From Body Tracking Interaction in Floor Projection Displays to Elderly Cardiorespiratory Training Through Exergaming

Afonso Gonçalves[1,2(✉)], Filipa Nóbrega[2], Mónica Cameirão[1,2], John E. Muñoz[1,2], Élvio Gouveia[1,3], and Sergi Bermudez i Badia[1,2]

[1] Madeira Interactive Technologies Institute, Funchal, Portugal
{afonso.goncalves,monica.cameirao,john.cardona, sergi.bermudez}@m-iti.org

[2] Faculdade de Ciências Exatas e da Engenharia, Universidade da Madeira, Funchal, Portugal

[3] Faculdade de Ciências Sociais, Universidade da Madeira, Funchal, Portugal
erubiog@uma.pt

Abstract. The opportunity to develop new natural user interfaces has come forward due to the recent development of inexpensive full body tracking sensors, which has made this technology accessible to millions of users. In this paper, we present a comparative study between two natural user interfaces, and a cardiorespiratory training exergame developed based on the study results. The focus was on studying interfaces that could easily be used by an elderly population for interaction with floor projection displays. One interface uses both feet position to control a cursor and feet distance to trigger activation. In the alternative interface, the cursor is controlled by forearm ray casting into the projection floor and interaction is activated by hand pose. These modes of interaction were tested with 19 elderly participants in a point-and-click and a drag-and-drop task using a between-subjects experimental design. The usability, perceived workload and performance indicators were measured for each interface. Results show a clear preference towards the feet-controlled interface and a marginally better performance for this method. The results from the study served as a guide to the design of a cardiorespiratory fitness exergame for the elderly. The game "Grape Stomping" uses ground projection and mapping to display real-size winery elements. These virtual elements are used to simulate, in a playful way, the process of grape maceration through repeated stomping. A playtest session with nine elderly users was completed and its insights are presented in addition to the description of the game.

Keywords: Large display interface · Floor projection · Elderly · Exergames natural user interface · Kinect

A. Holzinger et al. (Eds.): PhyCS 2016–2018, LNCS 10057, pp. 58–77, 2019.
https://doi.org/10.1007/978-3-030-27950-9_4

1 Introduction

Developed countries' populations are becoming increasingly older, with estimates that one-third of the European citizens will be over 65 years old by 2060 [1]. With the aging process, visual perception is commonly negatively affected [2] and the effects of sedentary lifestyles become more prominent. A computer system that could lighten such problems using large dimension displays and motion tracking interfaces could prove advantageous. More concretely, applications targeting engagement and physical fitness would provide extensive health benefits in older adults [3].

Meanwhile, the release of low-cost body tracking sensors for gaming consoles has 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 [4], and Kinect V2, having 3.9 million units bundled and sold along with Xbox One consoles by Jan. 2014 [5]. The popular 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 at reflecting user skills and taking full advantage of their capacities to fit their task and context demands from the moment they start interacting [6]. In addition to the body tracking sensors' unique interface capabilities, they also provide exciting possibilities for automatic monitoring of health-related problems through kinematic data analysis. For example, automated systems for assessing fitness indicators in elderly [7, 8], automatic exercise rehabilitation guidance [9], or diagnosis and monitoring of Parkinson's disease [10].

The coupling of body tracking depth sensors, such as Kinect, and projectors enable systems to not only track the user movements relative to the sensor but also to map virtual content on 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 [11].

In this paper, we present a comparative study of two interaction modalities for floor projections, and, based on the results of the comparison an exergame developed for cardiorespiratory training of the elderly. In the study, we combined floor projection mapping with whole body tracking to provide two modalities of body gesture NUIs in controlling a cursor. One modality is based on 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. The insights provided by the study results served as a preliminary step in the development of a senior exergame for cardiorespiratory fitness training. The game is projected on the ground and makes use of one-to-one mapping to project real-size winery elements. These virtual elements are used to mimic, in a fun and playful way, the process of grape maceration through stomping.

This work is an initial and important step in the development of content for a mobile autonomous robotic system designed to assist elderly in keeping an active lifestyle through adaptable exergames. The platform, equipped with a micro projector and depth sensor will be able to identify users and provide custom exergames through live projection mapping, or spatial augmented reality. While the results from this experiment guided the exergame interaction design, they will also help in the future development of a gesture interface for such mobile platform.

This paper is an extended version of the conference paper "Evaluating Body Tracking Interaction in Floor Projection Displays with an Elderly Population" [12], presented at the 3rd International Conference on Physiological Computing Systems PhyCS 2016. While it maintains the same structure and most of the original paper content, Sect. 6 was added as it represents work that was in progress at the time of the publication and a direct consequence of the original study results.

2 Related Work

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

An area where several in-air gesture interfaces have been proposed is in pan-and-zoom navigation control. In [13] the authors investigated the impact three interaction variables had in task completion time and navigation overshoots when interacting with a wall-sized display. The variables were: uni- vs. bi-manual, linear vs. circular movements, and 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). An NUI for controlling virtual globes is introduced in [14]. 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 an 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 relative position of the hands is used to control the virtual globe. For street view control, it makes use of gestures that mimic the 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 [15], 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 in the motorcycle metaphor. A handlebar like motion controlled turning and right-hand tilting acted as 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 in 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 form of a bird/airplane avatar. Descriptive statistic results showed high levels of usability and presence for both systems, with higher values for the latter. The use of the 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 [16]. 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 modeled 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 [17], 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 useful in interventional radiology. The majority also found it easy to moderately difficult to accomplish the tasks. Similarly, in [18] the authors introduced a solution for interaction with these systems using inertial sensors instead. 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 [19] the authors found, by running an experiment in a Wizard of Oz set-up, that participants would adopt the point-and-click mouse metaphor when asked to perform tasks in a large display. In [20], participants were asked to propose gestures for common TV functions. The gesture agreement was assessed for each command and a set of guidelines proposed. Contrary to what was shown in [13] for pan and zoom gesture, here one hand gesturing was preferred. Hand posture naturally emerged as a way of communicating intention for gesture interaction.

When designing an 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 [6]. Effective ways of countering the "live mic" problem are to reserve specific actions for interaction or reserve clutching mechanism that will disengage the gesture interpretation. The review made by Golod et al. [21] 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 [22], from a Wizard of Oz design, the authors tried to identify gestures for pan, zoom, rotate and tilt control. 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 [23] used the hand palm facing the screen for activating cursor control. [24] 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 [25] the authors describe an interactive floor prototype, controlled by body movement and mobile phones, which was set-up on a large public library hall. This arrangement enabled them not only to take advantage of the open space, filled by the large projected interface but also from its public function of promoting social interaction. These types of interfaces were proposed as an alternative to interactive tabletops [26], useful for not being as spatially restraining as the latter. In their study, the authors also explored the preferred methods of activation for buttons in these floors, being feet *tap* their final choice of design.

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

- When designing an 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 goals of an interface, 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 him to mentally project their arm into the floor. Therefore, we hypothesized that differences for 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. We expected that raycasting would provide better results as it is more widely used for interaction with large displays and requires little physical effort by the user.

Next, we engaged in the development of an exergame for cardiorespiratory fitness training for the elderly. The design of exergames to promote physical activity in senior adults has been characterized for the lack of focused game design methodologies which can include appropriate content, real needs of the senior population and adapted interfaces for a natural interaction [27, 28]. To overcome these limitations, several investigations point at the need of including older adults in early design stages and to constantly evaluate playable prototypes to include real field data in the exergame design process [29]. For instance, Gerling and colleagues carried out a study showing how the frail elderly population may not be suitable for playing Wii games since the interface and the navigation through menu structures produce inadequate feedback [30]. Thus, the design of exergames oriented to cover individual fitness levels and user needs is essential to deliver positive user experiences that maximize the health benefits of training with this technology. A study with 170 senior adults also showed how a good knowledge of game preferences and motivations as well as an active participation of the target population within the design of novel games might produce more satisfactory experiences, which will facilitate a long-term adoption of this technology, one of the cornerstones for elderly exergaming [31]. Consequently, we adopted user-centered design methodologies to design, develop and partially evaluate interaction techniques in an exergame based on floor-projection to promote physical activity in a group of senior adults. By using these game design techniques, we aimed to: (a) provide a use case scenario for exercise promotion to integrate the previously studied interaction techniques, and (b) create better exergame experiences based on field tests to facilitate acceptance of such a technology.

The main contributions of this manuscript lies in the description of methods and results to: (a) evaluate body tracking interaction in floor projection with a group of active senior adults, aimed at elucidate the best interaction technique based in usability, workload and task performance measurements (Sects. 3 and 4); and (b) the integration of such results for the design and evaluation of a novel spatial augmenting reality exergame (Sect. 6).

3 Methods for Evaluating Body Tracking Interaction in Floor Projection Displays with an Elderly Population

The first experiment aimed at understanding the differences in terms of usability, workload levels, and task performance between two interaction modes in floor projections: forearm ray casting and feet interaction. This was evaluated using two different tasks: the point-and-click and the drag-and-drop.

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 20 cm apart. For the second mode of interaction, named "*arm*", the forearm position and orientation is treated as a vector (from elbow to wrist) and raycasted onto the floor plane, the cast controls the position of the cursor (as schematized in Fig. 1), while activation is done by closing the hand. Due to low reliability of the Kinect V2 sensor 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.

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 for the common drag-and-drop.

In both tasks, the participant controls a circular cursor (ø 17 cm) with 1 s activation duration, meaning that the activation gesture (feet together or hand closed) must be sustained for 1 s for the cursor to interact with the virtual element it is positioned on.

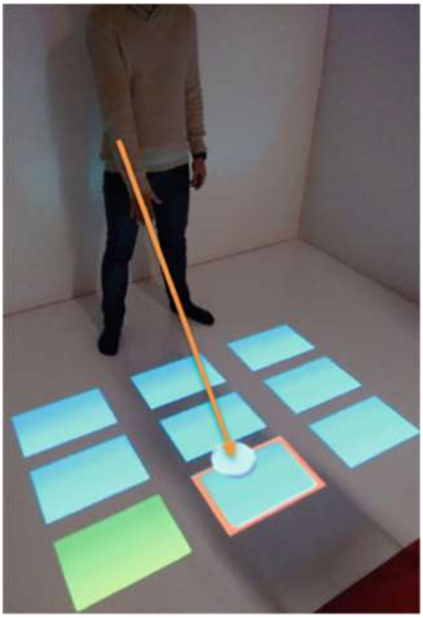

Fig. 1. Controlling the cursor position through forearm ray casting, extracted from [12].

This activation is represented on the cursor itself, which changes color in a circular way proportionally to the duration of the gesture.

Point-and-Click Task. In the point-and-click task, a set of 9 rectangles (40 cm 25 cm) are projected in the floor, on a 3 by 3 configuration, separated 12 cm laterally and 8 cm vertically as shown in Fig. 2. Out of the 9 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.

Drag-and-Drop Task. In the drag-and-drop task, 4 rectangles (40 cm × 25 cm) are projected on the ground, spaced 70 cm horizontally and 40 cm vertically, 3 of which are blue distractors and one is the target (green). In the center, a movable yellow rectangle (30 cm × 19 cm) is initially shown, as presented in Fig. 3. The participant can "*grab*" the yellow rectangle by activating it. Once it has been "*grabbed*", it can be

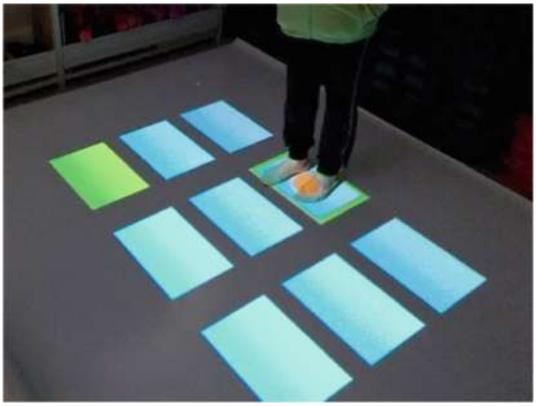

Fig. 2. Point-and-click task being performed with the "*feet*" interface, extracted from [12]. (Color figure online)

dropped by activating it again (joining the feet or closing the hand, depending on the mode of interaction). The purpose of the task 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 background (neutral drop); and drop yellow on distractors (incorrect drop). 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 frame of the yellow object changes to green indicating that is being dragged by the cursor. Dropping

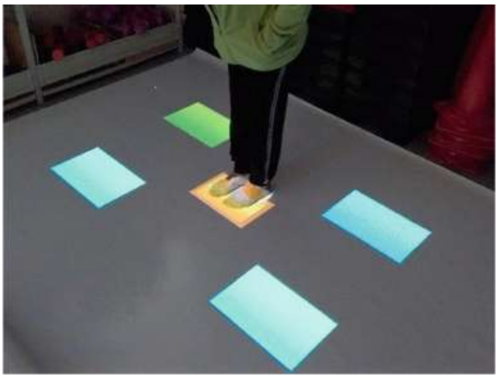

Fig. 3. The drag-and-drop task being performed with the "*feet*" interface, extracted from [12]. (Color figure online)

it on a distractor will create a red frame around the distractor, and dropping it on a target will show a green frame around it.

3.3 Technical Setup

The hardware was set up in a dimly illuminated room and a white PVC 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 × 90 cm). A Microsoft Kinect V2 was placed horizontally next to the projector, facing the projection area (Fig. 4).

Fig. 4. Experimental setup diagram, extracted from [12].

3.4 Sample

The target population of the study were 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 [32]);
3. Do not present low physical functioning (assessed by the Composite Physical Function scale [33]).

The experiment took place over the course of 2 days. Nineteen participants (16 females; ages: M = 70.2 SD = 5.3) volunteered and provided written informed consent. The participants were randomly allocated to one of the two conditions, 10 being assigned to the "*feet*" and 9 to the "*arm*" conditions of interaction.

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 individual participant trial, the point-and-click task was explained and shown being performed through example according to the participant experimental condition. This was followed by a training period and then by a 2 min' session while performance metrics were recorded. Lastly, participants were asked to fill the System Usability Scale (SUS) [34] and NASA-TLX (TLX) [35] questionnaires. After it, the same procedure was followed for the drag-and-drop task.

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 Subsect. 3.2. Normality of the data distributions was assessed using the Kolmogorov-Smirnov test for measurements concerning performance. The variables that showed a normal distribution are highlighted in Tables 1 and 2. For the pairs (between conditions) of measurements that fitted the assumption of normality, parametric t-tests were used When significant differences in the pairs variances were present, shown by the Levene's test, equal variances were not assumed. All the others 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.

4 Results in Evaluating Body Tracking Interaction in Floor Projection Displays with an Elderly Population

4.1 Point-and-Click Task

For the "*feet*" condition in the point-and-click task, the descriptive statistics are presented in Table 1. 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 1. Higher values of neutral and incorrect clicks are visible compared to the previous condition. Similarly, it can be seen a decrease in the median of the SUS usability score and an increase of the TLX workload index.

Results revealed significant higher System Usability Scale scores for the participants interfacing with their feet compared to 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, $U = 24.5$, $p > .05$ (Fig. 5). 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 (Fig. 6).

Table 1. Descriptive statistics of the measurements for the point-and-click task, extracted from [12].

Variable	"Feet" interface		"Arm" interface	
	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[a]	15
Neutral	1.00	2	4.00[a]	7
Incorrect	0.00	1	2.00[a]	3

[a]Normally distributed

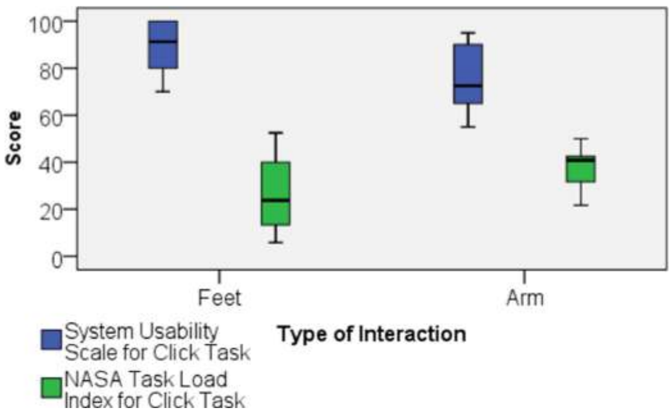

Fig. 5. System Usability Scale and Nasa-Task Load Index scores for the point-and-click task, extracted from [12].

4.2 Drag-and-Drop Task

The descriptive statistics for the "feet" condition, in the drag-and-drop task are presented in Table 2, 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 2, 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.

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 (Fig. 7). There were no significant differences in correct grabs, neutral grabs, and correct drops,

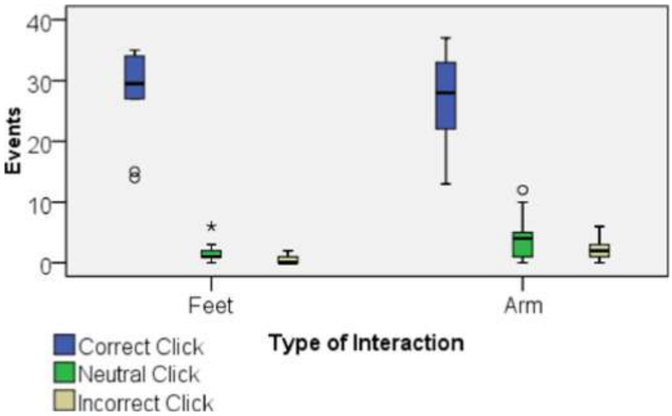

Fig. 6. Participants' performance on the point-and-click task (circles represent outliers and stars extreme outliers), extracted from [12].

Table 2. Descriptive statistics of the measurements for the drag-and-drop task, extracted from [12].

Variable	"Feet" interface		"Arm" interface	
	Median	Interquartile range	Median	Interquartile range
SUS	93.75	16.25	41.67	21.25
TLX	22.50	16.46	11.00[a]	22.50
Correct	14.50[a]	8	10.00[a]	9
Neutral	13.50[a]	4	10.00[a]	9
Incorrect	14.00[a]	7	1.00[a]	10

[a]Normally distributed

$t(17) = .565$, $t(17) = .863$ 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, $U = 44.5$, $p > .05$ (Fig. 8).

5 Discussion Regarding the Evaluation of Body Tracking Interaction in Floor Projection Displays with an Elderly Population

For both the point-and-click and drag-and-drop tasks we identified a significant impact on system usability, being the "*feet*" interaction method preferable in both cases. The "*feet*" modality achieved high levels of usability, scores over 90, while the "*arm*" had levels of usability around 71, very close to the standard lower limit of good, 68. In the case of perceived workload indexes, for the point-and-click there were no significant differences found between the conditions. For drag-and-drop, the "*feet*" interface was

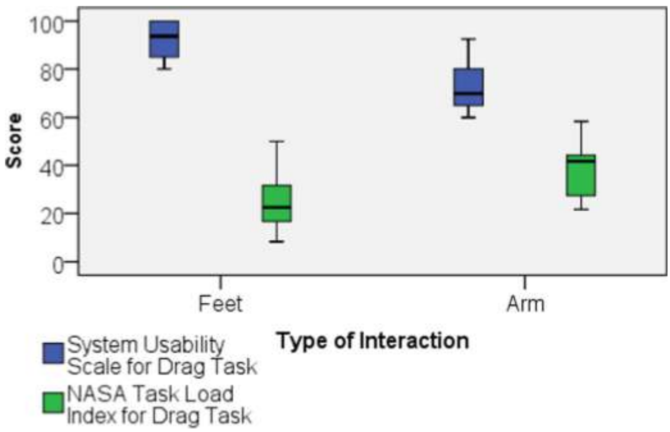

Fig. 7. System Usability Scale and Nasa-Task Load Index scores for the drag-and-drop task, extracted from [12].

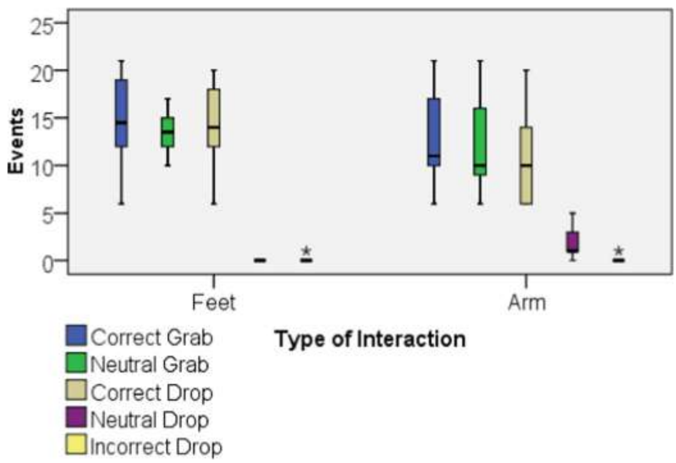

Fig. 8. Participants' performance on the drag-and-drop task, extracted from [12].

significantly less demanding for the participants. In both conditions, workload indexes for the "*feet*" were around 23 while for the "*arm*" the values were around 41. Although interfaces similar to our "*arm*" method have been the focus of previous research [13, 23] and shown to be a method that participants naturally display [19, 20, 22], in our experiment we found sufficient evidence that an alternative way of interacting with projected floor elements is preferred by elder people. This preference by the participants for the "*feet*" interface might be linked to the simpler mapping of the cursor control provided, which is known to have a lowering effect on cognitive load [16, 36]. In terms of performance, for the point-and-click task very low numbers of neutral and incorrect clicks (although significantly higher for the "*arm*") and a comparable number of correct clicks were observed. Similar results were found in the drag-and-drop task, with low numbers of neutral and incorrect drops for both methods and analogous 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 in the interpretation of these results as proof of a clear performance advantage provided by any of the interfaces.

From the observed results, we can summarize a take-home message. When elderly interact with floor projection, a direct activation of the virtual elements with the feet is preferred for both usability and performance. In contrast, pointing and hand pose are considered less usable and more cumbersome.

6 Design of "Grape Stomping": A Cardiorespiratory Training Exergame for the Elderly

This section focuses on describing the "Grape Stomping" exergame, a cardiorespiratory training game using floor projection and feet interaction developed based on the observations from the previous section. The story and aesthetics were rooted in the traditional winemaking activities of the Douro region, an important activity for the development of the Portuguese economy, being the Douro Region the most famous and traditional region for wine production [37]. One of the most characteristic activities around the wine culture in Portugal is the grape stomping. The grapes are placed in large tanks, people tread knee-deep in grapes, arms linked as they stomp and dance on the grapes underfoot following the beats of the traditional folk music. This activity was recreated using a real-size floor projection of an ancient Portuguese winery, see Fig. 9. In this virtual world, there are two main elements: a row of three open-top tanks that can hold grapes and be used to stomp them, and a conveyor belt that continuously bring grapes into the play area. The tanks have two small footprints which are used to provide feedback over the position of the user and correct stepping height. Background sound with a folk music accompanies the interaction and sound effects were added to the stomping and tank filling events. GUI elements in the exergame include a countdown visualizer as well as individual counters for each tank showing the number of times it has been successfully filled.

The grape stomping activity was chosen taking two foundations in consideration, that general repeated stepping is the recommended exercise for cardiorespiratory fitness assessment [38, 39], and our interaction study results take-home message. In order to facilitate the interaction with the exergame, we used projection mapping technology which allows the augmentation of real-world spaces using simple projections instead of special displays [40]. With this technology, users are able to physically stomp the virtual grapes placed in the tanks projected on the floor in real scale. This game mechanic was directly inspired by the results (and product of the developed technology) presented in Sect. 4. Specifically, the direct mapping of the grape stepping action was inspired on the Point-and-Click task using the *feet* modality, while for the upper limb mechanics we avoided direct mapping and opted instead for gesture motion-based interaction.

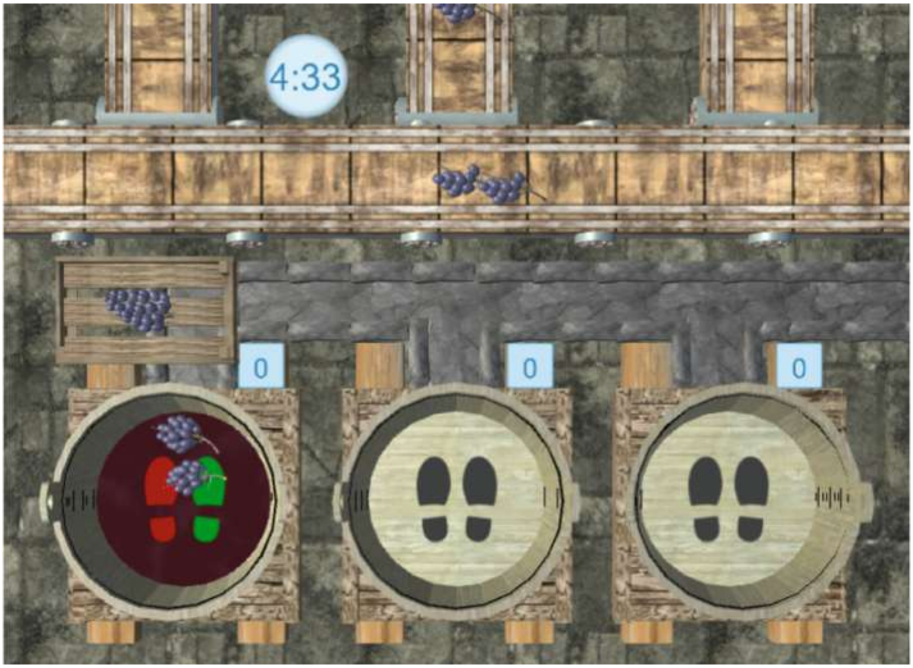

Fig. 9. A screenshot of the grape stomping exergame.

The goal of the exergame is to stomp on the grapes inside the tanks, pulled from the conveyor belt into the tank by small virtual baskets through repeated arm extension/flexion. The stomping process requires users to raise the knees to a pre-defined height when there are grapes in a tank, producing the grape juice. After filling up one of the tanks, users are free to move to the neighboring tanks to continue the exercise. The experience ends once the pre-established training time finalizes, presenting the total score in form of liters of wine produced, number of steps performed and number of grape bunches pulled.

An exergame playtest session was conducted in a local senior gymnasium in Madeira, Portugal, see Fig. 10. Nine older adults (8 females; ages M = 62.3, SD = 6.2) participated and two exercise instructors in the senior gymnasium were interviewed. From their feedback, we can enclose the main insights in the following items:

- Social aspect: senior elders enjoy playing exergames mainly for two reasons: they like to win competitions and they find here an opportunity to socialize (skills and experiences). Multiplayer playability is a key factor to improve technology adoption.
- In-time feedback: the lack of past experiences with gaming technologies obstructs a fluid interaction. Improving the quality and frequency of the feedback provided in the videogame facilitates the understanding of what to do, how to do and when to do it.

Fig. 10. "Grape Stomping" playtest at a local gym.

- Automatic movements: options for the exergame personalization must include well-defined strategies to facilitate the interaction for people with several motor disabilities. Since the exergame includes 3 body movements (e.g. step, side step, and arm extension), health professionals should be able to activate/deactivate individual body gestures depending on users' abilities.
- Cognitive tasks: the inclusion of more cognitive-demanding activities in conjunction with the physical exertion in the Exergaming might enhance the health benefits and increase the likelihood of the long-term adoption of this technology.
- Control parameters: the inclusion of multiple game parameters might help to facilitate the personalization of activities' in terms of the fitness domains and training dimensions. By defining a set of game parameters, the difficulty of the exergame will be controllable allowing a more precise adaptation to specific motor and/or cognitive skills.

With these insights in mind, several features were added to the game. Cognitive challenges were added using game mechanics around wine recipes to stimulate visual processing and concentration. Three types of grapes are used: green, maroon and damaged (distractors). Users are encouraged to grab specific grape bunches following a recipe (e.g. 10 greens and 5 maroons) avoiding damaged bunches which appear randomly. Thus, the damaged bunches are used as distractors and their percentage can be defined in the initial menu. Optional tutorial and in-time instructions were also added. While the interactive video tutorial explains how to play the game at the start of the game, in-game videos are triggered if the game detects inactivity or unsuccessful movements being performed by the player. Thus, reinforcing the learning process of each gesture separately when the exergame is running. Moreover, a multiplayer option allows up to three users to be side by side in individual barrels to either collaborate or compete.

The final version of the game includes the following experience personalization parameters: duration (minutes), number of players and multiplayer mode (collaborative, cooperative), stepping (yes/no), pulling (yes/no), treadmill velocity (grape bunches/s), step height (cm), recipes (yes/no) and distractors (%).

7 Conclusions

Due to the growing number of elderly in developed countries and their specific needs we tried to get an insight of the desirability of different modes of controlling interaction in interactive floors. A medium which, by being scaled easily, can mitigate the visual perception deficits associated with old age, and can promote physical activity. Thus, in this work, 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 on either controlling the cursor with the direct mapping of feet position onto the projection surface or, alternatively, 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 an insight into the preferred modes of interaction for the elder population. Contrary to our initial guess, the results showed that from the two proposed methods the "*feet*" interface was superior in all the domains measured. It was shown that this method was perceived as more usable in both the tasks tested and at least less demanding in terms of workload for the drag-and-drop task. In terms of performance, a marginal advantaged was shown also for the "*feet*" method. This insight, triggered the development of a floor projection exergame focused on cardiorespiratory fitness training. The game was designed around the winemaking traditions of the Douro region in Portugal and the main interaction method used was virtual grape stomping. A game playtest session with the end-users provided feedback necessary to take the game further and allowed the addition of supplementary features, considered significant by both the target population and elderly sports professionals. The most important additions where difficulty/exertion parameterization, multiplayer support and guidance instructions.

Acknowledgment. The authors thank Funchal's Santo António municipal gymnasium for their cooperation, Teresa Paulino for the development of the experimental tasks, Fábio Pereira for his help during the data collection process, and Diogo Freitas and John Sousa for their contributions in the development of the first prototype of "Grape Stomping".

This work was supported by the Fundação para a Ciência e Tecnologia through the AHA project (CMUPERI/HCI/0046/2013) and LARSyS – UID/EEA/50009/2013.

Contributions. Afonso Gonçalves designed and carried out the floor projection interaction study, designed the "Grape Stomping" game and wrote the paper. Filipa Nóbrega developed the game and contributed to the writing of the paper. Mónica Cameirão supervised the study, contributed to the game design and writing of the paper. John Muñoz, Élvio Gouveia and Sergi Bermudez i Badia contributed to the game design and writing of the paper.

References

1. European Commission, Economic and Financial Affairs: The 2012 Ageing Report (2012)
2. Fozard, J.: Vision and hearing in aging. In: Birren, J.E., Schaie, K.W. (eds.) Handbook of the Psychology of Aging, pp. 143–156. Academic Press, San Diego (1990)
3. World Health Organization: Global recommendations on physical activity for health (2010). http://www.who.int/dietphysicalactivity/publications/9789241599979/en/
4. Microsoft News Center: Xbox Execs Talk Momentum and the Future of TV (2013). http://news.microsoft.com/2013/02/11/xbox-execs-talk-momentum-and-the-future-of-tv/
5. Microsoft's Q2: Record $24.52 billion revenue and 3.9 million Xbox One sales. http://www.theverge.com/2014/1/23/5338162/microsoft-q2-2014-financial-earnings
6. Wigdor, D., Wixon, D.: Brave NUI World: Designing Natural User Interfaces for Touch and Gesture. Elsevier, Amsterdam (2011)
7. Chen, C., Liu, K., Jafari, R., Kehtarnavaz, N.: 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), pp. 4135–4138 (2014)
8. Gonçalves, A., Gouveia, É., Cameirão, M., Bermúdez i Badia, S.: Automating senior fitness testing through gesture detection with depth sensors. In: Proceedings of the IET International Conference on Technologies for Active and Assisted Living (TechAAL 2015). Institution of Engineering and Technology, London (2015)
9. Da Gama, A., Chaves, T., Figueiredo, L., Teichrieb, V.: Guidance and movement correction based on therapeutics movements for motor rehabilitation support systems. In: 2012 14th Symposium on Virtual and Augmented Reality (SVR), pp. 191–200 (2012)
10. Spasojević, S., Santos-Victor, J., Ilić, T., Milanović, S., Potkonjak, V., Rodić, A.: A vision-based system for movement analysis in medical applications: the example of Parkinson disease. In: Nalpantidis, L., Krüger, V., Eklundh, J.-O., Gasteratos, A. (eds.) ICVS 2015. LNCS, vol. 9163, pp. 424–434. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-20904-3_38
11. Jones, B., 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, pp. 637–644. ACM, New York (2014)
12. Gonçalves, A., Cameirão, M.: Evaluating body tracking interaction in floor projection displays with an elderly population. In: Proceedings of the 3rd International Conference on Physiological Computing Systems - Volume 1: PhyCS, Lisbon, Portugal, pp. 24–32 (2016)
13. Nancel, M., Wagner, J., Pietriga, E., Chapuis, O., Mackay, W.: Mid-air pan-and-zoom on wall-sized displays. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 177–186. ACM, New York (2011)
14. Boulos, M.N.K., Blanchard, B.J., Walker, C., Montero, J., Tripathy, A., Gutierrez-Osuna, R.: Web GIS in practice X: a Microsoft Kinect natural user interface for Google Earth navigation. Int. J. Health Geogr. **10**, 45 (2011)
15. Francese, R., Passero, I., Tortora, G.: Wiimote and Kinect: gestural user interfaces add a natural third dimension to HCI. In: Proceedings of the International Working Conference on Advanced Visual Interfaces, pp. 116–123. ACM, New York (2012)
16. Roupé, M., Bosch-Sijtsema, P., Johansson, M.: Interactive navigation interface for Virtual Reality using the human body. Comput. Environ. Urban Syst. **43**, 42–50 (2014)
17. Tan, J.H., Chao, C., Zawaideh, M., Roberts, A.C., Kinney, T.B.: Informatics in radiology: developing a touchless user interface for intraoperative image control during interventional radiology procedures. RadioGraphics. **33**, E61–E70 (2013)

18. Bigdelou, A., Schwarz, L., Navab, N.: An adaptive solution for intra-operative gesture-based human-machine interaction. In: Proceedings of the 2012 ACM International Conference on Intelligent User Interfaces, pp. 75–84. ACM, New York (2012)
19. Fikkert, W., van der Vet, P., van der Veer, G., Nijholt, A.: Gestures for large display control. In: Kopp, S., Wachsmuth, I. (eds.) GW 2009. LNCS (LNAI), vol. 5934, pp. 245–256. Springer, Heidelberg (2010). https://doi.org/10.1007/978-3-642-12553-9_22
20. Vatavu, R.-D.: User-defined gestures for free-hand TV control. In: Proceedings of the 10th European Conference on Interactive TV and Video, pp. 45–48. ACM, New York (2012)
21. Golod, I., Heidrich, F., Möllering, C., Ziefle, M.: Design principles of hand gesture interfaces for microinteractions. In: Proceedings of the 6th International Conference on Designing Pleasurable Products and Interfaces, pp. 11–20. ACM, New York (2013)
22. Lee, S.-S., Chae, J., Kim, H., Lim, Y., Lee, K.: 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, pp. 617–626. ACM, New York (2013)
23. Bragdon, A., DeLine, R., Hinckley, K., Morris, M.R.: Code space: Touch + Air gesture hybrid interactions for supporting developer meetings. In: Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces, pp. 212–221. ACM, New York (2011)
24. Hopmann, M., Salamin, P., Chauvin, N., Vexo, F., Thalmann, D.: Natural activation for gesture recognition systems. In: CHI 2011 Extended Abstracts on Human Factors in Computing Systems, pp. 173–183. ACM, New York (2011)
25. Krogh, P., Ludvigsen, M., Lykke-Olesen, A.: "Help Me Pull That Cursor" A collaborative interactive floor enhancing community interaction. Australas. J. Inf. Syst. **11**, 75–87 (2004)
26. Augsten, T., 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, pp. 209–218. ACM, New York (2010)
27. Larsen, L.H., Schou, L., Lund, H.H., Langberg, H.: The physical effect of exergames in healthy elderly—a systematic review. GAMES Health Res. Dev. Clin. Appl. **2**, 205–212 (2013)
28. Molina, K.I., Ricci, N.A., de Moraes, S.A., Perracini, M.R.: Virtual reality using games for improving physical functioning in older adults: a systematic review. J. Neuroeng. Rehabil. **11**, 156 (2014)
29. Vanden Abeele, V.A., Van Rompaey, V.: Introducing human-centered research to game design: designing game concepts for and with senior citizens. In: CHI 2006 Extended Abstracts on Human Factors in Computing Systems, pp. 1469–1474. ACM (2006)
30. Gerling, K., Masuch, M.: When gaming is not suitable for everyone: playtesting Wii games with frail elderly. In: 1st Workshop on Game Accessibility (2011)
31. Sayago, S., Rosales, A., Righi, V., Ferreira, S.M., Coleman, G.W., Blat, J.: On the conceptualization, design, and evaluation of appealing, meaningful, and playable digital games for older people. Games Cult. **11**, 53–80 (2016)
32. Folstein, M.F., Folstein, S.E., McHugh, P.R.: Mini-mental state. J. Psychiatr. Res. **12**, 189–198 (1975)
33. Roberta E. Rikli, C.J.J.: The Reliability and Validity of a 6-Minute Walk Test as a Measure of Physical Endurance in Older Adults (1998)
34. Brooke, J.: SUS-A quick and dirty usability scale. Usability Eval. Ind. **189**, 194 (1996)
35. Hart, S.G., Staveland, L.E.: Development of NASA-TLX (Task Load Index): results of empirical and theoretical research. In: Hancock, P.A., Meshkati, N. (eds.) Advances in Psychology, pp. 139–183. North-Holland, Amsterdam (1988)

36. Mousavi Hondori, H., Khademi, M., Dodakian, L., McKenzie, A., Lopes, C.V., Cramer, S. C.: Choice of human-computer interaction mode in stroke rehabilitation. Neurorehabil. Neural. Repair. **30**, 258–265 (2015)
37. Cunha, C.A., Cunha, R.: Culture and Customs of Portugal. ABC-CLIO (2010)
38. Oja, P., Tuxworth, B., et al.: Eurofit for adults: assessment of health-related fitness. Council of Europe (1995)
39. Rikli, R.E., Jones, C.J.: Development and validation of a functional fitness test for community-residing older adults. J. Aging Phys. Act. **7**, 129–161 (1999)
40. Bimber, O., Raskar, R.: Spatial Augmented Reality: Merging Real and Virtual Worlds. CRC Press, Boca Raton (2005)