balance, aiming to maximize physical activity benefits in older adults. Specifically, regarding exercise intensity for active older adults, the ACSM states that exercising at moderate-to-vigorous PA (MVPA) intensities can produce greater benefits than less intense training [7], [8] and recommends 150 minutes per week of moderate-intensity exercise.

Information and communications technologies (ICT) offer us an affordable option for preventing sedentarism through exercise games (exergames). Exergames are videogames that require physical exercise to be played, tying PA to game performance. They aim to encourage players to actively move in order to achieve in-game success, providing a genuinely fun strategy to promote PA in older adults [9]. Longitudinal and qualitative studies have been carried out with older adults revealing important usability patterns such as engagement, flow, adherence, enjoyment, and motivation, demonstrating the effectiveness of exergames to persuade players to keep exercising for long periods [10], [11]. Beyond the physical activity that exergames elicit on its users, another crucial aspect is the game design itself. A review of past efforts has concluded that the most effective serious games are the ones that are customized to both the target population and behavior that they want to promote/change, benefiting from a focus on game theories and behavioral prediction theory [12]. In terms of motivation, mastery over the game is an essential factor in players' preferences [13], and in the elderly, a clear preference for gesture-based controllers, such as Kinect, was identified [14]. The relation between skill and challenge as a motivating factor has been proposed [15] and extended to exergames by adding the fitness/intensity dimension [16]. Other identified factors that improve the engagement in games are the addition of music, guidance, lack of negative consequences for underperformance, multiplayer cooperation, and adaptive challenge difficulty [17] as well as clear goal-setting and progression levels [18]. Besides these specific factors, exergames should not be seen as an extraordinary activity but instead offered as an accessible and enjoyable activity akin to traditional exercise [18] while ensuring they cater to the user needs with appropriate content, interface design, and exercise demands [9].

Given the negative impact of sedentary behaviors in the elderly population, together with the clear potential benefits that exercise through gaming can deliver, this thesis's primary goal is to design a set of custom exergames capable of generating fitness benefits in the elderly. This general goal is achieved in three main steps and sub-goals, which can be further divided by their scientific contributions:

- Develop and evaluate a system for full-body interaction with large projection displays

    - Development of a broad scope software, KAVE, for virtual reality (VR) mediation using image projection and whole-body tracking.
    - Evaluation of the system's tracking reliability, sense of presence, and cybersickness, in the context of immersive VR
- Create exergames for elderly fitness
    - Feasibility assessment of using a commercial body tracking solution in fitness evaluation
    - Understand interaction preferences of elderly users with the proposed system on floor projections
    - Design of exergames customized to the Portuguese elderly population
- Evaluate the effects of the custom exergames in the elderly during prolonged training
    - Evaluation of the effectiveness of the games in eliciting desired exertion levels during training
    - Assessment of fitness benefits in a randomized controlled trial

The thesis is organized in three main sections, each corresponding to a sub-goal, and comprises seven scientific articles. Five papers that have been published and two that are in press. The first section describes the development and evaluation of VR technology using projection mapping, where each subsection corresponds to an individual paper. The second section focuses on the design of exergames for the elderly, supported by the technology, divided into three sections, each corresponding to a different publication. The third section evaluates the effects on the elderly fitness of the exergames described before. Lastly, there is a section for the conclusions of the thesis.

1 Development and Evaluation of Software for Low-Cost Virtual Reality Surround-Screen Projection Systems

This section describes the efforts in developing and evaluating a technology whose initial purpose was to facilitate exergaming in the elderly population by integrating full-body interaction with large projection displays.

While acceptance of new technological systems by the elderly can be challenging, a comprehensive meta-analysis on the relation between age and technology acceptance has shown that it is mainly mediated by the perceived ease-of-use [19]. In it, the authors point to the seemingly social consensus on the negative attitudes and adoption resistance towards new technology by the elderly [20]–[24], for which there are both supportive and contradictory studies [19]. By aggregating and meta-analyzing the results of 144 studies relatively to a technology acceptance model (TAM) [25], they provided a clearer view of the conditions in which age drives technology adoption. Age correlates negatively with perceived ease-of-use, perceived usefulness, and intention to use technology; however, the last two are fully mediated by the first [19]. Furthermore, the role of age was only relevant when the technology did not address the predominant needs of older adults, social and emotional [19]. This led to three implications: When using technology to address the prevalent needs of older adults, their age can be a benefitting factor in accepting the technology, thus lowering the need for age-sensitive interventions in this context. Second, there should be a prioritization of a technology perceived ease-of-use in this population to increase its acceptance. Direct experience, practical exercises, gradual learning, and using peer trainers should be favored. Finally, regardless of the first two, interaction with technology by this population has specific needs that should be addressed through user-centered design and usability considerations, such as large-font sizes, color contrast, visual aids, simple navigation, concise instruction, and avoiding technical terminology [19]. It was therefore our goal to develop an easy-to-use technology that could accommodate the exergame design considerations for enhanced elderly usability.

As in any interactive system, we were bound to the design of both input and output modalities. We faced a choice between three main systems, flat-panel displays, video projectors, or head-mounted displays (HMD) for video output. Flat-panel displays were perhaps the most familiar option. In terms of fundamental display features, such as resolution, brightness, contrast, display latency [26], their rapidly changing range of products can generally offer better performance than video projectors. However, they represent a physical limitation in screen size, which scales together with weight and installation workload, meaning that a large image requires an equally large screen with reduced mobility due to weight, size, and format. For example, Samsung's Q800T 82'' TV features a 7680x4320 pixel resolution (8K UHD), with a 207cm screen diagonal length, for 5000€. However this implies 183cm x 105cm x 2.6cm in size, and weighting 41.5kg [27]. In contrast, the video projector used in Section 3 of this thesis, the Optoma GT760 (500€ in 2017), while featuring a lower resolution, 1280x720 pixel resolution (720p), can provide much larger images, from 81cm to 773cm diagonal (at projector distances of .4m to 3.5m); this, while having only 10cm x 29cm x 24cm in size, and weighting 2.7kg [28]. The major drawback is that, with 3400 Lumens, low ambient light is required for projections bigger than 5m, luminance < 75cd/m^2. Head-mounted displays represented the third option for video output, a wearable technology with significant developments in recent years. Devices such as the Oculus Rift or HTC Vive can provide highly immersive VR experiences by natively combining six-degrees-of-freedom tracking with video output, thus providing a field of regard that completely surrounds the user; furthermore, they are easily portable and have a reduced floor print [29]. While these devices are made to be small and lightweight, they still are encumbering, uncomfortable to wear, and hard to accommodate eyeglasses [29]; with a weight around 500g, they can affect the users' perceived physical load [30]. Additionally, the lag between tracking and video output is a real concern and a leading cause of cybersickness, their field of view is restrictive, and prolonged use can lead to eye strain [29]. Furthermore, this type of HMDs is occlusive, meaning that it isolates the user from their surroundings, including other humans [31].

Since the technology's aim was exergaming for fitness, user input necessarily needed some sort of physical tracking beyond keys, buttons, or computer mice. Common options for interaction in this context are:

- Floor mat or platform game controllers (e.g., Power Pad and Dance Dance Revolution) – Provide a simple 2D board of N buttons to interact with the feet; they deliver N degrees of freedom, each with two discrete values.
- Force plates (e.g., Wii Balance Board) – Track the user's center of pressure over the board, provide 2 degrees of freedom, each of continuous input in a range of values (limited to the size of the board).
- Cameras (e.g., PlayStation Eye Toy, Kinect v1 and v2) – Exist in many distinct formats, from simple color webcams to more advanced RGB-D or stereo cameras, providing simple motion detection or full-body tracking and gesture recognition. For example, the Kinect v2 can offer continuous input by tracking 25 body joints, each with 6 degrees of freedom (3 of position and 3 of orientation), of up to 6 players at the same time.
- Tracked hand controllers (e.g., Wii Remote, PlayStation Move, Vive, and Rift controllers) – Exist with different capabilities. From the use of gyros and accelerometers for acceleration and orientation detection, up to precise tracking of 3D position and orientation in space. When worn in the player's hands, they can provide buttons, joysticks, or trackpads for interaction and haptic output.
- Infra-Red marker-based Motion Capture (e.g., Vicon, Optitrack) – These systems can track the whole body or parts of it by wearing appropriate suits with IR markers attached or by attaching said markers to the player's body.
- Exercise equipment (e.g., treadmills, stationary bicycles) – Offer interaction through a single specific function and provide the system with input (or output) concerning that function, e.g., speed, inclination, pace.

While exercise equipment can provide exertion levels beyond what is achieved through unassisted body movements or calisthenic exercise, their use would constrict the technology in its use. Additionally, any system using this technology would be tough to transport. The combined use of tracked hand controllers with floor mats or force plates would provide tracking of the upper limb extremities and interaction through all the limbs. However, the elderly have shown a preference for entirely passive gesture interaction (camera-based) to the detriment of hand controllers, even when they also provided passive tracking [14]. Additionally, players with mobility disabilities or decreased hand function would find themselves unable to use the system properly. The last remaining options were motion capture through cameras with computer vision algorithms or via wearable IR makers. They are both passive methods but with distinct characteristics. A two-camera Vicon system provides millimetric accuracy at 12500$ (2014) [32]. However, the multiple-camera setup and the need to wear IR markers make it unsuitable for our purpose. Meanwhile, an RGB-D camera such as the Kinect v2 is explicitly designed for exergaming and provides 25 body-joints tracking, with centimeter-level accuracy at 150€ (2016). An interesting parallel can be drawn between video projectors and cameras. Just like a small projector can create a large projection, a small camera can provide a wide tracking area for interaction; after all, it is just a matter in which direction the light is traveling. This makes such devices very mobile and easy to setup. In fact, the combination of these small form factor technologies can lead to the augmentation of whole rooms to become interactable spaces, such as RoomAlive [33].

The final choice of technologies fell on video projectors for imaging output and motion-sensing cameras for input, specifically the Kinect v2. Further arguments for this choice are: Projections are easily adaptable to spaces, can be projected on walls, floors, ceilings, or even objects. They can be made quite large. The Kinect is perceived as the easiest-of-use by the elderly relative to other controllers [34], which we already saw is fundamental in the intention-to-use-technology in this age [19]. The possibility to have automatic fall detections [35]. Finally, surround-screen projection systems with passive tracking are cybersickness-free, self-explanatory, and easy to use by seniors [36].

In the following two subsections, 1.1 and 1.2, we present the technology developed and evaluated using the chosen input and output modalities. However, given its potential, it goes beyond simple exergame interaction and evolved to become an open-source software capable of providing immersive virtual reality experiences. The creation of exergames for elderly fitness with this technology materializes later in Section 2.

1.1 Development of the KAVE: Methods for Surround-Screen Projection Management, Motion Parallax and Full-Body Interaction Support[1]

1.1.1 Introduction

In [38], the CAVE (CAVE Automatic Virtual Environment) was described for the first time, a stereo multi-screen virtual reality system where the images are projected on the inside walls and floor of a cubic structure. The system used head-tracking and off-axis projections to simulate visual presence in the virtual world. In their work, the authors refer to the eight visual cues that enable us to perceive depth: 1) Occlusion, 2) Perspective projection, 3) Atmospheric effects, 4) Lighting and Shadows, 5) Binocular disparity, 6) Convergence, 7) Motion parallax, and 8) Eye Accommodation (eye lens control reflex for focusing light from different distances). Of these, cues 1, 2, 3, and 4 can be provided by conventional displays, 5 and 6 can be achieved with stereo graphics (such as stereo glasses or Head Mounted Displays (HMD)), and 7 with head tracking. As with most Virtual Reality (VR) systems, their system could produce all of them except 8.

The possible applications of the CAVE in data visualization were first explored by inviting experts and professionals from different fields to experience the system; this included the visualization of architecture, cosmic, fractal, weather, and molecular dynamics data and models; the visualization of jobs being executed on the multiple parallel processors of a supercomputer, for algorithm execution optimization; and visualization of real-time brain activity overlaid on a digital head model [39]. Some other CAVE systems were used for visualization of geophysical simulation data [40], underground cave systems [41], computational fluid dynamics and neuroscience [42]. However, the use of CAVEs has been mostly constrained to data visualization roles, in laboratories or companies and their potential to be used in other VR applications such as gaming has been rarely explored. In this section, we present the specific methods we employed to add a monocular motion parallax effect, through head position-dependent perspective projection, to Unity applications using the Kinect v2 (Microsoft, Redmond, USA) as the head tracking sensor. The addition of these effects to computer graphics increases the number of depth cues it can provide, and it is a fundamental building block in CAVE systems. These methods were bundled in an open-source plugin, the KAVE, which represents a solution, for VR developers and researchers, in the development and conversion of virtual environments (VE) and games to immersive multiple large-screen projection systems. Using the KAVE, the presented methods were employed in several proof-of-concept configurations to showcase their flexibility, additionally a few example applications are presented to demonstrate what a user can expect from an VE powered by this tool.

1.1.2 Related Work

The methods for generating the motion parallax effect and head position-dependent perspective projection have been used consistently in computer graphics to design different three-dimensional displays such as HMDs, head coupled displays, and CAVEs [38], [43], [44], our technical implementation of these methods can

[1] The content of this section has been published in the proceedings of the ACM on Human-Computer Interaction – EICS [37]

be consulted in section 1.1.3.1. Although these methods have not changed through the years, many recreations of the original CAVE have been developed, focusing mainly on technical differences for both scientific and commercial purposes. Nevertheless, the costs for such a system remain very high. The CAVE [38] and CAVE2 [45] systems are estimated to have cost 2,000,000$ and 926,000$ respectively [45]. While low-cost solutions exist, implementing a low-cost CAVE, including the physical construction and software, can still be prohibitive (19,300€) [46]. In [46], they created a CryEngine2 game engine mod for games using that same engine. However, because it lacked user tracking, it assumed a static head position of the user and supported neither motion parallax nor stereo vision. Another low-cost CAVE was presented in [47], based on rolling screens suspended from the ceiling with a cost of 9,500$, excluding the computer and the necessary software for their custom setup. Their configuration is easy to use and does not permanently occupy floor space. For tracking sitting users, it employed marked glasses and infrared cameras. The immersive virtual content was provided by adapting individual Unity applications to their visualization paradigm. Unfortunately, the code was not made available. Another CAVE made using off-the-shelf hardware and the author's own software revealed to be an alternative low-cost solution for fire brigade training in simulation [48]. This solution was developed for Unity using a Kinect as a sensor; however, it lacks the desired replicability as it is not clear whether the software can be used in different setups other than the one for which it was developed.

While building a CAVE, a large part of the budget is generally allocated to the tracking technology, ranging from 8,000$ for a six low-end IR camera motion capture setup, up to 40,000$ for high-end cameras, from suppliers such as NaturalPoint Inc. OptiTrack or Vicon Motion Systems Ltd UK. Additionally, middleware software to render across multiple displays, integrate sensors, and applications is complex and must be bought for several thousands of dollars from specialized companies, such as MiddleVR [49] or custom-developed in-house. Like the Uni-CAVE [50], a plugin for managing surround-screen projection in Unity applications, with the drawbacks of lacking full-body tracking and requiring the CAVE setups to be defined before the VE (Virtual Environment) compile time.

Beyond CAVE setups, Microsoft Research has consistently used Kinect cameras to implement spatial augmented reality (SAR) applications through projection mapping in arbitrary scenarios. In LightSpace [51], the authors report their interactive installation that combines multiple depth cameras and projectors registered to a single 3D space. This unified space across projector/camera units (procam) units enables them to project graphics into the depth mapped surfaces and the users to interact with the projected elements with their bodies through the bodies' 3D mesh obtained from the Kinect sensor. With MirageTable [52], a similar but scaled-down system, single Kinect and projector, added support for non-flat surface projection and head tracking for head position-dependent perspective projection. In this system, the use of a single Kinect for both tracking the user, online, and mapping the surface, offline, requires that during calibration, the Kinect has to be first installed in a position and orientation that can capture the whole projection surface and then installed in a way that it can capture both the user and part of the projection. Again with a single Kinect and projector, IllumiRoom [53] shows how the same technology can be used to augment the arbitrary space surrounding a television using the depth data from a Kinect. The combination of these projects led to RoomAlive [33]. The RoomAlive toolkit joins together tools for the automatic calibration of procam and tools for authoring augmented content in Unity. This tool supports full-body tracking for interaction and head position-dependent perspective projection and projection mapping to arbitrary spaces using Kinect sensors. These capabilities enable developers to create SAR experiences in any space, theoretically in CAVEs also. The system chains together the information from multiple procam units to automatically calibrate and register the whole setup in the same 3D space reference frame, requiring one Kinect per projector, that their fields of view overlap with each other and include the intended user tracking area.

The works stated above provide tools to create VR experiences tied to a specific setup since development or focus on depth mapping of arbitrary spaces, which require elaborate setups for calibration and to run. In this work, we hope to contribute in a significant way to improve two different "situation, task and users" (STU) contexts [54] in the development and use of VEs for CAVEs. First, our contribution can separate VE developers from users of a VR CAVE application, creating the STU of programmers (tool users) developing VEs (task) for CAVE users (end users) that will use the VEs for their own purposes (end task). A separation that is hard to make with the related work tools, where setup configuration must be known at compile-time or obtained live from depth mapping. Starting with the second context, concerning CAVE users, when compared with alternative solutions, the applications built with our tool have the capability of reaching a much larger population due to the simplicity of the required setup, as the users do not require access to the source of the application, therefore taking exclusivity of CAVE-like system away from large commercial or research institutions. As for VE developers', our contribution gives them the flexibility of developing and compiling an application that can be deployed to a diverse range of user-dependent setups.

In the following sections, we describe the KAVE, our flexible CAVE software for the Unity 3D game engine (Unity Technologies, San Francisco, USA). In the form of a Unity plugin, our free, open-source software supports up to 8 displays (both projectors and screens) in any physical configuration. It is versatile regarding hardware as it supports different screens, projectors models, resolutions, or aspect ratios at the same time with arbitrary positions and orientations; and combined with the use of a low-cost tracking sensor makes entry-level CAVE-like systems much more affordable and easy to use by the community of game, simulation, VR developers and researchers. KAVE relies on a 140$ Kinect v2 tracking sensor, which has a remarkable accuracy of landmark movements [55] and provides full-body tracking of up to 6 simultaneous users, enabling unobtrusive full-body interaction in VR. Additionally, to complement our Unity plugin, we make available a versatile tool for the easy setup and calibration of CAVEs and describe our CAVE physical setup.

1.1.3 Methods

1.1.3.1 KAVE Plugin, Adding Motion Parallax and Whole-Body Support to Unity Using Kinect

An essential feature of a CAVE is the motion parallax effect produced. This is achieved by keeping continuous tracking of the user's point of view and automatically adjusting the projection of images consistent with that perspective. Figure 1A & B illustrate the phenomenon in a 2D monocular diagram. By knowing the positions of the user viewpoint, projection plane, and virtual elements to represent (Figure 1A), we can project images on that plane coherent with what the user would see if the virtual elements were in the real world (Figure 1B). By generalizing this effect to 3D and the use of multiple projection surfaces, we obtain the basic functioning of a CAVE, as shown in Figure 1C, where the blue lines represent the walls and floor of a CAVE (projection surfaces), the orange icon the user head position and the yellow elements represent the virtual objects in their virtual positions relative to the user. As a result, the CAVE must display the green images (planar images on the projection surfaces) to emulate the virtual objects' visual effect if they were real.

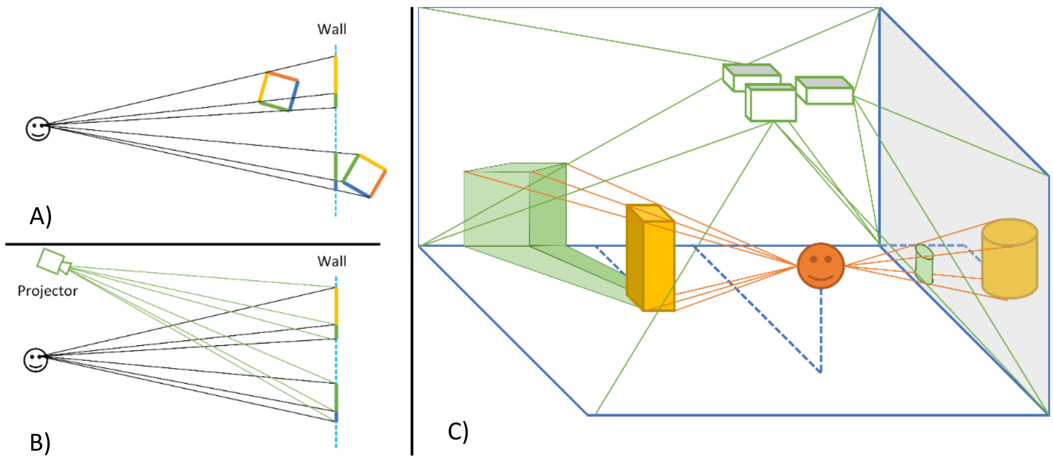


Figure 1: A) Finding out how virtual objects (colorful squares) should be seen by the user; B) Projecting on the wall what the user should see. C) Projections (solid green) of virtual elements (yellow) in real surfaces (blue) adjusted to the perspective of the user (orange) point-of-view by the CAVE projectors.

The parallax effect's creation depends on the precise knowledge about the size, position, and orientation of the projection surfaces, body tracking coordinate system, and correct mapping of the CAVE projectors' images to the surfaces. A calibration process must provide this information, as it is necessary to model the CAVE in Unity mathematically. Hence, the calibration process enables the correct mapping of the projections into the physical CAVE walls and to get an accurate location of the user's point of view relative to them.

Our calibration process creates a CAVE virtual replica description file with its origin located on the CAVE floor center. Consistent with this calibration file, the following elements are created at runtime by a Unity application using our plugin (Figure 2):

- CAVE Kinect v2: A Unity game object with the position and orientation matching the real sensor reads up to 6 users' 25 joints skeleton positions and orientations from the Kinect SDK in its coordinate system. It instantiates a transparent (by default) avatar for each user; these avatars can have their colliders activated, thus allowing users to interact with virtual elements through Unity's physics simulator;
- CAVE Screens or Surfaces: Invisible Unity quads/planes representing flat-panel displays (such as TVs or computer screens) or the CAVE projection surfaces' position, orientation, and size, they serve as rectangular reference targets for the User View Cameras;
- User View Cameras: Unity cameras attached to the user's head. The CAVE Kinect v2 object provides the head position. In the case of multiple users, the closest person to the sensor is chosen. One camera per CAVE Surface/Screen is created, each with its principal axis perpendicular to that surface/screen. The camera projection matrix values are calculated in real-time to ensure the camera frustum's coincidence with the target surface/screen edges, meaning that the surface/screen rectangle exactly frames each camera image. Depending on the target of the camera (screen or projector), the views from each camera are either sent directly to the screen display or turned to render textures and sent to the corresponding CAVE Projector for further warping;
- CAVE Projectors: Unity cameras that receive the images from the User View Cameras and warp them to achieve a correct mapping into the desired real-world projection surfaces. This warp is done by mapping each of the four corners of the image to new viewport locations corresponding to the actual projection wall corners. This makes possible configurations where the projectors are poorly aligned with the walls, such as when they are at an angle or when they project on more than one wall (usual when the aspect ratios of the projector and wall are different). The projection outside of the new image boundaries is black, eliminating any overlapping projections between adjacent surfaces. A

projector can receive multiple User View Cameras simultaneously and apply a different warping mask to each; this enables a projector to display on several surfaces at once. This warping process removes the burden of physically calibrate projectors, replaced by a more straightforward software calibration.

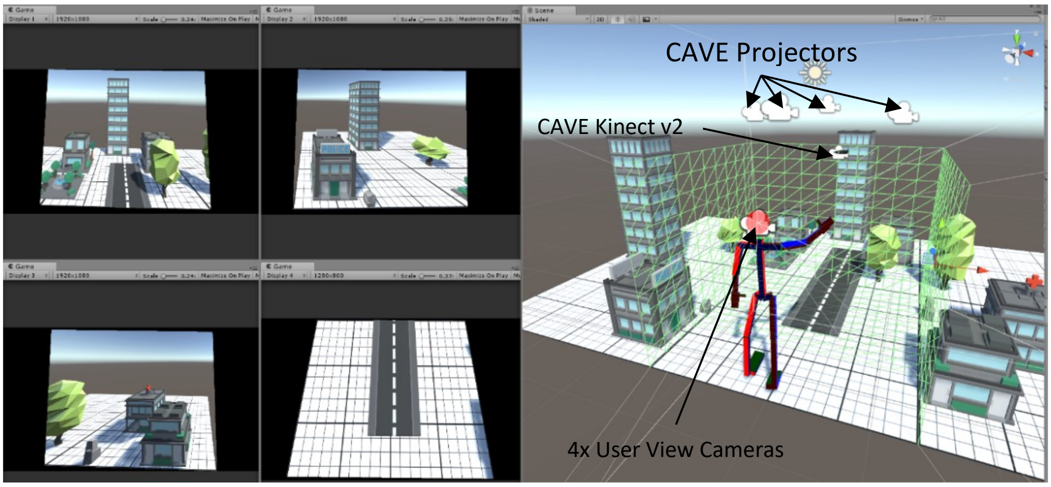


Figure 2: The virtual KAVE generated at runtime in the Unity editor. CAVE Surfaces are highlighted in green, the virtual equivalents of the four ceiling mounted projectors are seen as cameras on top, the four user view cameras (and their frustums) centered on the user's head, and the Kinect sensor is seen as a black bar above the central wall. On the left are shown the four views that the CAVE Projectors display, warped to match the CAVE walls.

While there are multiple ways of setting the User View Camera projection matrix to fit our goals, we took advantage of Unity's existing camera parameters and functions. At each new frame, the User View Cameras' position is updated per the CAVE Kinect v2 sensor data, while the orientation remains perpendicular to the corresponding CAVE Surface/Screen. Then, after resetting the camera projection matrix, the vertical field of view (FOV, α) is calculated per Equation 1, followed by the manipulation of the camera's principal point offset by setting the unity camera matrix elements $m_{1,3}$ and $m_{2,3}$ as described in Equations 2 and 3.

$$\tan\frac{\alpha}{2} = \frac{h/2}{dy} \equiv \alpha = 2\tan^{-1}\frac{h/2}{dy} \quad (1)$$

$$m_{1,3} = \frac{dz}{h/2} \quad (2)$$

$$m_{2,3} = \frac{dx}{w/2} \quad (3)$$

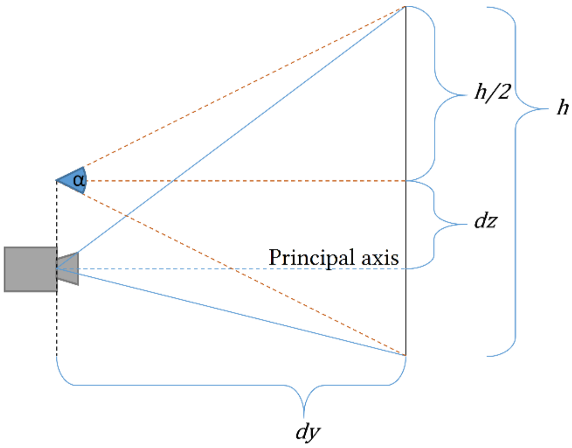


Figure 3: Setting a User View Camera frustum (solid blue) coincident with a CAVE surface (solid black), side view.

Here h and w are the height and width of the surface/screen, dy is the distance from the camera to it, and dz and dx are the camera's principal axis vertical and horizontal offsets relative to the surface/screen center. In

Unity, this results in having a camera frustum that passes through the edges of our projection planes/screens, as shown in Figure 3.

Our plugin uses the methods described above to make existing Unity 2017 projects CAVE compatible. The plugin is versatile regarding the physical setup and hardware, dealing with sets of heterogeneous projectors or flat screens in arbitrary positions and orientations. While initially developed to create CAVEs, it supports any other configuration of up to 8 displays, such as interactive walls, floors, or any other flat surfaces combination, always providing a parallax effect. It works by instantiating at runtime a virtual equivalent of the real CAVE parameterized in an XML configuration file. The virtual CAVE is composed of a set of prefabs from the types described above. In short, cameras attached to the user's head, thanks to the Kinect tracking, view the virtual world while having their frustum continuously adjusted to match the correspondent CAVE wall/screen edges. Their images are either streamed directly to display screens or to the virtual CAVE projectors that warp the images to match the real ones; these warped images are projected into the real-world CAVE walls, resulting in the correct perspective projection.

After importing the KAVE plugin into a Unity project, existing camera components should be deactivated and the prefab "CAVE Manager" added as a child of the player game object. This game object will take care of everything, including loading the calibration file and generating all the KAVE elements at runtime. The "CAVE Manager" position inside the game matches the reference frame origin of the CAVE calibration, usually on the floor. It is also possible to customize the cameras that are now instantiated at runtime by editing the "User View Camera" prefab parameters, such as culling mask and clipping planes or adding visual effects like it is possible in any Unity camera. Adding this plugin to an existing unity project to make it CAVE compatible takes less than five minutes. The KAVE plugin, source code, and demonstration project are fully available in its development repository at https://bitbucket.org/neurorehablab/kave.

1.1.3.2 KAVE Calibrator

A standalone tool to calibrate the KAVE was developed in Unity. The purpose of this tool is to allow the user to generate a configuration file that represents the physical CAVE setup, which the KAVE plugin can then load to recreate it in virtual space. It provides a 3D virtual environment (see Figure 4A), similar to the Unity editor where up to 8 surfaces and projectors can be dynamically added, individually adjusted in position and orientation, and resized to match the real-world configuration. Surfaces are associated with projectors to ensure no projection overlaps and to allow a projector to project on several walls at once (such as a corner). The mapping of the projections into the surfaces is done visually on the CAVE walls themselves. Each projector projects a 4-sided red polygon over the wall they are facing, and the user then drags with the mouse each of the corners of this mask to match the corners of the real wall (see Figure 4B). The viewport coordinates of the four corners are saved. Once these steps are done, the tracking sensor is added. Its horizontal position and azimuth angle relative to the CAVE reference frame are set on the calibrator. The height to the floor, pitch, and roll angles are automatically obtained from the floor plane's Kinect SDK estimation. The software saves the information in an XML file which is later loaded by the KAVE plugin or by the calibrator again for further adjustments.

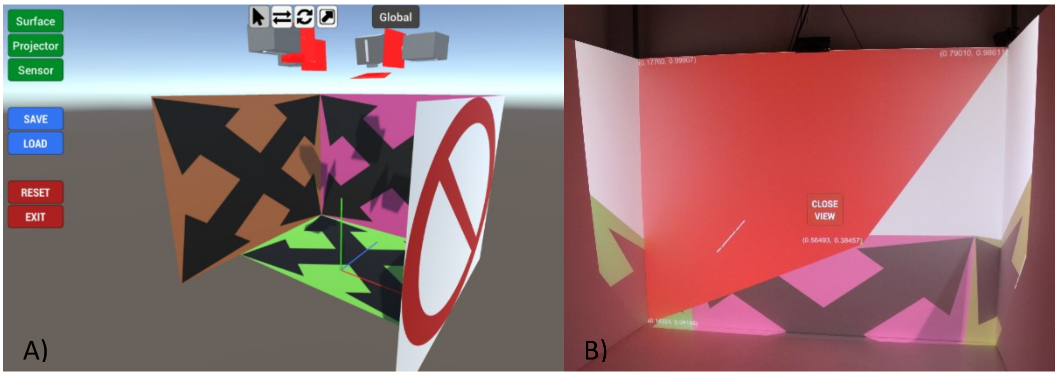


Figure 4: A) KAVE Calibrator being used to create a CAVE with three walls and floor along four projectors and their projection masks. B) Using the calibrator to find the wall corner coordinates in the projector viewport reference frame, the bottom right corner is still missing.

1.1.4 KAVE Configurations and Example Applications

Although the KAVE plugin was initially developed with a CAVE paradigm in mind, its configuration versatility can support many other setups. This section presents a limited list of example configurations and applications implemented with the KAVE. The diverse examples of supported configurations showcase the flexibility of the tool [54], as the same compiled application can be deployed in a multitude of setups simply by providing the corresponding calibration file at runtime. The example applications demonstrate the power it provides to the users, ranging from supporting simple VR visualization, by adding motion parallax, to UI and VR interaction through full-body tracking.

1.1.4.1 NeuroRehabLab CAVE

The KAVE is currently being used to power applications in our CAVE. The NeuroRehabLab CAVE has a configuration of three adjacent walls at 90 degrees with each other and a floor; the walls are 2.8 meters wide by 2.2 meters tall and the floor 2.8 by 2 meters, allowing up to four projections simultaneously (Figure 5). The projectors are four Optoma GT1080, with a 1080p image resolution and throw ratio of 0.5:1. The low throw ratio reduces user shadows inside a CAVE of these dimensions, and the price (around 800€) is on par with market values. For tracking, it uses a single Kinect v2 sensor located on the top of the center wall. In our effort to reduce the costs associated with CAVE ownership, we custom-made our structure to support the projectors, sensors, and projection surfaces. The metallic structure is made of galvanized iron, allows projectors' adjustments to floor and wall distances through sliding supports and individual orientation. Each wall is made of 3 wooden sheets (1.2 cm thick), plastered and painted white, while the floor is made of a rigid white vinyl sheet. Besides providing support to the Kinect sensor, the structural beams can also support other types of sensors. A set of two HTC base stations for the HTC Vive head-mounted display are also installed. The price of the CAVE (structure, projectors, and Kinect) was just under 3,800€ (projectors costing 3,200€), while the accompanying computer (Quad Core 3.4GHZ processor, 8GB of RAM, and Radeon RX 580 8GB graphics card) cost was 1,200€. The calibration of this setup using the KAVE Calibrator described in section 1.1.3.2 is straightforward.

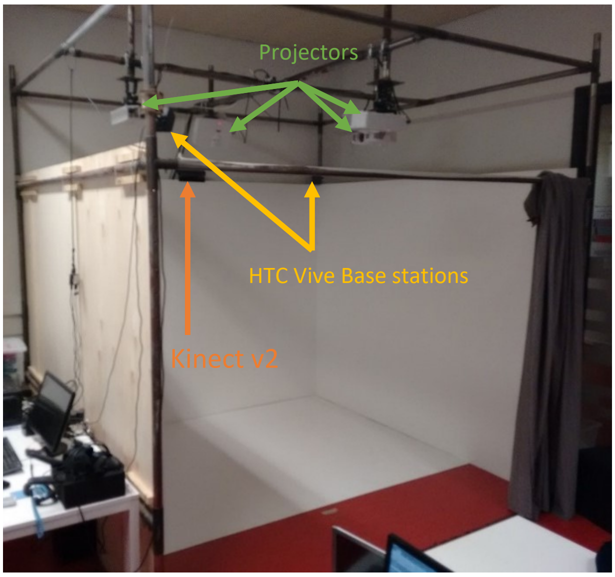


Figure 5: NeuroRehabLab CAVE.

1.1.4.2 Interactive Floor

A simple configuration of one projector facing a surface can turn it into a large gaming/visualization platform supporting direct interaction and the feeling of depth due to motion parallax. By attaching a projector vertically to a wall or ceiling, a large projection area can be made. An interactive floor can be implemented in this fashion by placing the Kinect sensor facing the user (Figure 6). This configuration has been used to develop a set of Portuguese themed exergames [39], which are the focus of Section 2.3 and Section 3 of this thesis. Calibrating a one surface/one projector setup such as this one is the most straightforward calibration possible, requiring only the floor dimensions and orientation, projector viewport coordinates of the 4-floor corner points, and the Kinect position, all relative to the center of the floor.

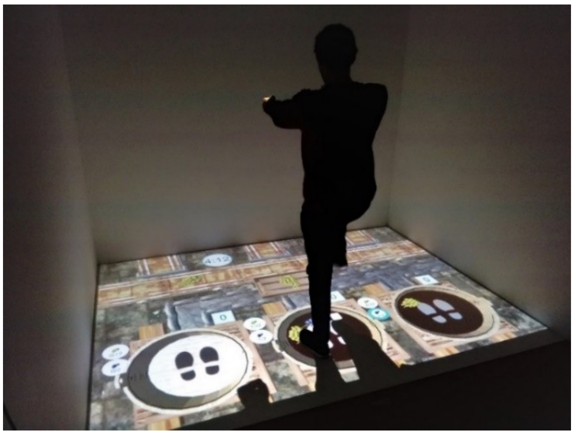


Figure 6: Interactive floor being used to play a grape stomping game [56] on the NeuroRehabLab CAVE floor.

1.1.4.3 Interactive Wall / PowerWall

Multiple projectors can project over a single large surface (such as a long wall). This is possible because the plugin can seam multiple projections into a single view over adjacent areas. Figure 7 shows one example of such a setup where 2 Acer S5200 (1024x768, 3000 Lumens) projectors are used to create a 2.84 m + 2 m wide by 2.6 m tall projection wall. The parallax effect is provided by setting the Kinect sensor behind the user. Having the Kinect placed behind the user does create a problem for interaction as this sensor is not trained to recognize people from the back and instead assumes they are facing the sensor. Thus, the accuracy in body joint estimation is reduced and the body is seen as inverted by the system; still, the head can be tracked. Alternatively, the Kinect can be placed in a better front-facing position. Up to 8 projectors with arbitrary resolutions can create a high-resolution interactive wall with the KAVE plugin in this configuration. The main calibration problem with this setup concern the precise way in which the user defines the boundaries between adjacent projections, the seam. First, the user must ensure that two adjacent projections overlap, then physically mark the desired corners of each rectangular "virtual wall" on the real wall. Finally, the calibrator can be used in the same manner as the previous two configurations, including to obtain the position of the markers in each projector viewport coordinates, just like they would map the four corners of individual walls in a CAVE.

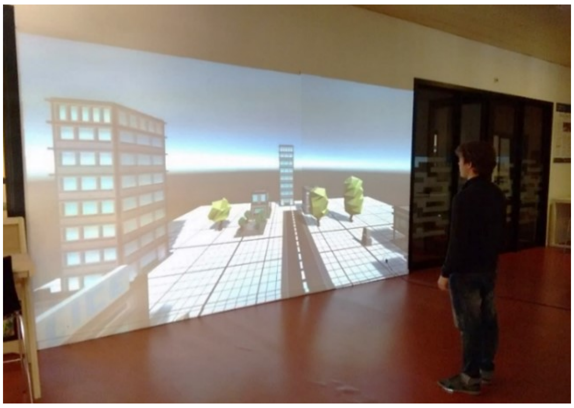


Figure 7: Interactive Wall, 4.84 m by 2 m of 2 projectors.

1.1.4.4 Corner Projection

A single projector can project multiple surfaces simultaneously, such as the corner of a room. This could be used to create a CAVE using two walls and floor projections with just one projector. An example of this is shown in Figure 8, where one portable LG Minibeam PW800 (1280x800, 800 Lumens) projector creates a 1.2 m x 1.2 m, 0.9 m tall CAVE on a corner. The corner CAVE concept can be made larger by simply using a low throw ratio projector covering a larger area. The main issues regarding this setup are where to place the Kinect relative to the user and the reduced brightness and resolution that results from having a projector facing multiple projection surfaces at high angles. Regarding calibration, as in the previous example, temporary markers on the walls are needed to serve as reference for each corner of every wall.

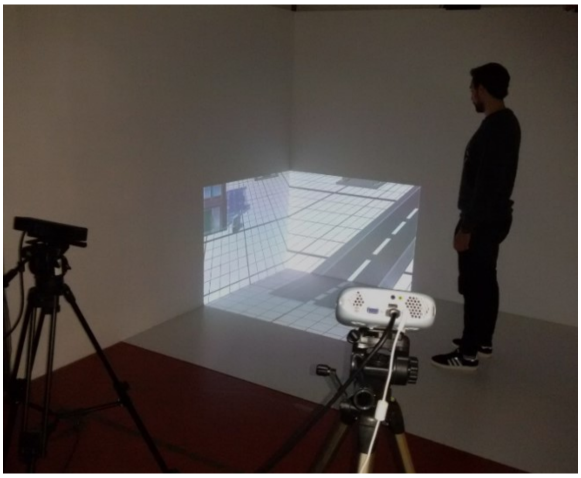


Figure 8: Improvising a minimal CAVE with a corner projection.

1.1.4.5 Parallax Screens

Since the plugin supports display screens and projectors, it can also be used for screen-based CAVEs or even living room and desktop applications. Here we illustrate it on a desktop screen with parallax effect (Figure 9). By using multiple or larger screens, a more immersive setting can be created. Even though the KAVE plugin supports this configuration, the KAVE calibrator does not yet support screens. Instead, this calibration was achieved by manually editing the calibration file to match the screen dimensions and orientation and the Kinect position relative to it.

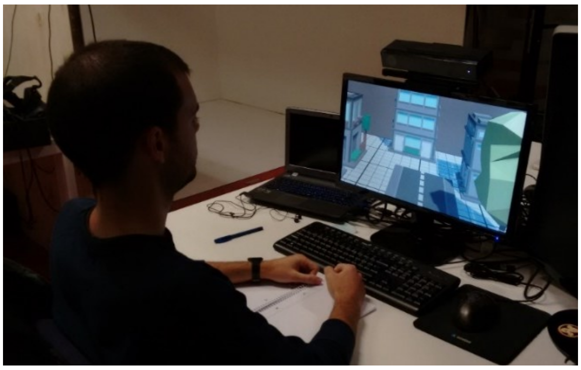


Figure 9: The Parallax Screen acts as a window into a virtual scene, or virtual elements can pop out from it.

1.1.4.6 Exclusive Use of the Plugin for Body Tracking

By creating a calibration file deprived of screens and projectors/surfaces instances, the plugin will only generate the Kinect v2 interface instance, responsible for avatar creation and control. Although this foregoes all the perspective control and projection management that the plugin provides, it can still help users who have other means of visualization in mind and are only interested in having their full-bodies represented or interacting with a virtual scene through Kinect tracking. An example of such an application would be integrating with HMD to provide virtual bodies to its users, allowing players to see, in first-person, their VR bodies and others around them, as shown in Figure 10. This is possible just by making sure that both HMD and KAVE reference frames are coincident.

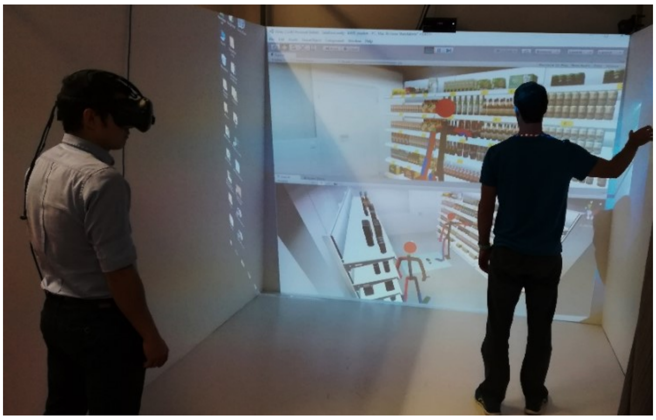


Figure 10: The KAVE plugin enabling an HMD wearer to see his own virtual body and others tracked by the Kinect. The scene's HMD view and perspective are being exceptionally projected on the front wall.

1.1.4.7 Example Applications

The NeuroRehabLab CAVE powered by the KAVE plugin gave the authors' research institute a powerful yet affordable VR system. A simple application was used during the plugin development to test the motion parallax effect generated. This application consists of a set of static 3D low polygon count assets and the KAVE plugin prefab "CAVE Manager" placed among them. The machine ran Windows 10, had an Intel processor I7-6700 3.4GHZ, 8GB of RAM, and a Radeon RX 580 8GB graphics card connected to a CSV-5400 quad monitor Multi-Stream Transport (MST) Hub. This setup achieved a stable performance of 60fps while displaying 8.2944Mp (4x 1920x1080p), averaging 4 ppcm, and can be seen in Figure 11.

The KAVE plugin was integrated into two already existent games, developed by institute collaborators, and incorporated in an embodied interaction course project.

“Explornesia: The Battle” is a real-time multiplayer sea battle game featuring the age of sail ships controlled from a 3rd person point of view. The conversion to a CAVE playable version was as easy as replacing the standard Unity game camera with the KAVE plugin game object. Two prototypes were made by editing the KAVE’s position and scale relative to the player’s ship. One where the player controlled the ship from the deck in first-person, see Figure 12A, and another maintaining the original third person perspective where the player’s ship was displayed as a 1-meter long ship floating on the CAVE floor.

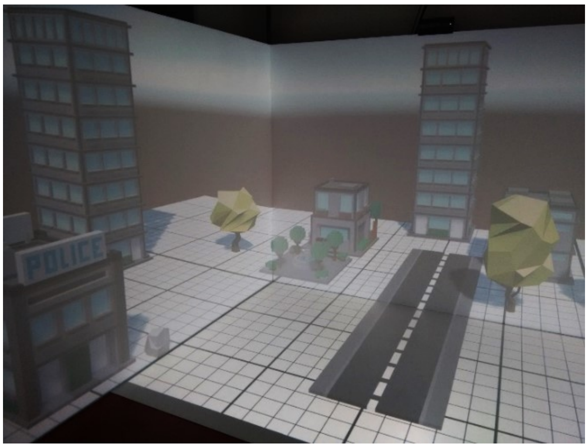


Figure 11: User perspective of the KAVE demo project running in the NeuroRehabLab CAVE.

The second game was “Keeper of Intheris,” an online, turn-based strategy game designed to be played on a computer screen with a keyboard and mouse from an angled perspective. The developers modified the game and added a non-playable proof-of-concept first-person observer view scene with the KAVE plugin (Figure 12B). It supports navigation through a handheld thumb joystick. Additionally, full-body interaction with some elements of the game environment was enabled through physics simulation. The observer can interact with them via body motion, such as kicking virtual rocks. Development of this scene required: creating a standalone non-playable scene of the game, where the game characters are in idle animation; adding rigid body and collider components to rocks; removing the scene original camera and adding the “CAVE Manager” to the scene; activating the Kinect body colliders; and add joystick support.

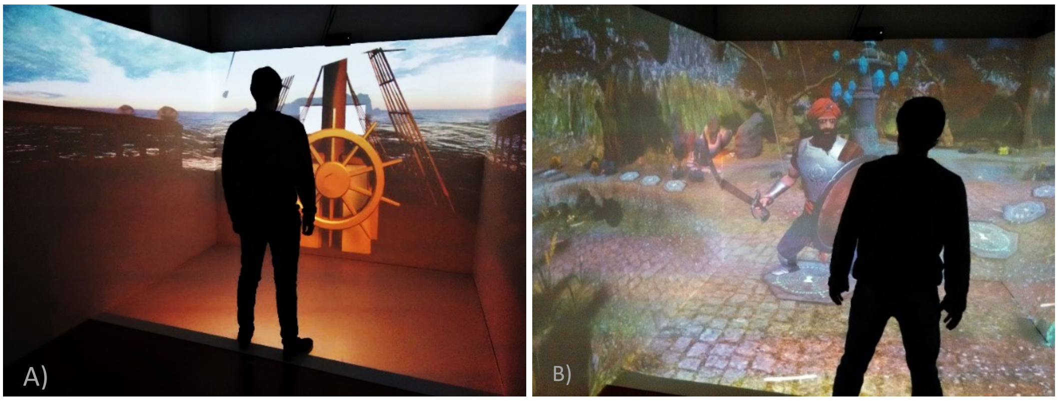


Figure 12: Using the KAVE to visualize games from the “inside”: A) “Explornesia” being played from the deck of a ship. B) An observer inside “Keepers of Intheris”.

Over a college course of embodied interaction, a group of students developed a proof-of-concept prototype for atmospheric and oceanic data visualization. Observing the work and interviewing researchers from the “Oceanic Observatory of Madeira” (OOM), a set of ideas to facilitate natural data visualization was put into practice with the KAVE. The resulting prototype featured a VR presentation of the topographical data in the center of the CAVE, which the user could explore by moving around it. The interaction was done on the wall-sized UI, either through buttons or sliders, all activated solely by the hand position of the users provided by the KAVE.

1.1.5 Discussion

Even though CAVEs have been around for a considerable amount of time, they remain hard (and costly) to implement and maintain. To this end, we developed an easy to use plugin for Unity developers that adds CAVE support and motion parallax effect to virtual environments across multiple displays. This, together with our configuration and calibration tool, provides all the required elements for a functional monocular CAVE system. A summary of strengths, weaknesses, opportunities, and threats is presented in Table 1.

Table 1: SWOT Analysis of our Contribution

Strengths: • Free and open-source software • Low-cost tracking • Versatile in multiple hardware and screen configurations • Space geometry calibration independent from VE development and compilation, all KAVE elements created at runtime • Unobtrusive interaction • Software easy and fast to use	**Weaknesses:** • Limited tracking accuracy • Monocular • Price of the projectors • Tracking area limited to the field of view of one sensor
Opportunities: • Large community of Unity developers • Very competitive compared to regular CAVE cost • Emergence of new game modalities supporting full-body interaction in immersive environments • Alternative middleware solutions for CAVEs are expensive • Popularization of VR	**Threats:** • Dependency on Kinect v2 for Windows • High investment in hardware for the average developer • Dominance of immersive VR by Head-Mounted Displays • Lack of dedicated content for CAVEs

Compared to the Uni-CAVE [50], noncommercial software with similar goals of powering CAVEs, the KAVE distinguishes itself from it for supporting full-body tracking and interaction, and runtime instantiation of all its elements. This means that the VE developer can compile a KAVE project independently from its users and their projection setups. This single built project can be distributed to different users, which only need to add their calibration files to experience the same VE correctly mapped to their setup. With the KAVE, contrary to the Uni-CAVE, there is no need for the users ever to touch the VE development project or for the developer to build a custom project for every individual user setup. This difference is also applicable to RoomAlive [33], a toolkit for SAR, which requires the static room geometry automatically obtained from its calibration tool to be loaded into the Unity editor project, thus tying VE development to a specific space geometry. Although RoomAlive also supports the dynamic surfaces to be obtained online from Kinect's depth stream, this is intended for projection mapping on moving physical objects. In both the Uni-CAVE and RoomAlive, the application user and the application developer are the same, while with the KAVE, they can remain independent. Also, to power a CAVE with RoomAlive, multiple sensors (each running on a dedicated PC) would be needed to cover the whole projection setup and user interaction space, dramatically increasing the cost and complexity of what would be otherwise a straightforward setup. It is of note that calibrating RoomAlive for planar surfaces, such as CAVEs, requires the procam units to be temporarily oriented towards non-planar surfaces due to their nature autocalibration procedure. Lastly, neither of these two tools support screen displays.

Some difficulties in the adoption of this system exist. Among them, the Kinect v2 is less accurate than the gold standard IR marker-based sensors, with a 1 cm mean error and standard deviation for the “head joint” [55]. However, the software can be upgraded later to support new sensors, such as the Kinect Azure, as they become available. Nevertheless, it has been reported that human head movements can average up to 2cm when standing and looking at static images on screens [57]; this might indicate that such sensor accuracy can be acceptable. Regardless of our solution, a CAVE still requires a minimal investment in projectors and physical structure, but it can be as low as 5,000€, as shown by the NeuroRehabLab CAVE. Our low-cost KAVE solution presents an excellent alternative to these systems’ traditionally high cost, namely by reducing the software, hardware, and sensor costs by using low-cost off-the-shelf equipment. In particular, our solution can have high acceptance by the industry and academia, as it offers most of the features of high-end commercial CAVE systems with a ready to use software bundle for Unity. While the chosen platform and the ease of setup potentiate the adoption by the large community of Unity developers, it will also contribute to the promotion and growth of CAVEs as research, professional and gaming platforms, and the creation of novel content exploiting both full-body interaction and immersive VR. The fact that the Kinect v2 sensor provides tracking of 25 body joints per user is an additional clear advantage over traditional marker-based systems with the potential of developing new non-obtrusive interaction modalities in VR and games.

Finally, alongside CAVEs, the plugin in its more distilled configuration form, of just one projection surface, can also be used to provide simple interactive experiences for public spaces, such as imaginary windows to virtual worlds, interactive walls, floors or ceilings, facilitating and opening opportunities to develop applications for these settings. Thus, this plugin’s extreme versatility makes it go beyond the purpose of CAVE creation and allows both developers and users to explore and set up different projection mapping scenarios easily. Besides the addition of motion parallax to Unity projects, the most exciting aspect of this plugin is the easy conversion of a Unity game into an immersive experience with the potential of natural user interaction. By having a virtual skeleton matching the user’s own body inside the virtual environment, the whole body becomes an interface in Unity. That is the exact scenario explored in this thesis from Section 2 onwards, where an interactive floor was used as a large projection display for our exergames.

1.1.6 Conclusions

This section presented and described the KAVE Unity plugin and the principles that make it work. The plugin is a publicly available tool for creating surround-screen projections, supporting up to 8 projectors or screens in any combination or geometrical arrangement. Motion parallax and full-body tracking are featured using affordable off-the-shelf components. It differentiates itself from commercial-related work by being free and open-source, and from academic-related work by its versatility. By untying the physical setup calibration phase from the VE development, we distinguish the KAVE from the previous software solutions where each application was tailored before compile time to a specific lab setup. We provide specific examples of the KAVE versatility in both applications and screen setups. Two examples of VEs being converted into CAVE compatible demos are provided, and five other variations in terms of hardware elements configuration, all of them accessible without requiring VE source code access. Given its ease to use and low requirements, we expect the plugin to be a valuable tool for research laboratories and VR enthusiasts.

1.2 Evaluation of the Tracking Accuracy, Sense of Presence, and Cybersickness of a Surround-Screen Projection Systems Powered by the KAVE[2]

1.2.1 Introduction

Some of the features that make Virtual Reality (VR) an exciting tool to tackle the challenges of this thesis is that it can provide computer-generated environments and digital simulations that provide a real-world-like experience, both from the interaction and exploration perspective [58], [59]. While there are different ways to produce VR, each system provides its specific level of immersion. According to Slater et al. [60], [61] immersion describes to what degree a system is extensive, surrounding, inclusive, vivid, matching, and self-representative. Consequently, immersion is associated with the number of sensory channels involved (extensive), the directionality of the stimulation and the natural modes such as stereopsis (surrounding), the number of sensory systems that are disengaged from reality (inclusive), the variety and richness of information (vivid), the match between our proprioceptive system and the information provided (matching), and the provision of a virtual body (self-representative). While these features are objective technological characteristics of a VR system and contribute to the realism of the experience [62], the personal traits can also modulate how users experience a virtual environment (VE) or a VR system and promote differences among them. For this reason, it is also essential to measure the subjective perception experienced by users while using our KAVE system; the sense of presence being one of the most relevant factors. Presence has been described as the awareness of being immersed in a VE while ignoring the technology that mediates the experience, a sense of "being there" in the VE, instead of merely perceiving it [59].

To produce VR experiences of a high immersive degree, different display systems have been developed, including head-mounted displays (HMDs) [63] and stationary surround VR displays, such as cave automatic virtual environments (CAVEs) [38], [64]. HMDs are headsets that provide visual stimulation with eye disparity and, in the last years, also provide built-in head tracking. In the CAVE and similar systems, screens are fixed in the physical world and surround the user; external optical tracking systems usually provide head tracking. Information about the head pose is used to match the images displayed to the user's perspective. While recent technological developments and entertainment interest in HMDs have facilitated their commercial availability at reasonable costs, CAVE-like systems are still costly solutions, both in money and space, for data visualization in research laboratories and companies. In the previous section (and corresponding paper [37]), we presented the KAVE software for managing monoscopic surround-screen projection with motion parallax, which uses the low-cost RGB-D camera Kinect v2 to provide head tracking, inspired by the work of J. Lee [65]. The KAVE aimed at providing a low-cost alternative and overcome some of the barriers that prevent CAVE-like systems' widespread use, such as proprietary software (and its costs), expensive tracking hardware, and strictness of the physical and hardware setup.

Although the KAVE software has shown to be a feasible alternative to create an immersive system exclusively using off-the-shelf, low-cost devices [37], the system's capability to elicit presence and cybersickness remains unexplored, as well as the quantification of the accuracy and precision of its tracking technology. To study some of the circumstances in which this system could be advantageous or comparable to the competition, we set out to explore our system in two controlled experiments, which allowed us to directly compare with a previous study by Borrego et al. [66]. In the first, we measured the head tracking's accuracy and precision compared to a laboratory-grade optical tracking system in a controlled static scenario. In the second, we measured the sense of presence and cybersickness felt by participants. With this, we aimed at answering the following two research questions: How accurate and precise is the Kinect v2 in estimating the user's head

[2] The content of this section has been accepted for publication at the IEEE Transactions on Visualization and Computer Graphics journal

position in our KAVE-powered CAVE? And, in a simple VR action space search task, to what extent can the KAVE induce presence while remaining cybersickness-free in a representative sample of healthy adults?

After occlusion, stereopsis (binocular disparity) and motion parallax are the two most relevant cues when estimating objects' depth in the observer's personal space (less than 2 m). While for distances less than 1 m, stereopsis outweighs motion parallax, at larger distances, motion perspective due to parallax is a better source for depth estimation [67]. However, both sources of information decline linearly with distance, and at 30 m are no longer reliable enough compared to other sources, thus defining the action space (2 m to 30 m) [67]. Most competing VR systems feature stereoscopic imaging, and therefore depth perception through stereopsis. Due to its monoscopic nature, this depth cue is not available in the KAVE. Because of this difference, we choose a task that does not depend on stereopsis to enable a valid comparison with systems featuring stereopsis. Thus, we focus our study on a visual search task happening mainly in the action space (slightly beyond arm's reach). Lastly, by replicating a previous experiment that compared a laboratory-grade CAVE and an HMD walking VR system, using the KAVE and a modern HMD instead, we could compare the KAVE to three alternative technologies. Comparing it with a modern HMD is relevant as their significant evolution in the past decade has changed the balance in the relation between presence and VR mediums, which previously favored CAVEs [68]. The comparison with a laboratory-grade CAVE allows examining our solution's advantages and disadvantages relative to this type of systems' state-of-art. Finally, the use of a large area walking VR draws the contrast with the other three technologies, tested in much more confined interaction areas, and allows for insight into how this can affect results.

1.2.2 Methods – Accuracy & Precision of the Head Tracking

To evaluate the accuracy and precision of the Kinect v2-based head tracking, we measured its position error and jitter in a 2.8 m by 2.8 m area, relative to a linear, four-camera ARTTRACK2 system (Advanced Realtime Tracking GmbH, Weilheim in Oberbayern, Germany), serving as the gold standard. Since the Kinect tracks human shapes and the ARTTRACK2 tracks IR markers, a medical upper-body mannequin with an IR marker placed in the space between the eyebrows (glabella) was used to provide a static tracking target for both systems. Measurements were registered by placing the mannequin in twelve intersection points of a 4x3 grid, with 60 cm x 80 cm spacing, at sitting (1.4 m), and standing height (1.7 m) (Figure 13). The Kinect v2 was positioned 66 cm away from the ARTTRACK2 camera plane, aligned with the center of the grid at 80 cm from its closest point, at the height of 2.2 m, tilted down 30°. The two closest corners of the grid were discarded as the Kinect v2 did not see them.

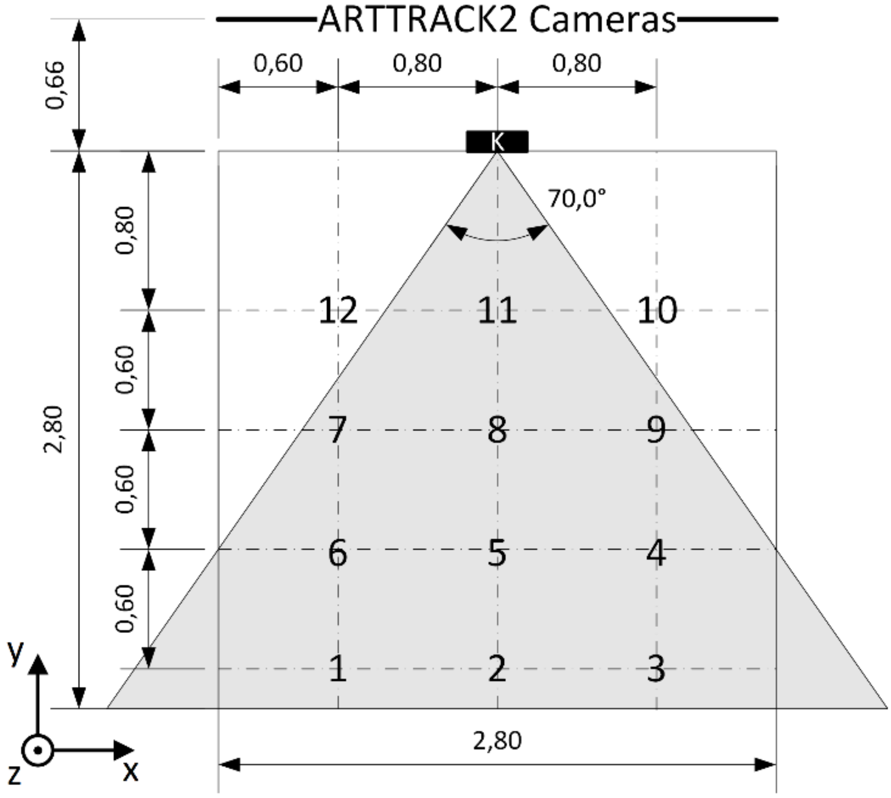


Figure 13: Placement of the Kinect v2, its field of view (simplified), the grid positions (1 to 12) and their spacing in meters.

The mannequin was sequentially positioned at the different points, and for each of them, its position was registered with both tracking systems for 5 seconds at a sampling frequency of 30Hz.

Because the two tracking systems had different reference frames, the transformation matrix (translation and rotation) between them was estimated using Procrustes analysis [69]. This process finds the linear transformation that minimizes the sum of squared errors between two configurations of points where there is a correspondence between points. The transformation was applied to the Kinect v2 data, converting all measurements to the same reference frame and eliminating any systematic difference (offset) between the estimated positions of the marker attached to the mannequin and the head joint identified by the Kinect v2.

Accuracy of the measurements (*e*) was calculated as the mean difference, during the five seconds, between the mean position of the IR marker estimated by the laboratory-grade tracking system (X) and the position of the head joint estimated by the Kinect v2 at sample *i* ($\tilde{X}_i$), as per (4), where N is the total number of samples. Jitter (*j*) was calculated as the standard deviation of the Kinect v2 measurements (5), where $\bar{X}$ is the mean value of $\tilde{X}$.

$$e = \frac{1}{N}\sum_{i=1}^{N}\left|X - \tilde{X}_i\right| \quad (4)$$

$$j = \sqrt{\sum_{i=1}^{N}\frac{\tilde{X}_i^2}{N} - \bar{X}^2} \quad (5)$$

1.2.3 Methods – Sense of Presence & Cybersickness

To determine the sense of presence and cybersickness that could be elicited by a surround-screen projection system using the KAVE software, we compared the reported experiences of a sample of healthy adults in two conditions. One after interacting with a low-cost implementation of a traditional CAVE system powered by the KAVE [37], the other after interacting with a modern HMD, HTC Vive. Additionally, we compared our results with the data from a previous study following the same procedure [66], in which participants interacted with two other VR setups, a laboratory-grade CAVE system, and an HMD-based walking system.

1.2.3.1 Instrumentation

Two experimental setups were used in this study; they are described in sections 1.2.3.1.1 and 1.2.3.1.2. Additionally, we described the setups tested by Borrego et al. [66] in sections 1.2.3.1.3 and 1.2.3.1.4. A comparison of the four immersive solutions' characteristics is provided in Table 2.

Table 2: Immersive characteristics of the four VR systems under study.

Immersive characteristics	Low-cost CAVE with KAVE	Head-mounted display – HTC Vive	Head-mounted display-based walking system	Laboratory-grade CAVE
Extensive	Visual, Auditive	Visual, Auditive	Visual, Auditive	Visual, Auditive
Surrounding	Monoscopic, 270° H x 128° V @ center	Stereoscopic, 360° H x 180° V, FOV 110°	Stereoscopic, 360° H x 180° V, FOV 100°	Stereoscopic, 270° H x 120° V @ center
Inclusive	Partially visual	Visual, Auditive, Wearable	Visual, Wearable	Partially visual, 3D Glasses
Vivid	4 x 1120 x 880, ≈4 p/cm	1080 x 1200	960 x 1080	4 x 1868 x 1200, ≈5.5 p/cm
Matching	Body tracking	Head tracking	Head tracking	Head tracking
Self-representation	Full body	None	None	Full body
Cost	≈5.000€	≈600€ + PC		≈200.000€

1.2.3.1.1 Low-cost CAVE powered by the KAVE software

A practical deployment of the KAVE software was done on a low-cost CAVE [37] of standard configuration, three walls, and floor (shown in Figure 14). Two walls on the sides and one at the front, describing 90-degree angles between them and with the floor. The walls are 2.8 m wide by 2.1 m tall, and the floor is 2.8 m wide by 2 m long, covering the area up to 2 meters from the front wall. Four Optoma GT1080 projectors (Optoma, New Taipei, Taiwan), with a 1080p image resolution and throw ratio of 0.5:1, deliver monoscopic front-projection over the four surfaces. The system provides a horizontal FOV of 270° and a vertical FOV of 128° at its center point, and has a resolution of 1120 x 880 pixels per wall with a pixel density of 4 pixels per cm. A Kinect v2 is centered over the front wall and tilted at a downward angle of 30°. It provides head tracking and human pose estimation, consisting of 25 joints, at 30 Hz, with an estimated latency of 66.66 ms [70]. The user's head position is used in real-time by the KAVE software, which applies it to the VE cameras' position, and continuously adjusts their projection matrices so that their frusta are perpendicular and framed with the respective projection surfaces. The projection of the images captured by the virtual cameras on the walls and floor generates motion parallax coherent with what the user would see if they were present and moving in the VE [37]. This sensor has a field of view of 70° and can track users up to 4.5 m. The tracking area of this sensor is presented in gray shading in Figure 13. The summary of the immersive characteristics of this CAVE – KAVE is presented in Table 2. One single computer with a Quad-Core 3.4GHZ processor, 8GB of RAM, and Radeon RX 580 8GB graphics card runs the system. For improved readability of the thesis, the "Low-cost CAVE powered by the KAVE software" is simply referred to as KAVE in the following sections.

Figure 14: Low-cost CAVE powered by the KAVE software.

1.2.3.1.2 Commercial head-mounted display – HTC Vive

The HTC Vive (Figure 15) was considered representative of modern HMD technology. It is an off-the-shelf HMD with a 1080 x 1200 per eye pixel resolution and a 110° FOV. It provides built-in head orientation and head position tracking through a laser-based inside-out tracking system, called Lighthouse. This tracking technology has been estimated to have an accuracy ranging from 0.19 cm to 1.22 cm, a position jitter of 0.03 cm, orientation jitter of 0.02°, and latency ranging from 4.44 ms to 22 ms [63], [71], [72].

In this study, the same computer was used to operate both the HTC Vive and the KAVE.

Figure 15: HTC Vive Head Mounted Display.

1.2.3.1.3 Head-mounted display-based walking system

The system consists of an Oculus DK2 (Oculus VR, Menlo Park, USA), a PlayStation Eye Camera (Sony Corporation, Tokyo, Japan) attached to it pointing upwards, and a 3.78 m by 5.7 6m pattern of 442 fiducial markers installed on the ceiling [66], depicted in Figure 16. The system runs on a laptop with an 8-core Intel Core i7 Haswell 2.50 GHz, 8 GB of RAM, and an NVIDIA GeForce GTX 860M 2GB. It uses the Oculus DK2 HMD native orientation tracking while the position tracking is done inside-out through the camera input, *i.e.,* the HMD tracks its position relative to the markers. When the user is at standing height, this system has a mean positional accuracy of 0.94 cm, a mean jitter of 0.10 cm, and a mean latency of 120 ms. The display specs are the standard for an Oculus DK2, a FOV of 100°, 960 x 1080 pixels per eye resolution, and stereoscopy.

Figure 16: Walking VR System, image adapted from [66].

1.2.3.1.4 Laboratory-grade CAVE

A CAVE, shown in Figure 17, with four projection surfaces (three 3.5 m wide by 2 m high walls and floor), was used (Barco N.V., Kortrijk, Belgium). Stereo images with an 1868 x 1200 resolution are provided by four projectors F35 AS3D WUXGA (ProjectionDesign, Fredrikstad, Norway). Images are back-projected on the vertical walls and mirror-projected on the floor. Stereoscopic immersion is provided through 3D glasses, the Crystaleyes 3 (StereoGraphics, San Rafael, USA), which have a constellation of infrared reflective markers attached to its frame. The glasses' position and orientation are estimated through four infrared ARTTRACK2 tracking cameras (Advanced Realtime Tracking GmbH, Weilheim in Oberbayern, Germany). The system runs on five high-end graphics computers equipped with Intel Xeon CPU ES-2620 @ 2.00 GHz, 16 GB of RAM, and NVIDIA Quadro 5000 graphics cards.

Figure 17: Laboratory-grade CAVE, image adapted from [66].

1.2.3.2 Procedure

The virtual environment from Borrego et al. [66] was used with some modifications. It consists of a supermarket aisle with two 4 m long by 2 m high shelves (6 racks) filled with 72 different sodas and corresponding price tags. The distance between shelves was adjusted from 1.5 m to 2.5 m relative to the original environment so that users could ambulate laterally. This VE provides a surrounding, generic scenario with lots of information in the participant action space, which was used for a visual search task. Two instances of the same Unity 3D VE were built for this study, one with the KAVE plugin that uses the projectors for imaging and the Kinect for tracking, another with Steam VR that uses the HTC Vive for imaging, together with two Lighthouse base stations for tracking. The physical space for both conditions was inside the 2.8 m x 2.8 m walled KAVE. Therefore, a 1.5 m x 1 m semi-transparent blue area with two footprints was added to represent the area where the user could freely move, as shown in Figure 14 and Figure 18. This served two purposes, ensure the KAVE users would not move forward (outside of the Kinect frustum) and avoid HMDs wearers from bumping into the real walls. Like in the study of Borrego et al. [66], a one-handed controller thumb analog joystick was used to virtually move the participant along (translation only) the longitudinal axis of the 4 m long supermarket aisle. As the aisle was longer than the experimental space (2.8 m), the above-mentioned blue area virtually moved together with the participant, remaining static relative to the participant's point of view.

The study was designed as a two-condition repeated measures experiment where every participant experimented once in each condition, KAVE, and HTC Vive. Participants were counterbalanced to avoid the order effect; thus, half of them were randomly chosen to experience one condition first and the other on a later date, the other half experienced the conditions in the opposite order.

At the beginning of the first session, participants were informed about the procedure, provided informed consent to participate in the study, and answered a brief questionnaire about their age, gender, education level, and video game experience, the latter in a 10-point Likert scale. Immediately before the experiment, they were guided to the center of the KAVE space. In the HTC Vive condition, they were assisted in setting up the interpupillary distance and putting the HMD. Then, participants had 5 minutes of free interaction with a basic version of the virtual environment, where the sodas and price tags were removed from the shelves. After this training time, the experiment started, and participants were asked to find the price of different items in the complete VE, which had the shelves filled with drinks and the price tags. Specifically, participants were given an item name, short description and were asked about its price. Once the correct price was answered, a new item was named, up to either a maximum of 5 items or 5 minutes had elapsed. Participants could move naturally inside the designated tracking area and use the joystick to move further along the aisle.

After the first assigned experimental condition, participants were asked to rate their cybersickness on a 7-point Likert scale. Then reported their sense of presence in the original Slater-Usoh-Steed Questionnaire [73] (SUS), a brief questionnaire with three questions rated on a 7-point Likert scale. Finally, participants answered a modified version of the Presence Questionnaire [74] (MPQ), a 21-item questionnaire, also rated on a 7-point Likert scale, as described previously [66]. They were then asked to return on a later date to repeat the same procedure with the other experimental condition.

Figure 18: The virtual environment used in the study.

1.2.3.3 Participants

Healthy participants over 18 years old with no motor or cognitive impairment were recruited for this study from the University of Madeira's faculty and student body. Thirty-two participants were recruited, of which 30 completed the study (Table 3, Current Study column). All provided informed consent before participation in the study. To guarantee this sample's equivalence with the participants in a previous study (Table 3, Previous Study column), we compared them across their characteristics. There were no significant differences between samples except for the years of schooling, with the participants from this study having slightly less (Median = 19) than the previous (Median = 21), U = 444.0, z = -2.55, $p < .05$.

Table 3: Characteristics of the participants from each study.

	Current Study	Previous Study [66]
Gender ratio	14 ♂ / 16 ♀ (0.87)	26 ♂ / 21 ♀ (1.24)
Age (years)	28.25 ± 5.6	28.1 ± 5.3
Years of Schooling (years)	19.83 ± 3.7	22.1 ± 4.4
Experience with videogames [1-10]	6.83 ± 2.8	5.8 ± 3.3

1.2.3.4 Analysis

Questions of the Presence Questionnaire were divided into four components: visual aspects, interaction, consistency with the real world, and subjective factors [66]. Wilcoxon's T-tests were used to compare the repeated measures results from the KAVE and HTC Vive conditions, whereas Kruskal-Wallis was used to find differences between the independent measures of the KAVE, CAVE, and Walking VR. The Mann-Whitney's U test, with Bonferroni's correction, was then used for the posthoc analysis between the KAVE and the other two systems. The significant difference level was set at 0.05, two-tailed. SPSS Statistics, version 22 (IBM®, Armonk, NY, USA) was used to analyze the data.

1.2.4 Results

1.2.4.1 Accuracy & Precision of the Head Tracking

The mean, standard deviation, and norm of these values are shown in Table 4 and Table 5.

Table 4: Mean ± standard deviation values of accuracy across the ten valid grid positions at the two tested heights for the Kinect v2.

Accuracy error (cm)	Kinect v2, Head @ 140 cm	Kinect v2, Head @ 170 cm
X axis	0.73 ± 0.57	0.63 ± 0.46
Y axis	1.07 ± 0.75	1.16 ± 0.78
Z axis	1.04 ± 0.71	0.29 ± 0.18
Norm	1.66 ± 1.18	1.35 ± 0.92

Table 5: Mean ± standard deviation values of jitter across the ten valid grid positions at the two tested heights for the Kinect v2.

Jitter (cm)	Kinect v2, Head @ 140 cm	Kinect v2, Head @ 170 cm
X axis	0.02 ± 0.01	0.02 ± 0.01
Y axis	0.03 ± 0.01	0.02 ± 0.01
Z axis	0.02 ± 0.02	0.01 ± 0.00
Norm	0.04 ± 0.02	0.03 ± 0.01

1.2.4.2 Sense of Presence & Cybersickness

This section first presents the experiment results measuring both cybersickness and the elicited sense of presence in two new conditions, the KAVE and HTC Vive HMD. Then we present the comparison of our results with the ones obtained by Borrego et al. [66], where the conditions CAVE and Walking VR correspond to a laboratory-grade CAVE and the authors' modified HMD.

1.2.4.2.1 Comparison between KAVE and HTC Vive

When reporting sickness levels, users felt significantly less sick in the KAVE (Median = 1) than in the HTC Vive (Median = 2), T = 22.5, z = -2.614, $p < 0.05$, r = -.48. Presence SUS levels were 2 points lower in the KAVE (Median = 14) than in the HTC Vive (Median = 16). This difference was found to be significant, T = 62, z = -3.221, $p < 0.05$, r = -.59. Presence Questionnaire levels in the KAVE (Median = 100) did not differ significantly from those in the HTC Vive (Median = 103), T = 202.5, z = -.325, p = .745. The same was true for all its components. For comparison with the previous study results, the means, standard deviations, and graphical representations of the response distributions can be seen in Table 6 and Figure 19.

Table 6: Mean and SD of the dependent variables measured in the KAVE and Vive study conditions, significant differences indicated.

Variable		KAVE	Vive	p
Sickness	**[1 - 7]**	1.73 ± 1.17	2.47 ± 1.63	p < 0.05
Presence SUS	**[3 - 21]**	13.23 ± 2.75	15.73 ± 2.90	p < 0.05
M. Presence Q.	**[21 - 147]**	99.30 ± 0.05	100.73 ± 9.23	.745
Visual	**[1 - 7]**	4.47 ± 0.53	4.47 ± 0.54	.386
Interaction	**[1 - 7]**	4.19 ± 0.52	4.21 ± 0.68	.691
Consistency	**[1 - 7]**	5.17 ± 0.97	5.44 ± 0.79	.147
Subjective	**[1 - 7]**	5.47 ± 0.84	5.57 ± 0.63	.897

1.2.4.2.2 Comparison between KAVE and CAVE & Walking VR

The data from the 47 participants of Borrego et al. [66] experiment, regarding presence from the SUS and Modified Presence Questionnaires, were reanalyzed for comparison with our experimental data. Their means and the standard deviation are presented in Table 7.

Table 7: Mean and SD of the subjective parameter measured in Borrego's et al. [66] *study.*

Variable		Walking VR	CAVE
Presence SUS	**[3 - 21]**	17.57 ± 2.42	14.61 ± 4.42
M. Presence Q.	**[21 - 147]**	116.68 ± 14.24	104.46 ± 22.74

Presence was significantly affected by the VR system used, as measured both by the SUS H(2) = 29.084, $p < 0.05$, and the MPQ H(2) = 24.767 $p < 0.05$. Follow up tests showed that presence SUS levels were significantly lower in the KAVE (Median = 14) than in the Walking VR (Median = 18), U = 174.5, z = -5.57, $p < 0.025$, r = -.63. This was also the case for Presence Questionnaire, where they were significantly lower in the KAVE (Median = 100) than in the Walking VR (Median = 116), U = 200.5, z = -5.27, $p < 0.025$, r = -.60. Presence SUS levels in the KAVE (Median = 14) did not differ significantly from those in the CAVE (Median = 15), U = 551, z = -1.62, p = .106. Neither did the Presence Questionnaire levels, the KAVE (Median = 100) did not differ significantly from the CAVE (Median = 102.5), U = 601, z = -.95, p = .344. These differences are shown in Figure 19.

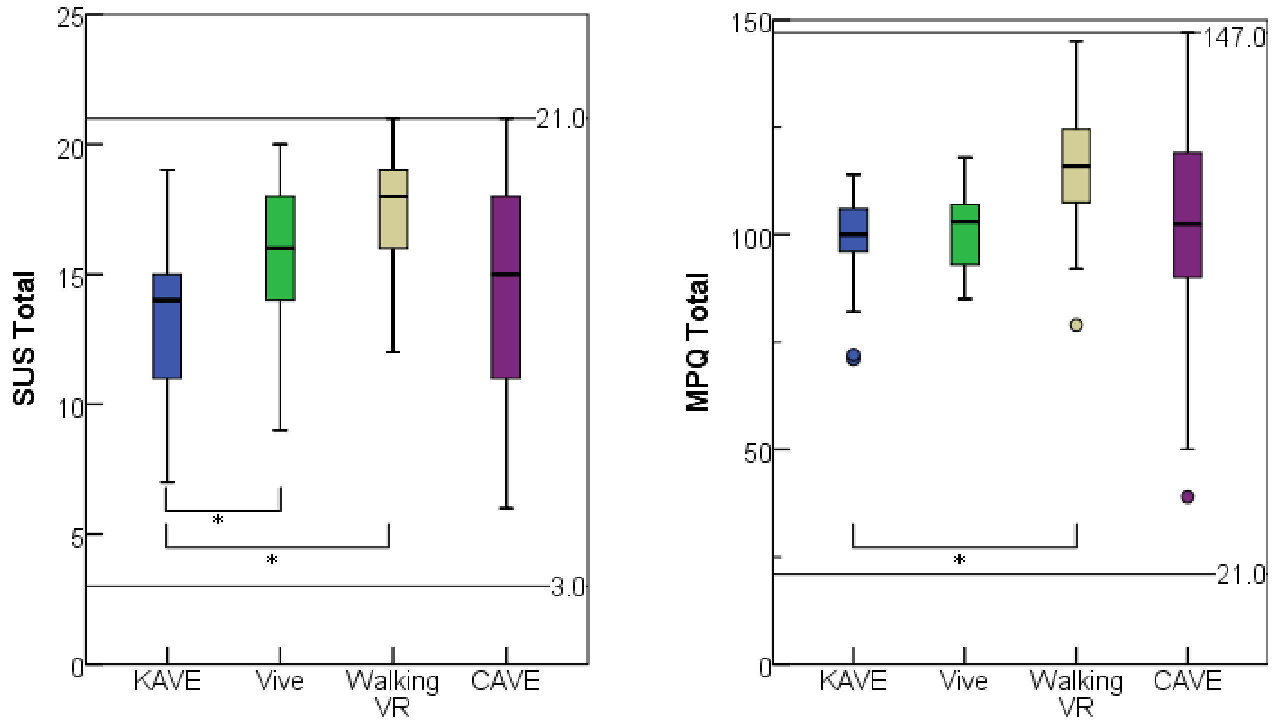


Figure 19: The Slater-Usoh-Steed and Presence questionnaire scores for each of the VR systems used. The asterisk indicates significant differences between the KAVE and the other systems.

1.2.5 Discussion

1.2.5.1 Accuracy & Precision of the Head Tracking

The accuracy of the Kinect v2 varied from 1.35 cm to 1.66 cm, slightly worse than the worst reported case for the HTC Vive, 0.19 cm to 1.22 cm; the same applies to its standard deviation that gives a notion about the difference of accuracy between positions on the grid (±1.18 cm). Increased accuracy at standing height can be observed, particularly in head height estimation, which had an average accuracy of approximately 3 mm. Anterior-posterior accuracy presented the worst results, and the mean accuracy on the horizontal plane (XY) was at centimeter-level, 0.63 cm to 1.16 cm. The measurements revealed a half-millimeter level jitter, which indicates a precise estimation of the head position in our KAVE, similar to the other systems [63], [66]. The Kinect head position accuracy results of one to two centimeters are consistent with what was reported by Otte et al. [55], who evaluated the Kinect v2 sensor tracking accuracy during landmark movements for possible clinical use. Considering that natural human head motion can average 2 cm when standing and looking at static images [57], the impact of a 2 cm offset and 1 mm jitter in head position estimation is most likely acceptable for perspective projection control.

1.2.5.2 Feeling of Cybersickness

Participants reported significantly lower levels of sickness in the KAVE than in the HTC Vive. This difference can be a consequence of two main characteristics of these systems. First, nausea can result from the latency between head movement and scene update (motion to photon) [29]. Although this was not measured in the study, in surround-screen displays such as the KAVE, scenes essentially do not change due to head rotations, and they are already projecting close approximations of the next frame around the user [29]. This increased immediacy of the system at the cost of rendering images that might be outside of the user's FOV is responsible for lowering the apparent latency to head rotation and therefore minimizing one of the leading causes of cybersickness. Second, motion sickness is caused by inconsistencies between visual and vestibular stimulation regarding the existence of movement [29]. According to the rest frame hypothesis [75], it is provoked when what is perceived as a rest frame provides cues that conflict with this resting state and do not match the users' physical inertial environment. Because the KAVE is only partially vision-inclusive in its immersive characteristics (usual for CAVE-like systems), it lets users see part of the real world in their peripheral vision, such as the top structure, ceiling, and back wall curtain. Hence, allowing participants to use the real world as a rest frame lowers the amount of visual information that contradicts their vestibular system. This result, together with the fact that presence was found lower for the KAVE, is consistent with work suggesting that presence is related to the degree to which a selected rest frame is determined by the virtual interface [75]. In the KAVE, the real-world rest-frame visual cues could have affected both, lowering the feeling of presence and suppressing sickness.

1.2.5.3 Sense of Presence

During our experiment, participants stated a lower sense of presence with the Slater-Usoh-Steed questionnaire [73] (SUS) in the KAVE system than with the HTC Vive HMD, with a large effect size being detected. However, the Presence Questionnaire differences were non-significant. This could be due to its questions being more numerous and broader in scope, leading to finer granularity results in a 21 to 147 score range than the SUS 3 to 21 score range. These results are consistent with Borrego et al. [66], which point to a lower sense of presence felt by the participants in a laboratory-grade CAVE than in a room-scale tracked HMD. In their case, this was identified on both questionnaires. In summary, the Walking VR was better than both CAVE systems, and the HTC Vive better than the KAVE in one of the two questionnaires.

Despite the high levels of presence promoted by all four systems, the question of why there are differences is important. Since the experimental task and virtual environment were the same across conditions, differences in the feel of presence must have originated in each system's immersive characteristics and how they relate with the particular visual search task used. Table 8 summarizes the main differences between systems in terms of immersive characteristics, where the dark shade represents the better and the lighter the worse. Considering significant differences found in this study, we can argue that differences were the cumulative effect of stereoscopy, full surround, no shadows and visual inclusiveness that the HTC Vive added that made it superior to the KAVE. Contrasting our experimental data with Borrego et al. [66] allowed the comparison between a KAVE-powered CAVE system and a professional laboratory-grade CAVE. The results revealed that the KAVE was no different from a laboratory-grade CAVE regarding the participants' feeling of presence elicited. Hence, it seems to indicate that just the stereoscopy and lack of shadows (due to back-projection) that the CAVE provided was not enough to differentiate it from the lower cost KAVE. While this seems to point that stereopsis was not a significant immersive characteristic driving presence in this task, we must keep in mind that the task took place outside of the personal space and that beyond the arms reach, the relative importance of binocular disparity is diminished. This result is very encouraging as both systems are entirely different regarding technical specifications and cost. They both extend to the visual and auditory senses (audio was not included in the experiment) and provide similar limited surrounding immersion. Their field of view (FOV) over the VE depends on the user's head position. Concerning inclusiveness, neither of them completely shuts down any senses. The KAVE has the rare advantage of not requiring the user to wear any device or tracker. Matching in the KAVE system is defined by the capability of the Kinect v2 sensor to track 25 joints of the user's body, complemented by the user self-representation inside the VE. This means that if the VE affords it, the user's body can interact with virtual elements. In this experiment, no interaction other than ambulation was considered. Hence, if other types of interaction were required, it could positively influence the KAVE feeling of presence. The KAVE – CAVE costs around 5.000€ to build with commercial-grade hardware (3.200€ for the four projectors, 1.200€ the computer, 600€ Kinect v2 and physical wall structure) and runs on our KAVE free and open-source software, thus avoiding software license costs. Although the estimated cost of the CAVE used in our study [76] (200.000€) is far from the estimated cost of the original CAVE and CAVE2 systems (2.000.000$ and 926.000$ [38], [45]), it is still considerably above alternative low-cost CAVEs, such as the CryVE (19.300€ [46]). Here we have pushed this limit by presenting an even cheaper solution for CAVEs, and despite its lower specifications, it produced comparable subjective responses from the experiment participants.

Table 8: Different immersive characteristics of the systems tested.

Main Differences	Low-Cost CAVE with KAVE	Head-mounted display – Vive	Head-mounted display-based walking system	Laboratory-grade CAVE
Stereo	No	Yes	Yes	Yes
Field of view	Variable	110	100	Variable
Surround	Partial	Full	Full	Partial
Shadows	Yes	No	No	Floor
Inclusiveness	Partial	Yes	Yes	Partial
Wearable	No	HMD	HMD	Glasses
Area	2.8 m x 2.8 m	2.8 m x 2.8 m	3.78 m x 5.76 m	3.5 m x 3.5 m
Control	Controller	Controller	Ambulation	Controller

Given the results, our second research question was answered, and the hypothesis was partially confirmed. In a simple VR action space visual search task, the KAVE can elicit similar presence levels to traditional CAVEs while keeping cybersickness low. However, results are not conclusive when comparing it to HMDs.

It is essential to ground the results and discussion to the task used in this study, as a different task could have led to different results. The two most important elements that defined the task were: First, the space in which it took place, virtual elements were always 1 m to 4 m away from the user. We speculate that longer distances would not produce different results between the systems. Keeping the virtual elements far enough from the user limits interaction and how they are perceived, as the binocular disparity is decreased. Second, the VE's short and linear nature, a 4 m long narrow aisle, only required exploration of the participant's field of regard. Hence, no mental representations of the environment had to be built, like when navigating a maze.

These results build the knowledge about our KAVE implementation and help us understand under what conditions the use of this system might be desirable. Our KAVE is an implementation of a low-cost CAVE and therefore maintains the specific advantages of CAVE systems over HMDs. Some of these advantages are no encumbrance of wearing a display mounted on the head, the use of peripheral vision due to large FOV, and the possible addition of real elements mixed with the virtual, such as a vehicle cockpit or control console of a machine [77]. Also, in the case of experiments required to be shared by multiple people seeing the same virtual elements or when wearing an HMD is not practical. The use of such a system is justifiable in the same way a CAVE is, except for experiences that require virtual elements to be present in the user's personal space. Meanwhile, in the context of VR, the addition of whole-body tracking offers an increased interaction potential, which was not studied in this work. However, we expect these features to be adopted to develop VR exergames (highly immersive exercise video games) where Kinect-like sensors can be used for its intended purpose of a game controller and natural user interface (NUI), such as in [78], [79].

1.2.6 Conclusions

In this section, the results from two experiments involving the KAVE are presented and discussed. We intended to evaluate a Kinect v2 sensor's accuracy in head-tracking and validate a KAVE implementation of a CAVE in eliciting the feeling of presence in a VE exploration task constrained to the action space. This CAVE implementation was done using the KAVE projection management software introduced before [37] and low-cost, commercial-grade hardware. This system uses a Kinect v2 sensor for body tracking, and therefore is non-intrusive, and no equipment needs to be worn by users. In the first experiment, we found that the Kinect v2 is accurate enough to provide real-time head position estimation to the KAVE software, with a 1.66 cm average error, and it can be used to track the user's head while driving a VR experience. In the second study, the KAVE was tested for its cybersickness and presence-inducing capability directly against an HTC Vive in a repeated measures experiment and indirectly against a laboratory-grade CAVE and a room-scale Oculus DK2 in an independent samples study. The experimental task consisted of a local visual search of the VE at 1 m to 4 m distances, where the importance of binocular vision is reduced and featured no interaction. Results show that, while both CAVEs produced lower feelings of presence than HMDs, the KAVE was no different from a laboratory-grade CAVE. Together with the lower cybersickness produced, this result shows that our immersive surround-screen VR system solution (KAVE) is a feasible alternative to the traditional CAVEs in research when dealing with similar conditions. Without losing the feel of presence, relatively to other CAVEs, it adds three main advantages: it provides an opportunity to the budget-constrained scientific community due to the low implementation cost, the user does not have to wear any apparatus, and it supports full-body tracking. While the first advantage is convenient, the other two can be instrumental in further lowering the ease-of-use of such systems [36], an essential aspect of technology adoption in the elderly [19]. Beyond those advantages, its versatility to other configurations besides CAVEs, such as interactive VR walls, floors, and screens, makes it a tool that facilitates access, prototyping, trial, and improvisation of what are traditionally permanent and complex installations.

In this thesis's context, these two experiments provided us with valuable data concerning the use of this technology to develop exergames. First, exergames powered by the KAVE, using the Kinect v2 sensor for tracking, can provide accurate and precise head position estimations. Information that can be used for real-time posture evaluation; and opens the possibility of developing highly immersive interactive experiences featuring motion parallax, although this might not be desired with an elderly population. Finally, visually setting the games in the players' action space seems to be an appropriate design choice to promote the feeling of presence and avoid cybersickness.

1.2.7 Limitations

The KAVE software evaluation was limited by being tested in only one setup, a low-cost CAVE made of consumer electronics grade and gaming devices. The same software could have achieved better results if used in the laboratory-grade CAVE (retro-projection with higher resolution projectors) or used the ARTTRACK2 system instead of the Kinect v2 camera. A similar limitation is valid for the HTC Vive HMD, which was not used up to its full potential. The HTC Vive affords much larger tracking spaces than what was used and can be made wireless since the Vive Wireless Adapter is available. Given this wireless capability and a tracking space large enough to accommodate the whole VE, to not depend on a joystick for navigation along the aisle, we speculate that a higher level of presence, comparable to or higher than those of the HMD-based walking VR system, could be achieved.

Concerning the study design, using two samples of users for four systems, one sample for each pair of systems, is a limiting factor. In ideal conditions, either one sample per system should have been used, or one sample should have tested all systems. However, that was not possible, and the data analysis was performed to minimize interference. The differences between KAVE and HTC Vive were analyzed first in a classical within-subjects study. However, in comparing the KAVE with existing data from the CAVE and Walking VR HMD, the latter two shared the same participants, which could have had an effect. The randomization of conditions mitigated the order effect that could have been present in comparing the KAVE to the CAVE, given that half the participants only experimented with HMDs after. Additionally, this experiment's participants are representative of a healthy young adult population, limiting our conclusions from generalizations to the target population of this thesis.

2 Designing Large Projection Exergames for Elderly Fitness – Computerization of Fitness Assessment, Evaluation of Interaction and Content Preferences

This section leverages the technology for full-body interaction with large projection displays to design customized exergames to promote fitness in older adults. It is comprised of three main steps, with corresponding subsections. First, we evaluate if low-cost motion capture with the Kinect v2 camera can deliver reliable evaluations of senior fitness, a contribution for automated fitness assessment, and validate this sensor's viability in capturing critical motions, later used in our exergames. Second, we study our target population preferences in interacting with an instance of our technology, a large floor projection. Lastly, we describe a detailed use case of applying human-centered design methodologies in the gamification of fitness training routines, from which the resulting exergames are evaluated in the subsequent section of the thesis.

2.1 Automating Senior Fitness Testing through Gesture Detection with Depth Sensors[3]

2.1.1 Introduction

The older adult population is diverse, with a broad range of needs, deficits, and co-morbidities. This makes exergame customization to each individual a fundamental strategy in maximizing their health benefits, in opposition to a generic and poorly personalized experience [9], [81]–[84]. When setting up preventive and rehabilitation exercise programs, it becomes essential to assess physical fitness [85]. Physical performance testing provides baseline data, which educates participants on their status relative to their cohorts; helps develop individual exercise prescriptions by targeting the identified functional limitations associated with poor health; allows the evaluation of progress and monitoring; and motivates participants by setting reasonable goals [85]. In older adults, the assessment of multiple physical function dimensions is commonly done using Senior Fitness Tests (SFT) [86]. These tests assess several physical parameters such as muscle strength, agility and dynamic balance, and aerobic endurance. Parameters that have high impact on people's ability to live independently, which according to Fleg et al. "*is dependent largely on the maintenance of sufficient aerobic capacity and strength to perform daily activities*" [87]. The administration of SFT requires particular training and elevated levels of concentration by a single test administrator who needs to simultaneously guide the elderly through the tests, evaluate the quality of their movements, keep test scores, and ensure safety.

Advances in information and communication technologies (ICT), specifically in the area of affordable and reliable motion tracking technology based on depth sensors, create new opportunities in the field of kinematic-based assessment such as SFT. These systems can now be adapted to assess the quality of movement execution and measure task performance in a non-invasive manner, allowing movement kinematics and their quality to be quantified objectively and reliably through machine-based metrics. Such an approach reduces the workload from the health and fitness professionals, but it also allows administration by non-experts and increases the results' accuracy, with a potential for home and autonomous use. This section presents a system for the automated administration of SFT that uses the Kinect v2 sensor for body tracking and gesture detectors to evaluate lower body strength, agility and dynamic balance, and aerobic endurance.

[3] The content of this section has been published in the proceedings of the IET International Conference on Technologies for Active and Assisted Living (TechAAL 2015) [80]

2.1.2 Related Work

In human body motion tracking, the most relevant sensing technologies are marker-based optical systems, inertial and magnetic systems, and marker-less infrared systems. Marker-based optical systems are the most precise, with most of their usage being in motion capture applications. However, the requirement of an elaborate and expensive multi-camera setup plus the use of markers distributed over the body makes them impractical for domestic or low-cost applications. Motion sensing through inertial or magnetic systems uses accelerometers, gyroscopes, or magnetometers attached to the body. These systems have the advantage of being able to work independently from an external setup. However, their main disadvantages are the presence of drift errors in the measurements and the fact that the sensors need to be "worn" by the user, which can be cumbersome or impractical. These systems were used in [88], who developed a novel automatic tracking device for weight training and calisthenics. The system uses a 3-axis accelerometer and 3-axis gyroscope installed in an armband. It automatically segments periods of activity, recognizes the exercise, and counts the repetitions, presenting high accuracy rates for both offline and online feedback. A different study compared the adhesion of the elderly to a fall prevention program when it was done through wearable sensor exergames instead of the traditional instructional booklet approach; results suggested that adherence is improved in exergames through increased levels of engagement [89]. In [90] the authors presented the lessons learned from developing games for stroke rehabilitation using the Nintendo Wii™ inertial remote and discussed what makes games playable, fun, challenging, and useful from a therapeutically perspective. However, the body of research on assessing fitness indicators themselves is much more reduced than for exergames. One exception is evaluating standing balance using the Wii Balance Board [91][92]. Comparing Balance Board-based exergames' scores and fitness indicators [93] showed significant correlations between game scores and aerobic fitness.

Marker-less infrared systems present the lowest cost alternatives. The adoption by Microsoft of this technology in their mass-produced motion controller Kinect has contributed to the widespread availability of such sensors. These specific devices can estimate the human body poses by analyzing the 3D depth information from a scene, also requiring minimal setup and no markers. The main disadvantage is the lower accuracy of the measurements when compared with the marker-based optical systems. Still, Kinect v1 is accurate enough to be used in rehabilitation [94], and improved accuracy has been shown for Kinect v2 [95]. These devices have been widely used in research, for example, for designing full-body interactions in exergaming for older adults [96]; for motion tracking in gait evaluation [97]–[99]; as guidance, correction and scoring prototype for shoulder abduction exercises [100]; for gesture detection associated with muscle and joint pain, common in older adults [101]; or as a tool to assist in the medical diagnosis and monitoring of Parkinson's disease through movement analysis [102].

There is, however, little work on the assistance or automation of SFT. To our knowledge, only one case has used such an approach [103], exploring the feasibility of a home-based solution through the combined use of a Kinect and inertial sensors to detect the correct performance of the SFTs. Hence, this gap in the application to fitness assessment in the elderly combined with the existance of higher resolution depth sensors (Kinect's v2) which provide a more accurate estimation of 25 skeleton joints [95], offers new opportunities for innovation.

2.1.3 Methods

2.1.3.1 Fitness Tests

The Senior Fitness Test (SFT) [86] is a valuable tool for evaluating and identifying risk factors, planning and assessing training programs, educating, setting goals, and motivating clients to be more active. The SFT is designed to be easy to administer by health and fitness professionals in standard community settings without extensive time (20-30 minutes), equipment, or space requirements. In this study, we considered the following domains and subtests of the SFT:

1. *Lower body strength* is an essential aspect of muscular fitness concerning health, namely, in retaining proficient functioning in most daily activities, especially with advancing age. It can be measured through the 30-second Chair-stand Test consisting of counting the number of times a participant can fully stand and sit from a chair, with the arms crossed, during a 30 seconds interval [86].
2. *Aerobic endurance* or *Cardio-respiratory Fitness* (CRF) is another critical component of health-related fitness. Low levels of CRF have been associated with a markedly increased risk of premature death, while high levels are associated with higher levels of habitual physical activity and many health benefits [85]. This fitness component is assessed with the 2-minute Step Test [86]. The test consists of having the participant step in place for 2 minutes, raising the knees to a height marker placed halfway between the knee and hip levels. The number of times each knee reaches the target height is the score of the test.
3. *Agility* (the ability to move the body and change direction quickly) and *dynamic balance* (maintaining postural stability while moving) are good predictors of recurrent falls and independent living [104]. It can be measured with the 8-foot Up-and-go Test [86]. In this test, starting from a seated position, the user stands on a "go" signal, walks 2.4 m, turns around, walks back to the chair and sits. The participant practices once and then perform two trials. The score is the fastest time to the nearest tenth of a second of the two trials.

2.1.3.2 Setup

Our system was developed using the Unity 3D game engine (Unity Technologies, San Francisco, USA), using Kinect's v2 plugin for Unity, Kinect's SDK, and its API. The Kinect v2 – an RGB-Depth sensor capable of tracking 25 body joints, per person, of up to 6 people simultaneously at a frequency of 30 Hz – was placed horizontally (no tilt angle) at the height of 0.74 m and facing a wall at 4.22 m distance (Figure 20). The Visual Gesture Builder from Kinect's SDK, which uses AdaBoostTrigger and RFRProgress detection technologies, was used to train gesture detection databases.

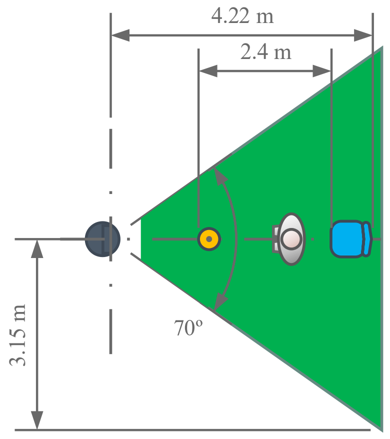


Figure 20: Top view of Kinect's v2 tracking area in green and chair and marker placement.

2.1.3.3 Gesture Detection Based Assisted SFT

In total, five gesture detectors were trained to compute the scores of the 3 SFT previously described, two discrete and three continuous detectors, as follows:

1. For the 30-second Chair-stand Test, two gesture detectors were trained: a discrete detector of “arms crossed” to monitor the correct arms pose and the continuous progress “sit-stand” detector to trigger the counting of full stands.
2. To measure the 2-minute Step Test score, two continuous “step” progress gesture detectors were used, 1 for each leg, to monitor and count the number of steps.
3. The 8-foot Up-and-go Test timing was attained via the “sit” discrete detector acting as a trigger to start and stop.

2.1.3.3.1 Discrete Detectors

The discrete gesture detectors use Adaptive Boosting (AdaBoost) [105] to construct a “strong” classifier as a linear combination of “weak” classifiers. The classifier output is a binary detection result of a gesture and its confidence level. For our system, two of these detectors were built:

1. “Arms crossed” detector, trained (ignoring the lower body: knees, ankles, and feet) with 12 videos of 3 different people. Totaling 34658 labeled frames (obtained at 30 Hz) with a ratio of 1:1.132 positives to negatives.
2. “Sit” detector, trained (ignoring the arms: elbows, wrists, and hands) using eight videos of 1 person. Totaling 21300 labeled frames and a ratio of 1:0.319 positives to negatives.

2.1.3.3.2 Continuous Detectors

Built with Random Forest Regression [106], the continuous detector’s output is a continuous measure of progress according to a regression model constructed from training data. For our system, three continuous detectors were built:

1. “Sit-stand” progress detector, trained (ignoring the arms) with 12 videos of three people. Totaling 36126 labeled example frames.
2. Two “step” progress detectors (one for each leg), both trained (ignoring the arms) with four videos of 3 people. Totaling 10974 and 10976 labeled examples for the left and right steps, respectively.

2.1.3.3.3 Automated SFT Score Computation

2-minute Step Test Automated Scoring: The system compares each step’s progress with two thresholds, one for detecting the knee raised to the target height, the other for detecting the foot on the ground. It requires the alternate triggering of right/left steps. Every time the right knee is detected above the top threshold and the opposite foot below the bottom threshold, the step count is incremented. From visual inspection of training data, the detection thresholds were set to the situations mentioned above' regression values, respectively 0.8 and 0.1.

30-second Chair-stand Test Automated Scoring: The output of the “sit-stand” progress detector is compared against two thresholds, a bottom one representing the regression output of a seating pose and a top one representing the output of a standing pose. Every time the detector outputs alternate from sit to stand, the stand count is incremented. For this test, two different sets of thresholds were established. The first, 0.25 and 0.7, set according to the values given by the “sit-stand” detector when the subjects were considered to be on the limit of being correctly seated/standing (by SFT definition) using the training data. The second set, 0.584 and 0.8, was defined after the experiment using an expert’s evaluation as ground truth for what were considered valid sit/stand poses.

8-foot Up-and-go Test: In this chronometry test, our system measures the time a subject takes to get up, walk 2.4 m forth and back, and sit. With the user seated, the timer starts counting as soon as the sit gesture is no longer detected (user got up from the chair), the system then tracks the subject’s waist, and once it detects that the user walked at least 2.4 m forward, a validity variable is activated. The timer is stopped when sitting is once again detected if the 2.4 m walk was considered valid.

2.1.3.4 Participants

This study included 22 volunteers (15 females), 65±5.6 years old, who gave their informed consent. Participants were recruited from a physical activity program in a senior center in Funchal, Portugal. The inclusion criteria were: (1) to be a community-dwelling older adult, aged 60 to 85 years old; (2) being able to walk independently and autonomy to perform everyday activities; and (3) absence of reported medical problems considered contraindications to exercise.

2.1.3.5 Experimental Protocol

Participants answered demographic and fitness questionnaires before the experimental session. Subsequently, they executed the three previously described SFT in the following order: 30-second Chair-stand Test; 8-foot Up-and-go Test; and 2-minute Step Test. These were administered and scored in real-time as follows:

- For the 30-second Chair stand test, a chair was placed against a wall facing the Kinect in the center of its FOV.
- The 2-minutes Step test was performed with the participant stepping in place centered in Kinect's field of view at a distance of around 3 m from the sensor.
- In the 8-foot Up-and-go test, a chair was placed as in the 30-second Chair stand test. A cone was placed 2.4 m from the chair's front edge. The administrator accompanied the participant by his/her side through the trial to ensure his/her safety but without getting between him/her and the sensor.

In this study, to maximize the assessment procedures' consistency, all the assessments were performed by the same trained fitness professional.

2.1.3.6 Data Analysis

Collected SFT score data consisted of three datasets: (1) "Traditional" – the standard assessment done by the professional live on-site; (2) "Recordings" – a posterior assessment done by the same professional five weeks later by carefully replaying the Kinect recorded data (in a blinded and randomized fashion); and (3) "Automated" – the assessment done by our proposed system when replaying the Kinect recordings (emulated as real-time data). The expert tagged video data for positives (correct movement patterns) and negatives (incorrect movement patterns). These data were then compared using Mathworks software MatLab R2013b, with the automated system's detection outputs.

A within-subjects design was used to compare the conditions. Normality of the distributions of differences was assessed using a Kolmogorov-Smirnov test. Because data deviated from normality, nonparametric statistical tests were used. For assessing the overall difference between assessments, a Friedman test was used on each dependent variable. For further pairwise comparisons, the Wilcoxon's T matched-pairs signed-ranks test was used. For all pairwise comparisons, a Bonferroni correction was used to account for the number of comparisons. Additionally, the inter-rater reliability was measured via Intraclass Correlation. All statistical testing was done using IBM software SPSS Statistics 22.

2.1.4 Results

Here we present a comparative analysis, including the overall scoring performance metrics and a comparison between automated movement detection and expert tagged data. The later comparison resulted in the identification of correct detections (True Positives – TP), correct non-detections (True Negatives – TN), incorrect detections (False Positives – FP), and incorrect non-detections (False Negatives – FN). The TP and TN detection rates represent the ratio of correct detections to the respective total number of positives or negatives. FP and FN detection rates are the ratios of incorrect detections to the total number of both positives and negatives.

2.1.4.1 30-second Chair-stand Test

In the 30-second Chair-stand Test case, the system was exceptionally tested with two different sets of parameters. Here, "Laboratory trained system" refers to the thresholds obtained from laboratory training data and "Expert trained system" to the thresholds derived from the expert tagged experimental data, as explained in section 2.1.3.3.3.

The scores, assessed as the number of repetitions, are presented in Table 9. The number of counted full stands did not differ across assessments (Traditional: 18; Recordings: 19; Laboratory Trained System: 18.5; Expert Trained System: 19), $\chi^2(3) = 5.723$, $p > .05$. The Intraclass Correlation Coefficient, for absolute agreement definition, was ICC(3,1) = 0.858 and ICC(3,4) = 0.960, high values meaning that the different measurement methods agree and are reliable between themselves.

Table 9: 30-second Chair-stand Test scoring for the different assessment methods.

Method	Median	Percentile 25	Percentile 75
Recordings	19.00	15.75	24.00
Traditional	18.00	16.00	21.00
Lab. Trained Sys.	18.50	15.75	20.50
Exp. Trained Sys.	19.00	15.75	23.25

For gesture detection, the positive gestures were fully seated and fully standing. Negatives were all other gestures, including positives of the opposite detector, *i.e.,* a positive sit is a negative stand, and a positive stand is a negative sit. Results of detection rates for both the Laboratory Trained System and the Expert Trained System show a high detection performance, above 95% for the Laboratory Trained and 98% for the Expert Trained, with false detections never exceeding 2% (Table 10).

Table 10: 30-second Chair-stand Test detection rates for both the Laboratory Trained and Expert Trained systems.

		Lab. Trained Sys.		**Exp. Trained Sys.**	
		True %	False %	True %	False %
Stand	**Positive**	98.39	0.45	98.62	0.45
	Negative	95.16	0.78	98.90	0.67
Sit	**Positive**	95.98	0.45	99.55	0.45
	Negative	97.50	2.01	97.73	0.22

2.1.4.2 2-minute Step Test

The next results encompass data from 21 (out of 22) subjects as one dataset was corrupted and excluded from the analysis. For this test, the number of counted complete steps differed significantly between methods used for counting (Traditional: 97; Recordings: 101; Automated: 96), $\chi^2(2) = 13.156$, $p < .05$ (Table 11). Interestingly, pairwise comparisons showed that the number of counted complete steps was significantly higher in the recordings than in the traditional measurements, $p = 0.001$, $T = 8.50$, effect size $r = -0.498$. However, no significant differences between the recordings and our automated system were found, $p = 0.109$, $T = 15$, $r = -0.248$, as well as between the traditional and the system method, $p = 0.359$, $T = 64.50$, $r = -0.142$. Average differences of about four steps (3.9%) can be seen between assessment methods. For the absolute agreement definition, Intraclass Correlation was measured at ICC(3,1) = 0.790 and ICC(3,3) = 0.919, indicating an agreement between the measuring methods.

For individual detector performance, positive gestures consisted of detecting knee elevation (one detector for each leg) up to the target height, and negatives all remaining cases. Step detection performance was above 95%, with less than 2% false detections (Table 12).

Table 11: Descriptive statistics for the different assessment methods and different tests.

	2-minute Step Test (nr rep)			8-foot Up-and-go (all Trials) (s)			8-foot Up-and-go (Score) (s)		
Method	**Median**	**Percentile 25**	**Percentile 75**	**Median**	**Percentile 25**	**Percentile 75**	**Median**	**Percentile 25**	**Percentile 75**
Recordings	101.00	91.50	116.50	4.70	4.13	5.10	4.65	4.08	5.03
Traditional	97.00	88.50	112.50	4.80	4.20	5.10	4.75	4.15	5.00
Automated	96.00	87.00	112.00	3.95	3.53	4.48	3.90	3.30	4.40

Table 12: 2-minute Step Test detection rates for the automated system.

Step		True %	False %
Right	**Positive**	98.67	0.16
	Negative	95.85	0.66
Left	**Positive**	96.15	0.55
	Negative	97.60	1.90

2.1.4.3 8-foot Up-and-go Test

As explained in section 2.1.3.1, the 8-foot Up-and-go Test requires the execution of two individual timed trials, where only the fastest is considered for the assessment. This enables us to analyze the results in two different ways. One where all the 44 timed trials for the 22 participants are considered. The second where only the fastest trial per subject is considered.

2.1.4.3.1 All trials

Measurements of execution time for each trial of the 8-foot Up-and-go Test differed significantly with the method used for timing (Traditional: 4.80; Recordings: 4.70; Automated: 3.95), $\chi^2(2) = 71.268$, $p < .05$ (Table 11). The average differences between traditional and computer-mediated assessment amount to approx. 0.7 sec., which represents a 15% difference. Time measured did not significantly differ between the traditional measurements in situ and the one performed by inspection of the recordings, $p = 0.077$, $T = 173$, $r = -0.188$. However, both “traditional” and “recording” were significantly higher than our system, $p < 0.001$, $T = 0$, $r = -0.619$, and $p < 0.001$, $T = 0$, $r = -0.618$ respectively. The Intraclass Correlation for absolute agreement definition was ICC(3,1) = 0.661 and ICC(3,4) = 0.854, whereas using the consistency definition was ICC(3,1) = 0.957 and ICC(3,4) = 0.985. The low correlation values for the absolute agreement definition indicate that timing methods did not provide reliable absolute measures between themselves. However, according to the consistency definition, very high Intraclass correlation values indicate that the methods were precise, although inaccurate due to the system’s delay.

2.1.4.3.2 Fastest trial

The fastest measured time for performing both trials (for each participant) of the 8-foot Up-and-go Test also differed significantly with the method used for timing (Traditional: 4.75 s; Recordings: 4.65 s; Automated: 3.90 s), $\chi^2(2) = 35.877$, $p < .05$ (Table 11). The measured time did not significantly differ between the traditional measurements in situ and the one performed by inspection of the recordings, $p = 1.000$, $T = 60$, $r = 0$. However, and consistent with the previous data, time was found significantly higher in both “recordings” and “traditional” methods than by our system, with $p < 0.001$, $T = 0$, $r = -0.622$ in both cases. Intraclass Correlation for the absolute agreement definition was ICC(3,1) = 0.674 and ICC(3,4) = 0.861, when calculated using the consistency definition was ICC(3,1) = 0.962 and ICC(3,4) = 0.987. Values are identical to the previous case.

2.1.5 Discussion and Conclusions

In this section, we developed and evaluated a low-cost system to support health and fitness professionals in assessing physical function in the elderly population, with the potential to be used autonomously and at home by non-experts. Not many researchers have addressed this issue, particularly in real scenarios and with end-users [103]. We presented a comparative study with 22 elderly community-dwelling participants using three standard SFT performed in a real-world scenario. The results confirmed that the proposed system could score as accurately as an expert in two of the three tests. The 8-foot Up-and-go Test presented a systematic error, underestimating time due to our experimental setup. Our system would only measure the time to perform the actions and not the instructions' reaction time. Thus, this error would not exist if the system itself had given the go signal. Low Intraclass Correlation for the absolute agreement but very high consistency values support the possibility of a systematic delay. The system's overall performance in gesture detection was very high, with TP and TN rates over 95%. This individual gesture detection results are better (3-6%) than what was presented in [101], and while the comparison is hard to make, for exercise repetition counting, they are similar to what was presented in [88].

In the case of the 30-second Chair Stand Test, we observed that despite very high rates for TP and TN with the Laboratory Trained System (≈95%), results were further improved with the Expert Trained System (≈99%). This sensitivity of score accuracy relative to the threshold values, confirmed by our data, shows the importance of using training data collected in realistic settings instead of laboratory conditions. This adaptation can be done by a large enough training sample in real scenarios (which is very demanding and time-consuming) or alternatively by introducing a system calibration phase immediately before each individual subject assessment. Our system's scores were not significantly different from the traditional assessment done in situ or by inspecting the recordings. This was corroborated by a high Intraclass Correlation Coefficient (ICC(3,4) = 0.960), indicating a high absolute agreement.

We identified a higher rate of FN for the left step detector in the 2-minute Step Test. The main contributor to this asymmetry was the occlusion introduced by a height marker the test administrator used in front of the subjects' left knee. A significant difference was found between the traditional assessment and that of the recordings. This could be attributed to the high attention levels this test requires from the administrator when performed live without a tally counter. It is challenging for a test administrator to evaluate every single step validity correctly (at rates sometimes over 2Hz), count them mentally, and ensure the participant's safety, further supporting the need for a system such as the one proposed here. In fact, our system produced scores that did not differ from those of the expert, indicating an accuracy equivalent to that of the post-analysis of the video recordings.

Our results support the use of the proposed technology and methods for automating senior fitness assessment. This can be viewed both as a standalone contribution and a proof of concept for its integration in exergames' development. In the first case, these methods could assist in traditional SFT assessment scoring, as they were in this section. Taken a step further, by enhancing them with guidance, there is potential to use them in remote and completely automatic assessment of seniors, such as at-home scenarios. Alternatively, the high accuracy obtained in the detection of repeated fundamental body movements, an established form to evaluate fitness, confirms that these methods are a viable option to be integrated into the development of exergames, as games could use the validated detectors to mediate exergame interaction, or possibly to evaluate fitness parameters during gameplay.

2.2 EVALUATING BODY TRACKING INTERACTION IN FLOOR PROJECTION DISPLAYS WITH AN ELDERLY POPULATION[4]

2.2.1 Introduction

With older age, visual perception is commonly negatively affected [108], and the effects of sedentary lifestyles become more prominent. A computer system that could alleviate such problems through large dimension displays and motion tracking interfaces could prove advantageous. More concretely, applications targeting physical fitness would provide extensive health benefits in older adults [109]. In this context, floor projections facilitate matching between the physical activity mechanics and the display. One of its main advantages for interaction is the possibility to use large existing surfaces, thus, avoiding spatial constraints [107] and the limitations of using a screen in terms of image size and interaction modalities. Also, by using a floor projection with a one-to-one mapping, we can guarantee that players' attention is focused on the same space they interact with. This is important as it lowers the need for spatial awareness that other means, such as wall projection or TV, would create, and splitting attention between those displays and the play area would increase the risk of fall or collision with obstacles

Low-cost body-tracking sensors for gaming consoles have made it possible for gesture detection to be present in millions of homes. Sensors like the Kinect v1, of which more than 24 million units were sold by Feb. 2014 [110], and Kinect v2, having 3.9 million units bundled and sold along with Xbox One consoles by Jan. 2014 [111, p. 2]. The widespread access to this technology opens the way for more *user-natural* ways of interacting with computing systems. Natural user interfaces (NUI), where users act with and feel like *naturals*, aim to reflect user skills and take full advantage of their capacities to fit their task and context demands from when they start interacting [112]. As shown in the previous section, the body tracking sensors' unique interface capabilities also provide exciting possibilities for the automatic monitoring of health-related problems through kinematic data analysis. For example, automated systems for assessing fitness indicators in elderly [80], [103], automatic exercise rehabilitation guidance [100], or diagnosis and monitoring of Parkinson's disease [102].

The coupling of body tracking depth sensors, such as Kinect and projectors, enables systems to track the user movements relative to the sensor and the mapping of the projection surfaces. In a well-calibrated system, where the transformation between the sensor and projector is known, this allows for immersive augmented reality experiences, such as the capability of augmenting a whole room with interactive projections [33].

This section presents the combination of floor projection mapping with whole-body tracking to provide two modalities of body gesture NUIs in controlling a cursor. One modality is based on the feet position over the display, while the other uses forearm orientation (pointing). We assessed the interfaces with an abstraction of two common interaction tasks, the point-and-click and drag-and-drop, on an elderly population sample. The differentiation was done by evaluating the systems in terms of usability, perceived workload, and performance. This work was an initial and essential step in developing a usable system designed for the elderly to keep an active lifestyle through adaptable exergames. This experiment's results help to understand best practices for gesture interface in such system and exergame interaction design.

2.2.2 Related Work

While gesture-based interaction is not a requirement for a NUI, it is an evident candidate for developing such an interface.

An area where several in-air gesture interfaces have been proposed is in pan-and-zoom navigation control. In [113], the authors investigated the impact three interaction variables had on task completion time and

[4] The content of this section has been published in the proceedings of the 3rd International Conference on Physiological Computing Systems - Volume 1: PhyCS [107]

navigation overshoots when interacting with a wall-sized display. The variables were: uni- vs. bi- manual, linear vs. circular movements, and the number of spatial dimensions for gesture guidance (in zooming). Panning was controlled by ray casting the dominant hand into the screen and activated by device clicking. Results showed that performance was significantly better when participants controlled the system bimanually (non-dominant hand zooming), with linear control and 1D guidance (mouse scroll wheel for zooming). A NUI for controlling virtual globes is introduced in [114]. The system uses a Kinect sensor to provide pan, zoom, rotation, and street view navigation commands to Google Earth. The system presents an interesting possibility for a NUI as in-air gestures follow the same logic as common multi-touch gestures. Hand poses (open/close) are used to activate commands, while the hands' relative position controls the virtual globe. For street view control, it uses gestures that mimic human walk; swinging arms makes the point-of-view move forward while twisting the shoulders rotates it. The use of metaphors that make computer controls relate to other known controls is not uncommon. In [115], two different approaches for interfacing with Bing Maps were tested for their usability, presence, and immersion. Using a Wiimote, the authors built a navigation interface inspired by the motorcycle metaphor. A handlebar-like motion-controlled turning and right-hand tilting acted as the throttle. Additionally to the metaphor, altitude over the map was controlled by left-hand tilting. The alternative approach used the Kinect to provide control and feedback inspired by the bird metaphor. Raising the arms asymmetrically enables turning, both arms equally raised or lowered from a neutral position control altitude, and moving the hands forward makes the user advance; the controls are enhanced by providing feedback in the form of a bird/airplane avatar. Descriptive statistic results showed high usability and presence for both systems, with higher values for the latter. The use of torso angle to control an avatar in a virtual reality city and how this control method affected the user understanding of size proportions in the virtual world was investigated in [116]. The system uses forward/backward leaning and shoulder turning to move and turn in the respective direction. It was tested on participants chosen for their knowledge in urban planning and building design, and compared to the common first-person-shooter mouse/keyboard interface. The results show that the system navigation was perceived as both easier and less demanding than the mouse/keyboard, and that it gave a better understanding of proportions in the modelled world.

Beyond navigation interface, gesture NUIs have been studied in the context of controlling computerized medical systems. This is particularly important in the surgery room, where doctors must maintain a sterile field while interacting with medical computers. In [117], the authors present their Kinect-based system for touchless radiology imaging control. It replaces the mouse/keyboard commands with hand tracking controls where the right-hand controls the cursor, and the left hand is used for clicking. The activation of the system was done by standing in front of the Kinect and waving. Tested for its qualitative rating with radiologists, 69% considered that the system would be helpful in interventional radiology. The majority also found it easy to moderately challenging to accomplish the tasks. Similarly, in [118], the authors introduced a solution for interaction with these systems using inertial sensors. Here the activation of the gesture detection was made by using a physical switch or voice commands.

Several exploratory research studies have been made to find the common gestures that naïve users would naturally perform. In [119], the authors found that by running an experiment in a Wizard of Oz set-up, participants would adopt the point-and-click mouse metaphor when asked to perform tasks in a large display. In [120], participants were asked to propose gestures for standard TV functions. The gesture agreement was assessed for each command, and a set of guidelines was proposed. Contrary to what was shown in [113] for pan and zoom gestures, one hand gesturing was preferred. Hand posture naturally emerged as a way of communicating the intention for gesture interaction.

When designing a NUI that supports in-air gestures, one must be aware of the "live mic" issue. As the system is always listening, if not mitigated, this can lead to false-positive errors [112]. Effective ways of countering the "live mic" problem are to reserve specific actions for interaction or reserve a clutching mechanism that will disengage the gesture interpretation. The review made by Golod et al. [121] suggests a *gesture phrase* sequence of gestures to define one command, where the first phase is the activation. The activation serves as the segmentation cue to separate casual from command gestures. Some example guidelines are the definition of activation zones or dwell-based interactions. In [122], from a Wizard of Oz design, the authors tried to identify pan, zoom, rotate, and tilt control gestures. More importantly, by doing so, they identified the natural clutching gestures for direct analog input, a subtle change from open-hand to semi-open. Similarly, the system proposed in [123] used the hand palm facing the screen for activating cursor control. [124] proposed two activation techniques: holding a remote trigger and activation through gaze estimation. These two activating techniques plus the control (trigger gesture of showing the palms to the screen) were tested for their hedonic and pragmatic qualities. Results showed that both the trigger gesture and remote trigger scored neutral on their hedonic and pragmatic scales. However, gaze activation scored high on both scales, achieving a "desired" rating.

Although much less common than vertical displays, interactive floors and floor projected interfaces possess unique features. In [125], the authors describe an interactive floor prototype controlled by body movement and mobile phones set up on a large public library hall. This arrangement enabled them to take advantage of the open space, filled by the large projected interface, and promote social interaction from its public function. These interfaces were proposed as an alternative to interactive tabletops [126], useful for not being as spatially restraining as the latter. The authors also explored the preferred methods of activation for buttons on these floors in their study, being feet *tap* their final design choice.

Even though the NUI literature is extensive, our review shows that most research has been made with exploratory or pilot design and could be advanced by validation studies. Furthermore, while most studies target the general population, their samples are usually not representative of the elderly, thus ignoring their specific impairments and needs. We focused our research on large interactive floors to generally address their visual perception impairments and support their physical activity and engagement needs. In order to better understand how this population can interact with such an interface, we proposed the following question:

- When designing a NUI to be used by an elderly population in floor projection displays, what interaction is best?

This was narrowed down by limiting the answers to two types of interface control: arm ray casting, commonly studied for vertical displays, and a *touch screen* like control, where the user activates interaction through stepping on the virtual elements. Considering the interface goals, we chose three elements to be rated: usability, workload, and performance. As one method would provide clear mapping at the expense of increased physical activity (stepping), the other would free the user from such movements while requiring them to project their arm into the floor mentally. Therefore, we hypothesized that differences in each of the three evaluation elements would exist when considering the two NUIs proposed. To test this hypothesis, the two proposed modes of interface control were developed and tested on an elderly population sample for two types of tasks that were evaluated in terms of usability, perceived workload, and performance. We expected that ray casting would provide better results as it is more widely used for interaction with large displays and requires little physical effort by the user.

2.2.3 Methods

2.2.3.1 Modes of Interacting

Two modes of interacting with the computer were developed based on the kinematic information provided by a Kinect v2 sensor and a display projection on the ground. In the first, henceforth named "*feet*", the cursor position is controlled by the average position of both feet on the floor plane; activation upon the virtual elements by the cursor is performed by placing the feet less than 20cm apart. For the second mode of interaction, named "*arm*," the forearm position and orientation is treated as a vector (from elbow to wrist) and ray cast onto the floor plane, the cast controls the position of the cursor (as schematized in Figure 21), while activation is done by closing the hand. Due to the Kinect v2 sensor's low reliability in detecting the closed hand pose, during the experiment, this automatic detection was replaced by the visual detection done by the researcher in a Wizard of Oz like experiment.

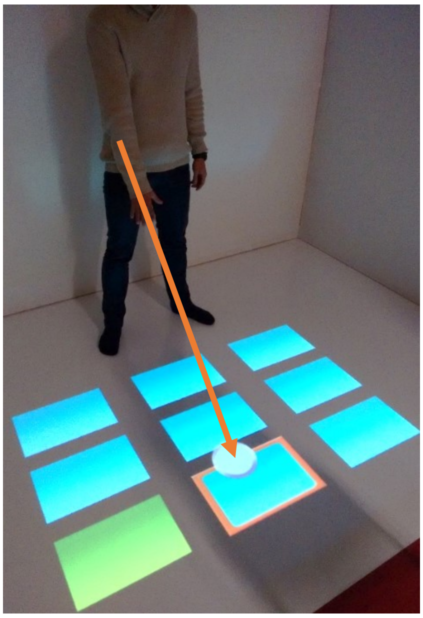


Figure 21: Controlling the cursor position through forearm ray casting.

2.2.3.2 Experimental Tasks' Description

The interfaces were tested in two different tasks to give a broader insight into what kind of interactions with computers our two systems would impact. A task to mimic the traditional point-and-click and another the common drag-and-drop.

In both tasks, the participant controls a circular cursor (ø 17 cm) with a 1-second activation duration, meaning that the activation gesture (feet together or hand closed) must be sustained for 1 second for the cursor to interact with the virtual element it is positioned on. This activation is represented on the cursor itself, which changes color in a circular way proportionally to the gesture's duration.

2.2.3.2.1 Point-and-Click Task

In the point-and-click task, a set of 9 rectangles (40 cm x 25 cm) are projected on the floor, on a three-by-three configuration, separated 12 cm laterally and 8 cm vertically, as shown in Figure 22. Out of the nine rectangles, 8 are distractors (blue), and one is the target (green). Every time the target is selected, it trades places with a distractor chosen on a random sequence (the same random sequence was used for all participants). The purpose of the task is to activate the target repeatedly while avoiding activating the distractors. Performance is recorded in this task as a list of events and their time tags, the possible events being: target click (correct click), background click (neutral click), and distractor click (incorrect click). In this task, maintaining the activation pose while moving the cursor from inside a rectangle to outside, or vice versa resets the activation timer.

Live feedback is given by drawing different colored frames around the rectangles. An orange frame is drawn around the rectangle over which the cursor is located. Upon activation, the frame changes color to red if the rectangle was a distractor or green if it was the target. This frame remains until the cursor is moved off the rectangle.

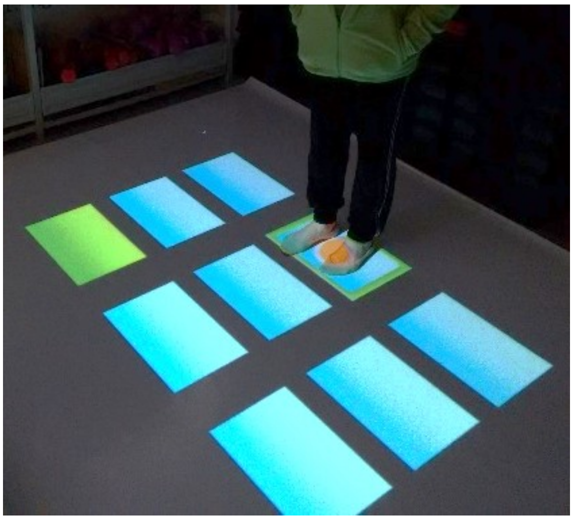


Figure 22: Point-and-click task being performed with the “feet” interface.

2.2.3.2.2 Drag-and-Drop Task

In the drag-and-drop task, four rectangles (40 cm x 25 cm) are projected on the ground, spaced 70 cm horizontally and 40 cm vertically, three of which are blue distractors, and one is the target (green). A movable yellow rectangle (30 cm x 19 cm) is initially shown in the center, as presented in Figure 23. The participant can “*grab*” the yellow rectangle by activating it. Once it has been “*grabbed,*” it can be dropped by activating it again (joining the feet or closing the hand, depending on the interaction mode). The task’s purpose is to “*grab*” the yellow rectangle and “*drop*” it onto the target repeatedly. Every time this is done successfully, the yellow rectangle is reset to the center, and the target changes places with one of the distractors in a random sequence (the sequence was kept constant across all participants). Performance is recorded as a list of events and their time tags. The possible events for this task are: grab yellow (correct grab), attempt to grab anything else (neutral grab), drop yellow on target (correct drop), drop yellow on the background (neutral drop), and drop yellow on distractors (incorrect drop). As before, maintaining the activation pose while moving the cursor from a rectangle to outside, or vice versa resets the activation timer. Likewise, a set of colored frames are used to give live feedback to the users. An orange frame highlights any rectangle under the cursor. Once activated, the yellow object’s frame changes to green, indicating that the cursor is dragging it. Dropping it on a distractor will create a red frame around the distractor; oppositely, dropping it on a target will show a green frame around it.

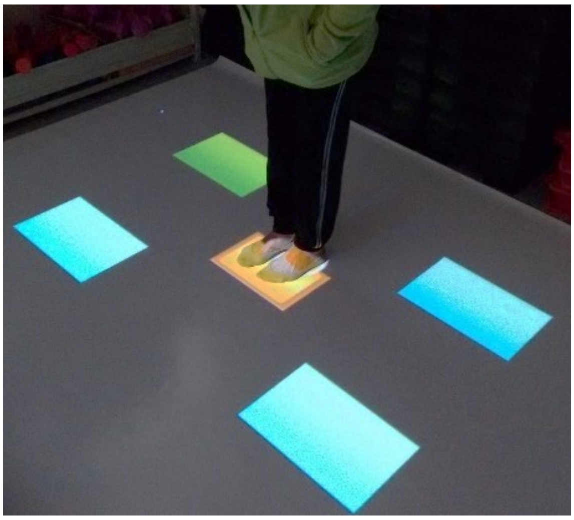


Figure 23: Drag-and-drop task being performed with the “feet” interface.

2.2.3.3 Technical setup

The hardware was set up in a dimly illuminated room, and a white plastic canvas was placed on the floor to enhance the reflectivity of projection. A Hitachi CP-AW100N projector was positioned vertically to face the floor. This arrangement enabled a high contrast of the virtual elements being projected and an area of projection greater than what our tasks needed (150 cm x 90 cm). A Microsoft Kinect v2 was placed horizontally next to the projector, facing the projection area (Figure 24).

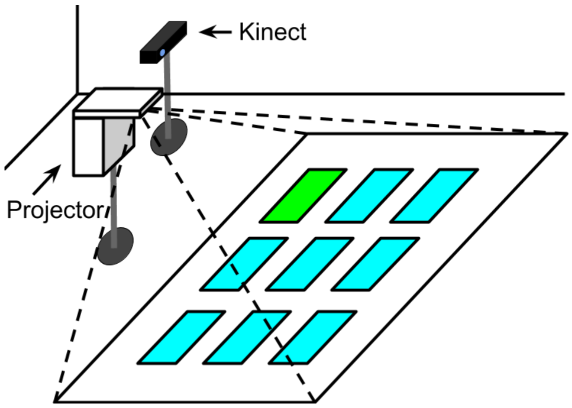


Figure 24: Experimental setup diagram.

2.2.3.4 Participants

The target population of the study was community-dwelling elderly. A self-selecting sample of this population was recruited at Funchal's Santo António civic center with the following inclusion criteria:

1. Being more than 60 years old;
2. Do not present cognitive impairments (assessed by the Mini-Mental State Examination Test [127], MMSE > 22);
3. Do not present low physical functioning (assessed by the Composite Physical Function scale [128]).

The experiment took place over two days at the facilities of the civic center municipal gymnasium for the elderly. Nineteen participants (ages: M = 70.2 SD = 5.3) volunteered and provided written informed consent, 3 males and 16 females. The participants were randomly allocated to each condition, ten being assigned to the "*feet*" and 9 to the "*arm*" condition of interaction.

2.2.3.5 Experimental Protocol

The experiment followed a between-subjects design. The participants were asked to answer questionnaires regarding identification, demographical information, and level of computer use experience. They were evaluated with the Composite Physical Function Scale and Mini-Mental State Examination Test. During each participant trial, the point-and-click task was explained and shown being performed through example, according to the participant's experimental condition. This was followed by a training period and then by a 2 minutes' session while performance metrics were recorded. Lastly, the participant was asked to fill the System Usability Scale (SUS) [129] and NASA-TLX (TLX) [130] questionnaires. After it, the same procedure was followed for the drag-and-drop task.

2.2.3.6 Analysis

For each participant, data consisted of SUS score and TLX index (both measured from 0 to 100) and task-related performance, as described in sub-sections 2.2.3.2.1 and 2.2.3.2.2. The normality of the data distributions was assessed using the Kolmogorov-Smirnov test for measurements concerning performance. The variables that showed such a distribution are highlighted in Table 13 and Table 14. For the pairs (between conditions) of measurements that fitted the normality assumption, parametric t-tests were used; when significant differences in the pair's variances were present, shown by Levene's test, equal variances were not assumed. All the other pairs were tested with Mann-Whitney's U test. Differences in the SUS and TLX scores (ordinal variables) between conditions were also tested with Mann-Whitney's U test. All statistical testing was done using 2-tailed testing at α .05 with the IBM software SPSS Statistics 22.

2.2.4 Results

2.2.4.1 Point-and-Click Task

For the "*feet*" condition, in the point-and-click task, the descriptive statistics are presented in Table 13, where we can observe very low values of incorrect clicks and high median scores for the SUS, which is considered to be a good value when over 68. The descriptive statistics for the "*arm*" condition are also presented in Table 13. Higher values of neutral and incorrect clicks are visible compared to the previous condition. Similarly, a decrease in the median of the SUS usability score and an increase in the TLX workload index can be seen.

Table 13: Descriptive statistics of the measurements for the point-and-click task.

	"Feet" Interface		"Arm" Interface	
Variable	**Median**	**Interquartile Range**	**Median**	**Interquartile Range**
SUS	91.25	21.25	72.50	25.00
TLX	23.75	27.71	40.83	18.33
Correct	29.50	10	28.00^{n}	15
Neutral	1.00	2	4.00^{n}	7
Incorrect	0.00	1	2.00^{n}	3

^{n} Normally distributed

Results revealed significantly higher System Usability Scale scores for the participants interfacing with their feet than the participants interfacing with their dominant arm, U = 18.5, p < .05, with effect size r = -.4997. The Task Load Index scores were not significantly different for both interfaces, U = 24.5, p > .05 (Figure 25). The number of correct and neutral clicks was not significantly different for both interfaces, U = 40.5 and U = 29.0, p > .05, respectively. However, it was found that there was a lower number of incorrect clicks for the participants interfacing with their feet compared to the participants interfacing with the arm, U = 15.0, p < .05, r = -.5863 (Figure 26, where circles represent outliers and stars extreme outliers).

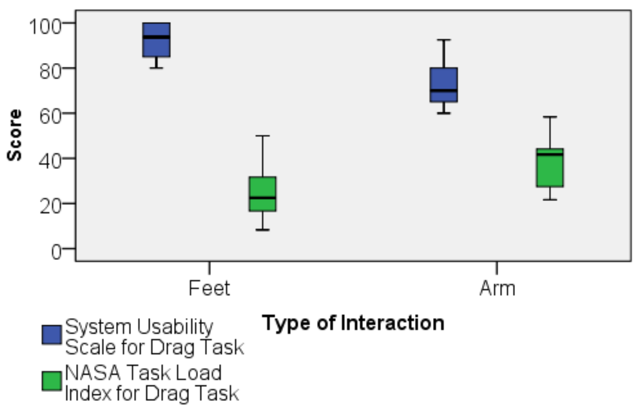


Figure 25: System Usability Scale and Nasa-Task Load Index scores for the point-and-click task.

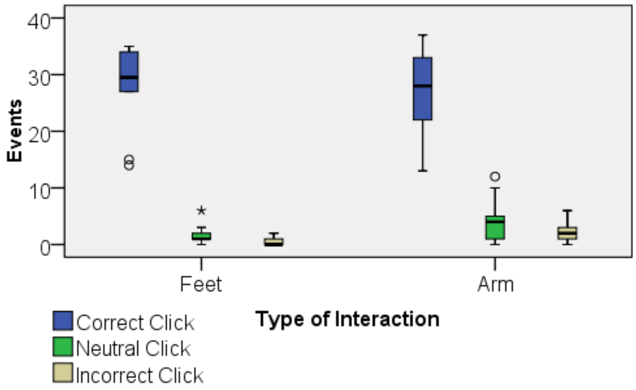


Figure 26: Participants' performance on the point-and-click task.

2.2.4.2 Drag-and-Drop Task

The descriptive statistics for the "*feet*" condition in the drag-and-drop task are presented in Table 14, where we can observe low values of incorrect drops and no neutral drops (accidental drops). The values of usability are very high, and workload moderately low. In the "*arm*" condition of the drag-and-drop task, we can see, in Table 14, a marginally good value for the SUS usability score, barely over 68. The TLX workload has relative medium levels, and neutral drops (accidental) are present.

Table 14: Descriptive statistics of the measurements for the drag-and-drop task.

	"Feet" Interface		"Arm" Interface	
Variable	**Median**	**Interquartile Range**	**Median**	**Interquartile Range**
SUS	93.75	16.25	70.00	21.25
TLX	22.50	16.46	41.67	22.50
Correct Grab	14.50^n	8	11.00^n	9
Neutral Grab	13.50^n	4	10.00^n	9
Correct Drop	14.00^n	7	10.00^n	10
Neutral Drop	0	0	1.00^n	3
Incorrect Drop	0.00	0	0.00	0

^{n} Normally distributed

The results indicated again a significantly higher System Usability Scale score and lower Task Load Index score for the *Feet* interaction condition, with U = 9 and U = 17, $p < .05$, effect size r = -.6777 and r = -.5247 respectively (Figure 27). There were no significant differences for the normally distributed data with correct grabs, neutral grabs, and correct drops, t(17) = .565, t(17) = .863 and t(17) = 1.336, $p > .05$, respectively. Neutral drops were significantly higher in the "*arm*" interaction condition, U = 10, $p < .05$, r = -.7595 and there were no significant differences between the number of incorrect drops, U = 44.5, $p > .05$ (Figure 28).

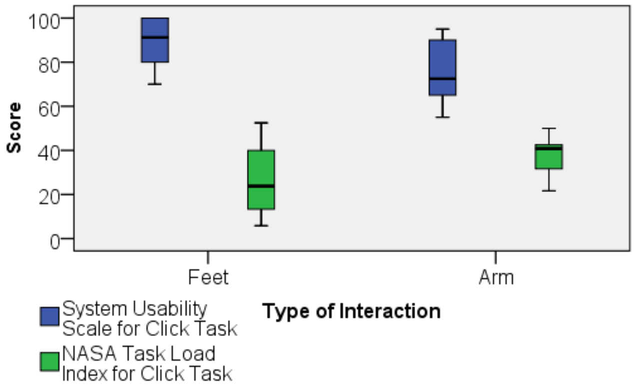


Figure 27: System Usability Scale and Nasa-Task Load Index scores for the drag-and-drop task.

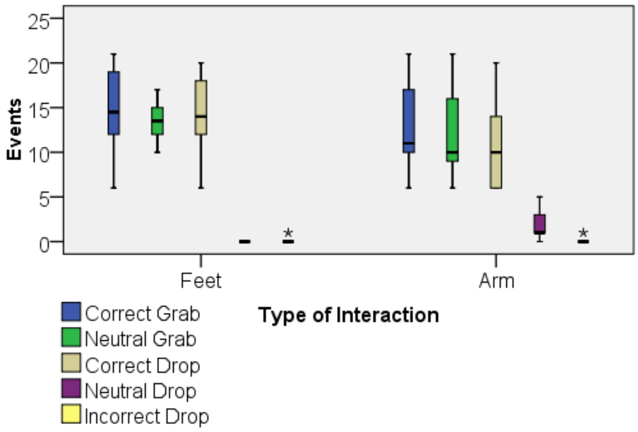


Figure 28*: Participants' performance on the drag-and-drop task.*

2.2.5 Discussion

For both the point-and-click and drag-and-drop tasks, it was found that there is a significant impact on system usability, being the "*feet*" interaction method preferable in both cases. With the "*feet*" modality achieving high usability levels, scores over 90, while the "*arm*" had usability levels around 71, very close to the standard lower limit of good, 68. In the case of perceived workload indexes, there were no significant differences between the conditions for the point-and-click. While for drag-and-drop, the "*feet*" interface was significantly less demanding to use by the participants. In both cases, the "feet" workload indexes were around 23, while for the "*arm,*" the values were situated around 41. Although interfaces similar to our "*arm*" method have been the focus of previous research [113], [123] and shown to be a method that participants naturally display [119], [120], [122], in our experiment, we found sufficient evidence that older people prefer an alternative way of interacting with projected floor elements. This preference by the participants for the "*feet*" interface might be linked to the more straightforward mapping of the cursor control provided, which is known to have a lowering effect on cognitive load [116], [131]. Finally, in terms of performance, the point-and-click task observed very low numbers of neutral and incorrect clicks (although significantly higher for the "*arm*") and a comparable number of correct clicks. Similar results were found in the drag-and-drop task, with low numbers of neutral and incorrect drops for both methods and identical values of correct grabs, neutral grabs, and correct drops. Still, the "*feet*" interface was again better, with the number of neutral drops being significantly lower than in the "*arm*" interface. Albeit these differences, the remaining performance indicators were shown not to be significantly different. Therefore, caution is advised to interpret these results as proof of a clear performance advantage provided by any of the interfaces.

2.2.6 Conclusions

Due to the increasing number of older adults in developed countries and this population's specific needs, we tried to understand the desirability of different modes of controlling interaction in interactive floors. By being easily scaled, this medium can mitigate the visual perception deficits associated with old age and promote physical activity. Thus, in this section, two methods of interacting with virtual elements projected on the floor were developed and tested for differences in their usability, perceived workload, and performance ratings by an elderly population. The interfaces consisted of either controlling the cursor with the direct mapping of feet position onto the projection surface or by mapping the cursor position to the participant's ray-casted forearm on the surface. These interfaces were tested on two different tasks, one mimicking a point-and-click interaction, the other a drag-and-drop. Although the NUI research field is extensive, there is a lack of studies that approach the floor projected interfaces, and studies with the elderly are even rarer. This study gives a successful insight into the preferred modes of interaction for this elderly population.

Contrary to our initial guess, the results showed that the "feet" interface was superior in all the domains measured from the two proposed methods. It was shown that this method was perceived as more usable in both tested tasks and at least less demanding in terms of workload for the drag-and-drop task. In terms of performance, a marginal advantage was also shown for the "*feet*" method. This insight contributes to HCI knowledge and is of general usefulness in developing systems aimed at providing full body NUI for this population. Moreover, it informed us that to design easy to use and less tiresome floor projected exergames that meet older adults' needs and expectations, a direct mapping of the feet to the projection should be used.

2.3 LESSONS LEARNED FROM GAMIFYING FUNCTIONAL FITNESS TRAINING THROUGH HUMAN-CENTERED DESIGN METHODS IN PORTUGUESE OLDER ADULTS[5]

2.3.1 Introduction

Information and Communication Technologies (ICT) show the potential to improve behavioral change concerning physical activity through global positioning systems, wearable devices, persuasive technologies, and interactive video games [132]. Combining game technologies, applied research, and healthcare professionals allow the extension of ICTs' application domains to topics such as training, health education, disease prevention, and aging. According to Deterding et al., gamification uses game elements to motivate and engage people in non-game contexts [133]. In this regard, the so-called exergames aim to promote physical activity by enhancing the experience via gamification of elements that engage players through interactivity, game challenges, points, rewards, and entertainment. This engagement may enable longer-term adoption and better adherence to the training program than conventional methods [134], [135]. In fact, exergames have been under investigation in recent years and show a high potential to promote physical activity in multiple environments and populations [132]. Research has shown encouraging findings for their role in motivation for exercising [83], improvements in cognitive function [136], [137], and physiological energy expenditure [138]. Interestingly, the use of exergaming in the older population has shown likely positive impacts on physical and mental health [139]. Physical, balance and postural control benefits have been widely reported as critical areas where exergaming practice has a clear positive impact [81], [140], [141]. Other functional fitness domains, such as cardiorespiratory or muscle strength, have been less explored [9], [81]. Nevertheless, some studies reported significant increases in physical activity participation [142], overall functional fitness function (e.g., flexibility, strength, endurance) as measured by standardized tools [143] and motivational aspects [83] in older adults under multidimensional training programs delivered through exergames. Regarding cognitive improvement, exergames have shown positive results in selective attention[144], visuospatial skills [145], executive functions, and processing speed [136]. Only one study reported cognitive benefits of multidimensional physical training with exergames specifically in older adults; improvements in shifting attention, working memory, and long-term visual memory after a 6-month intervention were demonstrated [146]. Other mental effects of exergaming interventions have been observed, including differences across interventions in enjoyment, immersion, and flow [147], critical game user experience variables that reinforce the fun aspect of playing exergames. Fundamentally, exergaming aims at overcoming some of the previously identified [148], [149] common barriers for PA by i) providing a real-time feedback of users' performance [137], [150], [151]; ii) allowing a more appealing and engaging way to carry out exercise while avoiding monotonous experiences [140], [152]; iii) recording the progress, providing measurable and quantifiable effects of working out [153], [154]; iv) remotely connecting health professionals with patients to facilitate supervised activities and promoting telerehabilitation programs, thus solving issues of transportation for older adults with limited mobility [9], [143], [155].

Still, several reviews in the field point out the need for additional research on the long-term adherence to exergames specifically designed for the older population [9], [156], [157]. A crucial aspect of ensuring both adherence to exergames and their effectiveness in producing health benefits is a proper design process. Researchers frequently use commercial exergames to explore multiple effects of exergaming programs [18], [158], but this repurposing approach raises concerns. Older adults' care applications have specific requirements for user interfaces, exercises, and game design that are generally not considered in commercial exergames, thus limiting the impact of such interventions [159]. Besides, since the older adult population is very heterogeneous in terms of needs, deficits and co-morbidities, exergames need to be specifically tailored to each individual to maximize the impact on their health [9], [81]–[84]. Research in exergaming for older

[5] The content of this section has been published in the Games for Health journal [79]

adults which demonstrates clear and measurable social wellbeing benefits is still lacking, due to the absence of experiments that compare the effects of exergames against conventional approaches, lack of strictness in research protocols and evidence of long term adoption as well as limitations in the sample sizes of the experiments [160]. Particularly, more research performed directly in the seniors' natural context, performed with a more careful observation process of the users' everyday lives and social interactions, has been suggested in multiple meta-reviews [84], [160], [161].

2.3.1.1 The rapid contextual design approach

In the Human-Computer Interaction (HCI) field, one of the most used methods for designing highly contextualized, personalized, usable, and useful interactive systems is the human-centered design approach [162]. Through this method, the system's final users are involved in all steps of the design and development process, thus supporting a solid understanding of users' needs and motivations, as well as allowing the possibility of keeping the solutions in a constant state of dynamic iteration. The research presented here adopted the contextual design process, a popular human-centered design method that focuses on collecting data about users in the field to drive the creation of a technical product [163]. This method has been adapted to exergaming by following a structured participatory design process, where the creation of more contextually enriched experiences might enhance the final impact of the developed technologies [164]. This method has been simplified and restructured in what is called the Rapid Contextual Design (RCD) [165] and can be divided into three main stages:

- Requirements gathering: the target user is characterized through a field data gathering process aimed at finding population needs, habits, issues, activities, and potential opportunities. Here, specific surveys and questionnaires must be designed to collect as much relevant information as possible.
- Activity redesign: after being processed, the above data is used to introduce technology that supports, complements, or replaces the existing activities to redesign life practices. This is done through several data interpretation sessions to identify and characterize each of the interviewed users and capture patterns, tasks, and needs. A work modeling process is then used to investigate the structure of the user's routine tasks by representing descriptive models of physical spaces and activity sequences. Finally, an affinity diagram is used to gather and organize all the information collected from users, depicting all issues and insights using a hierarchical logic. Affinity diagrams simplify the data clustering process through their natural relationships' analysis. As a result, users' Personas, representing real users, are created and modeled.
- User experience design: the new technological component is refined through iterative testing with users.

In this section, we used RCD to create an integrative model of our physically active older adults as the target population and applied this model to develop highly engaging exergames for exercise promotion and fitness training. Mainly, we used a specific RCD process called focused-RCD, which is intended to be carried out in 6 to 10 weeks with a sample of 8 to 12 users [165].

2.3.1.2 Recommendations for Physical Activity in Older Adults

The American College of Sports and Medicine (ACSM) and the American Heart Association (AHA) established a set of general recommendations for physical activity which describe the amounts and types of physical activity that promote health and prevent diseases and functional limitations in older adults [7]. The training areas should include aerobic fitness (cardiorespiratory), musculoskeletal function, flexibility, and balance (or neuromotor) to deliver sufficient and efficient physical activity. The recommended exercise features are typically described using the FITT (Frequency, Intensity, Time, Type) model. For instance, the older population's recommended activity intensity is moderate (5-6 in a 0-10 physical effort scale), with an accumulated time of 150 minutes per week [166]. These guidelines apply to all adults aged 65+ years as well as to adults aged 50-64. It is worth mentioning that literature reviews in the domain of exergames for older adults highlighted the importance of including multiple physical functions in a structured way that follows health standards and well-defined protocols [9], [18]. Past investigations with a large group of older adults (n=802) established a positive association between functional fitness and physical activity, allowing a preliminary identification of the dimensions that should be trained through exergaming [167]. Therefore, we aim to provide specific examples of the development process of exergames for functional fitness training in older adults and guidelines that express the lessons learned along the process. Our goal was to create a set of exergames that can meet the ACSM recommendations in the different fitness domains by offering a high level of training personalization with a fine-grained game parametrization.

2.3.1.3 Our approach

This section describes our effort to introduce a structured and systematic methodology for designing exergames targeted at Portuguese older adults. We designed and developed a set of exergames, which combine standardized fitness training programs and a contextually rich user modeling process to overcome current limitations in using exergames by older adults. We introduce a design framework that addresses critical aspects of the process, such as:

- Human-centered approaches in different stages of the game design process;
- Emphasis on fun experiences through playful interaction and gamification elements (story, aesthetics, mechanics, and technology);
- Integration of health professionals in the design process to identify health-related standards to be incorporated in the exergames;
- Rapid prototyping and fast iteration with end-users.

This section presents the whole design process from the definition of the health requirements to the final polishing stage of the proposed exergames. Our experimental scenario is a senior gymnasium frequently used by active adults to exercise who socialize with their peers and participate in several cultural activities. We aim at inquiring into methodologies for designing enjoyable and scientifically valid exergames, which promote exercise in older adults by adopting human-centered design processes. Specifically, we wanted to answer the following research question: how should human-centered design methods and insights be used to design highly contextualized exergames for exercise promotion in older adults (aged 55+)? By answering this question, we will help adopt exergaming technologies that might enhance the quality and quantity of physical activity, improve training personalization, and diversify work-out practices [168], [169]. Additionally, we wanted to assess the feasibility of combining iterative game design with human-centered design methodologies to design a fitness training program based on exergames for older adults.

The following sub-section briefly describes the state-of-the-art in the use of human-centered design approaches for exergame design and existent guidelines for exergaming design.

2.3.2 Related Work

2.3.2.1 Human-Centered Design and Exergames

Human-centered design has been used to inquire about the benefits and barriers to older adults playing video games [170], gaming preferences, personalities, and motivators [171], [172], the design of user-oriented exergames for fitness promotion [173], gamifying systems for children's nutrition and fitness education [174], and entertainment [175]. Research shows that older adults are widely different from younger cohorts in their game preferences in terms of playability, challenges, and motivators [148], [176]. Moreover, methods such as focus groups have highlighted the importance of perceived benefits, difficulty, and relevance for engaging with new technologies [177]. Since the personalization of exergames has been reported as one of the major bottlenecks in probing this technology's effectiveness when applied to seniors [178], the use of participatory and inclusive design methodologies can facilitate technology adoption by them. Indeed, multiple researchers have identified the human-centered design approach as adequate and preferable in designing compelling, useful, and usable exergames for exercise promotion [173], [179]. For instance, Uzor et al. demonstrated how empowering seniors with design tools and techniques could boost the likelihood of creating enjoyable rehabilitation tools for fall prevention [180]. Participants have been modeled using Personas, facilitating the understanding and communication of their characteristics and needs [162], [181]. The fact that senior adults tend to show different technological literacy levels and video game experience must also be considered [182], [183]. This presents an exciting challenge and opportunity to develop novel assistive technology for the older population [184]. Because the current Portuguese elder population has little to no experience with video games, starting a design process by inquiring about preferred video game mechanics, aesthetics or technologies is not adequate. Multifaceted and participatory approaches have been widely recommended by physicians and sports scientists, stressing the need for clearly defining an appropriate fit of game design elements for each target group [157].

2.3.2.2 Guidelines and Frameworks for Exergame Design

Exergaming research is growing fast, and results from several studies have defined valuable design guidelines. Table 15 shows five different sets of design guidelines for exergames found in the literature, each created for specific goals and populations. The guidelines included in the table were chosen considering four criteria: i) being exergame design guidelines, as opposed to recommendations for development or validation, ii) their previous or potential use with older adults, iii) their origin in field studies instead of theoretical approaches, and iv) their clarity in defining comprehensible items that could be easily transformed into actionable items.

These guidelines provide a helpful starting point for the development of novel exergames. Nevertheless, practical examples on how to incorporate users' feedback in the design process and how to use the insights from human-centered inquiries to improve in-game elements, such as mechanics or aesthetics, and provide valuable recommendations for such process are still scarce.

Table 15: Design guidelines for the development of exergames

Name	Goal	Population	Guidelines
Designing for movement quality in exergames [175].	Fall prevention exercises	Senior citizens	1. Weight shift: Elicit weight shift in players by motivating them to move around a larger physical area and displace their center of mass. 2. Temporal variation: Provide temporal variation in movements by offering adaptive changes in-game speed. 3. Step length variation: Promote step length variation by offering variation in exergame tasks. 4. Visual independence: Elicit variation in movement direction during gameplay.
Guidelines for the design of movement-based games [176].	Entertainment	Overall	1. Movement requires special feedback: "…instead of trying to remove this (the movement) ambiguity, work with it: players enjoy surfing uncertainty and trying to figure out optimal strategies in a somewhat messy system." a. Embrace ambiguity b. Celebrate movement articulation c. Consider the movement's cognitive load d. Focus on the body 2. Movement leads to bodily challenges: "…intend fatigue when using it as a game challenge but avoid it when it is not part of the game." e. Intend fatigue f. Exploit risk g. Map imaginatively 3. The movement emphasizes certain kinds of fun: "Movement becomes easier with a beat, so support players in identifying a rhythm to their movements." h. Highlight rhythm i. Support self-expression
A methodology for the design of effective and safe therapeutic exergames [177].	Therapeutic (balance and posture, neglect rehabilitation)	Poststroke patients	1. "A clear identification of all the exercise requirements, not only in terms of goals of the therapy but also in terms of additional constraints." 2. "…Discussion between clinical and ICT (Information and Communication Technologies)

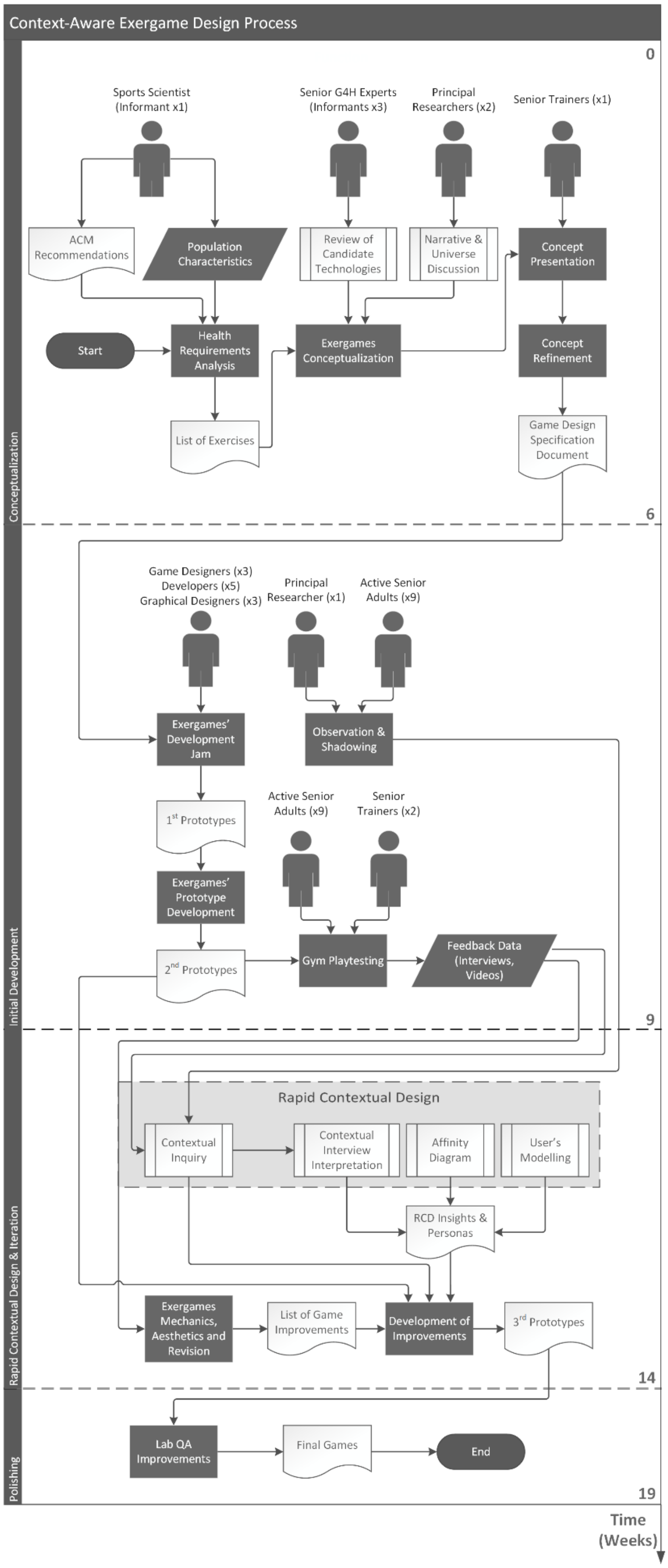
			teams allows maximizing the effectiveness of exergames implementation." 3. "The final exergame is realized by introducing on top of the exercise all the game elements suggested by good game design to maximize entertainment."
Exergame design guidelines for elderly [178].	Enhance physical and social activities	Elderly	1. Mind the physical condition: "... consider the physical condition of elderly, since age-related processes may have an impact on the ability to move." 2. Use appropriate gestures: "It should not be required that the players have to perform a gesture to trigger another action." 3. Avoid small objects. 4. Give visual and auditory feedback.
FitForAll approach [142]	Exercise and maintenance/advancement of healthy physical status and wellbeing	Senior citizens	1. Physical condition considerations. The range of motion, adaptability. Continuous player support. 2. Avoid small/fast-moving objects. Clean, attractive, and friendly user interface. 3. Suitable topics. Provide audiovisual feedback. Provide opportune feedback. 4. Adjustable difficulty. Exertion management. Record/display users' past behavior. 5. Encourage social interaction. 6. Simple setup.

2.3.3 Procedure

2.3.3.1 Exergames Design and Development

We started the process by designing a set of exergames based on the examination of health requirements. A complete process diagram can be seen in Figure 29, which depicts all the conceptualization steps until the polishing stage.

Figure 29: The flowchart diagram of the design process describing people involved, inputs, deliverables, and design stages. The process's duration was 19 weeks, divided into four main stages: conceptualization, initial development, rapid contextual design and iteration, and polishing.



2.3.3.1.1 Exergames Conceptualization, Requirements, and Input from Experts

Two different sources of information were used for the health requirements analysis: a) population characteristics (health screening) and b) assessment of functional fitness status [167]. We identified that the relevant components to be trained and measured through exergaming were: motor ability (balance, agility, and flexibility), aerobic endurance, and muscular strength (lower and upper limbs and trunk). Additionally, we identified three levels of training profiles: low, moderate, and high functioning. These profiles allow for balancing the intensity of training to users' fitness levels, also called the effectiveness loop in the dual flow model [16]. We included sport science professionals with experience in older adult's fitness training and assessment for the video game design process. They were encouraged to illustrate several exercises, movements, and activities that users need to perform to train each of the different training profiles' domains.

The game design process's biggest challenge was in the ideation of an exergame universe that would embrace the specified health requirements. How can we design thoroughly fun exergames starting from pre-defined fitness/exercise recommendations avoiding the chocolate-covered broccoli effect [185]?

2.3.3.1.1.1 Technology selection

First, we gathered with three senior researchers of games for health to define an initial set of candidate technologies that would be adequate for a training program based on exergaming. Motion tracking and physiological computing technologies were considered for designing an integrated solution that could be used in both controlled (e.g., laboratory) and non-controlled (e.g., senior gymnasiums) environments. The Kinect v2 motion tracking controller was chosen because of both its non-invasiveness and its low cost. Additionally, the Kinect v2 sensor allows an accuracy sufficient for the current application [55], tracking 25 joints distributed along the whole body for standing and sitting postures [95], and accuracy in detecting key motions, Section 2.1. The use of floor projections facilitates game mechanics and physical activity matching to the display in all the exergames; it also avoids spatial constraints and provides a sizeable interactable projection. The setup includes a PC, a projector, a white PVC surface (2.5m m x 3.0 m), and the Kinect v2 sensor located ≈3m in front of the users (Figure 30). To facilitate interaction, we also explored spatial augmented reality, which allows the augmentation of real-world objects and spaces using simple projections instead of unique displays [186]. With this technology, users can physically interact with the objects placed in virtual environments and projected on the real scale floor.

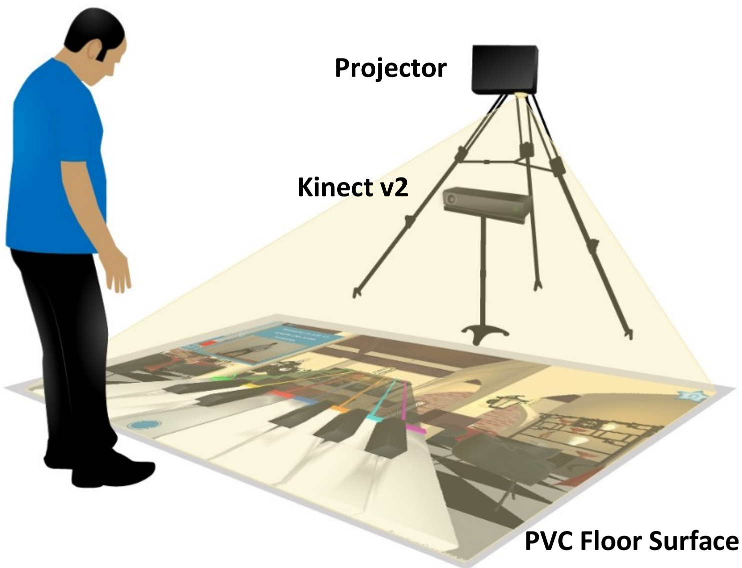


Figure 30: Setup for the exergames consisting of a Kinect v2 sensor and a virtual environment projected on a floor surface. PVC, polyvinyl chloride.

2.3.3.1.1.2 Game ideation

For the exergames ideation stage, we used brainstorming [187]. We focused on creating multiple video game stories enclosed in imagined virtual scenarios in which the users' playful interactions would be concealed exercises. These scenarios would define a narrative and a universe for the exergames. After approximately ten short (20 min) brainstorming sessions involving game designers and sports professionals, we came out with five categories of activities, each with multiple possible scenarios, as follows:

- Fantasy: space exploration, medieval life, maritime activities, and aviation.
- Sports: Olympic Games, traditional Portuguese games, and martial arts.
- Professional manual labor: agriculture, food preparation, animal care, mineral exploration, gardening, hospital activities, automotive mechanic, child-care, and construction work.
- World traveling: continental or national tourism.
- Creative oriented activities: painting, sculpting, dancing and fashion design.

These ideas were combined with the pre-defined exercises for physical training, and, ultimately, we decided to choose the world-traveling category to develop a virtual tour in Portugal. We based our choice on the hypothesis that computer-illiterate users would feel more identified and engaged with the exergame experiences via culture-specific elements. Preliminary ideas involved developing exergames that would feature iconic activities specific to Portuguese culture. The idea was then presented to and discussed with a fitness trainer in a local senior gymnasium, and the concept was refined by identifying exciting places and activities that could engage the aged audience (see *Concept Presentation* and *Concept Refinement* in Figure 29). A game design and a specification document were produced detailing initial ideas for possible exercises taking place in the imagined virtual universe (as recommended by game design frameworks [188], [189])**.** Once the core of the story component was established, we engaged in developing the game mechanics in a game jam [190].

2.3.3.1.2 Exergames Development Jam

After the conceptualization, we used a rapid iterative prototyping approach [187], which was implemented in a local game jam. The jam was carried out in a research laboratory facility and lasted about 40 hours along five consecutive days. The first activity completed was a visit to the local senior gymnasium where participants could observe conventional exercise sessions with a group of around 30 senior adults. Three video game designers, three graphical artists, two psychologists, two sport science professionals, five programmers and three games-for-health experts participated actively. The jam teams were divided and given different fitness domains to cover, following the health requirements described previously. We emphasized and prioritized the importance of making exergames appealing, meaningful (regarding health and social benefits), and playable for older people with different preferences [191]. Since this population is very heterogeneous, developing a generic and poorly personalized experience will end up as a minimal solution for exercise promotion [9].

Three different presentations of playable demos took place during the exergame jam for evaluation of the prototypes. Two health professionals from a senior gymnasium and three games for health experts evaluated the video games in different stages. At the end of the development jam, the first versions of the exergames prototypes had a high percentage of the game story completed while a moderate percentage of the game mechanics and a low percentage of the game aesthetics were completed. At this stage, we had a set of virtual experiences covering four different places in Portugal that addressed all fitness domains. Prototypes were further developed (see *exergame's prototype development* in Figure 29) and iterated until they were robust enough to carry out a field test in a local senior gymnasium in Madeira, Portugal.

2.3.3.1.3 An Exergaming Experience through Portuguese Traditions

At the end of the one-week jam, our vision to create a virtual tour in Portugal materialized in a set of exergames covering several touristic activities in different places (Figure 31, Table 16).

Figure 31: Screenshots of the initial prototype exergames. (A) Grape stomping, (B) Rabelos, (C) Exerfado, and (D) Toboggan Ride.

Table 16: Description of each individual exergame, Portuguese tradition, goal, fitness domains addressed, and related movements

Exergame (Portuguese Region) description	**Goal and score**	**Fitness domain and movements**
Grape Stomping (Douro Region): winemaking through crushing grapes barefoot to release the juice.	Goal: Grape maceration to produce wine by stepping on virtual grapes placed into open tanks. Score: liters of wine produced.	Cardiorespiratory training: Flexion-extension arm movements are needed to drag the grapes into the tanks. Then, the stomping process starts, requiring users to raise the knees repeatedly. After filling up one tank, users can move to the others to continue the exercise.
Rabelos (Porto City): wine barrel transport in cargo boats navigating in the Douro river.	Goal: Transport barrels of wine along the Douro river by rowing in a boat, avoiding rocks and collecting more barrels in the riverbanks. Score: Quantity of barrels collected.	Upper limb muscular strength: Perform circular arm movements simulating rowing gestures to move the boat forward. For lateral displacements of the boat, standing users can use either lateral movements or trunk leaning; or they can be seated and use trunk leaning only. To navigate the rapids, users must avoid rocks by rapid lateral movements. To pick up the barrels, users must turn the trunk to the dock and make an extension/flexion movement of the elbow.
Exerfado (Lisbon City): Fado music playing by a melodic guitar soundtrack.	Goal: Catch the correct notes by selecting of the frets in a virtual guitar (GuitarHero-like game).	Lower limb flexibility: Players collect notes by performing rapid lateral movements to select the correct guitar's fret. Players should move their entire bodies to collect the musical notes since the paddle reflects the hip movement.

	Score: Number of successfully collected musical notes.	Once empty, the music score can be refilled using a “turn the page” or swipe gesture with the left hand extended.
Toboggan Ride (Madeira Island): a downhill journey in a sliding sled.	Goal: Control the direction and acceleration of the car to collect items placed on the street. Score: Number of successfully collected banana bunches.	Balance, trunk muscular strength: The sled direction is controlled through trunk lateral flexions, while the acceleration is controlled through trunk flexion/hyperextension movements. Flexion movements are used to increase the vehicle's acceleration, while hyperextensions are used to deaccelerate it.

2.3.3.2 *Applying Human-Centered Design Methods for Inquiring about Exergaming Preferences: Hands-on with the RCD Process*

After the game jam, the subsequent efforts focused on testing the second version of the prototypes with end-users and collecting structured feedback. Following the focused RCD process, we designed three different surveys: a demographics form, a video game experience questionnaire, and a final survey from which we collected feedback related to the exergaming experience.

2.3.3.2.1 Demographics

Participants for this study were selected from the older adult population exercising (at least two hours per week) in the senior gymnasium. The gym trainers invited participants to collaborate in a structured interview, which included interaction with a gaming system under development that would potentially be used to complement their exercise routines. Volunteers included nine (9) older adults (8 females, M=62.3 years old, SD=6.2 years old; 1 male) (Table 17). Participants signed informed consent before participation.

Table 17: Characteristics of the participants

Subject	Gender	Age	Schooling (years)	Previous or current occupation	Time in the gym (years)	Hours of physical activity (weekly)
1	Male	56	12	Accountant	3	3
2	Female	57	9	Homemaker	7	2
3	Female	56	12	Homemaker	15	8
4	Female	67	0	Homemaker	9	7
5	Female	70	0	Homemaker	2	4
6	Female	57	9	Factory worker	7	4
7	Female	67	4	Telephone operator	5	10
8	Female	59	4	Cleaning services	10	5
9	Female	72	9	Retired	2	3

Table 18: Sport and leisure time indexes computed from the Modified Baecke Questionnaire

Subject No.	1	2	3	4	5	6	7	8	9
Sport Index (min 1–max 4)	3.3	3.0	3.3	2.3	2.8	3.0	3.3	3.5	3.0
Leisure Time Index (min 1–max 4)	2.9	2.9	2.7	2.6	3.1	2.3	3.1	3.4	2.7

We used a modified version of the Baecke questionnaire [192] to query users about their physical activity levels. The Baecke questionnaire is a tool that assesses qualitative and quantitative indices of dimensions such as weekly physical exercise practice and leisure time activities during the last 12 months. The hours-per-week estimate for the different activities is multiplied by the respective metabolic rates (expressed as metabolic equivalent- METs) to obtain energy expenditure during sports and active leisure time. From that, we computed the Sport Index - which is a global measure of exercise practice embracing frequency and intensity - and the Leisure Time Index. Both indexes are descriptive of the participants' activity level and are expressed in MET values (Table 18).

2.3.3.2.2 Contextual inquiry

Since attitudes, behaviors, and personal opinions are usually challenging to capture, observation and shadowing of users [193] during the work-out time was performed on different days of the two weeks preceding the data collection process. We conducted informal conversations with end-users to identify relevant contextual information; the data was used to complement the user modeling process.

The video game experience questionnaire was used to investigate two main topics:

- Patterns of digital gameplay: frequency, intensity, history, social aspects, and gameplay's perceived skillfulness.
- Motivators to play: how they started to play, favorite games and video games, attractive elements for video games, interest to start/continue playing.

The interviews were semi-structured questionnaires with one interviewer and one note-taker present. A pre-determined set of questions were asked, and the inquiries were tailored with follow-up questions based on each interviewee's answers. This strategy was made to encourage interviewees to expand on their initial comments. Data were collected in individual sessions lasting around 1 hour per participant in a designated room at the senior gymnasium, where the system was set up. After a complete explanation of the procedure, participants signed the consent form, and the demographic information was collected. Posteriorly, the videogame experience questionnaire was used to collect game literacy information. The system setup and exergame prototypes were introduced to participants who had the opportunity to interact individually with them and provide feedback right after each playtest.

2.3.3.2.3 Contextual interview interpretation and users' Personas.

Following the focused RCD methodology, we conducted:

- *Contextual interview interpretation*: we analyzed the exercise routines of the older adults in the gymnasium, their socialization, and their relationship with colleagues and fitness instructors. We focused on identifying intrinsic motivators, main barriers, and possible facilitators for exergaming our local population in their gymnasium.
- *Work modeling*: we focused on developing multiple sequence models of daily exercise practice in the gymnasium, representing the ordered steps in which a senior performs the prescribed activities. Subsequently, we designed possible interaction scenarios with the exergames inside the gymnasium and represented them as sequence models of activities. This helped us envision future user experience issues and contrast them with the information collected from the real interaction using the digital prototypes (see section 3.3.2).
- *Affinity diagram*: during the observation and shadowing processes, detailed field notes were taken, coded, and subsequently merged with the notes from the contextual inquiry and exergaming experience. Since we were interested in viewing the opportunities to integrate exergaming practices in the senior gymnasium, an affinity diagram was used as a starting point for constructing design considerations. We coded the information based on users' preferences towards exercising at the gymnasium and behaviors that might reflect barriers to integrating the exergames in their daily life activities. Also, we clustered some definitions used by participants to describe their notions about videogames. Finally, the diagram revealed information regarding the users' habits and behaviors during exercise practice in the gymnasium as well as connections between technical elements and the

users' perception of video games. The affinity diagram was used to a) categorize older adults, bearing in mind their opinions and experiences with technology, and b) compile a list of insights and design ideas to modify the exergames considering the feedback gathered from the users and instructors.

- *User Personas*: users' archetypes for the active senior population were modeled according to the treated information. We focused on their willingness to integrate video games into their lives. We identified two main types of information to create our users' Personas.
  - Personal habits and physical activity practice: information related to the activities during the leisure time, ways ofexercising, preferences regarding activities of daily life and barriers.
  - Technology literacy and game experience: information regarding past experiences with gaming devices such as video game consoles, PCs, and mobile devices were considered in order to create a technological curriculum for the seniors. We focused on the motivators for playing exergames such as social elements [194], health issues, and curiosity levels.

The models for the user's Personas contained the following information:

- Basic information: a picture of the processed archetype, a small quote reflecting a personal opinion about video games, hobbies and preferred activities (such as traveling, exercising and having fun with children).
- Personal information: a short-biography describing some characteristics of the user's personality, such as preferences and habits, together with the physical activity agenda and representative goals and frustrations.
- Technology literacy and gaming willingness: technology literacy is primarily expressed in terms of past experiences with computers, mobile phones, tablets, and video game consoles. Six bar scores denote specific motivations to play exergames: fun and pleasure, escape from the daily routine, social interaction, physical health, mental health, and curiosity about new experiences.

Two realistic and one ideal user Personas were modeled. The realistic models are a Skeptic and a Curious archetype (see Figure 32 and Figure 33), referring to the senior's willingness to use video games in their daily routines. A profound mistrust over the possible health and social benefits of using video games characterizes the skeptic persona. This view is firmly based on the limited past experiences some older adults have with gaming technologies, which is reflected as a simplistic vision of playing video games: a waste of time. More than half of the interviewed users were identified as Skeptic users.

Figure 32: Model of the realistic senior Skeptic Persona.

Figure 33: Model of the realistic senior Curious Persona.

The Curious Persona is not indifferent to the incorporation of technology into her daily life activities. She has had some previous experiences with video games, and it had a positive impact on the perception of how helpful video games can be concerning social interaction and fun. Curious users are more prone to accept the use of exergames for exercising.

Finally, the ideal user is the Enthusiastic. Although we did not find active senior gamers in this research, we wanted to model the ideal user to understand the conditions for the long-term adoption of exergaming. The Enthusiastic archetype has a medium-high ability to interact with computers, mobile devices, or video game consoles. This user uses video games in her daily life routines, firmly believing in their health and social benefits (see Figure 34).

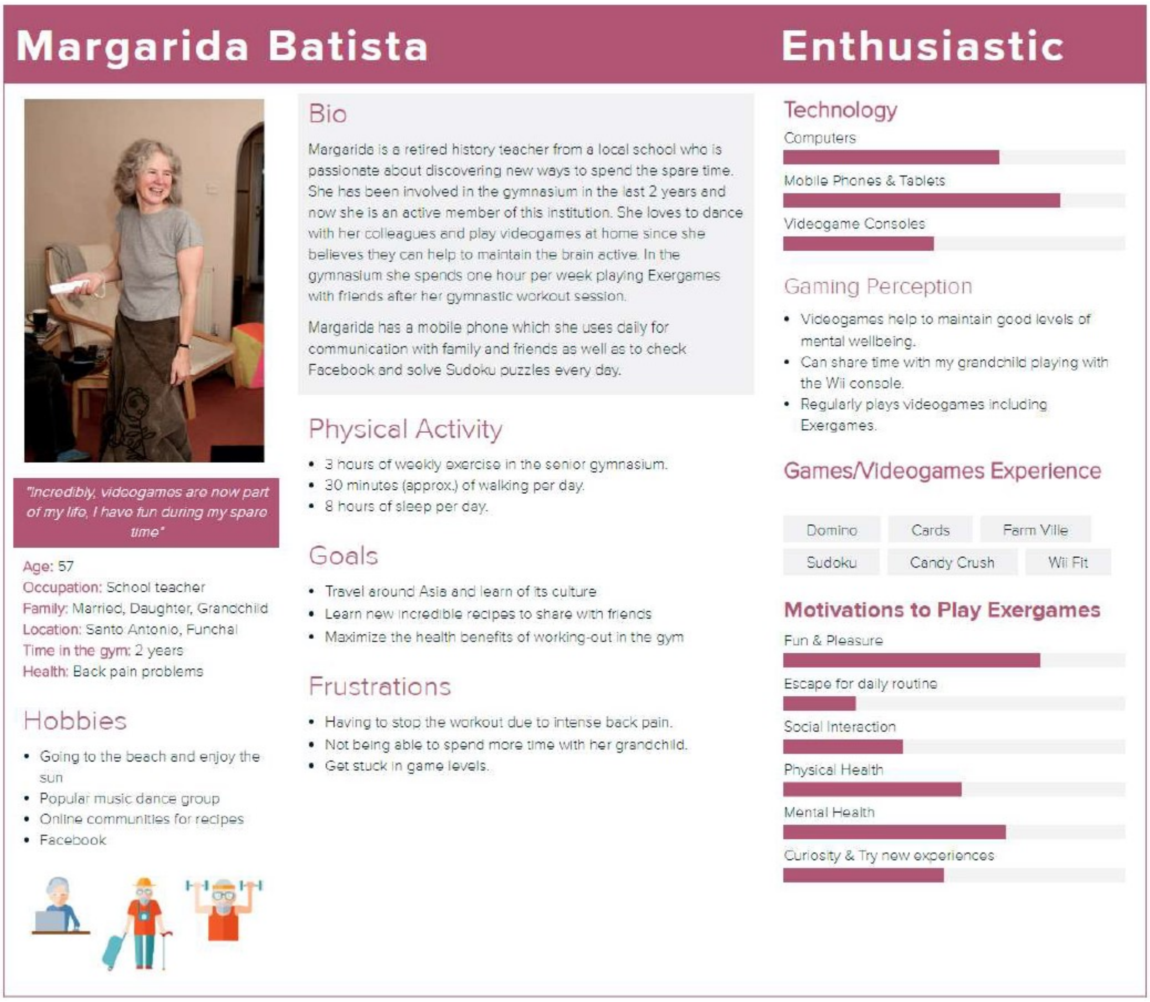


Figure 34: Model of the idealistic senior Enthusiastic Persona.

The consolidated models and the revealed affinities were evocative representations of the interviewed older adults, covering an amalgam of elements such as fitness habits and characterization, social behaviors in the gymnasium, personal perception about the use of technology for exercise, and characteristics of senior needs in terms of engagement and motivation to make exercise routines.

2.3.3.3 Iterations and Feedback

2.3.3.3.1 Playtesting with the Target Population

With the second version of the exergames prototypes, we conducted a first playtest with the target population at the senior gymnasium with the same users who participated in the contextual inquiry (see *Gym Playtesting* in Figure 29). The interaction with each exergame took about 5 minutes, and then users were asked to enumerate: a) three video game elements they liked, b) three video game elements they did not like, and c) three things they would like to add/remove from each exergame. Finally, we asked their opinion on how to integrate the exergames in the gymnasium. Videos from the experience were recorded to be analyzed afterward (see *Feedback Data* in Figure 29). Participants' interactions with the exergames were recorded to be analyzed by the principal investigators and game programmers to identify problems with the game mechanics, particularly the gesture interaction. Figure 35 shows the playtesting scenario with the technical setup.

Figure 35: Playtesting scenario depicting a user playing Grape Stomping.

We also conducted informal interviews with two exercise instructors from the gymnasium to understand their needs and reactions to the exergames. They supported the exergames' mechanics, the story theme and provided specific requirements such as:

- Include strategies to teach movements gradually, one at a time: for instance, while interacting with the Rabelos exergame, one of the users claimed to have "difficulties in coordinating the movements and understanding the movements," suggesting the need for more explicit feedback for learning.
- Increase the challenge to enhance workout effects: for instance, while interacting with the Exerfado exergame, one of the users stated wishing for "more rhythmic music" in order to get a more dynamic gameplay experience. Additionally, while playing the Toboggan Ride exergame, one of the users claimed to need "more obstacles" to increase the game difficulty.
- Allow broader customization of the video game parameters to facilitate difficulty adjustment and the inclusion of seniors with movement limitations: for instance, while interacting with the Grape Stomping exergame, one of the users claimed that in the game "it should be easier to push grapes since the movement can be bad for the shoulder," illustrating how game difficulty perception may vary between subjects.

After gathering the feedback from the users and instructors, we carried out multiple development iterations on the prototypes aiming to improve the mechanics and aesthetics and the usability of the system (*see exergames mechanics, aesthetics and usability revision* in Figure 29). A list of improvements was merged with the users' Personas models and the RCD process's insights to generate the 3rd generation of the exergames.

Final polishing tasks were done in the research laboratory before releasing the exergames' final version after 19 weeks.

Merging the information from both RCD and exergaming feedback, we observed:

- Social aspect: older adults enjoy playing mainly for two reasons, they like to win competitions, and they find it to be an opportunity to socialize (skills and experiences). Multiplayer playability is a crucial factor in improving technology adoption.
- In-time feedback: for both the skeptic and curious profiles, the lack of past experiences with gaming technologies hinders a fluid interaction in different stages. It is imperative to have high quality and frequent feedback in the video game to facilitate the understanding of what to do, how to do it, and when.
- Customization of movements: personalization options must include clear strategies to facilitate the interaction for people with diverse motor abilities. Since each exergame includes a set of various body movements (e.g., drag and step, row and navigate), health professionals should be able to activate/deactivate individual body gestures depending on users' abilities.
- Cognitive tasks: the inclusion of more cognitively demanding activities in conjunction with physical exertion might enhance health benefits and increase engagement, and consequently, the likelihood of long-term adoption of this technology [136].
- Parametrization: the multiple game parameters facilitate the personalization of activities regarding fitness domains and training dimensions. By defining a set of game parameters, the difficulty will be controllable, allowing a more precise adaptation to the specific motor and cognitive skills.

2.3.3.3.2 Modifications to the Exergames

Several changes were made in each exergame, which aimed to incorporate both the *RCD insights and Personas* as well as the feedback data obtained from the playtests. The most relevant changes are the addition of multiplayer options, instructional videos for in-time feedback, and the parameterization of each game. Table 19 summarizes the changes made in each game after the iteration process.

Table 19: Changes carried out along the iteration process with the Portugal tour exergames

Exergames	**First implementation**	**Implementation after iterations**
Grape Stomping	a) Single player. b) No instructions supplied. c) Pull and step. d) Only physical challenge. e) Nonparameterized winemaking process.	a) Single and multiplayer (collaborative or competitive). Users are situated in different tanks. b) A tutorial providing in-time videos for feedback in the drag, stepping and lateral movements. c) Pull and/or step. d) Physical and cognitive challenges. Cognitive challenges were added using wine recipes. e) The treadmill velocity and the number of steps for winemaking parameters were added for difficulty modulation.
Rabelos	a) Single player. b) Instructions used images and were displayed at the beginning. c) Rowing and boat lateral displacement. d) Nonparameterized course. e) Predefined calibration for the floor.	a) Single and multiplayer (collaborative or competitive). Users embody different boats. b) A tutorial providing in-time video feedback for the rowing, lateral displacements, and barrels pick up gestures. c) Optional rowing and/or boat direction. Either can be automated by the system. d) Docks and rocks separation distance was added as game parameters for difficulty modulation.

		e) A calibration module was added to the project allowing an initial configuration of the spatial parameters for the position tracking.
Exerfado	a) Single joint interaction. b) Only one song was used. c) Nonparameterized gestures and song-tracks. d) No clear audiovisual feedback supplied.	a) Waist and feet tracking. b) MIDI files can be added. c) Four parameters addressing arm extension, the time between notes, note sliding time, and song track can be used to personalize the experience. d) A swipe gesture visual feedback was added in conjunction with audio feedback for punishment and reward.
Toboggan Ride	a) Single player. b) No instructions supplied. c) Direction control only with trunk lateral flexions. d) Nonparameterized course. Only objects to collect. e) Predefined calibration for the floor. f) Male avatar.	a) Single and multiplayer (collaborative or competitive). Users placed in different toboggans. b) A tutorial providing in-time videos for feedback on the car's acceleration and deacceleration as well as for the car's direction. c) Lateral displacements were added to control the direction of the car. d) We added elements to avoid in the scene and a control parameter for the distance between the objects to collect. e) A calibration module was added, allowing an initial configuration of the spatial parameters for the position tracking. f) Male and female avatars.

After this iteration process, the graphical elements of each exergame were improved. The GUI was made uniform across all the menus, including the in-time feedback for gesture guidance and a panel for text and video instructions. We included English and Portuguese languages for the initial configuration menu, a tutorial, in-time instructions (they appear sequentially to support users along the experience), and a final screen with the game metrics. Figure 36 shows screenshots of the final version of each exergame.

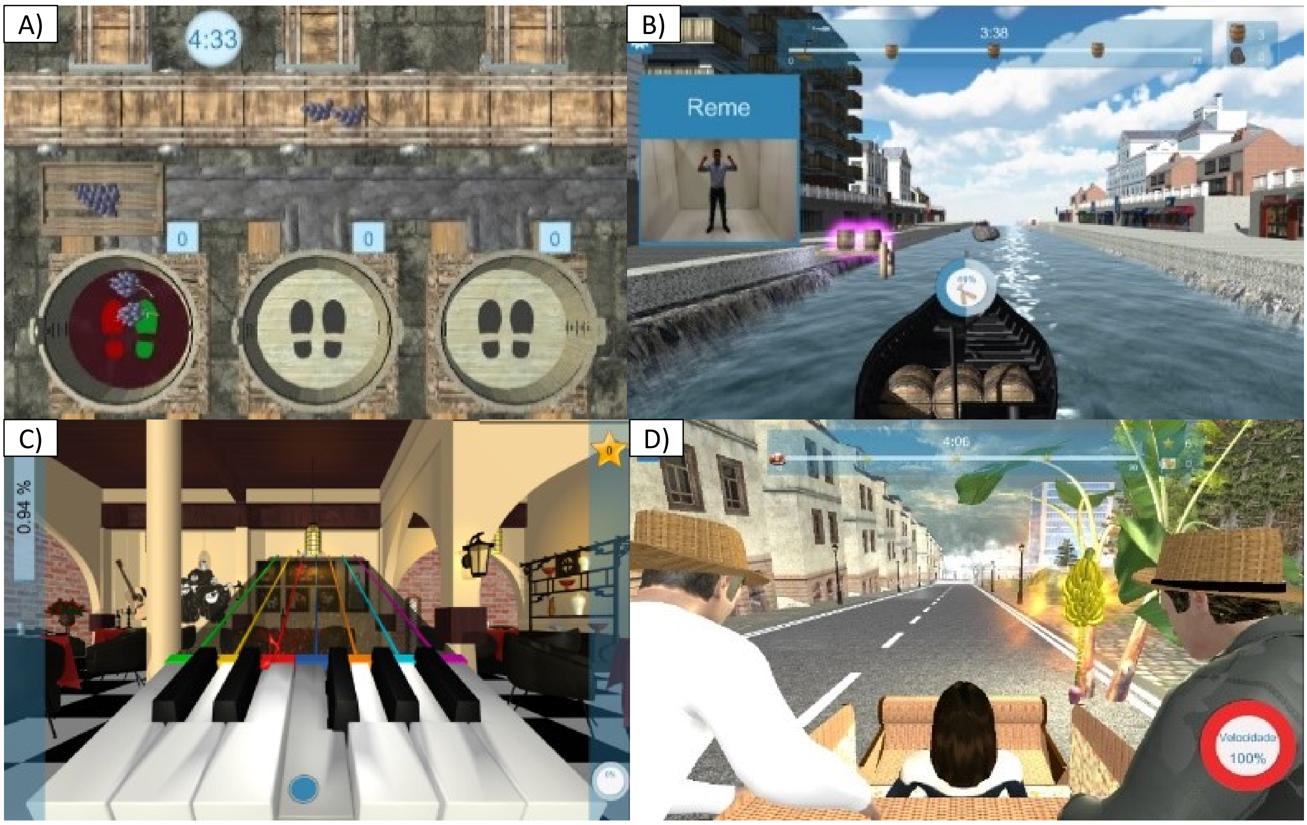


Figure 36: Screenshots of the final exergames. (A) Grape Stomping, (B) Rabelos, (C) Exerfado, and (D) Toboggan Ride

Finally, Table 20 summarizes the exergame parameters that can be set to adjust the fitness training dimensions' difficulty. The table shows how specific sets of game parameters cover all fitness domains.

Table 20: Summary of the complete set of exergame parameters, which can be modified to cover the motor ability, cardiorespiratory, and muscular strength fitness domains

	Motor Ability				Muscular Strength		
Exergame	**Balance**	**Agility**	**Flexibility**	**Cardio-respiratory**	**Upper**	**Lower**	**Trunk**
Grape Stomping	Pose, stepping, step height, side movements	Treadmill velocity, recipes, distractors (%)	Pose, step height	Duration, pose, stepping, step height, dragging	Dragging	Pose, number of steps for bunch	None
Rabelos	Pose, boat direction	Pose, separation of rocks	Pose, boat direction, rowing mode, separation of docks	Duration, boat direction, rowing mode, separation of rocks	Rowing mode, separation of docks	Pose, boat direction, separation rocks.	Pose, separation of the rocks
Exerfado	Tracking mode	Time between notes, note sliding time	Arm extension, tracking mode	Duration, boat direction, rowing mode, separation of rocks	Arm extension	Tracking mode	None
Toboggan Ride	Pose, car direction mode	Pose, objects separation	Pose, objects separation	Duration, car direction mode, objects separation	None	Pose, car direction mode, objects separation	Pose, car direction mode

The following is a summary description of the final version of the games:

- Grape Stomping: replicates the ancient traditional methods of grape pressing used in wine production, where people repeatedly tread the grapes with their bare feet to extract the juice that will be fermented to produce wine. This game presents three half barrels and a conveyor belt that continuously brings new grape bunches into the play area (see Figure 36A). The player must physically step into the projection and chose a barrel where to stand, then, by flexion-extension of the arms, they pull the grapes from the conveyor into their barrel. Once there are grapes in the barrels, the player can step on them to tread the grapes, which are converted into a rising level of juice contained on the barrel. Each grape bunch has a limit of steps it takes to be successfully processed. Once that that limit has been reached, more grape bunches are needed to be pulled from the conveyor belt into the barrel to raise the juice level higher. As soon as the grape juice hits the top, that barrel starts emptying its contents through a channel and becomes unavailable to play for the duration, forcing players to move laterally into another barrel. The grape bunches can come in three kinds, green, red, and rotten (distractors). The game can be configured to have an extra cognitive difficulty layer by having each barrel require specific amounts of red and green grape bunches. If these limits are broken, that barrel becomes unplayable for some seconds. Pulling rotten grapes into any barrel has the same consequence. The game's goal is to produce as much grape juice as possible, which elicits a stepping-in-place exercise, typically used in aerobic training. This is combined with the arm pulling motion for extra variety. The height that the feet must be raised from the ground to stomp a grape successfully can be configured at startup, together with the percentage of rotten grapes (distractors), the grape type requirements per barrel, and the time between new grape bunches on the belt, this allows an adjustment of the exercise difficulty.
- Rabelos: For centuries, the Douro river valley has been a wine-producing region of Portugal, famous for the port wine. In the past, to transport the wine barrels downstream, from the vineyards to the city's cellars, wooden cargo boats named Rabelos were used. This game replicates those voyages. The player is in charge of navigating a Rabelo boat downriver, avoiding obstacles, and docking on the margins to collect barrels. The game takes a third-person perspective from behind the boat, where the riverbanks are aligned with the projection's lateral edges, see Figure 36B. The lateral boat position on the river is controlled through the player waist position, which is directly mapped to the projection, or alternatively by leaning of the trunk. To move the boat forward, an arm rotation gesture that replicates the rowing activity is required. In this manner, this exergame aims at exercising the upper limbs. As the players row the boat downriver, they encounter rocks that they have to avoid through lateral movement and barrel-filled docks at the river's margins. These docks must be approached, and their barrels collected via an elbow extension-flexion motion. The goal of the game is to collect as many barrels as possible while avoiding the river rocks. The difficulty is set by adjusting the rowing mode (light or hard), the distance between docks, and the number of rocks on the river.
- Exerfado: this game reproduces the environment of a typical Fado house from Lisbon, Portugal, where people go at night to eat, drink, and listen to live music. Inspired by the Guitar Hero video game, the Exerfado resorts to music's potential as a physical activity stimulator. The projection renders a keyboard on the floor with seven keys aligned with the projection's bottom, see Figure 36C. The player stands at the bottom of the projection, and both feet control which key they want to activate. Vertically aligned with each key is a colored track over which musical notes travel downwards. The goal is to have the correct key activated when a musical note hits it. Therefore, the player must play the piano with their feet in synchrony with the visual cues. Over which track each note appears depends on the pitch of the music being played, with low pitch-making notes spawn at the keyboard's left keys and high pitch at the right. Additional to this, there are special notes that can be activated by an arm "swiping" movement. This activity is intended to train agility in both upper and lower limbs.

Missed musical notes lead to distortion of the song being played, producing negative audio feedback. The music to be played can be chosen from an extensive list of midi files, each with different durations. By setting, at startup, the falling notes' speed and the time between consecutive notes, the difficulty is controlled.

- Toboggan: on the island of Madeira, Portugal, a unique way of transportation was used in the past. Wicker toboggans, driven by two people, would carry passengers downhill, from the hills to the city center. This tradition is now kept alive as a touristic activity. In the game, this activity is recreated virtually. The toboggan and player are presented in the center of the screen from a third-person perspective, as seen in Figure 36D. Lateral movement is controlled just like Rabelos, by moving sideways along the bottom edge of the projection. The speed is adjusted through trunk inclination; leaning the trunk forward accelerates the toboggan while leaning backward deaccelerates it. Over the path, there are pedestrian crossings and car intersections to force the player to slow down. While the game's goal is to drive as far as possible in the allotted time, there are also obstacles to avoid, bonuses to collect, and speed limits to keep.

2.3.4 Guidelines for Context-Aware Exergame Design

Past paradigms for developing fully engaging video games aimed at exercise promotion and rehabilitation failed to provide clear elements for sustainable motivation [178]. They failed at both system personalization [83] and contextual information integration [195], thus limiting exergames' impact. Our study emphasizes the need for a paradigm shift that will generate exergames for older adult populations that are more engaging and usable through the combination of multiple HCI techniques.

After analyzing our main results in designing exergames through human-centered design and iterative approaches, we propose a set of guidelines that intend to summarize our efforts in constructing a methodology for exergame design and development for exercise promotion in older adults.

a) Focus on human-centered approaches
   - Use agile software development and standardized methodologies (e.g., RCD).
   - Create user models to aid the communication process (e.g., Personas).
   - Go for the short and frequent field inquiries instead of a long and single interview.
   - Investigate the target population's game preferences and interaction limitations to enrich the game design process frequently.

b) Increase emphasis on fun experiences through gamification elements
   - Balance attractiveness and effectiveness by interweaving storytelling with game mechanics [196].
   - Exploit the familiarity of contextualized activities and factors (e.g., the wine industry in Portugal).
   - Consider elaborating a game design and specifications document before starting the development process.

c) Improve collaboration with sports and healthcare professionals
   - Include professionals actively in each project stage.
   - Ask healthcare professionals to summarize and list the health requirements for the gamified system.
   - Promote spaces and situations that boost interactions between technologists and sports and healthcare professionals (e.g., game jams).
   - Include strategies to integrate technology in sports facilities and healthcare institutions beyond the research phase.

d) Rapid digital prototyping and fast iteration
   - Investigate existing software and methodological tools to accelerate the prototyping process before starting a development process from scratch.

- Include multiple playtesting sessions to evaluate game design elements with end-users before the research trials.
- Contemplate at least three generations of prototypes in the project timeline.

2.3.5 Discussion

Transversal inquiries that consult users about technology preferences and systems' usability issues are not enough to effectively model seniors' personalities. Research-centered approaches conventionally start from an extensive scientific-grounded review of the evidence, moving to system development, and finishing with a strict evaluation process with end-users. On the contrary, an immersion into daily routines and preferences of end-users is usually the first step in a human-centered intervention (also known as empathizing) [197], [198]. User models of behaviors, preferences, habits, and needs may help reveal issues such as intrinsic (vs. extrinsic) motivations in using games as exercise and conditions, preferred time, and intensities [157]. Due to the vast heterogeneity of the elderly population, the design of highly personalized exergames for exercise promotion or rehabilitation must include rich contextual information obtained directly from real-life scenarios. Balancing research-centered approaches with human-centered methods might enhance the design of the interventions. The development of three different user Personas from our study confirms the need to increase the content's personalization, targeting a broader set of alternatives that can satisfy specific user's needs regarding personal preferences and health requirements. Since studies reflecting attitudes towards exergaming technologies adoption and gaming affinities in our targeted population (Portuguese older adults) are scarce [140], [199], we decided to create our own models to cover particular aspects such as willingness to play exergames in the gymnasium, personal predispositions with gaming technologies and motivators beyond pure enjoyment or novelty of the exergames.

Moreover, dynamic modifications of video game mechanics and aesthetics based on playtesting insights with final users might increase familiarity with the storyline, facilitating the recognition of specific video game elements (such as places and activities), which can elucidate memories of personal experiences or transport users to dreamed situations. Therefore, favoring user experiences for the elderly and increasing the possibility of long-term technology adoption. Concerning the adoption aspect, we ensured that all the exergames developed had flexibility regarding game parametrization, thus reinforcing the possibility of creating a more personalized game experience.

The convergence of both components from the dual flow model, effectiveness, and attractiveness, seems to be a dependable method for avoiding the development of exergames with high health demands and inferior engagement levels. Attractiveness has been demonstrated to be a bottleneck factor in the popularization of exergames among several populations [200]. The gamification of specific exercise activities demands strategies to overcome stuck points in the design and development process. Following some of the guidelines presented in Table 15 for exergame design [201], [202], our efforts to gamify touristic activities and recreate Portuguese spaces include:

- Embracing the variability of movements established in the health requirements makes them a part of the story instead of disconnected and isolated actions.
- Showing real-time feedback on movement quality in order to help senior users learn about movements and coordination.
- Designing the exergames as virtual visits to different Portuguese regions facilitates an imaginary journey, avoiding the classic static scenes for physical activity experiences. Also, the exergames were designed under a narrative umbrella that can enhance motivation and increase engagement [196] during gameplay.
- Enriching the video game experience via multimodal stimulation using thematic soundtracks for some exergames as well as using the music itself as a video game mechanic (Exerfado). Furthermore, the

spatial augmented reality used in the Grape Stomping exergame promises to become an exciting mediator between natural movements and the video game rules and goals. Past research with this technology shows how promising it is in assistive technologies for rehabilitation purposes [186], [203].

The conceptualization of the video game design stage as a collaborative and interdisciplinary process is required to guarantee a balance between engagement and video game effectiveness. We decided to use a robust theoretical basis in exercise prescription for our target population, the ACSM guidelines, which are the most widely used exercise guidelines for fitness training [7]. Theoretical notions such as defining exercise training in terms of specialized fitness domains, such as cardiorespiratory and muscular strength, as well as pre-defining exercise intensities and movements, allowed us to define the gamification strategies better from the beginning. Furthermore, these exercise guidelines were a critical common point between game designers, researchers, and sport science professionals, facilitating communication and workflow in general. Although current approaches in games for health development are mainly driven by the research community [204], we believe that a closer relationship between sports and health professionals working in the field and experienced video game designers/developers might make the difference for technology adoption. This intersectoral collaboration with health stakeholders should occur in a continuous timeline, transversally to the inquiry, design, and development processes [174], [205]. Video game designers must be sure to fully understand the health requirements to provide benefits through a playful digital experience. On the other side, health professionals must be aware of technological constraints to gamifying training or health-related activities. This reciprocal knowledge provides a rich scenario to balance both components in the dual flow model [16]. Ironically, after research trials, many video games developed do not become available, thus impeding their integration into healthcare units. Nevertheless, new web-based platforms have been promoting the sharing of novel serious games developed for health purposes, helping in more widespread use and adoption of these interactive technologies. Considering physicians', therapists,' and sports professionals' skepticism levels regarding the use of video games for well-being purposes [206], a collaborative approach might help the adoption of these technologies in their daily work routines.

The fast pace of technology development has created tensions in serious gaming research due to the risk that the hardware/software used in the video game will become outdated before its evaluation [207]. Typically, healthcare interventions using serious games are developed after strict literature review processes, protocol assessment, and short pilot studies with a small number of users and experts. Although exergame design requires developers to be familiar with motion tracking sensors and game engines, novel software and hardware tools have been developed to accelerate the creation of playable digital prototypes that can be used for initial tests [208], [209]. Then, as described in [210], healthcare interventions using information and communication technologies can utilize approaches focused on evaluating specific working mechanisms (such as video game mechanics) rather than the complete system's operability. Video game playtesting can be used to iterate over concrete elements, which will strongly affect the user experience. Agile software development methods such as Scrum can be applied to game development [211]. Here, video games are continuously tested with final users providing valuable feedback that drives the next steps. This iterative process requires active collaboration between health professionals, video game designers, developers, and final users. We focused on testing several video game mechanics for the development of our exergames, which was useful to find a good match between the required movements, the story proposed, and the affordable gestures for the Kinect sensor. Moreover, the exergame jam provided an optimal scenario to exchange ideas between the different actors and foster cohesiveness between the teams. Interestingly, past investigations have shown how game jam events are viable approaches to rapid video game prototyping and generate interest in health-related issues [212]. Finally, we highlight as strengths of our research approach: i) the active participation of final users during design stages, ii) game design decisions based on real insights and data from the field, iii) the use of an iterative refinement process which avoids the chocolate-covered broccoli effect and iv) "tunable" game parameters to improve training's personalization

2.3.6 Limitations

Despite the already mentioned strengths of the research carried out, it is evident that due to the several challenges faced and the dimension of our highly multidisciplinary approach, several limitations can be identified: i) time-demanding procedure, ii) input from several sources (e.g., researchers, designers, health professionals) can be challenging to manage, iii) the lack of a consistent, conclusive, fully reproducible, and straightforward method for contextually rich exergame design, v) the absence of clear theoretical frameworks for adopting human-centered exergame design and vi) a gender bias in our sample which might have impacted on the design decisions made. The gender bias arises from the fact that the clients of the seniors' gymnasiums where we conducted the study are predominantly females, which resulted in a heterogeneous sample in terms of gender.

2.3.7 Conclusion

Digital gameplay through exergaming can enhance the motivation for doing physical activity by providing enjoyment and fun and stimulating physical and cognitive capacities. Traditional exergame design processes often fail in systematically including end-users' gameplay preferences and primary motivators, thus reducing their likelihood of being adopted and extensively used. Also, practical examples of integrating end-users' feedback and contextual factors to enhance gamification strategies for exergaming are scarce. Here, we demonstrated that using human-centered design methods for creating contextually rich exergames allows a better understanding of underlying motivators and preferences for playing exergames in older adults. Specifically, we illustrated how these methods and the produced insights were rigorously used in the exergame design process, thus answering our main research question: how should human-centered design methods and insights be used to design highly contextualized exergames for exercise promotion in older adults? Firstly, we showed that together with health professionals' feedback, rapid game development responses, and a systematic iteration process, the contextual information could be effectively used to create more personalized and highly enjoyable exergames. Secondly, we demonstrated how specific insights such as preferred social aspects, feedback modalities, cognitively demanding mechanics, and system customization features could be successfully integrated into the arduous task of generating pleasurable and gamified exercise experiences. Finally, Figure 29 shows concrete examples of which techniques were used in each design process stage, allowing identifying inputs, processes, and deliverables. For instance, from the conceptualization stage, the game design specification document was vital to facilitate communication with the game developers. Furthermore, from iterations with playable prototypes in end-users, a significant amount of feedback data was collected and processed to identify improvements in the game elements. After almost 20 weeks, we passed through many stages using several methodologies, which channeled our desire to bring a realistic and complementary view of how to gamify fitness training beyond points, leaderboards, and badges. Contextual factors should be continuously inquired to move towards a harmonious balance between effectiveness and attractiveness. We demonstrated the feasibility of combining game design methodologies and human-centered design to produce a complete set of exergames to support seniors' fitness training. To conclude, we believe that due to its diverse nature and flexibility regarding game parameters, this set of exergames can be used to create training programs that can meet and exceed the physical activity guidelines proposed by the ACSM for older adults. There is a continuous need for more extensive research to disentangle the impact of novel exergame design methodologies on fostering health and well-being in aged populations.

3 EVALUATION OF CUSTOM LARGE PROJECTION EXERGAMES FOR ELDERLY FITNESS[6]

3.1 INTRODUCTION

In the preceding section, we presented four exergames for the elderly, intended to train several key functional fitness dimensions. The game design was steered by what was learned from the literature on serious games for health and made to the target population via a user-centered design approach [79]. A Kinect v2 is used as a game controller, this provides low-cost hands-free interaction, and with the help of sports professionals, the games' difficulty was parameterized to allow players with a wide range of functionality to remain competitive. This section of the thesis focuses on investigating the long-term effects on elderly functional fitness and health-related quality of life that our custom-made exergames produce in a continuous training program and the PA patterns associated with multi-dimensional training programs with these games.

Results from reviews on the quantification of PA during exergaming have exposed the potential of using active playing strategies to elicit physiological responses, as measured by heart rate (HR), oxygen consumption (VO2), and energy expenditure (EE) [84], [161], [213], [214]. Most of the studies with healthy older adults have used commercially available exergames (e.g., Nintendo Wii, PlayStation Move, Xbox 360 with Kinect) [18], [156]. For instance, energy expenditure in a group of 19 older adults was measured during boxing and bowling exergaming sessions (Nintendo Wii and Xbox 360 with Kinect) [138] and compared against rest. Three different measurements were used to quantify the energy expenditure: indirect calorimetry (respiratory gas exchange), two dual-axial accelerometers worn on the right hip and dominant wrist, and rating of perceived exertion using a 6-20 scale. Results revealed that those off-the-shelf exergames provided light-intensity exercises and elicited significantly higher energy expenditure compared to rest. Reported physical exertion (RPE) values were not significantly different between equivalent exergames played during 5-minutes intervals. Cardiac and electrodermal activities were also measured during an exergaming session that exhibited greater HR modulation (as measured through heart rate variability) and arousal responses compared to the non-active version of the same game [182]. Other studies compared exergaming training with closely matching exercises, showing, for instance, lower EE and HR responses in exergames when compared to aerobic exercise on a treadmill [215] and lower perceived exertion levels for exergaming conditions against non-gamified exercises [216]. No studies were found to report the effects of multidimensional training in the quality and quantity of movement measured during exergaming interventions, which is necessary to enable a more comprehensive interpretation of the benefits of training multiple physical functions [18]. One study concluded that exergames are a feasible alternative to traditional aerobic exercises for older adults, however, without comparing to other exercise modalities or resting state [217].

In longitudinal interventions, pre- and post-test methodologies have been used to investigate the physical effects of exergame-based training programs in older adults, quantifying the longitudinal physical effects of such exercises. In a 24-session intervention (1 hour per session, 14 weeks), Maillot and colleagues measured the fitness responses in a group of older adults while playing with the Wii Fit game. Cardiovascular responses measured through the mean HR, maximum HR, and HR reserve did not show significant differences either in the second, twelfth, or twentieth sessions [13]. The HR levels measured fell within moderate (and below) intensity ranges for older adults [7]. Using a more immersive, adaptive, and interactive setup, older adults who survived a coronary artery bypass grafting intervention were evaluated in a submaximal endurance training

[6] This section includes content both published in The Visual Computer journal [78] and accepted for publication on the Games for Health journal.

program with a controlled exergame played on a treadmill. Participants were subject to twenty training sessions (30 minutes per session, ten weeks) and were evaluated for maximum load, target oxygen consumption, and target HR during the intervention. Results revealed that by the end of the twentieth training session, 9 of 10 subjects that played the exergame had reached the recommended target HR. Moreover, all players also achieved their targeted metabolic costs after two training sessions [218]. In another study, wearable activity trackers were used in investigating the long-term use of a custom-made exergame targeting the improvement of balance skills in older adults. The average walking speed was used as the kinematic marker to measure an exergaming training program's effectiveness that lasted 12 weeks. On average, there was an improvement in the walking speed (5%) in the older adults that used the exergame compared against a reduction in the same variable with those who did not [89].

Similarly, functional fitness tests have been used to quantify the effects of PA using exergames in senior players, demonstrating significant differences in fitness domains such as balance, muscle strength, flexibility, and other important mobility factors [143], [158], [219]. It has been observed that exergames can be as effective or more than conventional exercise [18] with small but significant clinical effects outcomes [12], several studies reporting particular improvements in gait, balance, and cognitive function [220], [221]. Previous studies on the effects of exergaming have shown a positive impact on physical fitness and health in general [9], [12], [18], [81], [83], [84], [139], [140], [143], [161], [220]–[223]. As well as traditional exercise, exergaming requires more energy expenditure compared to resting [220] and facilitates moderate-intensity physical activity levels [84], [161]. Also, exergaming has been shown to be as effective or more than conventional exercise [18]. Regarding clinical effects, the outcomes have been shown to be small but significant [12], with several studies reporting particular improvements in gait, balance, and cognitive function [220], [221]. Exergames have also been found to reduce depressive symptoms while improving the quality of life [220] and positively affecting healthy lifestyles [12]. In the particular case of exergames for the elderly, it showed positive impacts on physical and mental health [139], with benefits on balance and postural control [81], [140], [222], and balance and mobility [9]. Increases to overall fitness [143] and motivation aspects [83] were observed with older adults participating in multidimensional training with exergames. Additionally, participants have reported forgetting time, place, and pain to a certain degree, which can benefit exergames' immersive nature [223].

Although there is an extensive body of research in this field and increasing efforts in assessing the effectiveness of exergames as a complementary (or alternative) way to deliver exercise training programs to the older population [156], there are still considerable efforts to be done if exergames are to become a scientifically proven modality for exercise. Most literature reviews point to a lack of methodological excellence in this field. According to a review of 149 publications, the average participants in exergames studies are teenage boys, with 40% of the research focusing on 10 to 20 years old [224]. This review also concluded that the demographics should be widened to include the elderly and females underrepresented. It was also observed that almost three-quarters of the studies were being conducted in a laboratory environment, which reduces the generalizability of results and feasibility in field settings. When focusing on exergames for older adults, a review of thirty articles showed that two-thirds of the studies targeted balance or fall prevention as primary outcomes [221], with a lack of studies covering other motor performance domains. A review of 60 studies targeting older adults only found that over two-thirds of them used commercial exergames [18]. Unfortunately, these games tend to be designed with other populations in mind and overlook older adults' specific needs, having many inappropriate design choices for this population. Additionally, they never capitalize on the potential that personalization to goals and performance levels bring, which custom-made games can deliver. Other reviews into this subject also report a lack of devotion and studies exploring the long-term effect of exergames [9], [18]. Also, beyond the clear potential in elderly physical health improvement provided by exergames [9], many approaches fail in the quantification of the PA through accurate metrics [161], where the use of full-body tracking sensors further expands the possibilities of gathering quantitative data relative to rehabilitation progress, caloric expenditure, and aerobic activity [225], [226]. To guarantee effectiveness in PA promotion through exergames, movement quantification, via activity trackers and specialized tools, accurately characterizes the exercise and defines their suitability to be used in the older population. However, the use of these technologies has been limited, and the investigation of the body responses during long-term interventions with exergames is still a work-in-progress [227].

There is a need for comparative and descriptive studies that report the differences in older adults' physical and perceived responses to custom-made exergames and conventional exercises as these data might reveal critical elements for the design of genuinely effective exergames [16]. The related work introduced above points to lower EE and RPE values in exergames sessions compared to conventional exercise sessions. Different studies measuring HR during exergaming have shown conflicting results, either failing or meeting the target HR. This leads us to expect that exergames' training sessions might be used in elderly training programs as a way to achieve an efficient MVPA level while having a lower EE and RPE. Therefore, one of this section's goals is to demonstrate the effectiveness of such training sessions in eliciting recommended levels of PA in older adults, measured by gold-standard activity trackers and RPE, and to assert if such sessions have an impact on the behavior of participants in long-term multidimensional training. The other goal is to validate the long-term effects of using these games in a continuous training program on elderly functional fitness and health-related quality of life.

To evaluate the four exergames' effectiveness and benefits for the elderly presented in the previous section [56], we performed a three-month-long study. It comprised a combined exercise program (exergames and conventional training) and an equivalent conventional exercise training program, acting as a control condition. With this design, we wanted to address some of the above limitations and answer the following questions:

1) What are the older adults' motor performance and quality of life benefits obtained by complementing custom-made (multidimensional) exergames with traditional exercise during longitudinal training compared to traditional training alone?
2) Can custom-made exergames be as effective as conventional exercise in achieving MVPA levels during elderly multidimensional training sessions? And if so,
    a) Will the EE in exergames be lower than in conventional exercise?
    b) Will the RPE in exergames be lower than in conventional exercise?
3) Does exergaming, used as a complement to exercise in long-term multidimensional training, affect how the elderly perform when exercising?

3.2 METHODS

3.2.1 Experimental Design & Study Protocol

We designed a 12 week long randomized controlled trial, where participants were randomly allocated to an experimental or control group. Both groups underwent two sessions per week consisting of warmup (10 minutes, stretching, muscular preparation), multidimensional physical training (40 minutes, intense exercise), and cooldown (10 minutes, muscle relaxation). Two sports professionals alternately led the sessions that were designed to be equivalent in Frequency, Intensity, Time, and Type (FITT) [132]. The training was structured following ACSM recommendations for multidimensional training for older adults [7]: 50% of aerobic training, 30% upper and lower limb strength, and 20% motor ability training. The difference between the groups lay in the exercise modality which was practiced at the sessions.

- *Exergames group* – Combined exergames and conventional training: Engaged once a week in individual exergames sessions. The other weekly session was a conventional group exercise.

- *Control group* – Conventional training: Engaged in conventional exercise group sessions two times per week.

The two conditions' exercise sessions were carried out in a suitable room of the local senior gymnasium. Conventional training sessions were based on the gym's exercise patterns, covering marching in place, step touches, stepping on pads, squats, and others. Odd-numbered sessions (1 to 23) of both groups were always conventional exercise, while the even sessions (2 to 24) were conventional for the *control* group and exergames for the *exergames* group.

PA levels of participants were measured at each session, and participants wore accelerometers that were configured, considering their age, gender, and weight. A first session was performed one week before the study to familiarize users with the exercise routines.

Participants were assessed for functional fitness and health-related quality of life at four different moments: Pre-intervention (0^{th} week), mid-intervention (6^{th} week), post-intervention (12^{th} week), and 1-month follow-up (16^{th} week).

As the research questions demand, this experimental design was done so that the long-term impact of complementing traditional exercise with exergames could be measured and compared to an equivalent training modality. More specifically, it enabled us to measure benefits over time of fitness, balance, and health-related quality of life for both exergames and traditional training, which we could compare. Moreover, measuring the same data at the 1-month follow-up enables us to assess if the effect of such an intervention would last.

3.2.2 System Setup

The system that was used to play the exergames consisted of a modern gaming computer with an i3-8100 CPU, 4 GB of RAM, and an Nvidia GTX 1050 2 GB graphics card (Nvidia, California, USA); an Optoma GT760 projector (Optoma, New Taipei, Taiwan) that was oriented to project the games on the floor, and a Kinect v2 sensor that was mounted under the projector to track the players' bodies, as already seen in the previous section Figure 30. A white PVC (PolyVinyl Chloride) floor surface was used to improve the graphics' contrast and facilitate the interaction. The 2.5m by 3.0m floor projection acted as a digital playground large enough to be able to demand measurable amounts of exertion in older users. The system was installed in a multi-purpose exercise room at the local senior gymnasium facilities; lighting conditions and privacy were controlled.

3.2.3 Exergames

Five custom-made exergames were used in the intervention; the four games presented in Section 2.3 of this thesis (Grape Stomping, Rabelos, ExerFado, and Toboggan Ride) [56], [79], and Exerpong, all shown in Figure 37.

Exerpong, Figure 37E, is an exergame adaptation of the classic games of Pong or Breakout, used to provide a fast-paced game experience aimed at train aerobic endurance [182]. In this game, the player controls a virtual paddle through lateral movements while a ball bounces around the walls. The player, who stands along the bottom of the projection, has his or her waist tracked, and the game matches the paddle location on the screen with it so that both player and paddle are always aligned. The ball bounces around the other three edges of the screen which are covered by walls; the player must then use the paddle by moving laterally along the bottom of the projection to prevent the ball from going through the lowest edge. A pattern of colorful bricks is represented at the center of the screen; these bricks get destroyed whenever the ball passes over them twice. The game's goal is to clear these bricks without letting balls pass through the bottom of the screen. The game difficulty can be adjusted by varying the paddle's width and the ball's size and velocity. Alternatively, the game can be set to adjust its own difficulty according to the player's performance by increasing the paddle width by a small amount every time a ball is lost, making it easier, and increasing the ball velocity when the player successfully hits it, making it harder.

All the exergames, except Grape Stomping, are played along the bottom edge of the projection, but with players far enough from it so they are not forced to look directly down to the floor in what could be an unnatural pose. By contrast, in Grape Stomping, where the player needs to stand over the projection to interact with it, its immersiveness is advantageous to players' perception and engagement. Because these games are played through full-bodied interactions, enabled by a natural user interface that relies on simplified motions and gestures, the need to acquire competencies before interaction and engagement is minimal.

The exergames' sessions were defined according to their training dimensions, ACSM guidelines (times and intensities), and characteristics from each game, as follows: i) Exerpong (aerobic) 10 minutes, ii) Rabelos (strength) 7 minutes, iii) Grape Stomping (aerobic) 10 minutes, iv) Exerfado or Toboggan (motor ability) 7 minutes. Breaks of 2 minutes were used for the transition between each exergame.

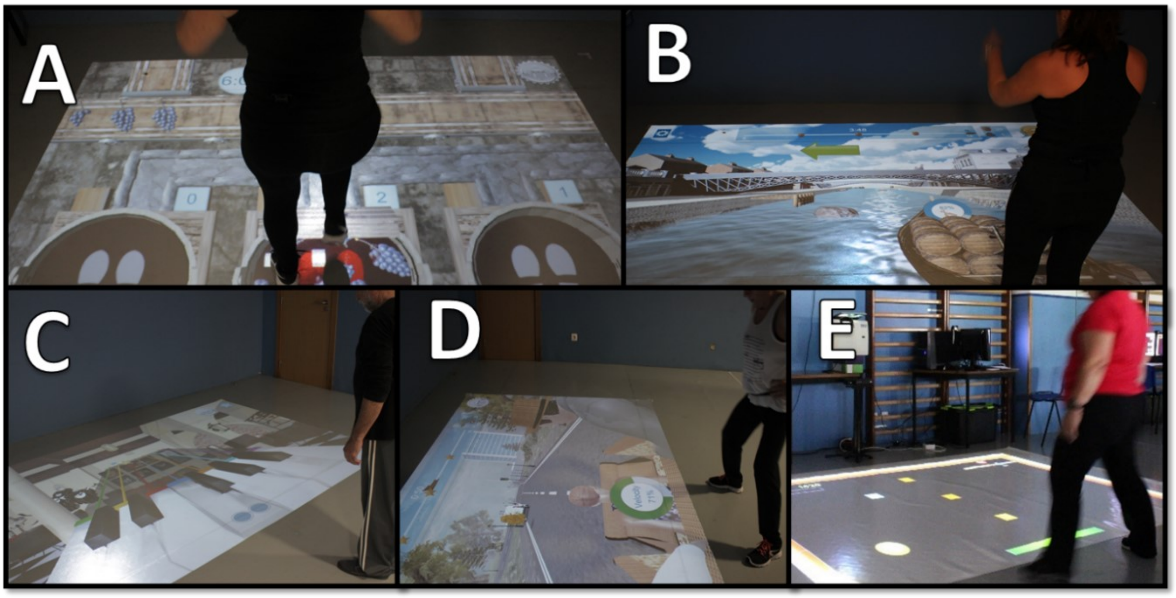


Figure 37: The set of exergames used in the Combined group. Grape stomping (A) and Exerpong (E) train aerobic fitness. Rabelos (B) trains upper and lower limbs strength while the Exerfado (C) and Toboggan Ride (D) train motor ability.

3.2.4 Measurements

To assess fitness, we applied the Senior Fitness Test (SFT) battery [86]. The SFT is a reliable and valid fitness assessment tool for older adults, designed to be easy to administer, covering the major fitness components needed to perform everyday activities in this population. The fitness domains that were covered were: lower-body strength, upper-body strength, lower-body flexibility, shoulder range of motion, agility and dynamic balance, and aerobic endurance. The six tests used, their respective domains, and scores are as follows:

- 30-s Chair Stand Test (CST) – Lower-body strength, number of complete sit-to-stand repetitions in 30 seconds.
- 30-s Arm Curl Test (ACT) – Upper-body strength, number of bicep curl repetitions with a weight in 30 seconds.
- Chair Sit-&-Reach (CSAR) – Lower-body flexibility, centimeters that the fingertips go past the toes when reaching the toes from a legs-extended sitting position.
- Back Scratch Test (BST) – Shoulder range of motion, hand overlap (or distance) in centimeters when meeting them behind the back.

- 8-Feet Up-&-Go (8FUG) – Agility and dynamic balance, time to stand, walk 2.4 meters, return, and sit. A higher score means lower performance.
- 6-Minute Walk Test (6MWT) – Aerobic endurance, distance, in meters, walked in 6 minutes.

To quantify fall risk, we used the Short Form Fullerton Advanced Balance Scale (FAB), which consists of 4 tasks that test static and dynamic balance [228]. Each item is scored on a 5-point ordinal scale (0-4); the sum of all item scores is the total FAB, rated 0 to 16. Finally, the 12-Item Short-Form Health Survey (SF-12) was used to evaluate the quality of life outcomes from the participants' perspective [229]. The score of this test is made of two components, mental and physical.

Finally, to quantify PA levels, we relied on both objective (measured) and subjective (RPE) data. Accelerometers quantified PA: the research-grade three-axial accelerometer ActiGraph WGT3X-BT (Actigraph, Florida, USA) was used to monitor the player's PA. The waist-worn sensor was set to register the complete routines of 40 minutes at 30 Hz sampling frequency and using epochs of 30 seconds. Using the manufacturers' standalone software (Actilife 6.10), we computed the time people spent in MVPA (in minutes). Besides, the software provides the EE (metabolic equivalent - METs). This sensor has been widely used and is considered a gold-standard tool to quantify PA in different populations [230] due to its accuracy in characterizing the human movement effectively. To collect subjective data of the levels of reported physical exertion after each exercise routine, we used a pictorial version of the 0-10 rating of RPE scale OMNI [8]. The final OMNI score for the exergaming sessions was the average of each game's reported values in that session.

3.2.5 Participants

The study was reviewed by the ethical council of the Faculty of Human Motricity, University of Lisbon (Review 14/2017), which confirmed it to comply with the national and international guidelines for scientific research with humans. It took place in a local (Madeira, Portugal) senior gymnasium where active community-dwelling older adults were recruited. The inclusion criteria were: 50 to 75 years old, able to read and write, members of the gymnasium for three or more months, able to understand the procedure, game rules and goals, no severe visual impairments, no impediment to exercise practice, no severe or unstable heart diseases, a FAB score higher than 9, and no falls over the past six months. A total of 37 volunteers were gathered, two volunteers did not receive the intervention, and four failed to show at follow-up. The remaining 31 completed the study (Figure 38). Sixteen participants were assigned to the *control* group (12 females, age avg. 69.1 SD 4.4), and 15 to the *exergames* group (10 females, age avg. 67.6 SD 5). All the participants gave their informed written consent.

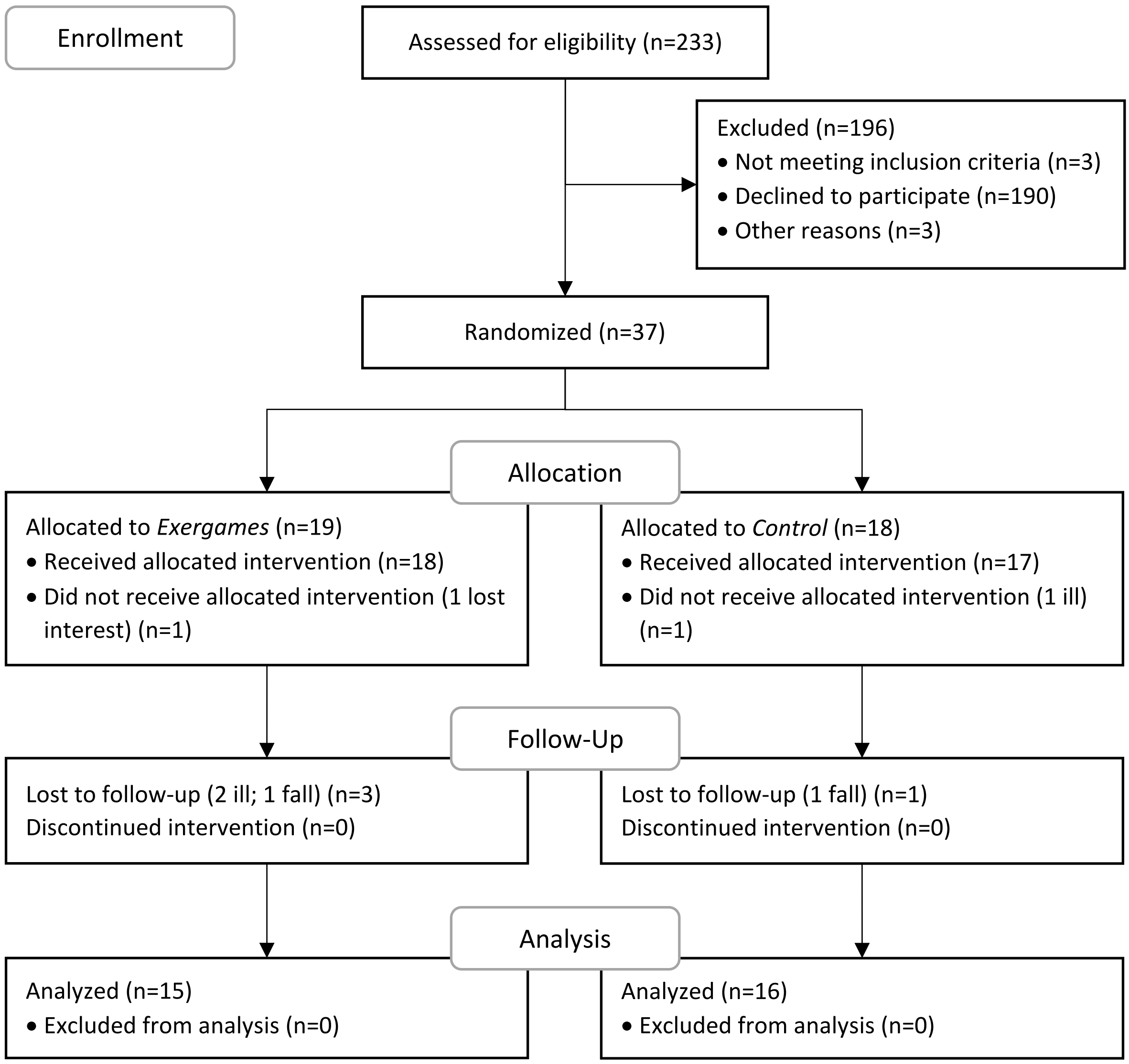


Figure 38: Participants flow through the phases of the randomized controlled trial.

The pre-intervention assessment of the participants' functional fitness through the SFT was used to confirm both groups' equivalency. No significant differences were found between the groups in any of the tests. See Figure 39 for more details.

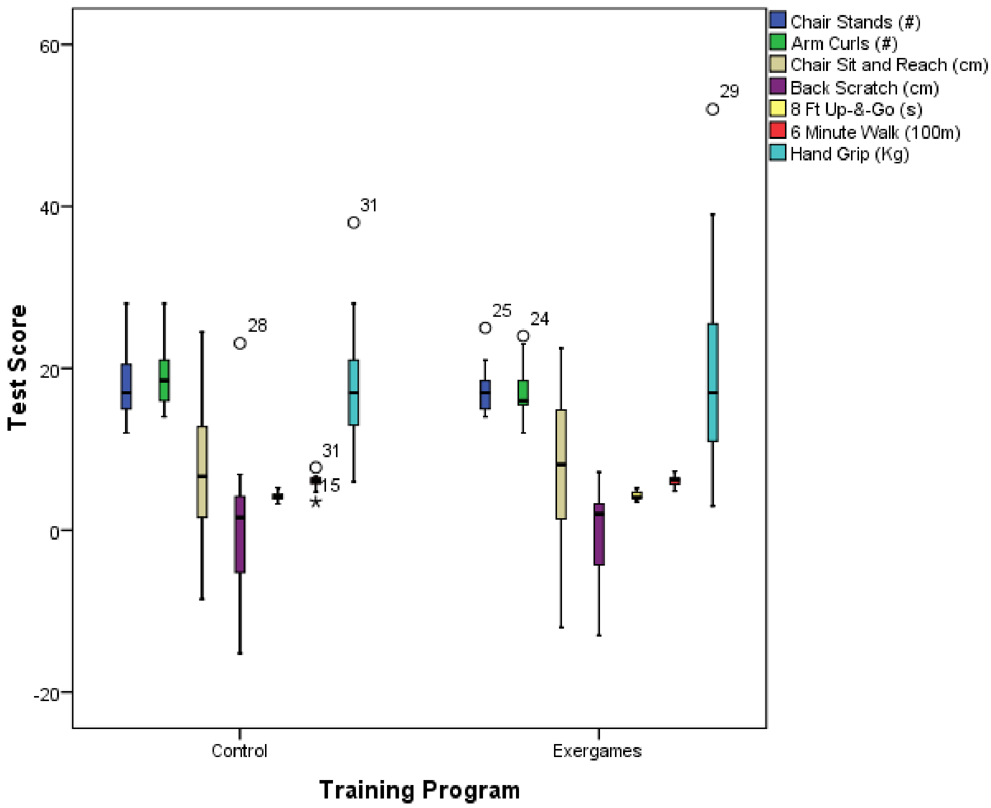


Figure 39: Tests' scores of the participants, from both groups, in the SFTs' battery of seven tests.

3.2.6 Data Analysis

Given the different nature of our data, we set up two different analyses, one concerning the four assessments done during the intervention, the other for the data originating from the individual PA measured at each session.

For the data relative to the four assessment moments of functional fitness and health-related quality of life, non-parametric tests were used due to either non-normal distributions or the data's ordinal nature. The normality of distributions was assessed using the Kolmogorov-Smirnov test. The Friedman test was used to detect significant within-group differences over time. The analysis focused on finding which variables had significantly improved during the intervention and which decreased from end to follow-up. For pairwise comparisons, the unidirectional Wilcoxon test was used to find if the increase from start to end (0^{th} to 12^{th} week) and decrease from end to follow-up (12^{th} to 16^{th} week) were significant. Next, a between-group analysis was done to understand if the *exergames* condition had significantly larger performance gains in each measurement than *control* from start to end and lower losses from end to follow-up. The difference in differences method was used [231], first calculating the differences between the 0^{th} and 12^{th} weeks, and 12^{th} and 16^{th} week, for both conditions. Then, using the unidirectional Mann-Whitney test to check for higher *exergames* gains from the 0^{th} to 12^{th} week and higher *control* losses from the 12^{th} to 16^{th} week.

Concerning PA, we divided the data into two, the odd sessions consisting of the conventional sessions of both conditions and the even sessions for both conventional and exergames. This allowed us to run the same analysis to 1) compare the differences between conventional exercise and exergames' sessions, and 2) see if the training program affected the users' response to exercise. For analysis of these data, a two-way Mixed MANOVA was used. The between-subjects factor was the training program each participant was allocated to (2 levels), and the within-subjects factor was program progression (session number). The dependent variables were RPE on the OMNI scale, METs spent, and minutes of MVPA. Separate ANOVAs were run for each dependent variable to ascertain which ones were genuinely affected by the training program.

There was incomplete data from accelerometry on the 1st, 2nd, 6th, 7th, 9th, and 11th weeks of the study. Thus, as the two-way mixed MANOVA requires complete data, we removed those weeks from the analysis. Additionally, data from the 5^{th} and 8^{th} weeks were also excluded because it failed Levene's test of equal variance. The remaining weeks, 3^{rd}, 4^{th}, 10^{th}, and 12^{th} were analyzed using the MANOVA with four levels of within-subjects factor.

The significance level used was α = 0.05, and Bonferroni's correction was used to correct for multiple comparisons. All analysis was done using IBM SPSS Statistics 22 (IBM, New York, USA).

3.3 Results

3.3.1 Senior Fitness Test

As mentioned in Section 3.2.4, to show the impact of exergames in elderly fitness, the Senior Fitness Test battery was used to evaluate the changes in several fitness dimensions over time. Each condition results are presented separately to determine which dimensions were influenced over time by each condition. This is followed by the difference between the impact of both interventions.

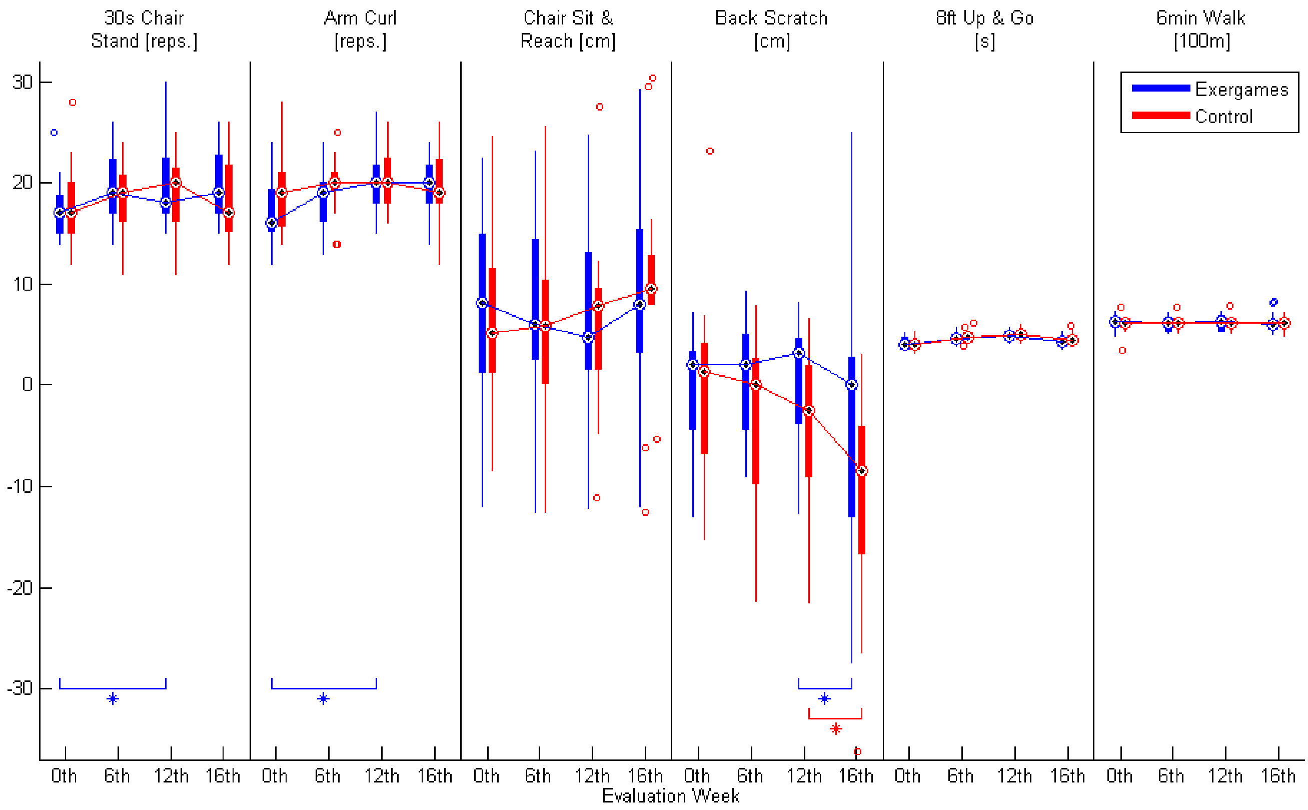


Figure 40: Results of the 6 SFTs over time for both conditions, significant differences highlighted with an asterisk.

3.3.1.1 Differences over Time

3.3.1.1.1 Exergames

There were significant differences over time, $p < .05$, in lower-body strength, $\chi^2(3) = 8.127$, upper-body strength, $\chi^2(3) = 11.553$, shoulder range of motion, $\chi^2(3) = 18.103$, and agility and dynamic balance, $\chi^2(3) = 14.534$. Post hoc analysis testing after Bonferroni correction showed a significant increase, $p < 0.05/2$, of lower-body, T = 6, r = -.44, and upper-body strength, T = 14.5, r = -.474, from 0th to 12th week, and a significant decrease of shoulder range of motion, T = 14, r = -.441, from 12th to 16th week (Figure 40). The agility and dynamic balance results were significant but did not validate our hypothesis due to the directionality of the differences.

3.3.1.1.2 Control

Only the shoulder range of motion, $\chi^2(3) = 25.591$, and agility and dynamic balance, $\chi^2(3) = 32.758$, were found to be significantly different over time. Post hoc analysis showed a significant decrease in shoulder range of motion between the end of the intervention and follow-up, T = 11, r = -.508 (Figure 40).

3.3.1.2 Differences Between Conditions of Gains/Losses Between Moments

When checking the difference in differences of the evolution of fitness from 0th to 12th and 12th to 16th between conditions, we found that upper-body strength and shoulder range of motion had significantly higher developments in the *exergames* group as compared to *control*, U=76, r=-.315, and U=67.5, r=-.373, from 0th to 12th week (Table 21).

Table 21: Descriptive statistics of the differences over time from pre to post-intervention and post-intervention to follow-up of both conditions for the SFTs and SF-12 results, and Mann-Whitney sig. differences of differences between conditions.

Measurement	Evaluation Week	Exergames Dif. Median	Exergames Dif. Interquartile Range	Control Dif. Median	Control Dif. Interquartile Range	Mann-Whitney Sig. Dif. in Dif.
30s Chair Stand (reps)	0th to 12th	1	5	1	3	No (p = .235)
	12th to 16th	-1	5	-1	4	No (p = .256)
Arm Curl (reps)	0th to 12th	3	4	1	5	**Yes** (p = .043)
	12th to 16th	-1	4	0	3	No (p = .357)
Chair Sit & Reach (cm)	0th to 12th	.90	7.0	-.4	8.6	No (p = .286)
	12th to 16th	1.35	9.7	2.8	9.2	No (p = .357)
Back Scratch (cm)	0th to 12th	1.00	2.5	-1.3	7.7	**Yes** (p = .019)
	12th to 16th	-4.15	8.6	-8.6	10.6	No (p = .153)
8ft Up & Go (sec)	0th to 12th	.73	1.36	.89	.56	No (p = .150)
	12th to 16th	-.50	.32	-.41	.42	No (p = .372)
6 min Walk (m)	0th to 12th	2.5	55	-5	45	No (p = .097)
	12th to 16th	-2.5	59	-5	60	No (p = .326)
SF-12 Mental	0th to 12th	11.91	28.57	2.38	22.63	**Yes** (p = .021)
	12th to 16th	0.00	15.47	-2.38	17.86	No (p = .157)
SF-12 Physical	0th to 12th	4.67	41.67	6.66	13.33	No (p = . 327)
	12th to 16th	0.00	25.33	0.00	14.67	No (p = . 066)

3.3.2 Short Form Fullerton Advanced Balance Scale

3.3.2.1 Differences over Time

3.3.2.1.1 Exergames

As measured by the FAB scale score, balance and risk of fall were found to have significant differences over time during the exergames intervention, $\chi^2(3) = 14.026$. The post hoc analysis showed that the FAB score increased significantly during the intervention, T = 2.5, r = -.499., but decreased in the following month, T = 3.5, r = -.488 (Figure 41).

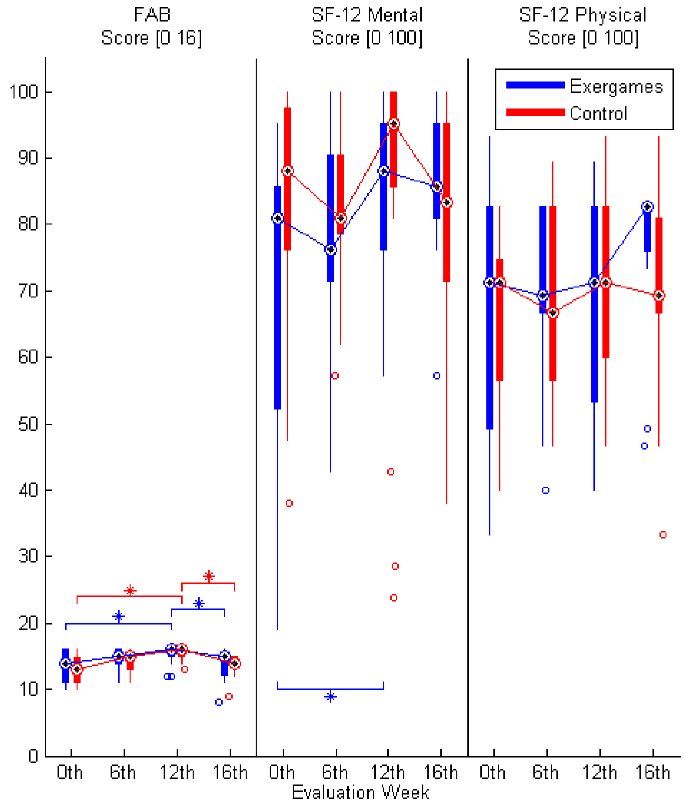


Figure 41: Results of the FAB and SF-12 evaluations over time for both conditions, significant differences highlighted with an asterisk.

3.3.2.1.2 Control

The FAB score was significantly different for *control* as well, $\chi^2(3) = 17.837$, revealing the effect of time in balance. The post hoc analysis showed that the FAB score increased significantly during the intervention, T = 0, r = -.497, and had significantly decreased one month after the intervention, T = 5.5, r = -.562 (Figure 41).

3.3.2.2 Differences Between Conditions of Gains/Losses Between Moments

There were no significant differences between both conditions in their differences over time.

3.3.3 12-Item Short Form Health Survey (SF-12)

3.3.3.1 Differences over Time

3.3.3.1.1 Exergames

When considering the health-related quality of life in the *exergames* participants, the mental component measured by the SF-12 questionnaire was found to have been significantly affected by time, $\chi^2(3) = 8.366$. Post hoc analysis testing showed a significant increase between the start of the intervention and the end, T = 10.5, r = -.515 (Figure 41). No effect was detected on the physical component.

3.3.3.1.2 Control

Neither the mental nor the physical components showed an effect during the intervention in the *control* group.

3.3.3.2 Differences Between Conditions of Gains/Losses Between Moments

When comparing conditions, only the mental component showed significant differences, with the e*xergames* having a significantly higher improvement (Mdn 11.91) than control (Mdn 2.38), U = 68.5, r = -.367 (Table 21).

3.3.4 Physical Activity – Comparison Between Different Program Sessions

When the conventional exercise sessions of the *control* group were compared with the corresponding weekly exergame sessions of the *exergames* program group, a statistically significant effect of the type of training program was identified, $F(3, 27) = 27.958$, $p < .05$; Wilks' $\Lambda = .244$. The univariate ANOVAs revealed significant differences on all three outcomes, with more METs spent on *control* sessions (M = 2.976, SD = .106) than *exergames* (M = 2.046, SD = .110), $F(1) = 37.138$, $p < .05$ (Figure 42A); but, on the other hand, MVPA, $F(1) = 11.044$, $p < .05$, and OMNI, $F(1) = 7.977$, $p < .05$, had higher marginal means for the *exergames* program (M = 36.183, SF = .545; M = 3.767, SD = .271) than *control* (M = 33.664, SD = .527; M = 2.703, SD = .262) (Figure 42B and Figure *43*A).

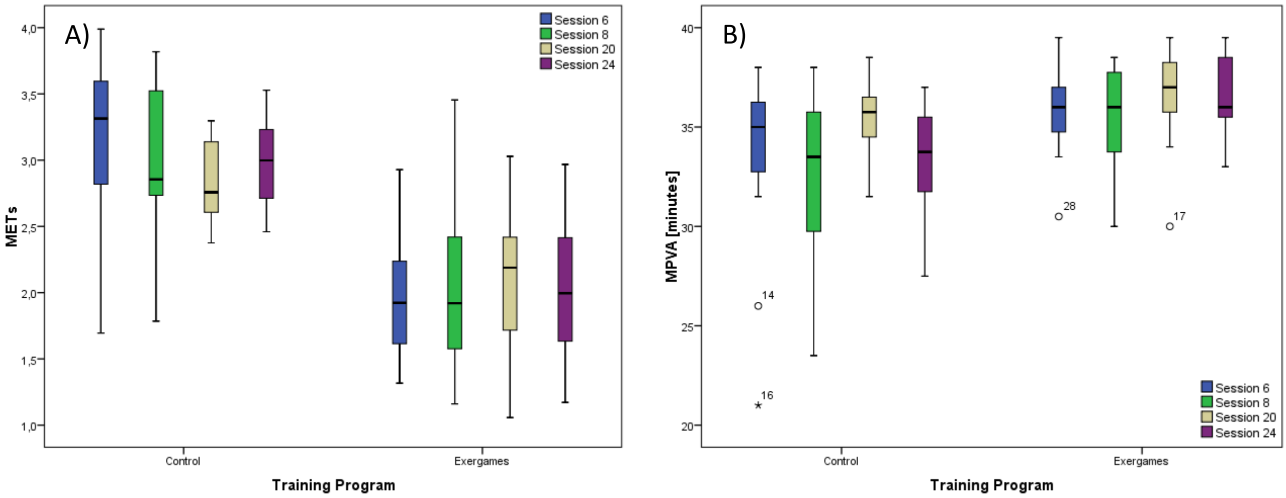


Figure 42: Total METs spent and minutes of MVPA during conventional sessions by participants in the control program and exergame sessions by the subjects in the exergames exercise program, at weeks 3, 4, 10 and 12.

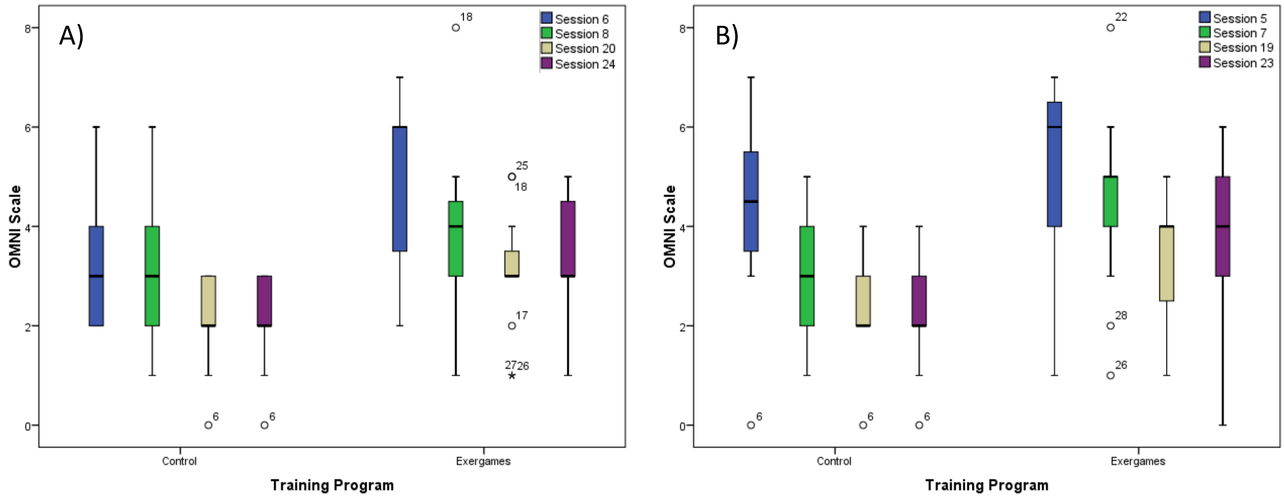


Figure 43: Self-reported exertion, on the OMNI Scale, weeks 3, 4, 10 and 12, A) at the end of the conventional exercise sessions by subjects in the control program and the end of the exergames sessions by the subjects in the exergames program, B) at the end of the conventional exercise sessions by participants of both the Conventional and Combined exercise program.

There was not a statistically significant interaction effect between the type of training program and time, $F(9, 21) = 2.106$, $p = .077$; Wilks' $\Lambda = .526$. The univariate ANOVAs of the outcome variables presented no significant differences for OMNI, $F(2.517) = 1.253$, $p = .296$, METs, $F(3.000) = 1.733$, $p = .166$, and MVPA, $F(2.911) = 1.198$, $p = .315$.

3.3.5 Physical Activity – Comparison Between Conventional Sessions of the Training Programs

The comparison between conventional exercise sessions of participants in the *exergames* exercise program with the equivalent (conventional) session by the *control* group revealed a statistically significant effect of the type of training program on the dependent variables, $F(3, 27) = 3.444$, $p < .05$; Wilks' $\Lambda = .723$. Separate univariate ANOVAs on the outcomes did not show significant differences in both spent METs, $F(1) = 2.280$, $p = .142$, and MVPA, $F(1) = .041$, $p = .841$. However significant differences were observed in the OMNI scale, $F(1) = 6.119$, $p < 0.05$, with higher score of RPE ($M = 4.117$, $SD = .302$) for the *exergames* exercise program than *control* group ($M = 3.078$ $SD = .292$) (Figure *43*B).

There was not a statistically significant interaction effect between the type of training program and time, $F(9, 21) = 1.810$, $p = .126$; Wilks' $\Lambda = .563$. Univariate analysis of the outcomes also failed to find significant difference in METs, $F(2.220) = 2.905$, $p = .057$, MVPA, $F(2.260) = .805$, $p = .465$, and OMNI, $F(2.596) = .830$, $p = .467$.

3.4 Discussion

This study aimed to determine the effect of exergames in functional fitness (e.g., balance, strength, flexibility) and health-related quality of life of older adults. A secondary goal was to investigate the PA patterns associated with multidimensional training programs with exergames. For that effect, custom-made exergames were used together with conventional exercise to create a combined multidimensional fitness training for healthy older adults. The equivalent conventional exercise was used as control. The study comprised a 3-months training intervention and a 1-month follow-up. In both conditions, we observed a global and positive change of the variables assessed in the participants.

3.4.1 Benefits on Strength

The functional fitness effects of commercial, non-customizable, exergames (e.g., Wii, Kinect) in longitudinal interventions with older adults have been widely questioned since those games have not been created considering the capabilities and limitations of aged players; thus, they likely fail in eliciting the desired responses [161]. The results obtained from our multidimensional training program with exergames revealed significant increases over time in lower-body and upper-body strength. No significant improvements were detected in the *control* group. Improvements of specific functional fitness domains such as strength in the upper and lower limbs can significantly impact elders' daily life activities [167], as well as an association with decreased physical impairments and functional limitations [232]. Only the upper body flexibility domain did not reveal sustained effects after follow-up for the *exergames* condition. The same was observed in the *control* condition. The reason for these results may be that the exergames used in this research involve lower-limbs in a more consistent and transversal fashion than upper-limbs. Thus, more features that encourage upper-limb movements, and training of other aspects besides strength, such as flexibility, should be incorporated in the next generation of these exergames. Research using similar methods with different training program duration and frequency showed similar results in the overall strength component [143].

3.4.2 Benefits on Balance

The FAB test was used to quantify training programs' effects on older adults' risk of experiencing injuries due to falls. Static and dynamic balance was assessed using the FAB scale, which covers aspects such as the participant's capabilities to stand on one leg and reach forward to retrieve an object. Both the *exergames* and c*ontrol* groups showed significant improvements in balance, revealing the importance of physical activity in supporting balance skills [7]. These findings are essential since improvements in balance can significantly reduce long-term functional impairments due to falls. Also, reducing the risk of falls has been associated with decreased mortality [233] and independence to perform daily life activities [234]. However, the positive effects were not maintained one month after the intervention. This can be due to the training program length, which limits a sustained impact [9]. These results support the idea that there is a continuous need for exercises that target this domain for maintaining a low risk of falls in older age.

3.4.3 Benefits on Health-Related Quality of Life

We observed improvements in perceived quality of life, particularly in the *exergames* group who showed a significant improvement in the quality-of-life mental component. These results are aligned with what was reported in exergames reviews: improvements in balance, cognitive function, quality of life, and health [9], [145], [220], [221]. So far, the relationship between exergaming training programs and perceived quality of life in older adults has not been well-established, and the existing literature does not provide conclusive insights [235]. This finding is vital for improving our understanding of the specific domains in which exergaming programs are better than conventional exercises [235].

3.4.4 Exergames & Conventional versus Conventional

Our *exergames* condition offered a complementary training of exergames and conventional exercise, in line with what is recommended [236], which was compared to an equivalent training of conventional exercise alone. The combined approach contrasts with the supplemental, adding extra exergames' sessions to conventional training, and the alternative approach, training with exergames only, predominantly studied in the literature. The results revealed that upper-body strength improvements in the *exergames* condition were significantly higher than those obtained for the *control*. This allows us to extract an important finding of this research: combining exergames with conventional exercises was more effective in impacting functional fitness compared with conventional exercises alone. Integrating exergames to the already existing conventional exercise programs seems to be a feasible and effective strategy to extend the use of these technologies beyond research trials in senior care facilities [237] and at home. Besides, one of the strengths of this study is the carefully selected *control* condition, which allowed us to carry out a fair comparison between conditions. Since past literature reviews of RCTs including exergames have pointed out the limitations of using non-equivalent control conditions, our decision to include conventional multidimensional training with a sports science coach allowed a more reliable scientific evidence towards defining specific and quantifiable effects of longitudinal exergaming [235].

3.4.5 Effects of Physical Activity

The results concerning PA show that older adults can use exergames to perform exercise sessions that meet the international recommendations of MVPA without altering their behavior, relative to conventional exercise, during a three-month training program.

As was hypothesized, the exergames' sessions were able to meet the MVPA goals, which even marginally surpassed the minutes of MVPA spent during conventional exercise. Differences between MVPA and METs reflect the need to interpret data of the activity trackers carefully. For example, more energy expenditure does not necessarily mean greater health benefits in the older population [132]. Having more time spent in MVPA and lower METs during the exergaming might be interpreted as participants in the conventional workout exercised with higher intensities but spent less time within the recommended levels when compared with exergaming. Therefore, participants during the exergaming were able to exercise with lower intensity levels but, at the same time exercising within the recommended levels for longer, being more efficient in their training, which answered our second research question. One possible explanation of why the participants spent more time in MVPA during exergames sessions than traditional exercise is that the games can keep players engaged with the activity, as the participants get absorbed by the individual stimulation of a game that reacts to them, which in conventional training would equate to personal training. In this sense, it could have also been the higher individualization of exercise through gaming that motivated people to engage for more extended periods. This might have meaningful impacts on the long-term adoption of exergaming technology in the older population, producing a firm notion of a safe environment for exercising [18]. Nevertheless, having a very similar MVPA in both multidimensional exergaming and conventional exercise illustrates how combined strategies can create enjoyable routines without losing efficiency in the PA. This study used research-grade activity trackers and perceived exertion scales in quantifying the PA levels of older adults with training that use custom-made exergames. Subjective data from the OMNI goes against what was expected, showing higher perceived exertion in the Exergames condition than *control*. However, it never exceeded the hard intensity (score = 8), which successfully meets the ACSM guidelines [238].

The fact that, over three months, there were no differences in the objective measurements of PA between conventional sessions of both training programs answers our third research question. Exposure to exergames had the same effect on PA as conventional exercise. Mainly, we focused on researching the long-term effects of engaging with custom-made exergames and their effects on PA's measured and perceived levels. Although some studies have reported the impact of exergaming in the time players spent exercising at MVPA levels in young adults [239] and children [240], to our knowledge, this is the first study reporting MVPA exertion in the older population.

Results emphasize the importance of accurately quantifying the PA and characterizing human movement during interventions with exergames to exploit them as an attractive option to promote exercise and an effective training modality for older adults [16]. While in our study the difficulty of the games was set by a trainer, given the player capabilities and needs, we envision that they can have specific difficulty levels for each fitness domain that would be set automatically according to fitness evaluations and live physiological monitoring of the player. Further studies comparing exergames with standardized training routines in older adults should reveal the quantitative differences of PA, leading to a more objective, consistent, and in-depth discussion about the real impact of playing while exercising beyond motivation and enjoyment.

We highlight the importance of using custom-made exergames rather than commercially-grade consoles to promote exercise in older adults. This has been mentioned as one of the most renowned limitations of exergames since the older population is diverse and complex [9], [18]. Our approach included a set of highly personalized exergames especially designed to cover multidimensional training in older adults.

Besides the positive results that we found in this study, a highlight is that we focused on the shortcomings referred in reviews of this research field by running a randomized controlled trial with a sample of our target population in field conditions, being evaluated by validated scales on a longitudinal fashion. This will positively help in building the argument on the benefits of exergaming.

3.5 Conclusions

Modern societies face both the demographic challenge of population aging and the health risk posed by increasing sedentary behaviors; exergames can be used to tackle these issues by fighting elderly sedentarism. Therefore, there were two aims in this section, the first was to determine the benefits of exergaming in the functional fitness, and health-related quality of life of older adults, and second, to measure the PA intensities achieved during sessions with our exergames. To that effect, the custom-made exergames were used together with conventional exercise to create a multidimensional fitness training for healthy older adults that was compared to conventional exercise. In both conditions, we observed a global and positive change of the variables assessed in the participants. Additionally, the exergames were shown to be an effective complement to training programs for the elderly, with the exergames' sessions promoting a higher percentage of time spent in MVPA than conventional exercise, which is a benefit for growth and preservation of functional aptitude. From the significant improvements observed in several of the dimensions measured in the *exergames* condition participants, we conclude that older adults' physical training benefits can be enhanced when the training includes exergames. Integrating carefully designed and highly personalized exergames in multidimensional fitness training can be attractive (e.g., more engaging and fun) for older adults and more effective in eliciting the desired physical improvements. The higher RPE by the participants while playing exergames shows that greater care to workload monitorization is required, strengthening the need for better sensor integration with the games.

Our research adds to the body of literature in exergaming for older adults by 1) showing results of a longitudinal intervention where custom-made exergames, instead of commercial, are used 2) providing data collected in a real-world setting through a randomized controlled 3-month longitudinal intervention, 3) using a quantifiable and multidimensional evaluation methodology to validate the use of exergames in real-life scenarios, using widely validated and accessible physical fitness tools and questionnaires, and 4) demonstrating the benefits of using exergames as complement to conventional exercise instead of as a supplement, as recommended by previous review papers [236].

3.6 LIMITATIONS

Despite the positive results, some limitations should be acknowledged. Identifying the multidimensional fitness training program's specific benefits' origin when combining several exergames with conventional exercise is impossible. Some of the combined training program gains declined one month after the intervention, showing the need for lasting exercise programs to provide durable benefits in older adults. Although the activity tracker used in this study is considered a very accurate and trustable device for PA quantification, studies have suggested that the cut-points which define the light, moderate and vigorous thresholds are dependent on the units of analysis (30 epochs and tri-axial in our case) and subjects variability [241]. However, there is still no consensus on the values for these cut points for the older population [242], [243]. We adopted a wide range of ages (50-75) in our sample; related literature typically reports studies with participants over 65 years of age, making it somewhat difficult to compare results. Additionally, our inclusion criteria guaranteed a healthy sample of participants, not representative of the general older adult population. We suggest following up this research with validation on frailer population, usually institutionalized under senior care, given our positive results. Finally, a technical limitation is the setup used, making this intervention hard to adopt in senior care facilities. For this reason, we have developed "PEPE: Portable Exergames Platform for the Elderly" to deploy our custom-made exergames [244].

4 CONCLUSIONS – MAIN FINDINGS AND CONTRIBUTIONS

As was pointed out throughout this thesis, there is a growing body of evidence that exercise through gaming can deliver several health benefits, particularly to older adults. This thesis's main goal was to capitalize on this potential by developing exergames and then validating the hypothesis that they could be used, in complement with traditional exercise, to generate significant fitness benefits in elderly players during longitudinal training.

The process described in the thesis consisted of three main steps. First, we developed a comprehensive and flexible low-cost technology for full-body interaction with large projection displays, which we validated in user studies showing that it performed at a comparable level to laboratory-grade immersive technologies. Second, we validated automatic methods to assess senior fitness and used the technology to study user preferences in interaction with floor projections. The knowledge gathered was complemented by a participatory design process that resulted in a set of exergames to promote elderly fitness. Lastly, the exergames were put to the test, in the field, in a longitudinal randomized control trial of 3 months, which revealed specific benefits in favor of the exergames group.

This thesis represents the culmination of a lengthy process; alongside the seven papers adapted and presented, seven others were written and published during its progress, see Appendix A – Curriculum Vitae: Publications. The work presented here involved 111 participants in total, from the development of projection technology, exergames' design, and finally to their validation. Perhaps the most significant testament to its relevance is that the exergames developed already crossed into non-academic scenarios are currently being used in elderly homes by private and public entities.

Further details on the specific contributions made at each step of this thesis are presented in the following subsections.

4.1 VIRTUAL REALITY SURROUND-SCREEN PROJECTION SYSTEMS

In our effort to design exergames that would positively affect the fitness of elderly players, we started by developing the technology that would enable large interactive projections. This led to one of the main technical contributions of this thesis, the KAVE open-source software. While its versatility is sufficient for implementing floor projection exergames, such as used in this thesis, it affords much more complex setups and immersive experiences. Its main advantage lies in managing VR experiences in monoscopic surround-screen systems such as CAVEs. Relatively to the standard CAVEs, it offers three main advantages. First, it is a free and open-source software with low hardware requirements and cost, which enabled us to build a surround-screen projection system for 5.000€. Second, it supports full-body tracking, meaning that the whole-body acts as a controller instead of hand controllers. Finally, it is entirely wearable-free, requiring neither glasses, wands, or markers to track the human head and limbs.

We evaluated our technology by using it in a CAVE system and conducting two tests, one to evaluate objective characteristics and another for subjective. In the objective evaluation, we wanted to answer the following research question: How accurate and precise is the Kinect v2 in estimating the user's head position in our KAVE-powered CAVE? For this, we measured the accuracy and precision of the Kinect v2 sensor. In the worst conditions measured, the sensor had an average error norm of 1.66 ± 1.18 cm and 0.04 ± 0.02 cm jitter. We found this accurate enough for real-time head position estimation, which is otherwise a limiting factor for immersive VR experiences. In evaluating subjective characteristics, we asked the following question: In a simple VR action space search task, to what extent can the KAVE induce presence while remaining cybersickness-free in a representative sample of healthy adults? For this, we measured the feeling of presence (through the Slater-Usoh-Steed Questionnaire and Presence Questionnaire) and cybersickness of 31 participants who interacted with our system. We then compared our results with the reports from interaction with a laboratory-grade CAVE and two HMDs. The KAVE was no different from a laboratory-grade CAVE in induced presence; however, it had significantly lower performance than HMDs. The KAVE produced significantly lower cybersickness than the HMD alternative. Our system was deemed to produce sufficient presence (as a CAVE) in tasks located in the user's action space.

4.2 DEVELOPMENT OF EXERGAMES FOR ELDERLY

After developing the appropriate projection technology for our needs, and before the design of the exergames themselves, we had to ensure two main elements: that the sensor intended to be used with the games was good enough to assess fitness accurately and reliably, and that we would be using the most effective way for elderly players to interact with a floor projection.

In the context of the first element, we developed an automated version of three senior fitness tests (SFT [86]), the 2-minute step test, the 30-second chair-stand test, and the 8-ft Up-and-Go test. The validation of such a system would ensure that it could quantify users' behaviors and consequently adapt games to the measured metrics. Therefore, we asked the question: Can a computerized tool be used to score specific SFTs? In a user test with 22 elderly participants (avg. age 65), our system was shown to be no different from the ground truth, a sports science professional, in scoring the 30-second Chair-stand Test and 2-minute Step Test. It provided very high rates of true detections of key body movements. In scoring the 8-foot Up-and-go Test, the score was deemed precise but inaccurate due to an offset, which can be controlled. Collectively, this ensured that we could design exergames relying on the correct detection of such body movements while using this sensor as a game controller.

Regarding the second element, we were able to contribute to NUI's body of knowledge by studying the performance and preferences of the elderly in interacting with floor projections, a population often overlooked. The research question asked was: When designing a NUI to be used by an elderly population in floor projection displays, what interaction is best? In an experiment with 19 elderly participants (avg. age 70.2), when measuring usability (System Usability Scale) and workload (NASA-TLX), a clear preference for the direct mapping of the feet to the display was observed relative to pointing. These results conditioned the development of our exergames.

Finally, we asked ourselves: How should human-centered design methods and insights be used to design highly contextualized exergames for exercise promotion in older adults (aged 55+)? We used those methods to understand better older adults' underlying motivators, preferences, and needs for playing exergames. Despite the low interest for this modality of exercise in the elderly population, we found strategies to raise it, such as promoting competition, use of real-time feedback to mitigate lack of experience, contextualization of the games' themes to familiar places and activities, parametrization and personalization of the game to the player limitations and performance, and inclusion of cognitive tasks during play. The design effort under those guidelines resulted in a set of four exergames, each representing a region and a distinguished cultural activity from it, under the context of a national tour. While each game is aimed at a particular fitness dimension, their difficulty and exertion demands can be adapted to meet the player's needs. This is done by controlling which game mechanics are needed to play or are automated for players with disabilities, while the difficulty is controlled by the virtual elements' speed and distance, and type and amount of player movement required.

4.3 EFFECT OF CUSTOMIZED EXERGAMES ON ELDERLY FITNESS

As the final contribution of this thesis, we aimed to quantify the impact of using our custom-made exergames for fitness training in elderly players; a population often overlooked in this field. We asked a set of three main questions:

- Can custom-made exergames be as effective as conventional exercise in achieving MVPA levels during elderly multidimensional training sessions?
- What are the older adults' motor performance and quality of life benefits obtained by complementing custom-made (multidimensional) exergames with traditional exercise during longitudinal training compared to traditional training alone?
- Does exergaming, used as a complement to exercise in long-term multidimensional training, affect how the elderly perform when exercising?

After a 3-month long randomized controlled trial with 31 elderly participants, we observed that exergame sessions successfully met the moderate-to-vigorous intensity goal of exercise recommended by ACSM to this age group, keeping them engaged longer than in traditional exercise sessions. Regarding the second question, the use of exergames in combination with traditional training was not only beneficial to motor performance, such as balance (assessed by the FAB scale [228]) but also produced benefits in dimensions that traditional training alone failed to affect, namely on lower and upper-body strength (assessed through the SFT [86]). This is valuable information as domains other than balance and fall prevention are not often studied in this population. Additionally, a positive evolution in the mental component of the quality of life, measured through the validated questionnaire SF-12 [229], was greater in the participants who played our exergames than those who solely engaged in traditional exercise. Furthermore, regarding the third question, we found no differences in objective measurements of PA between conventional sessions of both training programs.

In addition to the positive outcomes, it is essential to note that a substantial effort was made to avoid previous research's methodological pitfalls, efforts such as conducting our study in field conditions, the use of validated tests for senior fitness assessment, and accurate metrics for PA quantification, and the use of games purposely developed for elderly training. These support the feasibility of adopting this practice in senior gyms, its results' generalizability, and strengthen our conclusions.

4.4 Future Work

The KAVE software supports head-tracking through both the Kinect v2 SDK and any system sending UDP messages in the ARTTRACK2 format. However, the discontinuation of the Kinect v2 camera is an obstacle to the adoption of our software. For it to compete or be relevant, its future development should address this tracking limitation. Adding support for new off-the-shelf solutions such as Azure Kinect (Microsoft, Redmond, USA) or Intel RealSense cameras (Intel, Santa Clara, California) would provide an immediate solution with a possible but uncertain increased performance [245]–[247]. Alternately, creating an abstract layer between the software and the sensor could make it sensor agnostic. Nevertheless, the open-source nature of the KAVE makes it easily accessible and modifiable by the community.

Current and future uses of the research presented in this thesis are being developed and planned. Some of the exergames have been adapted to have dynamic difficulty, which changes during play according to the player bio-physiological signals, resulting in a closed loop control paradigm. Besides exergames, there is ongoing development of virtual environments for functional fitness assessment aimed at the elderly population. They benefit from the ecological validity of VR experiences by creating a high sense of presence, increasing the transfer of skills to the real world, and take the form of virtual bus rides to assess balance and nature treks to assess cardio-respiratory endurance [248], [249].

AFONSO RODRIGUES GONÇALVES

Phone: (+351) 963579049
afonso.r.goncalves@gmail.com

Rua Fernando Lopes Graça, N 400, 2º Esq
2775-570, Carcavelos, Portugal

Education

PhD Universidade da Madeira, — 2021
Informatics Engineering – Human-Computer Interaction

Thesis: "Fitness Applications for Healthy Older Adults Using Large Projection Displays – Methodology, design, assessment, and field validation"

Advisor: Prof. Sergi Bermúdez i Badia

MSc IST, Universidade Técnica de Lisboa, — November 2011
Aerospace Engineering – Specialization in Avionics

Thesis: "Aircraft Navigation and Attitude Determination using Multiple GNSS Receivers"

Advisor: Prof. Fernando Nunes

BSc IST, Universidade Técnica de Lisboa, — January 2011
Aerospace Engineering

Research Experience

Madeira Interactive Technologies Institute, Funchal (Portugal) — 2014 to 2020
Research Assistant, NeuroRehabLab

- Associated to the "AHA - Augmented Human Assistance" and "MACBIOIDI: Promoting the cohesion of Macaronesian regions through a common ICT platform for biomedical R & D & i" projects
- Planning, execution and documentation of research in the fields of:
    - Human Computer Interaction
    - Virtual Reality
    - Serious Games
- Software development:
    - Depth sensor based human motion analysis tools
    - Serious games for health
    - Full development of a surround screen projection system (KAVE)
- User studies:
    - Responsible for VR interaction and serious games' studies with adults and seniors
    - Support to VR studies using EEG

Institute for Systems and Robotics, IST, Lisbon (Portugal) — 2013 to 2014
Research Assistant, VisLab

- Associated to the "POETICON++" project
- Planning, execution and documentation of research in the field of humanoid robotics:

    - Learning of affordances
    - Perception systems
    - Action planning
    - Simulation
- Software development for the iCub humanoid robot

Work Experience

Sensei - SENSEIDATA, S.A., Lisboa (Portugal) 2021 to Present
AI Game Engine Software Engineer, Synthetic Datasets

EFACEC - Engenharia e Sistemas, S.A., Moreira da Maia (Portugal) 2012 to 2013
Engineer Intern, Space Activity Office

- Associated to the European Space Agency "ABPA - Assessment and Breadboarding of a Planetary Altimeter" project
- Modelling of the planetary altimeter:
    - Creation of the mathematical model and simulator
    - Definition of the model interfaces with users and subcontractors
    - Integration of the subcontractor's models

Publications

Book Chapter

Accoto, F., Vourvopoulos, A., **Gonçalves, A.**, Bucho, T., Caetano, G., Figueiredo, P., De Paolis, L. and Bermúdez i Badia, S., "The Effect of Vividness in CAVE VR Neurofeedback Training for Enhancing Working Memory," Technology-Augmented Perception and Cognition, Human–Computer Interaction Series, T. Dingler and E. Niforatos, Eds. Cham: Springer International Publishing, 2021, pp. 11–45.

Gonçalves, A., Nóbrega, F., Cameirão, M., Muñoz, J. E., Gouveia, É., and Bermúdez i Badia, S., "From Body Tracking Interaction in Floor Projection Displays to Elderly Cardiorespiratory Training Through Exergaming," Physiological Computing Systems, Cham, 2019, pp. 58–77, doi: 10.1007/978-3-030-27950-9_4.

Journal Publications

Gonçalves, A., Borrego, A., Latorre, J., Llorens, R., and Bermúdez i Badia, S., "Evaluation of a Low-Cost Virtual Reality Surround-Screen Projection System," Transactions on Visualization and Computer Graphics [In Press 2021], doi: 10.1109/TVCG.2021.3091485.

Gonçalves, A., Muñoz, J., Gouveia, E., Cameirão, M. and Bermúdez i Badia, S., "The Benefits of Custom Exergames for Fitness, Balance and Health-Related Quality of Life: A Randomized Controlled Trial with Community-Dwelling Older Adults," Games for Health Journal [In Press 2021].

Gonçalves, A., Montoya, M., Llorens, R., and Bermúdez i Badia, S., "A Virtual Reality Bus Ride as an Ecologically Valid Assessment of Balance: A Feasibility Study," Virtual Reality (2021), doi: 10.1007/s10055-021-00521-6.

Gonçalves, A., Muñoz, J., Gouveia, E., Cameirão, M. and Bermúdez i Badia, S., "Effects of Prolonged Multidimensional Fitness Training with Exergames on the Physical Exertion Levels of Older Adults," The Visual Computer (2019), doi: 10.1007/s00371-019-01736-0.

Muñoz, J., **Gonçalves, A.**, Gouveia, E., Cameirão, M. and Bermúdez i Badia, S., "Lessons learned from gamifying functional fitness training through human-centered design methods in Portuguese older adults," Games for Health journal (2019), doi: 10.1089/g4h.2018.0028.

Conference Papers

Sousa, H., Gouveia, E., Cameirão, M., **Gonçalves, A.**, Muñoz, J., Paulino, T., Simão, H., Nunes, R., Bernardino, A., and Bermúdez i Badia, S., "Custom-made exergames for older people: New inputs for multidimensional physical." In 2019 5th Experiment International Conference (exp. at'19), pp. 249-250. IEEE, 2019, doi: 10.1109/EXPAT.2019.8876569.

Gonçalves, A. and Bermúdez i Badia, S., "KAVE: Building Kinect Based CAVE Automatic Virtual Environments, Methods for Surround-Screen Projection Management, Motion Parallax and Full-Body Interaction Support," Proc. ACM Hum.-Comput. Interact., vol. 2, EICS, Article 10, 2018, doi: 10.1145/3229092.

Muñoz, J., **Gonçalves, A.**, Cameirão, M., Bermúdez i Badia, S. and Gouveia, E., "Measured and Perceived Physical Responses in Multidimensional Fitness Training through Exergames in Older Adults," Proceedings of the 10th International Conference on Virtual Worlds and Games for Serious Applications (VS-Games), Sept. 2018, doi: 10.1109/VS-Games.2018.8493433, pp. 1-4.

Gonçalves, A., Muñoz, J., Gouveia, E., Cameirão, M. and Bermúdez i Badia, S., "Portuguese Tradition Inspired Exergames for Older People," Proceedings of the 5th International Congress on Sports Sciences Research and Technology Support (icSports 2017), 2017.

Muñoz, J., **Gonçalves, A.**, Vieira, T., Cró, D., Chisik, Y. and Bermúdez i Badia, S., "Space Connection-A Multiplayer Collaborative Biofeedback Game to Promote Empathy in Teenagers: A Feasibility Study," Proceedings of the 3rd International Conference on Physiological Computing Systems, Jul. 27-28, 2016, pp. 88-97.

Gonçalves, A. and Cameirão, M., "Evaluating Body Tracking Interaction in Floor Projection Displays with an Elderly Population," Proceedings of the 3rd International Conference on Physiological Computing Systems, Jul. 27-28, 2016, pp. 24-32.

Gonçalves, A., Gouveia, E., Cameirão, M. and Bermúdez i Badia, S., "Automating senior fitness testing through gesture detection with depth sensors," Proceedings of the IET International Conference on Technologies for Active and Assisted Living (TechAAL), Nov. 2015, doi: 10.1049/ic.2015.0132.

Gonçalves, A., Abrantes, J., Saponaro, G., Jamone, L. and Bernardino, A., "Learning intermediate object affordances: Towards the development of a tool concept," Proceedings of the 4th International Conference on Development and Learning and on Epigenetic Robotics, Genoa, Oct 13-16, 2014, doi: 10.1109/DEVLRN.2014.6983027, pp. 482-488.

Gonçalves, A., Saponaro, G., Jamone, L. and Bernardino, A, "Learning Visual Affordances of Objects and Tools through Autonomous Robot Exploration," Proceedings of the IEEE International Conference on Autonomous Robot Systems and Competitions (ICARSC), Espinho, May 14-15, 2014, doi: 10.1109/ICARSC.2014.6849774, pp. 128-133.

Gonçalves, A., Jamone, L., Saponaro, G. and Bernardino, A, "Autonomous Learning of Tool Affordances," 19th Portuguese Conference on Pattern Recognition (RecPad), Lisbon, Nov. 1, 2013.

Computer Skills

Programming: Extensive experience with C#, C++, C and Matlab. Experience with Python, OpenGL, OpenCV and Arduino.

Applications: Extensive experience with game development in Unity 3D; statistical analysis with SPSS; systems' modeling, simulation and analysis with Matlab; computer-aided design and modeling in Solidworks; and Microsoft Office solutions.

Platforms: Windows, Linux.

Professional Service

Organizing Committee, Student Volunteers Chair
SutainIT, 2017

Languages

Portuguese: Native Language

English: Proficient (C2) Listening, Reading, Speaking and Writing

References

[1] World Health Organization, *World Report on Ageing and Health*. World Health Organization, 2015.

[2] Instituto Nacional de Estatística, "Distribution of resident population (%) by Age group," *Statistics Portugal*, Jun. 14, 2019. https://www.ine.pt/xportal/xmain?xpid=INE&xpgid=ine_indicadores&indOcorrCod=0001276&xlang=en&contexto=bd&selTab=tab2 (accessed Feb. 10, 2020).

[3] European Commission: DG Economic and Financial Affairs, "The 2012 Ageing Report: Economic and budgetary projections for the 27 EU Member States (2010-2060)," European Commission, 2012. [Online]. Available: https://ec.europa.eu/economy_finance/publications/european_economy/2012/2012-ageing-report_en.htm

[4] World Health Organization, *Global Recommendations on Physical Activity for Health*. Geneva: World Health Organization, 2010. Accessed: Feb. 10, 2020. [Online]. Available: https://www.ncbi.nlm.nih.gov/books/NBK305057/

[5] D. P. of E. and I. D. R. Detels, R. Detels, M. Gulliford, Q. A. Karim, and C. C. Tan, *Oxford Textbook of Global Public Health*. Oxford University Press, 2015.

[6] J. A. Wullems, S. M. P. Verschueren, H. Degens, C. I. Morse, and G. L. Onambélé, "A review of the assessment and prevalence of sedentarism in older adults, its physiology/health impact and non-exercise mobility counter-measures," *Biogerontology*, vol. 17, no. 3, pp. 547–565, Jun. 2016, doi: 10.1007/s10522-016-9640-1.

[7] American College of Sports Medicine and B. A. Bushman, *ACSM's Complete Guide to Fitness & Health*. Human Kinetics, 2017.

[8] W. J. Chodzko-Zajko *et al.*, "Exercise and physical activity for older adults," *Medicine & science in sports & exercise*, vol. 41, no. 7, pp. 1510–1530, 2009.

[9] L. H. Larsen, L. Schou, H. H. Lund, and H. Langberg, "The Physical Effect of Exergames in Healthy Elderly—A Systematic Review," *Games for Health Journal*, vol. 2, no. 4, pp. 205–212, Jun. 2013, doi: 10.1089/g4h.2013.0036.

[10] S. Lee, N. D. Myers, T. Park, C. R. Hill, and D. L. Feltz, "An Exploratory Study on the Köhler Effect and Flow in Long-term Exergaming," *Simulation & Gaming*, p. 1046878118776043, 2018.

[11] P. Oesch *et al.*, "Exergames versus self-regulated exercises with instruction leaflets to improve adherence during geriatric rehabilitation: a randomized controlled trial," *BMC geriatrics*, vol. 17, no. 1, p. 77, 2017.

[12] A. DeSmet *et al.*, "A meta-analysis of serious digital games for healthy lifestyle promotion," *Preventive Medicine*, vol. 69, pp. 95–107, Dec. 2014, doi: 10.1016/j.ypmed.2014.08.026.

[13] P. Maillot, A. Perrot, and A. Hartley, "Effects of interactive physical-activity video-game training on physical and cognitive function in older adults," *Psychology and Aging*, vol. 27, no. 3, pp. 589–600, 2012, doi: 10.1037/a0026268.

[14] T. P. Pham and Y.-L. Theng, "Game Controllers for Older Adults: Experimental Study on Gameplay Experiences and Preferences," in *Proceedings of the International Conference on the Foundations of Digital Games*, New York, NY, USA, 2012, pp. 284–285. doi: 10.1145/2282338.2282401.

[15] M. Csikszentmihalyi, *Flow: The Psychology of Optimal Experience*. New York, NY, USA: Harper & Row, 1990.

[16] J. Sinclair, P. Hingston, and M. Masek, "Exergame Development Using the Dual Flow Model," in *Proceedings of the Sixth Australasian Conference on Interactive Entertainment*, New York, NY, USA, 2009, p. 11:1-11:7. doi: 10.1145/1746050.1746061.

[17] F. Laamarti, M. Eid, and A. E. Saddik, "An Overview of Serious Games," *Int. J. Comput. Games Technol.*, vol. 2014, p. 11:11-11:11, Jan. 2014, doi: 10.1155/2014/358152.

[18] N. Skjæret, A. Nawaz, T. Morat, D. Schoene, J. L. Helbostad, and B. Vereijken, "Exercise and rehabilitation delivered through exergames in older adults: An integrative review of technologies,

safety and efficacy," *International Journal of Medical Informatics*, vol. 85, no. 1, pp. 1–16, Jan. 2016, doi: 10.1016/j.ijmedinf.2015.10.008.

[19] N. Hauk, J. Hüffmeier, and S. Krumm, "Ready to be a Silver Surfer? A Meta-analysis on the Relationship Between Chronological Age and Technology Acceptance," *Computers in Human Behavior*, vol. 84, pp. 304–319, Jul. 2018, doi: 10.1016/j.chb.2018.01.020.

[20] W. C.K. Chiu, A. W. Chan, E. Snape, and T. Redman, "Age Stereotypes and Discriminatory Attitudes towards Older Workers: An East-West Comparison," *Human Relations*, vol. 54, no. 5, pp. 629–661, May 2001, doi: 10.1177/0018726701545004.

[21] B. B. Neves and F. Amaro, "Too Old For Technology? How The Elderly Of Lisbon Use And Perceive ICT," *The Journal of Community Informatics*, vol. 8, no. 1, Art. no. 1, Mar. 2012, doi: 10.15353/joci.v8i1.3061.

[22] R. A. Posthuma and M. A. Campion, "Age Stereotypes in the Workplace: Common Stereotypes, Moderators, and Future Research Directions†," *Journal of Management*, vol. 35, no. 1, pp. 158–188, Feb. 2009, doi: 10.1177/0149206308318617.

[23] J. K. Oostrom, D. van der Linden, M. Ph. Born, and H. T. van der Molen, "New technology in personnel selection: How recruiter characteristics affect the adoption of new selection technology," *Computers in Human Behavior*, vol. 29, no. 6, pp. 2404–2415, Nov. 2013, doi: 10.1016/j.chb.2013.05.025.

[24] K. Chen and A. H. S. Chan, "Gerontechnology acceptance by elderly Hong Kong Chinese: a senior technology acceptance model (STAM)," *Ergonomics*, vol. 57, no. 5, pp. 635–652, May 2014, doi: 10.1080/00140139.2014.895855.

[25] V. Venkatesh and H. Bala, "Technology Acceptance Model 3 and a Research Agenda on Interventions," *Decision Sciences*, vol. 39, no. 2, pp. 273–315, 2008, doi: https://doi.org/10.1111/j.1540-5915.2008.00192.x.

[26] R. R. Hainich and O. Bimber, *Displays: Fundamentals & Applications, Second Edition*. CRC Press, 2016.

[27] "Smart TV 8K QLED Q800T 82'' | Samsung Portugal," *Samsung pt*. https://www.samsung.com/pt/tvs/qled-tv/q800t-82-inch-qled-8k-smart-tv-qe82q800tatxxc/ (accessed Mar. 12, 2021).

[28] "Optoma GT760 DLP Projector." https://www.projectorcentral.com/Optoma-GT760.htm (accessed Mar. 12, 2021).

[29] W. R. Sherman and A. B. Craig, *Understanding Virtual Reality: Interface, Application, and Design*. Morgan Kaufmann, 2018.

[30] K. Ito, M. Tada, H. Ujike, and K. Hyodo, "Effects of Weight and Balance of Head Mounted Display on Physical Load," in *Virtual, Augmented and Mixed Reality. Multimodal Interaction*, Cham, 2019, pp. 450–460. doi: 10.1007/978-3-030-21607-8_35.

[31] D. R. Mestre, "CAVE versus Head-Mounted Displays: Ongoing thoughts," *Electronic Imaging*, vol. 2017, no. 3, pp. 31–35, Jan. 2017, doi: 10.2352/ISSN.2470-1173.2017.3.ERVR-094.

[32] "What you need to know about 3D motion capture," *Engadget*. https://www.engadget.com/2014-07-14-motion-capture-explainer.html (accessed Mar. 13, 2021).

[33] B. Jones *et al.*, "RoomAlive: Magical Experiences Enabled by Scalable, Adaptive Projector-camera Units," in *Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology*, New York, NY, USA, 2014, pp. 637–644. doi: 10.1145/2642918.2647383.

[34] K. Tanaka, J. R. Parker, G. Baradoy, D. Sheehan, J. R. Holash, and L. Katz, "A Comparison of Exergaming Interfaces for Use in Rehabilitation Programs and Research," *Loading...*, vol. 6, no. 9, Feb. 2012, Accessed: Jan. 05, 2015. [Online]. Available: http://journals.sfu.ca/loading/index.php/loading/article/view/107

[35] E. E. Stone and M. Skubic, "Fall Detection in Homes of Older Adults Using the Microsoft Kinect," *IEEE Journal of Biomedical and Health Informatics*, vol. 19, no. 1, pp. 290–301, Jan. 2015, doi: 10.1109/JBHI.2014.2312180.

[36] P. Pala *et al.*, "Analysis of Street-Crossing Behavior: Comparing a CAVE Simulator and a Head-Mounted Display among Younger and Older Adults," *Accident Analysis & Prevention*, vol. 152, p. 106004, Mar. 2021, doi: 10.1016/j.aap.2021.106004.

[37] A. Gonçalves and S. Bermúdez, "KAVE: Building Kinect Based CAVE Automatic Virtual Environments, Methods for Surround-Screen Projection Management, Motion Parallax and Full-Body Interaction Support," *Proc. ACM Hum.-Comput. Interact.*, vol. 2, no. EICS, p. 10:1-10:15, Jun. 2018, doi: 10.1145/3229092.

[38] C. Cruz-Neira, D. J. Sandin, and T. A. DeFanti, "Surround-screen Projection-based Virtual Reality: The Design and Implementation of the CAVE," in *Proceedings of the 20th Annual Conference on Computer Graphics and Interactive Techniques*, New York, NY, USA, 1993, pp. 135–142. doi: 10.1145/166117.166134.

[39] C. Cruz-Neira *et al.*, "Scientists in wonderland: A report on visualization applications in the CAVE virtual reality environment," in *Proceedings of 1993 IEEE Research Properties in Virtual Reality Symposium*, Oct. 1993, pp. 59–66. doi: 10.1109/VRAIS.1993.378262.

[40] N. Ohno and A. Kageyama, "Scientific visualization of geophysical simulation data by the CAVE VR system with volume rendering," *Physics of the Earth and Planetary Interiors*, vol. 163, no. 1–4, pp. 305–311, Aug. 2007, doi: 10.1016/j.pepi.2007.02.013.

[41] P. Schuchardt and D. A. Bowman, "The Benefits of Immersion for Spatial Understanding of Complex Underground Cave Systems," in *Proceedings of the 2007 ACM Symposium on Virtual Reality Software and Technology*, New York, NY, USA, 2007, pp. 121–124. doi: 10.1145/1315184.1315205.

[42] T. W. Kuhlen and B. Hentschel, "Quo Vadis CAVE: Does Immersive Visualization Still Matter?," *IEEE Computer Graphics and Applications*, vol. 34, no. 5, pp. 14–21, Sep. 2014, doi: 10.1109/MCG.2014.97.

[43] I. E. Sutherland, "A Head-mounted Three Dimensional Display," in *Proceedings of the December 9-11, 1968, Fall Joint Computer Conference, Part I*, New York, NY, USA, 1968, pp. 757–764. doi: 10.1145/1476589.1476686.

[44] C. Ware, K. Arthur, and K. S. Booth, "Fish Tank Virtual Reality," in *Proceedings of the INTERACT '93 and CHI '93 Conference on Human Factors in Computing Systems*, New York, NY, USA, 1993, pp. 37–42. doi: 10.1145/169059.169066.

[45] A. Febretti *et al.*, "CAVE2: a hybrid reality environment for immersive simulation and information analysis," in *The Engineering Reality of Virtual Reality*, Burlingame, California, United States, Mar. 2013, vol. 8649. doi: 10.1117/12.2005484.

[46] A. Juarez, W. Schonenberg, and C. Bartneck, "Implementing a low-cost CAVE system using the CryEngine2," *Entertainment Computing*, vol. 1, no. 3–4, pp. 157–164, Dec. 2010, doi: 10.1016/j.entcom.2010.10.001.

[47] W. Stuerzlinger, A. Pavlovych, and D. Nywton, "TIVS: temporary immersive virtual environment at simon fraser university: a non-permanent CAVE," in *2015 IEEE 1st Workshop on Everyday Virtual Reality (WEVR)*, Mar. 2015, pp. 23–28. doi: 10.1109/WEVR.2015.7151691.

[48] F. Poschner, "Fire fighting and related simulations in a CAVE using off-the-shelf hardware and software," in *Proceedings of SIGRAD 2014*, Göteborg, Sweden, Jun. 2014, pp. 33–40. Accessed: Jan. 17, 2018. [Online]. Available: http://www.ep.liu.se/ecp/article.asp?issue=106&article=005

[49] "MiddleVR | MiddleVR: Improve reality!" http://www.middlevr.com/home/ (accessed Jan. 19, 2018).

[50] R. Tredinnick, B. Boettcher, S. Smith, S. Solovy, and K. Ponto, "Uni-CAVE: A Unity3D plugin for non-head mounted VR display systems," in *2017 IEEE Virtual Reality (VR)*, Mar. 2017, pp. 393–394. doi: 10.1109/VR.2017.7892342.

[51] A. D. Wilson and H. Benko, "Combining Multiple Depth Cameras and Projectors for Interactions on, Above and Between Surfaces," in *Proceedings of the 23Nd Annual ACM Symposium on User Interface Software and Technology*, New York, NY, USA, 2010, pp. 273–282. doi: 10.1145/1866029.1866073.

[52] H. Benko, R. Jota, and A. Wilson, "MirageTable: Freehand Interaction on a Projected Augmented Reality Tabletop," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, New York, NY, USA, 2012, pp. 199–208. doi: 10.1145/2207676.2207704.

[53] B. R. Jones, H. Benko, E. Ofek, and A. D. Wilson, "IllumiRoom: Peripheral Projected Illusions for Interactive Experiences," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, New York, NY, USA, 2013, pp. 869–878. doi: 10.1145/2470654.2466112.

[54] D. R. Olsen Jr., "Evaluating User Interface Systems Research," in *Proceedings of the 20th Annual ACM Symposium on User Interface Software and Technology*, New York, NY, USA, 2007, pp. 251–258. doi: 10.1145/1294211.1294256.
[55] K. Otte *et al.*, "Accuracy and Reliability of the Kinect Version 2 for Clinical Measurement of Motor Function," *PLOS ONE*, vol. 11, no. 11, Nov. 2016, doi: 10.1371/journal.pone.0166532.
[56] A. Gonçalves, J. Muñoz, É. Gouveia, M. Cameirão, and S. Bermúdez i Badia, "Portuguese Tradition Inspired Exergames for Older People," presented at the 5th International Congress on Sports Sciences Research and Technology Support (icSports 2017), Funchal, Portugal, 2017. Accessed: Mar. 26, 2018. [Online]. Available: https://www.researchgate.net/profile/Sergi_Bermudez_i_Badia/publication/320852961_Portuguese_Tradition_Inspired_Exergames_for_Older_People_Strategic_Tools_to_Promote_Functional_Fitness/links/59fdff68458515d0706ac05a/Portuguese-Tradition-Inspired-Exergames-for-Older-People-Strategic-Tools-to-Promote-Functional-Fitness.pdf
[57] L. F. Ciria *et al.*, "Head movement measurement: An alternative method for posturography studies," *Gait & Posture*, vol. 52, pp. 100–106, Feb. 2017, doi: 10.1016/j.gaitpost.2016.11.020.
[58] G. C. Burdea and P. Coiffet, *Virtual Reality Technology*, 2nd ed. John Wiley & Sons, 2003.
[59] J. Jerald, *The VR Book: Human-Centered Design for Virtual Reality*. Morgan & Claypool, 2015.
[60] M. Slater, V. Linakis, M. Usoh, R. Kooper, and G. Street, "Immersion, Presence, and Performance in Virtual Environments: An Experiment with Tri-Dimensional Chess," 1996, pp. 163–172. [Online]. Available: http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.34.6594
[61] M. Slater and S. Wilbur, "A Framework for Immersive Virtual Environments (FIVE): Speculations on the Role of Presence in Virtual Environments," *Presence: Teleoperators and Virtual Environments*, vol. 6, no. 6, pp. 603–616, Dec. 1997, doi: 10.1162/pres.1997.6.6.603.
[62] D. A. Bowman and R. P. McMahan, "Virtual Reality: How Much Immersion Is Enough?," *Computer*, vol. 40, no. 7, pp. 36–43, Jul. 2007, doi: 10.1109/MC.2007.257.
[63] A. Borrego, J. Latorre, M. Alcañiz, and R. Llorens, "Comparison of Oculus Rift and HTC Vive: Feasibility for Virtual Reality-Based Exploration, Navigation, Exergaming, and Rehabilitation," *Games for Health Journal*, vol. 7, no. 3, Jun. 2018, doi: 10.1089/g4h.2017.0114.
[64] C. Cruz-Neira, D. J. Sandin, T. A. DeFanti, R. V. Kenyon, and J. C. Hart, "The CAVE: Audio Visual Experience Automatic Virtual Environment," *Commun. ACM*, vol. 35, no. 6, pp. 64–72, Jun. 1992, doi: 10.1145/129888.129892.
[65] Johny Chung Lee, *Head Tracking for Desktop VR Displays using the WiiRemote*, (Dec. 21, 2007). Accessed: Mar. 06, 2020. [Online Video]. Available: https://youtu.be/Jd3-eiid-Uw
[66] A. Borrego, J. Latorre, R. Llorens, M. Alcañiz, and E. Noé, "Feasibility of a walking virtual reality system for rehabilitation: objective and subjective parameters," *Journal of NeuroEngineering and Rehabilitation*, vol. 13, p. 68, Aug. 2016, doi: 10.1186/s12984-016-0174-1.
[67] J. E. Cutting and P. M. Vishton, "Chapter 3 - Perceiving Layout and Knowing Distances: The Integration, Relative Potency, and Contextual Use of Different Information about Depth*," in *Perception of Space and Motion*, W. Epstein and S. Rogers, Eds. San Diego: Academic Press, 1995, pp. 69–117. doi: 10.1016/B978-012240530-3/50005-5.
[68] L. Rebenitsch and C. Owen, "Review on cybersickness in applications and visual displays," *Virtual Reality*, vol. 20, no. 2, pp. 101–125, Jun. 2016, doi: 10.1007/s10055-016-0285-9.
[69] G. A. F. Seber, *Multivariate Observations*. John Wiley & Sons, 2009.
[70] D. Webster and O. Celik, "Experimental evaluation of Microsoft Kinect's accuracy and capture rate for stroke rehabilitation applications," in *2014 IEEE Haptics Symposium (HAPTICS)*, Feb. 2014, pp. 455–460. doi: 10.1109/HAPTICS.2014.6775498.
[71] Oliver Kreylos, "Lighthouse tracking examined," *Doc-Ok.org*, May 25, 2016. http://doc-ok.org/?p=1478 (accessed Aug. 30, 2018).
[72] D. C. Niehorster, L. Li, and M. Lappe, "The Accuracy and Precision of Position and Orientation Tracking in the HTC Vive Virtual Reality System for Scientific Research," *Iperception*, vol. 8, no. 3, May 2017, doi: 10.1177/2041669517708205.

[73] M. Slater, A. Steed, and M. Usoh, “The Virtual Treadmill: A Naturalistic Metaphor for Navigation in Immersive Virtual Environments,” in *Virtual Environments ’95*, Springer, Vienna, 1995, pp. 135–148. doi: 10.1007/978-3-7091-9433-1_12.
[74] B. G. Witmer and M. J. Singer, “Measuring Presence in Virtual Environments: A Presence Questionnaire,” *Presence: Teleoperators and Virtual Environments*, vol. 7, no. 3, pp. 225–240, Jun. 1998, doi: 10.1162/105474698565686.
[75] J. Prothero and D. Parker, “A Unified Approach to Presence and Motion Sickness,” in *Virtual and Adaptive Environments*, 2003, pp. 47–66.
[76] “i3B UPV.” http://www.i3b.upv.es/en/ (accessed Aug. 29, 2018).
[77] J. J. LaViola, E. Kruijff, R. P. McMahan, D. Bowman, and I. P. Poupyrev, *3D User Interfaces: Theory and Practice*. Addison-Wesley Professional, 2017.
[78] A. R. Gonçalves, J. E. Muñoz, É. R. Gouveia, M. da S. Cameirão, and S. Bermúdez i Badia, “Effects of prolonged multidimensional fitness training with exergames on the physical exertion levels of older adults,” *Vis Comput*, Jul. 2019, doi: 10.1007/s00371-019-01736-0.
[79] J. E. Muñoz, A. Gonçalves, É. Rúbio Gouveia, M. S. Cameirão, and S. Bermúdez i Badia, “Lessons Learned from Gamifying Functional Fitness Training Through Human-Centered Design Methods in Older Adults,” *Games for Health Journal*, Aug. 2019, doi: 10.1089/g4h.2018.0028.
[80] A. Gonçalves, É. Gouveia, M. Cameirão, and S. Bermúdez i Badia, “Automating Senior Fitness Testing through Gesture Detection with Depth Sensors,” presented at the IET International Conference on Technologies for Active and Assisted Living (TechAAL), London, United Kingdom, May 2015. doi: 10.1049/ic.2015.0132.
[81] C. M. Bleakley, D. Charles, A. Porter-Armstrong, M. D. J. McNeill, S. M. McDonough, and B. McCormack, “Gaming for Health: A Systematic Review of the Physical and Cognitive Effects of Interactive Computer Games in Older Adults,” *J Appl Gerontol*, vol. 34, no. 3, pp. NP166–NP189, Apr. 2015, doi: 10.1177/0733464812470747.
[82] E. Loos, “Exergaming: Meaningful Play for Older Adults?,” in *International Conference on Human Aspects of IT for the Aged Population*, 2017, pp. 254–265.
[83] K. I. Molina, N. A. Ricci, S. A. de Moraes, and M. R. Perracini, “Virtual reality using games for improving physical functioning in older adults: a systematic review,” *Journal of NeuroEngineering and Rehabilitation*, vol. 11, p. 156, 2014, doi: 10.1186/1743-0003-11-156.
[84] J. Sween, S. F. Wallington, V. Sheppard, T. Taylor, A. A. Llanos, and L. L. Adams-Campbell, “The Role of Exergaming in Improving Physical Activity: A Review,” *J Phys Act Health*, vol. 11, no. 4, pp. 864–870, May 2014, doi: 10.1123/jpah.2011-0425.
[85] American College of Sports Medicine, *ACSM’s Guidelines for Exercise Testing and Prescription*. Wolters Kluwer, 2018.
[86] R. E. Rikli and C. J. Jones, *Senior Fitness Test Manual*. Human Kinetics, 2013.
[87] J. L. Fleg *et al.*, “Accelerated Longitudinal Decline of Aerobic Capacity in Healthy Older Adults,” *Circulation*, vol. 112, no. 5, pp. 674–682, Feb. 2005, doi: 10.1161/CIRCULATIONAHA.105.545459.
[88] D. Morris, T. S. Saponas, A. Guillory, and I. Kelner, “RecoFit: Using a Wearable Sensor to Find, Recognize, and Count Repetitive Exercises,” in *Proceedings of the 32Nd Annual ACM Conference on Human Factors in Computing Systems*, New York, NY, USA, 2014, pp. 3225–3234. doi: 10.1145/2556288.2557116.
[89] S. Uzor and L. Baillie, “Investigating the Long-term Use of Exergames in the Home with Elderly Fallers,” in *Proceedings of the 32Nd Annual ACM Conference on Human Factors in Computing Systems*, New York, NY, USA, 2014, pp. 2813–2822. doi: 10.1145/2556288.2557160.
[90] G. Alankus, A. Lazar, M. May, and C. Kelleher, “Towards Customizable Games for Stroke Rehabilitation,” in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, New York, NY, USA, 2010, pp. 2113–2122. doi: 10.1145/1753326.1753649.
[91] W. Young, S. Ferguson, S. Brault, and C. Craig, “Assessing and training standing balance in older adults: A novel approach using the ‘Nintendo Wii’ Balance Board,” *Gait & Posture*, vol. 33, no. 2, pp. 303–305, Feb. 2011, doi: 10.1016/j.gaitpost.2010.10.089.

[92] R. J. Reed-Jones, S. Dorgo, M. K. Hitchings, and J. O. Bader, “WiiFit™ Plus balance test scores for the assessment of balance and mobility in older adults,” *Gait & Posture*, vol. 36, no. 3, pp. 430–433, Jul. 2012, doi: 10.1016/j.gaitpost.2012.03.027.
[93] B. Haddock *et al.*, “Fitness Assessment Comparison Between the ‘Jackie Chan Action Run’ Videogame, 1-Mile Run/Walk, and the PACER,” *Games for Health Journal*, vol. 1, no. 3, pp. 223–227, Apr. 2012, doi: 10.1089/g4h.2012.0007.
[94] A. Fernández-Baena, A. Susin, and X. Lligadas, “Biomechanical Validation of Upper-Body and Lower-Body Joint Movements of Kinect Motion Capture Data for Rehabilitation Treatments,” in *2012 4th International Conference on Intelligent Networking and Collaborative Systems (INCoS)*, Sep. 2012, pp. 656–661. doi: 10.1109/iNCoS.2012.66.
[95] X. Xu and R. W. McGorry, “The validity of the first and second generation Microsoft Kinect™ for identifying joint center locations during static postures,” *Applied Ergonomics*, vol. 49, pp. 47–54, Jul. 2015, doi: 10.1016/j.apergo.2015.01.005.
[96] K. Gerling, I. Livingston, L. Nacke, and R. Mandryk, “Full-body Motion-based Game Interaction for Older Adults,” in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, New York, NY, USA, 2012, pp. 1873–1882. doi: 10.1145/2207676.2208324.
[97] M. Gabel, R. Gilad-Bachrach, E. Renshaw, and A. Schuster, “Full body gait analysis with Kinect,” in *2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Aug. 2012, pp. 1964–1967. doi: 10.1109/EMBC.2012.6346340.
[98] E. E. Stone and M. Skubic, “Passive in-home measurement of stride-to-stride gait variability comparing vision and Kinect sensing,” in *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBC*, Aug. 2011, pp. 6491–6494. doi: 10.1109/IEMBS.2011.6091602.
[99] A. Chaaraoui, J. Padilla Lopez, and F. Florez Revuelta, “Abnormal gait detection with RGB-D devices using joint motion history features,” presented at the 1st International Workshop on Understanding Human Activities through 3D Sensors (UHA3DS’15), Ljubljana, Slovenia, May 2015. Accessed: Oct. 13, 2015. [Online]. Available: http://eprints.kingston.ac.uk/31370/
[100] A. Da Gama, T. Chaves, L. Figueiredo, and V. Teichrieb, “Guidance and Movement Correction Based on Therapeutics Movements for Motor Rehabilitation Support Systems,” in *2012 14th Symposium on Virtual and Augmented Reality (SVR)*, May 2012, pp. 191–200. doi: 10.1109/SVR.2012.15.
[101] S. Saha, M. Pal, A. Konar, and R. Janarthanan, “Neural Network Based Gesture Recognition for Elderly Health Care Using Kinect Sensor,” in *Proceedings of the 4th International Conference on Swarm, Evolutionary, and Memetic Computing - Volume 8298*, New York, NY, USA, 2013, pp. 376–386. doi: 10.1007/978-3-319-03756-1_34.
[102] S. Spasojević, J. Santos-Victor, T. Ilić, S. Milanović, V. Potkonjak, and A. Rodić, “A Vision-Based System for Movement Analysis in Medical Applications: The Example of Parkinson Disease,” in *Computer Vision Systems*, Springer International Publishing, 2015, pp. 424–434. Accessed: Jul. 09, 2015. [Online]. Available: http://link.springer.com/chapter/10.1007/978-3-319-20904-3_38
[103] C. Chen, K. Liu, R. Jafari, and N. Kehtarnavaz, “Home-based Senior Fitness Test measurement system using collaborative inertial and depth sensors,” in *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Aug. 2014, pp. 4135–4138. doi: 10.1109/EMBC.2014.6944534.
[104] D. J. Rose, *Fallproof!: A Comprehensive Balance and Mobility Training Program*. Human Kinetics, 2010.
[105] Y. Freund and R. E. Schapire, “A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting,” *Journal of Computer and System Sciences*, vol. 55, no. 1, pp. 119–139, Aug. 1997, doi: 10.1006/jcss.1997.1504.
[106] L. Breiman, “Random Forests,” *Machine Learning*, vol. 45, no. 1, pp. 5–32, Oct. 2001, doi: 10.1023/A:1010933404324.
[107] Afonso Gonçalves and Mónica Cameirão, “Evaluating Body Tracking Interaction in Floor Projection Displays with an Elderly Population,” presented at the PhyCS 2016 - 3rd International Conference on Physiological Computing Syst, Lisbon, Jul. 2016.

[108] J. Fozard, “Vision and hearing in aging,” in *Handbook of the Psychology of Aging*, 3rd ed., Academic Press, 1990, pp. 143–156.
[109] World Health Organization, “Global recommendations on physical activity for health.” 2010. Accessed: Jan. 07, 2015. [Online]. Available: http://www.who.int/dietphysicalactivity/publications/9789241599979/en/
[110] Microsoft News Center, “Xbox Execs Talk Momentum and the Future of TV,” *News Center*, 2013. http://news.microsoft.com/2013/02/11/xbox-execs-talk-momentum-and-the-future-of-tv/ (accessed Jan. 19, 2016).
[111] “Microsoft’s Q2: record $24.52 billion revenue and 3.9 million Xbox One sales,” *The Verge*, 2014. http://www.theverge.com/2014/1/23/5338162/microsoft-q2-2014-financial-earnings (accessed Jan. 19, 2016).
[112] D. Wigdor and D. Wixon, *Brave NUI World: Designing Natural User Interfaces for Touch and Gesture*. Elsevier, 2011.
[113] M. Nancel, J. Wagner, E. Pietriga, O. Chapuis, and W. Mackay, “Mid-air Pan-and-zoom on Wall-sized Displays,” in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, New York, NY, USA, 2011, pp. 177–186. doi: 10.1145/1978942.1978969.
[114] M. N. K. Boulos, B. J. Blanchard, C. Walker, J. Montero, A. Tripathy, and R. Gutierrez-Osuna, “Web GIS in practice X: a Microsoft Kinect natural user interface for Google Earth navigation,” *International Journal of Health Geographics*, vol. 10, no. 1, p. 45, Jul. 2011, doi: 10.1186/1476-072X-10-45.
[115] R. Francese, I. Passero, and G. Tortora, “Wiimote and Kinect: Gestural User Interfaces Add a Natural Third Dimension to HCI,” in *Proceedings of the International Working Conference on Advanced Visual Interfaces*, New York, NY, USA, 2012, pp. 116–123. doi: 10.1145/2254556.2254580.
[116] M. Roupé, P. Bosch-Sijtsema, and M. Johansson, “Interactive navigation interface for Virtual Reality using the human body,” *Computers, Environment and Urban Systems*, vol. 43, pp. 42–50, Jan. 2014, doi: 10.1016/j.compenvurbsys.2013.10.003.
[117] J. H. Tan, C. Chao, M. Zawaideh, A. C. Roberts, and T. B. Kinney, “Informatics in Radiology: Developing a Touchless User Interface for Intraoperative Image Control during Interventional Radiology Procedures,” *RadioGraphics*, vol. 33, no. 2, pp. E61–E70, Mar. 2013, doi: 10.1148/rg.332125101.
[118] A. Bigdelou, L. Schwarz, and N. Navab, “An Adaptive Solution for Intra-operative Gesture-based Human-machine Interaction,” in *Proceedings of the 2012 ACM International Conference on Intelligent User Interfaces*, New York, NY, USA, 2012, pp. 75–84. doi: 10.1145/2166966.2166981.
[119] W. Fikkert, P. van der Vet, G. van der Veer, and A. Nijholt, “Gestures for Large Display Control,” in *Gesture in Embodied Communication and Human-Computer Interaction*, S. Kopp and I. Wachsmuth, Eds. Springer Berlin Heidelberg, 2009, pp. 245–256. Accessed: Sep. 30, 2015. [Online]. Available: http://link.springer.com/chapter/10.1007/978-3-642-12553-9_22
[120] R.-D. Vatavu, “User-defined Gestures for Free-hand TV Control,” in *Proceedings of the 10th European Conference on Interactive Tv and Video*, New York, NY, USA, 2012, pp. 45–48. doi: 10.1145/2325616.2325626.
[121] I. Golod, F. Heidrich, C. Möllering, and M. Ziefle, “Design Principles of Hand Gesture Interfaces for Microinteractions,” in *Proceedings of the 6th International Conference on Designing Pleasurable Products and Interfaces*, New York, NY, USA, 2013, pp. 11–20. doi: 10.1145/2513506.2513508.
[122] S.-S. Lee, J. Chae, H. Kim, Y. Lim, and K. Lee, “Towards More Natural Digital Content Manipulation via User Freehand Gestural Interaction in a Living Room,” in *Proceedings of the 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, New York, NY, USA, 2013, pp. 617–626. doi: 10.1145/2493432.2493480.
[123] A. Bragdon, R. DeLine, K. Hinckley, and M. R. Morris, “Code Space: Touch + Air Gesture Hybrid Interactions for Supporting Developer Meetings,” in *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces*, New York, NY, USA, 2011, pp. 212–221. doi: 10.1145/2076354.2076393.

[124] M. Hopmann, P. Salamin, N. Chauvin, F. Vexo, and D. Thalmann, “Natural Activation for Gesture Recognition Systems,” in *CHI ’11 Extended Abstracts on Human Factors in Computing Systems*, New York, NY, USA, 2011, pp. 173–183. doi: 10.1145/1979742.1979642.
[125] P. Krogh, M. Ludvigsen, and A. Lykke-Olesen, “‘Help Me Pull That Cursor’ A Collaborative Interactive Floor Enhancing Community Interaction,” *Australasian Journal of Information Systems*, vol. 11, no. 2, May 2004, doi: 10.3127/ajis.v11i2.126.
[126] T. Augsten *et al.*, “Multitoe: High-precision Interaction with Back-projected Floors Based on High-resolution Multi-touch Input,” in *Proceedings of the 23Nd Annual ACM Symposium on User Interface Software and Technology*, New York, NY, USA, 2010, pp. 209–218. doi: 10.1145/1866029.1866064.
[127] M. F. Folstein, S. E. Folstein, and P. R. McHugh, “‘Mini-mental state’: A practical method for grading the cognitive state of patients for the clinician,” *Journal of Psychiatric Research*, vol. 12, no. 3, pp. 189–198, Nov. 1975, doi: 10.1016/0022-3956(75)90026-6.
[128] C. J. J. Roberta E. Rikli, “The Reliability and Validity of a 6-Minute Walk Test as a Measure of Physical Endurance in Older Adults,” *Journal of aging and physical activity*, 1998.
[129] J. Brooke, “SUS-A quick and dirty usability scale,” *Usability evaluation in industry*, vol. 189, p. 194, 1996.
[130] S. G. Hart and L. E. Staveland, “Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research,” in *Advances in Psychology*, vol. 52, P. A. H. and N. Meshkati, Ed. North-Holland, 1988, pp. 139–183. Accessed: Oct. 21, 2015. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0166411508623869
[131] H. Mousavi Hondori, M. Khademi, L. Dodakian, A. McKenzie, C. V. Lopes, and S. C. Cramer, “Choice of Human-Computer Interaction Mode in Stroke Rehabilitation,” *Neurorehabil Neural Repair*, Jul. 2015, doi: 10.1177/1545968315593805.
[132] V. H. Heyward and A. Gibson, *Advanced Fitness Assessment and Exercise Prescription 7th Edition*. Human Kinetics, 2014.
[133] S. Deterding, D. Dixon, R. Khaled, and L. Nacke, “From game design elements to gamefulness: defining gamification,” in *Proceedings of the 15th international academic MindTrek conference: Envisioning future media environments*, 2011, pp. 9–15.
[134] W. Zhu and N. Owen, *Sedentary Behavior and Health: Concepts, Assessments, and Interventions*. Human Kinetics, 2017.
[135] Y. Oh and S. Yang, “Defining exergames & exergaming,” *Proceedings of Meaningful Play*, pp. 1–17, 2010.
[136] F. Zhang and D. Kaufman, “Physical and cognitive impacts of digital games on older adults: A meta-analytic review,” *Journal of Applied Gerontology*, vol. 35, no. 11, pp. 1189–1210, 2016.
[137] C. Anderson-Hanley *et al.*, “Exergaming and Older Adult Cognition: A Cluster Randomized Clinical Trial,” *American Journal of Preventive Medicine*, vol. 42, no. 2, pp. 109–119, Feb. 2012, doi: 10.1016/j.amepre.2011.10.016.
[138] L. M. Taylor, R. Maddison, L. A. Pfaeffli, J. C. Rawstorn, N. Gant, and N. M. Kerse, “Activity and energy expenditure in older people playing active video games,” *Archives of physical medicine and rehabilitation*, vol. 93, no. 12, pp. 2281–2286, 2012.
[139] A. K. Hall, E. Chavarria, V. Maneeratana, B. H. Chaney, and J. M. Bernhardt, “Health Benefits of Digital Videogames for Older Adults: A Systematic Review of the Literature,” *Games Health J*, vol. 1, no. 6, pp. 402–410, Dec. 2012, doi: 10.1089/g4h.2012.0046.
[140] A. Nawaz, N. Skjæret, J. L. Helbostad, B. Vereijken, E. Boulton, and D. Svanaes, “Usability and acceptability of balance exergames in older adults: A scoping review,” *Health Informatics J*, vol. 22, no. 4, pp. 911–931, Dec. 2016, doi: 10.1177/1460458215598638.
[141] M. Van Diest, C. J. Lamoth, J. Stegenga, G. J. Verkerke, and K. Postema, “Exergaming for balance training of elderly: state of the art and future developments,” *Journal of neuroengineering and rehabilitation*, vol. 10, no. 1, p. 101, 2013.
[142] K. A. Strand, S. L. Francis, J. A. Margrett, W. D. Franke, and M. J. Peterson, “Community-based exergaming program increases physical activity and perceived wellness in older adults,” *Journal of aging and physical activity*, vol. 22, no. 3, pp. 364–371, 2014.

[143] E. I. Konstantinidis, A. S. Billis, C. A. Mouzakidis, V. I. Zilidou, P. E. Antoniou, and P. D. Bamidis, “Design, Implementation, and Wide Pilot Deployment of FitForAll: An Easy to use Exergaming Platform Improving Physical Fitness and Life Quality of Senior Citizens,” *IEEE Journal of Biomedical and Health Informatics*, vol. 20, no. 1, pp. 189–200, Jan. 2016, doi: 10.1109/JBHI.2014.2378814.

[144] D. Schoene, T. Valenzuela, S. R. Lord, and E. D. de Bruin, “The effect of interactive cognitive-motor training in reducing fall risk in older people: a systematic review,” *BMC geriatrics*, vol. 14, no. 1, p. 107, 2014.

[145] E. Stanmore, B. Stubbs, D. Vancampfort, E. D. de Bruin, and J. Firth, “The effect of active video games on cognitive functioning in clinical and non-clinical populations: a meta-analysis of randomized controlled trials,” *Neuroscience & Biobehavioral Reviews*, 2017, [Online]. Available: http://www.sciencedirect.com/science/article/pii/S014976341730129X

[146] P. Eggenberger, V. Schumacher, M. Angst, N. Theill, and E. D. de Bruin, “Does multicomponent physical exercise with simultaneous cognitive training boost cognitive performance in older adults? A 6-month randomized controlled trial with a 1-year follow-up,” *Clinical interventions in aging*, vol. 10, p. 1335, 2015.

[147] S. Lee, W. Kim, T. Park, and W. Peng, “The Psychological Effects of Playing Exergames: A Systematic Review,” *Cyberpsychology, Behavior, and Social Networking*, vol. 20, no. 9, pp. 513–532, 2017.

[148] V. Baert, E. Gorus, T. Mets, C. Geerts, and I. Bautmans, “Motivators and barriers for physical activity in the oldest old: a systematic review,” *Ageing research reviews*, vol. 10, no. 4, pp. 464–474, 2011.

[149] K. A. Schutzer and B. S. Graves, “Barriers and motivations to exercise in older adults,” *Preventive medicine*, vol. 39, no. 5, pp. 1056–1061, 2004.

[150] C. Anderson-Hanley, A. L. Snyder, J. P. Nimon, and P. J. Arciero, “Social facilitation in virtual reality-enhanced exercise: competitiveness moderates exercise effort of older adults,” *Clinical interventions in aging*, vol. 6, p. 275, 2011.

[151] S. C. Barathi *et al.*, “Interactive Feedforward for Improving Performance and Maintaining Intrinsic Motivation in VR Exergaming,” in *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, 2018, p. 408.

[152] R. Mellecker, E. J. Lyons, and T. Baranowski, “Disentangling fun and enjoyment in exergames using an expanded design, play, experience framework: A narrative review,” *GAMES FOR HEALTH: Research, Development, and Clinical Applications*, vol. 2, no. 3, pp. 142–149, 2013.

[153] G. L. Dimaguila, K. Gray, and M. Merolli, “Person-Generated Health Data in Simulated Rehabilitation Using Kinect for Stroke: Literature Review,” *JMIR rehabilitation and assistive technologies*, vol. 5, no. 1, 2018.

[154] G. Bamparopoulos, E. Konstantinidis, C. Bratsas, and P. D. Bamidis, “Towards exergaming commons: composing the exergame ontology for publishing open game data,” *Journal of biomedical semantics*, vol. 7, no. 1, p. 4, 2016.

[155] M. Rasinaho, M. Hirvensalo, R. Leinonen, T. Lintunen, and T. Rantanen, “Motives for and barriers to physical activity among older adults with mobility limitations,” *Journal of aging and physical activity*, vol. 15, no. 1, pp. 90–102, 2007.

[156] Y.-Y. Chao, Y. K. Scherer, and C. A. Montgomery, “Effects of using Nintendo Wii™ exergames in older adults: a review of the literature,” *Journal of aging and health*, vol. 27, no. 3, pp. 379–402, 2015.

[157] J. Wiemeyer *et al.*, “Recommendations for the Optimal Design of Exergame Interventions for Persons with Disabilities: Challenges, Best Practices, and Future Research,” *Games for Health Journal*, vol. 4, no. 1, pp. 58–62, Dec. 2014, doi: 10.1089/g4h.2014.0078.

[158] L. Taylor, N. Kerse, J. Klenk, R. Borotkanics, and R. Maddison, “Exergames to Improve the Mobility of Long-Term Care Residents: A Cluster Randomized Controlled Trial,” *Games for health journal*, vol. 7, no. 1, pp. 37–42, 2018.

[159] D. Webster and O. Celik, “Systematic review of Kinect applications in elderly care and stroke rehabilitation,” *J. Neuroeng. Rehabil*, vol. 11, no. 1, p. 108, 2014.

[160] E. Loos and D. Kaufman, “Positive Impact of Exergaming on Older Adults’ Mental and Social Well-being: In Search of Evidence,” in *International Conference on Human Aspects of IT for the Aged Population*, 2018, pp. 101–112.
[161] W. Peng, J.-H. Lin, and J. Crouse, “Is Playing Exergames Really Exercising? A Meta-Analysis of Energy Expenditure in Active Video Games,” *Cyberpsychology, Behavior, and Social Networking*, vol. 14, no. 11, pp. 681–688, Jun. 2011, doi: 10.1089/cyber.2010.0578.
[162] D. Alan, F. Janet, A. Gregory, and B. Russell, “Human-computer interaction,” *England: Pearson Education Limited*, vol. 5, 2004.
[163] H. Beyer and K. Holtzblatt, *Contextual Design: Defining Customer-centered Systems*. Morgan Kaufmann, 1998.
[164] K. Caro, M. Tentori, A. I. Martinez-Garcia, and I. Zavala-Ibarra, “FroggyBobby: An exergame to support children with motor problems practicing motor coordination exercises during therapeutic interventions,” *Computers in Human Behavior*, 2015, Accessed: Apr. 11, 2017. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0747563215004380
[165] K. Holtzblatt, J. B. Wendell, and S. Wood, *Rapid contextual design: a how-to guide to key techniques for user-centered design*. Elsevier, 2004. Accessed: Oct. 19, 2016. [Online]. Available: https://www.google.com/books?hl=en&lr=&id=yzOrcSl6zvMC&oi=fnd&pg=PP1&dq=rapid+contextual+design&ots=aCNeUxHhDR&sig=BJzKYrC9XZonR1uH4dzmk9YaNmk
[166] M. E. Nelson *et al.*, “Physical activity and public health in older adults: recommendation from the American College of Sports Medicine and the American Heart Association,” *Circulation*, vol. 116, no. 9, p. 1094, 2007.
[167] É. R. Gouveia, J. A. Maia, G. P. Beunen, C. J. Blimkie, E. M. Fena, and D. L. Freitas, “Functional Fitness and Physical Activity of Portuguese Community-Residing Older Adults,” *Journal of Aging and Physical Activity*, vol. 21, no. 1, pp. 1–19, Jan. 2013, doi: 10.1123/japa.21.1.1.
[168] S. Czaja, S. Beach, N. Charness, and R. Schulz, “Older adults and the adoption of healthcare technology: Opportunities and challenges,” in *Technologies for active aging*, Springer, 2013, pp. 27–46.
[169] S. Hardy, T. Dutz, J. Wiemeyer, S. Göbel, and R. Steinmetz, “Framework for personalized and adaptive game-based training programs in health sport,” *Multimedia Tools and Applications*, pp. 1–23, 2014.
[170] D. Kaufman, L. Sauvé, L. Renaud, and E. Duplàa, “Benefits and Barriers of Older Adults’ Digital Gameplay.,” in *CSEDU (1)*, 2014, pp. 213–219. Accessed: Apr. 04, 2017. [Online]. Available: https://files.ifi.uzh.ch/stiller/CLOSER%202014/CSEDU/CSEDU/Information%20Technologies%20Supporting%20Learning/Short%20Papers/CSEDU_2014_46_CR.pdf
[171] K. A. Blocker, T. J. Wright, and W. R. Boot, “Gaming preferences of aging generations,” *Gerontechnology*, vol. 12, no. 3, pp. 174–184, 2014.
[172] A. F. A. de Vette, M. Tabak, H. J. Hermens, and M. M. R. Vollenbroek-Hutten, “Game preferences and personality of older adult users,” 2015. Accessed: Apr. 04, 2017. [Online]. Available: http://eprints.eemcs.utwente.nl/26351/
[173] A. Vaatanen and J. Leikas, “Human-centred design and exercise games: users’ experiences of a fitness adventure prototype,” *Design and use of serious games. Springer, Dordrecht*, 2009.
[174] C. Richards, C. W. Thompson, and N. Graham, “Beyond designing for motivation: the importance of context in gamification,” in *Proceedings of the first ACM SIGCHI annual symposium on Computer-human interaction in play*, 2014, pp. 217–226. [Online]. Available: http://dl.acm.org/citation.cfm?id=2658683
[175] V. A. Vanden Abeele and V. Van Rompaey, “Introducing human-centered research to game design: designing game concepts for and with senior citizens,” in *CHI’06 extended abstracts on Human factors in computing systems*, 2006, pp. 1469–1474. Accessed: Mar. 27, 2017. [Online]. Available: http://dl.acm.org/citation.cfm?id=1125721
[176] J. P. Salmon, S. M. Dolan, R. S. Drake, G. C. Wilson, R. M. Klein, and G. A. Eskes, “A survey of video game preferences in adults: Building better games for older adults,” *Entertainment Computing*, 2017, Accessed: May 03, 2017. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1875952117300381

[177] S. K. Jali and S. Arnab, “The Perspectives of Older People on Digital Gaming: Interactions with Console and Tablet-Based Games,” in *International Conference on Serious Games, Interaction and Simulation*, 2016, pp. 82–90. Accessed: May 17, 2017. [Online]. Available: http://link.springer.com/chapter/10.1007/978-3-319-51055-2_11
[178] R. Malaka, “How Computer Games Can Improve Your Health and Fitness,” in *Games for Training, Education, Health and Sports*, S. Göbel and J. Wiemeyer, Eds. Springer International Publishing, 2014, pp. 1–7. Accessed: Apr. 29, 2015. [Online]. Available: http://link.springer.com/chapter/10.1007/978-3-319-05972-3_1
[179] S. Hardy, A. El Saddik, S. Göbel, and R. Steinmetz, “Context aware serious games framework for sport and health,” in *Medical Measurements and Applications Proceedings (MeMeA), 2011 IEEE International Workshop on*, 2011, pp. 248–252. [Online]. Available: http://ieeexplore.ieee.org/abstract/document/5966775/
[180] S. Uzor, L. Baillie, and D. Skelton, “Senior designers: empowering seniors to design enjoyable falls rehabilitation tools,” in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2012, pp. 1179–1188. Accessed: Apr. 11, 2017. [Online]. Available: http://dl.acm.org/citation.cfm?id=2208568
[181] T. Miaskiewicz and K. A. Kozar, “Personas and user-centered design: How can personas benefit product design processes?,” *Design Studies*, vol. 32, no. 5, pp. 417–430, 2011.
[182] Muñoz J.E., Bermudez S., Rubio E., and Cameirao M., “Modulation of Physiological Responses and Activity Levels During Exergame Experiences,” in *2016 18th International Conference on Virtual Worlds and Games for Serious Applications*, 2016, p. In press.
[183] S. B. i B. Ana Lúcia Faria, “Benefits of virtual reality based cognitive rehabilitation through simulated activities of daily living: a randomized controlled trial with stroke patients,” *Journal of NeuroEngineering and Rehabilitation*, 2016, Accessed: Nov. 08, 2016. [Online]. Available: https://jneuroengrehab.biomedcentral.com/articles/10.1186/s12984-016-0204-z
[184] A. Sixsmith, “Technology and the Challenge of Aging,” in *Technologies for Active Aging*, Springer, 2013, pp. 7–25. Accessed: Oct. 21, 2016. [Online]. Available: http://sci-hub.cc/http://sci-hub.cc/http://link.springer.com/chapter/10.1007/978-1-4419-8348-0_2
[185] J. Torbeyns, E. Lehtinen, and J. Elen, *Describing and Studying Domain-Specific Serious Games*. Springer, 2015.
[186] O. Bimber and R. Raskar, *Spatial augmented reality: merging real and virtual worlds*. CRC press, 2005. Accessed: Nov. 03, 2016. [Online]. Available: https://www.google.com/books?hl=en&lr=&id=JEu3BgAAQBAJ&oi=fnd&pg=PP1&dq=spatial+augmented+reality&ots=Neeu8iJ6U-&sig=S8fU9d1T-LDGcFCeqEIFoSKHBg4
[187] K. Becker and J. Parker, “An overview of game design techniques,” in *Learning, Education and Games*, 2014, pp. 179–198. Accessed: Oct. 24, 2016. [Online]. Available: http://dl.acm.org/citation.cfm?id=2811158
[188] T. Fullerton, *Game design workshop: a playcentric approach to creating innovative games*. CRC press, 2008.
[189] J. Schell, *The Art of Game Design: A book of lenses*. AK Peters/CRC Press, 2014.
[190] A. Kultima, “Defining Game Jam,” in *Proc. of Foundations of Digital Games Conference*, 2015, vol. 15. Accessed: Oct. 21, 2016. [Online]. Available: http://sci-hub.cc/http://sci-hub.cc/https://www.researchgate.net/profile/Kultima_Annakaisa/publication/281748266_Defining_Game_Jam/links/55f729d908ae07629dc114bd.pdf
[191] S. Sayago, A. Rosales, V. Righi, S. M. Ferreira, G. W. Coleman, and J. Blat, “On the Conceptualization, Design, and Evaluation of Appealing, Meaningful, and Playable Digital Games for Older People,” *Games and Culture*, vol. 11, no. 1–2, pp. 53–80, 2016.
[192] M. A. Pols *et al.*, “Validity and repeatability of a modified Baecke questionnaire on physical activity,” *International journal of epidemiology*, vol. 24, no. 2, pp. 381–388, 1995.
[193] M. G. Helander, T. K. Landauer, and P. V. Prabhu, *Handbook of Human-Computer Interaction*. Elsevier, 1997.

[194] E. Brox, L. F. Luque, G. J. Evertsen, and J. E. G. Hernández, “Exergames for elderly: Social exergames to persuade seniors to increase physical activity,” in *Pervasive Computing Technologies for Healthcare (PervasiveHealth), 2011 5th International Conference on*, 2011, pp. 546–549.

[195] R. Malaka, M. Herrlich, and J. Smeddinck, “Anticipation in Motion-Based Games for Health,” in *Anticipation and Medicine*, Springer, 2017, pp. 351–363. Accessed: Nov. 10, 2016. [Online]. Available: http://link.springer.com/chapter/10.1007/978-3-319-45142-8_22

[196] A. S. Lu, “Narrative in exergames: Thoughts on procedure, mechanism, and others,” *Games for health journal*, vol. 4, no. 1, pp. 19–24, 2015.

[197] E. Brox, S. T. Konstantinidis, and G. Evertsen, “User-Centered Design of Serious Games for Older Adults Following 3 Years of Experience With Exergames for Seniors: A Study Design,” *JMIR Serious Games*, vol. 5, no. 1, p. e2, 2017.

[198] K. M. Gerling, C. Linehan, B. Kirman, M. R. Kalyn, A. B. Evans, and K. C. Hicks, “Creating wheelchair-controlled video games: Challenges and opportunities when involving young people with mobility impairments and game design experts,” *International Journal of Human-Computer Studies*, vol. 94, pp. 64–73, 2016.

[199] P. A. Silva, F. Nunes, A. Vasconcelos, M. Kerwin, R. Moutinho, and P. Teixeira, “Using the smartphone accelerometer to monitor fall risk while playing a game: the design and usability evaluation of dance! Don’t fall,” in *International Conference on Augmented Cognition*, 2013, pp. 754–763.

[200] J. Sinclair, P. Hingston, M. Masek, and K. Nosaka, “Testing an exergame for effectiveness and attractiveness,” in *2010 2nd International IEEE Consumer Electronics Society’s Games Innovations Conference*, 2010, pp. 1–8. Accessed: Nov. 10, 2016. [Online]. Available: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5716909

[201] F. Mueller and K. Isbister, “Movement-based game guidelines,” in *Proceedings of the 32nd annual ACM conference on Human factors in computing systems*, 2014, pp. 2191–2200. Accessed: Nov. 11, 2016. [Online]. Available: http://dl.acm.org/citation.cfm?id=2557163

[202] K. Isbister and F. “Floyd” Mueller, “Guidelines for the design of movement-based games and their relevance to HCI,” *Human–Computer Interaction*, vol. 30, no. 3–4, pp. 366–399, 2015.

[203] H. M. Hondori, M. Khademi, L. Dodakian, S. C. Cramer, and C. V. Lopes, “A spatial augmented reality rehab system for post-stroke hand rehabilitation.,” in *MMVR*, 2013, pp. 279–285. Accessed: Nov. 11, 2016. [Online]. Available: https://pdfs.semanticscholar.org/e183/7e399d3b8ecb2d8e7d13078c4a8c2c51743b.pdf

[204] T. M. Fleming *et al.*, “Maximizing the impact of e-therapy and serious Gaming: time for a Paradigm shift,” *Frontiers in psychiatry*, vol. 7, 2016, Accessed: Nov. 11, 2016. [Online]. Available: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4834305/

[205] K. Seaborn and D. I. Fels, “Gamification in theory and action: A survey,” *International Journal of Human-Computer Studies*, vol. 74, pp. 14–31, 2015.

[206] M. Papastergiou, “Exploring the potential of computer and video games for health and physical education: A literature review,” *Computers & Education*, vol. 53, no. 3, pp. 603–622, 2009.

[207] P. M. Kato, “The Role of the Researcher in Making Serious Games for Health,” *Serious Games for Healthcare: Applications and Implications*, pp. 213–231, 2012.

[208] S. Koenig, A. Ardanza, C. Cortes, A. De Mauro, and B. Lange, “Introduction to low-cost motion-tracking for virtual rehabilitation,” in *Emerging Therapies in Neurorehabilitation*, Springer, 2014, pp. 287–303. Accessed: Nov. 14, 2016. [Online]. Available: http://link.springer.com/10.1007/978-3-642-38556-8_15

[209] A. Vourvopoulos, A. L. Faria, M. S. Cameirao, and S. Bermudez i Badia, “RehabNet: A distributed architecture for motor and cognitive neuro-rehabilitation,” in *2013 IEEE 15th International Conference on e-Health Networking, Applications Services (Healthcom)*, Oct. 2013, pp. 454–459. doi: 10.1109/HealthCom.2013.6720719.

[210] D. C. Mohr *et al.*, “Trials of Intervention Principles: Evaluation Methods for Evolving Behavioral Intervention Technologies,” *J. Med. Internet Res.*, vol. 17, no. 7, p. e166, Jul. 2015, doi: 10.2196/jmir.4391.

[211] A. Godoy and E. F. Barbosa, “Game-scrum: An approach to agile game development,” *Proceedings of SBGames*, pp. 292–295, 2010.

[212] J. A. Preston, “Serious Game Development: Case Study of the 2013 CDC Games For Health Game Jam,” in *Proceedings of the 2014 ACM International Workshop on Serious Games*, 2014, pp. 39–43. Accessed: Oct. 21, 2016. [Online]. Available: http://sci-hub.cc/http://sci-hub.cc/http://dl.acm.org/citation.cfm?id=2656721

[213] S. Bronner, R. Pinsker, R. Naik, and J. A. Noah, “Physiological and psychophysiological responses to an exer-game training protocol,” *Journal of Science and Medicine in Sport*, 2015.

[214] K. C. O’Leary, M. B. Pontifex, M. R. Scudder, M. L. Brown, and C. H. Hillman, “The effects of single bouts of aerobic exercise, exergaming, and videogame play on cognitive control,” *Clinical Neurophysiology*, vol. 122, no. 8, pp. 1518–1525, 2011.

[215] L. E. Graves, N. D. Ridgers, K. Williams, G. Stratton, and G. T. Atkinson, “The physiological cost and enjoyment of Wii Fit in adolescents, young adults, and older adults,” *Journal of physical activity & health*, vol. 7, no. 3, pp. 393–401, 2010.

[216] D. Kappen, P. Mirza-Babaei, and L. Nacke, “Gamification of Older Adults’ Physical Activity: An Eight-Week Study,” 2018.

[217] B. Guderian *et al.*, “The cardiovascular and metabolic responses to Wii Fit video game playing in middle-aged and older adults.,” *The Journal of sports medicine and physical fitness*, vol. 50, no. 4, pp. 436–442, 2010.

[218] T.-Y. Chuang, W.-H. Sung, H.-A. Chang, and R.-Y. Wang, “Effect of a virtual reality–enhanced exercise protocol after coronary artery bypass grafting,” *Physical therapy*, vol. 86, no. 10, pp. 1369–1377, 2006.

[219] Y.-Z. Wu, J.-Y. Lin, P.-L. Wu, and Y.-F. Kuo, “Effects of a hybrid intervention combining exergaming and physical therapy among older adults in a long-term care facility,” *Geriatrics & gerontology international*, 2018.

[220] L. V. Klompstra, T. Jaarsma, and A. Strömberg, “Exergaming in older adults: A scoping review and implementation potential for patients with heart failure,” *Eur J Cardiovasc Nurs*, vol. 13, no. 5, pp. 388–398, Oct. 2014, doi: 10.1177/1474515113512203.

[221] K. E. Ravenek, D. L. Wolfe, and S. L. Hitzig, “A scoping review of video gaming in rehabilitation,” *Disability and Rehabilitation: Assistive Technology*, vol. 0, no. 0, pp. 1–9, Mar. 2015, doi: 10.3109/17483107.2015.1029538.

[222] M. van Diest, C. J. Lamoth, J. Stegenga, G. J. Verkerke, and K. Postema, “Exergaming for balance training of elderly: state of the art and future developments,” *Journal of NeuroEngineering and Rehabilitation*, vol. 10, no. 1, p. 101, Sep. 2013, doi: 10.1186/1743-0003-10-101.

[223] F. J. Keefe *et al.*, “Virtual reality for persistent pain: A new direction for behavioral pain management,” *PAIN®*, vol. 153, no. 11, pp. 2163–2166, Nov. 2012, doi: 10.1016/j.pain.2012.05.030.

[224] H. Kharrazi, A. S. Lu, F. Gharghabi, and W. Coleman, “A Scoping Review of Health Game Research: Past, Present, and Future,” *Games for Health Journal*, vol. 1, no. 2, pp. 153–164, Apr. 2012, doi: 10.1089/g4h.2012.0011.

[225] J. A. Garcia, K. Felix Navarro, D. Schoene, S. T. Smith, and Y. Pisan, “Exergames for the elderly: towards an embedded Kinect-based clinical test of falls risk,” *Stud Health Technol Inform*, vol. 178, pp. 51–57, 2012.

[226] A. E. Staiano and S. L. Calvert, “The promise of exergames as tools to measure physical health,” *Entertainment Computing*, vol. 2, no. 1, pp. 17–21, 2011, doi: 10.1016/j.entcom.2011.03.008.

[227] A. Velazquez, A. I. Martínez-García, J. Favela, and S. F. Ochoa, “Adaptive exergames to support active aging: An action research study,” *Pervasive and Mobile Computing*, vol. 34, pp. 60–78, 2017.

[228] D. J. Rose, N. Lucchese, and L. D. Wiersma, “Development of a Multidimensional Balance Scale for Use With Functionally Independent Older Adults,” *Archives of Physical Medicine and Rehabilitation*, vol. 87, no. 11, pp. 1478–1485, Nov. 2006, doi: 10.1016/j.apmr.2006.07.263.

[229] J. E. Ware, M. Kosinski, and S. D. Keller, “A 12-Item Short-Form Health Survey: Construction of Scales and Preliminary Tests of Reliability and Validity,” *Medical Care*, vol. 34, no. 3, pp. 220–233, 1996.

[230] A. H. Chu *et al.*, “Comparison of wrist-worn Fitbit Flex and waist-worn ActiGraph for measuring steps in free-living adults,” *PloS one*, vol. 12, no. 2, p. e0172535, 2017.
[231] P. J. Gertler, S. Martinez, P. Premand, L. B. Rawlings, and C. M. J. Vermeersch, *Impact Evaluation in Practice, Second Edition*. World Bank Publications, 2016.
[232] E. F. Binder *et al.*, “Effects of exercise training on frailty in community-dwelling older adults: results of a randomized, controlled trial,” *Journal of the American Geriatrics Society*, vol. 50, no. 12, pp. 1921–1928, 2002.
[233] A. J. Bell, J. K. Talbot-Stern, and A. Hennessy, “Characteristics and outcomes of older patients presenting to the emergency department after a fall: a retrospective analysis,” *Medical Journal of Australia*, vol. 173, no. 4, pp. 179–182, 2000.
[234] A. M. Tromp, S. M. F. Pluijm, J. H. Smit, D. J. H. Deeg, L. M. Bouter, and P. Lips, “Fall-risk screening test: a prospective study on predictors for falls in community-dwelling elderly,” *Journal of clinical epidemiology*, vol. 54, no. 8, pp. 837–844, 2001.
[235] M. Cacciata *et al.*, “Effect of exergaming on health-related quality of life in older adults: A systematic review,” *International journal of nursing studies*, vol. 93, pp. 30–40, 2019.
[236] E. Reis, G. Postolache, L. Teixeira, P. Arriaga, M. L. Lima, and O. Postolache, “Exergames for motor rehabilitation in older adults: an umbrella review,” *Physical Therapy Reviews*, vol. 24, no. 3–4, pp. 84–99, Jul. 2019, doi: 10.1080/10833196.2019.1639012.
[237] G. Pichierri, A. Coppe, S. Lorenzetti, K. Murer, and E. D. de Bruin, “The effect of a cognitive-motor intervention on voluntary step execution under single and dual task conditions in older adults: a randomized controlled pilot study,” *Clinical interventions in aging*, vol. 7, p. 175, 2012.
[238] C. J. Jones and D. J. Rose, *Physical activity instruction of older adults*. Human Kinetics, 2005. [Online]. Available: https://www.google.com/books?hl=es&lr=&id=ohfWKmzGDuIC&oi=fnd&pg=PR17&dq=Physical+Activity+Instruction+of+Older+Adults&ots=XWn0h_r6Vn&sig=dN3GJD7KSoiWceUtqhcmNMP4mAg
[239] C. Höchsmann, M. Schüpbach, and A. Schmidt-Trucksäss, “Effects of exergaming on physical activity in overweight individuals,” *Sports Medicine*, vol. 46, no. 6, pp. 845–860, 2016.
[240] J. P. Chaput *et al.*, “Objectively measured physical activity, sedentary time and sleep duration: independent and combined associations with adiposity in canadian children,” *Nutrition & diabetes*, vol. 4, no. 6, p. e117, 2014.
[241] N. Aguilar-Farías, W. J. Brown, and G. G. Peeters, “ActiGraph GT3X+ cut-points for identifying sedentary behaviour in older adults in free-living environments,” *Journal of Science and Medicine in Sport*, vol. 17, no. 3, pp. 293–299, 2014.
[242] J. L. Copeland and D. W. Esliger, “Accelerometer assessment of physical activity in active, healthy older adults,” *Journal of aging and physical activity*, vol. 17, no. 1, pp. 17–30, 2009.
[243] A. Barnett, D. van den Hoek, D. Barnett, and E. Cerin, “Measuring moderate-intensity walking in older adults using the ActiGraph accelerometer,” *BMC geriatrics*, vol. 16, no. 1, p. 211, 2016.
[244] H. Simão and A. Bernardino, “User Centered Design of an Augmented Reality Gaming Platform for Active Aging in Elderly Institutions,” in *Proceedings of the 5th International Congress on Sport Sciences Research and Technology Support*, Funchal, Portugal, 2017, vol. Volume 1: AHA, pp. 151–162. doi: 10.5220/0006606601510162.
[245] F. L. Siena, B. Byrom, P. Watts, and P. Breedon, “Utilising the Intel RealSense Camera for Measuring Health Outcomes in Clinical Research,” *J Med Syst*, vol. 42, no. 3, p. 53, Feb. 2018, doi: 10.1007/s10916-018-0905-x.
[246] R. A. Clark, B. F. Mentiplay, E. Hough, and Y. H. Pua, “Three-dimensional cameras and skeleton pose tracking for physical function assessment: A review of uses, validity, current developments and Kinect alternatives,” *Gait & Posture*, vol. 68, pp. 193–200, Feb. 2019, doi: 10.1016/j.gaitpost.2018.11.029.
[247] J. A. Albert, V. Owolabi, A. Gebel, C. M. Brahms, U. Granacher, and B. Arnrich, “Evaluation of the Pose Tracking Performance of the Azure Kinect and Kinect v2 for Gait Analysis in Comparison with a Gold Standard: A Pilot Study,” *Sensors*, vol. 20, no. 18, Art. no. 18, Jan. 2020, doi: 10.3390/s20185104.

[248] A. Gonçalves, M. F. Montoya, R. Llorens, and S. Bermúdez i Badia, "A virtual reality bus ride as an ecologically valid assessment of balance: a feasibility study," *Virtual Reality*, Apr. 2021, doi: 10.1007/s10055-021-00521-6.
[249] M. Ahmad, H. Sousa, É. Quintal, and S. Bermúdez i Badia, "Efficacy of Augmented Reality-based Virtual Hiking in Cardiorespiratory Endurance: A Pilot Study," Feb. 2021, vol. 2, pp. 575–582. doi: 10.5220/0010312405750582.