Augmented Reality for Older Adults: Exploring Acceptability of Virtual Coaches for Home-based Balance Training in an Aging Population

Fariba Mostajeran
Universität Hamburg
Hamburg, Germany
mostajeran@informatik.uni-hamburg.de

Frank Steinicke
Universität Hamburg
Hamburg, Germany
frank.steinicke@uni-hamburg.de

Oscar Javier Ariza Nunez
Universität Hamburg
Hamburg, Germany
ariza@informatik.uni-hamburg.de

Dimitrios Gatsios
University of Ioannina
Ioannina, Greece
dgatsios@cc.uoi.gr

Dimitrios Fotiadis
University of Ioannina
Ioannina, Greece
fotiadis@cc.uoi.g

ABSTRACT
Balance training has been shown to be effective in reducing risks of falling, which is a major concern for older adults. Usually, exercise programs are individually prescribed and monitored by physiotherapeutic or medical experts. Unfortunately, supervision and motivation of older adults during home-based exercises cannot be provided on a large scale, in particular, considering an ageing population. Augmented reality (AR) in combination with virtual coaches could provide a reasonable solution to this challenge.

We present a first investigation of the acceptance of an AR coaching system for balance training, which can be performed at home. In a human-centered design approach we developed several mock-ups and prototypes, and evaluated them with 76 older adults. The results suggest that older adults find the system encouraging and stimulating. The virtual coach is perceived as an alive, calm, intelligent, and friendly human. However, usability of the entire AR system showed a significant negative correlation with participants’ age.

Author Keywords
Health and well-being, augmented reality, older adults, balance training

CCS Concepts
•Human-centered computing → Mixed / augmented reality; •Applied computing → Consumer health;

INTRODUCTION
Falls are a major concern among older adults as approximately one out of three people above 65 and one out of two older adults above 80 years old fall annually [3]. The consequent injuries due to falls can vary from a scratch to hip fractures and the hospital stays due to falls can be long and even last for the rest of the patient’s life. As a matter of fact, more than 50% of all injury-related hospitalizations among older adults is due to falls [44]. In addition, post-fall syndrome can affect the quality of life and cause further restrictions such as fear of falling, immobilization, depression, and loss of autonomy [33]. The injuries caused by falls can also be fatal. They account, indeed, for the largest percentage of the deaths caused by unintentional injuries, which are the seventh leading cause of death among older adults [12].

These damages can be reduced through practicing recommended preventive interventions. Several health organizations such as the joint American and British Geriatric Society (ABGS) [35] and the National Institute of Clinical Excellence (NICE) UK have reported a set of clinical guidelines for assessment and prevention of falls in older adults [19]. According to these guidelines, customized exercise programs for strength, balance, gait, and coordination training are effective in reducing falls [46]. They are recommended to be prescribed individually and monitored by a trained professional. In one study, for instance, home-based exercise programs were successful in reducing falls and fall injuries as much as 35-45% and equally effective for older women and men [41]. These programs included walking, muscle strengthening, and balance training, which were individually prescribed by physiotherapists (in the first two trials) and trained nurses (in the second two trials) during home visits. However, limited availability of the experts to supervise every single exercise session per individual introduces yet another restriction to such preventive interventions.

The system proposed in this paper is designed to support clinical experts in prescribing and monitoring fall preventive exercises. It provides individualized exercise programs for home-based balance training based on the NICE guidelines. It covers sitting, standing, and walking exercises each in three progression levels [36]. For the older adults, the system consists of a holographic virtual coach that supervises the training at home.

A holographic representation of a virtual coach displayed via an augmented reality (AR) head-mounted display (HMD)
walking and talking at once puts older adults at a higher risk of falling [59, 8]. However, simultaneous training of cognitive and physical abilities, in particular primary multi-modal exercise programs in combination with multi-sensory secondary tasks can improve cognitive and motor-cognitive dual task performance of older adults [22, 53]. Based on these evidences, the proposed system delivers, in addition, cognitive training and exercise games as well as an auditory training application for improving auditory attention and memory.

Further modules are motion capture and wearable sensors which enable monitoring of the correct performance of the exercises and an activity planning application for self-monitoring physical and cognitive activity performance and achieved goals. For the health care professionals, the system provides a dashboard which demonstrates the collected and analysed data from the older adults’ homes and enables them to adjust the training program per individual.

In order to provide a system with optimal usability, the human-centered design approach was followed in this work. Through an iterative development circle, the requirements were analysed, different versions of the prototypes were implemented and evaluated by older adults. After each iteration, new insights were evolved which led to re-prototyping of the concepts.

In this work, we pursued the answers to the following research questions:

- **Q1**: How acceptable and usable is our AR-based balance and cognitive training solution for older adults?
- **Q2**: How is the virtual coach perceived?

Our evaluations show that the answer to Q1 depends on their age: the usability for 65-79 is OK to Good and for 80+ is the worst imaginable [6]. And for Q2, the virtual coach appears socially present in the physical room like an alive and intelligent human. Further contributions are (i) conceptual description of the AR-based balance training system and (ii) explanation of iterative implementations during the human-centered development process.

**RELATED WORK**

Technology-based fall prevention interventions have been deployed in different contexts [26]. In the context of fall detection, for instance, Du et al. [15] proposed a robot control system for remote fall risk assessment in home environments. Silva et al. [47] presented a game to assess the quality of
the older adults’ locomotion using the data from accelerometers built in a smartphone. Another smartphone-based fall detection system was proposed by Abbate et al. [2]. Their system monitored the movements of older adults and in case of detection of a fall notified the caregivers. Computer vision techniques have also been used to automatically detect falls of older adults in their homes [63]. Furthermore, Ogonowski et al. [34] proposed a Kinect based interactive TV system for fall prediction and prevention. Their system made use of personalized physical exercise programs, gamification, and wearable sensors (senior mobility monitor device) and showed the feasibility of integrating such systems into the daily life of older adults.

There have been a few prior studies, which have explored if and how older adults accept such fall detection systems. For instance, Wu and Munteanu [61] involved older adults in the design and development of a fall risk assessment wearable device (in form of a belt). They also employed a field study to investigate the acceptance of the final device. Their findings suggested that the combination of contextual information for fall risk assessment and practical fall prevention instructions can improve the acceptance of such assistive technologies by older adults. Tyagi et al. [57] found out that the patient attributes can determine the adoption of technology-based interventions such as telerehabilitation (TR). In their observations, for instance, those patients who preferred TR were relatively younger than those who chose to go to a day rehabilitation center. Thus, they recommended to include introductory videos in TR programs and provide technical support for older patients.

In addition, Harte et. al. [27] suggested a three-phase human-centered design methodology to enhance the usability and UX of health systems for older adults. The first phase of their methodology contains the construction of a context of use document to report use cases, mock-ups, and user feedback. The second phase suggests an expert usability inspection, and the third phase emphasizes regular UX testings to improve the final prototype. In the end, they reported a successful implementation of their methodology for the design and development of a system for fall detection for older adults.

In the context of decreasing the intrinsic risk factors of fall prevention intervention, prior studies have made use of different types of technology to retain balance and improve functional abilities of older adults [28, 23, 55, 18, 13, 14, 25, 1, 1, 17, 40, 48, 49, 50, 51, 54, 58]. For instance, Hardy et al. [1] created an exercise game to encourage the older adults to perform balance and gain training.

Games can as well mitigate cognitive impairments of older adults [29, 38, 43]. For instance, Schoene et al. showed that playing a step game at home for a period of eight weeks can improve physical and cognitive abilities of older adults. Besides, playing video games in the context of a programmed activity can be valuable for older adults in residential care. It promotes positive self-views by reintroducing challenge and fun in late life for independent older adults. However, these benefits could be inhibited by the degree of age-related changes and impairments for vulnerable older adults [24].

In addition, to enhance the mobility of older adults, Felberbaum et al. [16] identified three main categories of technological requirements. The first category is based on social aspects, which provides social companionship. Co-walking with a virtual companion is one of the recommendations in this category which was mentioned by the interviewees and was seen as a way to overcome the feeling of loneliness and making the physical activity more enjoyable for the older adults. The second category of requirements refers to the feedback given to the users to encourage mobility (e.g. setting goals and target destinations). The third is (remote) monitoring (e.g. automatic fall detection) to ensure a safe and at the same time independent mobility. We incorporated these three categories in our AR-based virtual coach.

Virtual coaches have been employed to assist older adults in their everyday life [56], for example, to commit to their medications [37] or to stay physically active [10, 4]. Albaina et al. [4] proposed a virtual coach in form of an animated flower to motivate older adults to walk. The users were involved in the design and development process through two focus groups and one field study. Their results suggested that the elderly users enjoyed interacting with the flower virtual coach and liked to use it for a longer time. Although their virtual coach was not proven to be critical for motivating people, it was shown to improve acceptability of the system. A human-like virtual coach displayed on a tablet in combination with a pedometer could as well increase the amount of walking among older adults in a 2-months period [10]. The users also expressed a high level of satisfaction from the virtual coach who could in their opinion help them to walk more.

Moreover, AR technology has been used by older adults in previous studies [62, 42, 31]. For example, Bianco et al. [9] proposed a tablet-based AR application for fall preventive home modifications such as installing more handrails at desired locations at home. Also, training with a projection-based 3D AR game has been shown to be effective in improving mental rotation ability of older adults [30]. However, despite the potentials and advantages that AR can offer to older adults, it has not been used for balance training yet.

CONCEPT
Based on interviews and focus groups with our cooperation partners from the medical domain in the project, we defined the following five technical modules, required to provide a home-based AR balance training system:

- **Balance Physiotherapy Hologram (BPH)**, which is an AR virtual coach that presents instruction, demonstrates correct performance of the balance exercises, and gives feedback to the users about their performance (see Figure 1(a)).

- **Cognitive Training and Exercise Games (CTEG)** provides holographic cognitive training games as well as gamified balance exercises (so-called *exergames*). Figure 1(b) shows the first progression of the walking exergame in which the user is asked to follow the path on the ground and
walk towards the deer. Figure 1(c) shows one of the cognitive games. In this game, the user should remember the sequence of the cards and recall it after it has been shuffled.

- **Motion Capture and Wearable Sensors (MCWS)** to track the user’s body movements, for instance, by depth cameras or pressure sensors. A heart rate sensor in addition measures the users’ cardiac responses.

- **Auditory Training Tool (ATT)** aims at training speech in noise detection as well as auditory memory. Figure 1(d) shows the speech in noise detection task whose user interface shows animal pictures and colored numbers. Following an audio instruction (such as “Show the duck where the blue five is.”), the user should click on the correct colored number (e.g. blue 5 button). The instructions contain different background noises (e.g. cafeteria type noise) to increase the difficulty of this task, which adapts on every question asked and according to the user’s correct/correct answers.

- **Activity Planning App (APA)**, which allows the users to see their performance by means of charts, pictograms, and badges (e.g. when the goals are achieved) (see Figure 1(e)). Moreover, the app contains a calendar which shows the training program as well as a social networking feature to communicate with other users.

- **Dashboard Module (DM)**, which is exclusively designed to be used by the clinical experts, which diagnose the older adults and prescribe personalized exercise programs (see Figure 1(f)).

Figure 2 depicts the architecture of the entire system. It is composed of the home-located devices (MCWS), cloud storage, and processing infrastructure. The data collected from the sensors are fused and analysed on the Edge Computer, which provides real-time feedback to the user articulated by the virtual coach. The results are then sent to the cloud to be monitored by the therapist. Based on the user’s performance, the therapist can adapt the exercise program.

**METHOD**

In this work, we followed the human-centered design approach to understand the context of use, specify user requirements, produce mock-ups and prototypes, and evaluate them against the requirements through semi-structured focus groups and interviews. This process was iterated two times and across three countries.

**Iteration I**

The initial requirements were defined through interviews and focus groups with our cooperation clinical partners. Based on these requirements the first versions of the prototypes and mock-ups were developed. These prototypes were then presented to the older adults in London to familiarize them with our concept and to collect their initial requirements. The BPH was, in this phase, displayed via a Meta 2 head-mounted display (HMD) and the motions of the users were tracked by a Kinect 2 sensor. The CTEG, AAT, and APA were ready to present to the users just before the second focus group. By then, the CTEG and AAT were functional prototypes, but APA was still a paper mock-up.

**Focus Group with Older Adults I**

In our first study to collect feedback of the older adults, five seniors (3 female) with an average age of 74.8 years old (SD = 6.14, min=69, max=84, Mdn=72) took part in a focus group in
London. Three participants (60%) did not have a balance disorder and 80% (4 persons) did not have prior experience with HMDs. All five participants used the then available system, which consisted of a Meta 2 HMD. While using the system, the older adults saw five profiles of the virtual coach (see [32]). The different visual alternatives were identified together with the clinical experts of our project to explore commonly used representations in commercial exergames (male and female cartoon avatars), in AAA games (realistic male and female avatars) or in high-end telepresence applications (polygonal 3D reconstruction of the experimenter created using RGBD cameras of a Kinect 2). Each virtual coach demonstrated the correct performance of an exercise (in sitting, standing, and walking postures) to the participants. The participants then performed one standing exercise (i.e. standing and bending over as if to pick up an object from the floor) and received feedback from the virtual coach about their performance. Following that, they fill-out questionnaires and answered our questions in a semi-structure interview.

The first questionnaire was the System Usability Scale (SUS) [11] questionnaire. The average SUS score was calculated as 72.5, which is considered to be above average.

The second questionnaire was a modified version of the Co-Presence and Social Presence in Virtual Environments questionnaire [39], which has 15 statements to measure four sub-dimensions of social presence using five-point Likert scales (1=strongly disagree - 5=strongly agree). This questionnaire was originally developed to measure the impression of the virtual audience during a public speaking task in VR. We modified the questionnaire to evaluate the impression of the virtual coach in our scenario. All four factors of this questionnaires were rated above medium: 1) Participant’s Reaction to the virtual coach (Mdn=4, M=3.55, SD=1.57), 2) Perceived virtual coach’s Reaction (Mdn=4, M=3.6, SD=1.57), 3) Impression of Interaction Possibilities (Mdn=3, M=2.8, SD=1.51), and 4) (Co-)Presence of the virtual coach (Mdn=3, M=3, SD=1.77).

Moreover, the participants were asked about their virtual coach preference. 60% of the participants (two females and one male) preferred the realistic male virtual coach. The mean score for the naturalness of the virtual coach’s movement was 3.4 (on a scale of 1 (very bad) to 5 (very good)) and the mean score for the naturalness of the virtual coach speech was 3.8 (on a scale of 1 to 5 ranging from very bad to very good).

We also received several comments resulting in the following user preferences: Addition of volume control for the virtual coach voice, a lighter and smaller HMD, “make the virtual coach more responsive”, provide more feedback and more interaction, “less instruction from the coach with more dialogue and conversation”, “UK English accent would be preferred”, “should not be a generic experience, should be more adaptive on a person basis”, “virtual coach mood should look more optimistic and happy”, a progress feedback on a daily basis, and detection of safety by the system.

Focus Group with Older Adults II

Our second focus group in London (see Figure 3(c)) was intended to verify if the previous findings of the first focus group were representative. Here, the participants did not use the system and saw only a live demonstration as well as pictures and videos of the system (i.e. all main modules: BPH, MCWS, CTEG, ATT, and APA). 47 older adults (29 females) aged between 60 to 84 years (M=73.61, SD=6.12, Mdn=72.5) took part in this focus group. The participants could ask questions, discuss the system with the technical and clinical experts, which were present during the focus group, and finally express their opinions through questionnaires as well as direct conversations with technical, clinical and usability experts.

Before presenting any module, participants answered a demographic questionnaire. The results of this questionnaire show that 13 individuals (28%) had balance disorder, 14 (30%) persons were not sure and answered the question with maybe, and the rest with no (42%, 20 people). After that, participants were asked about the frequency of experiencing a fall (min=0, max=14, M=1.07, SD=2.27) or near fall (min=0, max=24, M=3.26, SD=4.83) during the past 6 months, given the definition of each type of falls to them: a fall is defined as an event that results in a person coming to rest inadvertently on the ground or floor or other lower level; whereas a near fall is an event in which a person feels a fall is imminent, but avoids it by compensatory action, such as grabbing a nearby object or controlling the fall. In addition, they stated that they are physically active on a daily (81%) or weekly (15%) basis and only 4% indicated that they never do physical exercises. Besides, half of the participants had ≥ 3m², one third 2m², and
the rest (17%) had \( \leq 1 m^2 \) free space at home for performing exercises.

The number of participants, which use computer systems (including smart phones, tablet PCs, etc.) were quite high: 94% daily, 4% weekly, and 2% never use any type of computers. Besides, on a scale of 1 (very bad) to 5 (very good) they rated their confidence in using technologies on average 3.23 (SD = .91, Mdn=3). Around 94% of the participants claimed that they do have broadband internet at home. We solicited information about their means of communication with their medical doctors. While telephone was still the most frequent tool for communication (89%), newer technologies are used as well such as Email (38%), mobile phone (37%), SMS/text message (11%), instant messengers (e.g., WhatsApp) (8%), and social media (6%). Furthermore, we received comments such as “through online appointment system”, “doctor does not have mobile phone (I do have an iPhone)”, “almost impossible to contact”, and “increasingly difficult to make an appointment”. Hence, we can assume that the social media panel of the APA could be of use for the older adults as our participants were social media users (e.g. 64% use Facebook, 36% Twitter, and 21% WhatsApp).

Moreover, 62% of the participants stated that they did not have any previous experience with 3D holograms. Two persons also commented that “[No but I] know what it is though”. Also, 63% had no previous experience with HMDs. They gave us two comments as “[No but I have] read about them for films” and “yes and no (I am aware of them)”. Thus, 37-38% of the participants did have prior experience with 3D holograms and HMDs.

After that we asked whether they have ever used a fitness wrist band (e.g. Fitbit). 64% of our participants answered this question with No. Two people added their comments as “[No but I] know what it is though”. Also, 63% had no previous experience with HMDs. They gave us two comments as “[No but I have] read about them for films” and “yes and no (I am aware of them)”. Thus, 37-38% of the participants did have prior experience with 3D holograms and HMDs.

Next, they were asked about their experience in using exergames (they were provided with an explanation about this term). 85% had no previous experience with such technologies. Three people gave us their comments as “[Yes but] I got bit confused using it (pre use: no instruction seen)” I guess if I had it explained I would have been [doing] as I thought it was very versatile and easier to body”, “[No but I have used DVD for exercises]”, and “[No just] briefly with grandchildren”. Their previous experience using video games for cognitive training were also limited as 76% of them had no previous experience with such games.

In order to estimate the usability aspects of each module, a modified version of the SUS questionnaire with five statements was repeated after each module’s demonstration. Since participants had to fill out 99 questions in 31 pages we reduced the SUS questionnaire to the five most important items of the SUS for our case and reported their mean score. The Assistance question consisted of multiple choice questions. Thus, for each part of the system we could calculate how many percent of the participants needed what type of assistance. As it can be seen in Table 1, our participants were positive about using all modules frequently if it would reduce their risk of falling. All modules were also perceived as easy to use, with an exception of APA, which was rated towards a neutral opinion. Indeed, all modules were more advanced in the development than the APA, which existed at that time only as a paper mock-up. The responses to the next statement: “I would imagine that most people would learn to use this system very quickly” were also rather neutral for all modules. On the other hand the participants stated that they would feel confident using any of the modules presented to them. An assistance to use the system was not seen as requirement, unless it is a technical assistance, and as some comments clarify that such technical assistance would be beneficial in particular in the beginning: “Technical assistance to start with” and “to confirm I am using it correctly”.

Furthermore, the participants gave feedback and comments about the virtual coach. The most frequent comment was about the computer generated voice: e.g. “speech needs to be quite clear syllables, well-articulated, as older people ‘decode’ fast speech rather badly quite often”. As a result, 61% of the participants wished for a change in the voice of the virtual coach towards a more “natural voice” or a “computerised speech, but it must follow natural English rhythm”. Some other characteristics of the virtual coach that participants would like to personalize were gender (48%), age (41%, e.g. “age 35+, looking ‘mature’/knowledgeable”), clothes (28%), and size (26%, e.g. “not stick thin”). Some participants in addition asked for a change of the ethnicity of the virtual coach. The repetition of the instruction and the exercises was as well the most desired interaction function.

In order to receive older adults’ preferences in respect to the CTEG prototype, a self-made questionnaire was developed. Since this prototype featured realistic 3D models of different types of animals to be augmented in the real environment, i.e., older adults’ homes, questions about animals were the main theme of this questionnaire. Based on its results, 68% of the participants liked animals in general and 60% of them have had at least once a pet in their life. Household as well as forest animals were their most favourite types. Mouses/rats, snakes, and desert animals were on the other hand the least favourite ones.

The AAT prototype consisted of two tasks: 1) speech in noise detection task, which contained pictures of animals on its user interface and 2) auditory memory task, which played a short audio story to be listened carefully (to be able to answer the questions about the story afterwards). Again, to better understand the users’ preferences a self-made questionnaire was employed. The pictures of the first task’s interface were preferred to be means of transportation (55%), animals (36%), musical instruments (38%), fruits, and “sport, anti-capitalist or anti-racist leader/heroes”. The genre of the audio stories to be played in the second task was chosen to be comedy (60%), suspenseful (43%), crime, and “factual information”. The last two question of the AAT questionnaire focused on the length of the training session. Based on the responses, most of the participants stated that a 10-15 minutes auditory training per session for 3-4 days a week could be easily followed by them.
on a regular basis. For instance, a person commented “daily duties vary, most people fit it to their daily routines”. They (58%) disliked longer sessions on fewer days. In the end, we received quite a number of general comments and feedback summarized in the following paragraphs:

- “From what is shown very early days. I think the idea is good and could reach some people. The motivation issue is probably the most difficult thing to solve. This could be very interesting.”

- “I would like this to be extended to a group use eventually. I think this is a great preventative idea.”

- “I would be keen to have headset + sensors as compact + neat as possible, to take up least space. I would like the system to be easy to clean + maintain ([pressure sensing] socks do not appeal unless washable).”

- “This should not be a substitute for face-to-face contact and group advises, it might help those people who are isolated, but we are social beings and this should be encouraged.”

- “Inspirational, brilliant particularly as falls clinics are rare in our regions. You may fall, but not be referred even if you experience multiple falls.”

- “Interesting, early stages of develop, but suggest testing out in reality in each stage to ensure connectivity from one stage to the next.”

**Iteration II**

In the second iteration, the requirements and user preferences gathered from the first two focus groups were used to further iterate the prototypes. The most emphasized requirement was a lighter, more affordable, and cableless AR-based HMD. Thus, in the second iteration, the BPH and CTEG were optimized for Pixel 3 XL smartphone, which can be integrated in the lightweight Docooler AR Headset.

Moreover, in order to address the accessibility of our concept for the older adults in different locations, the third focus group was held in Freiburg (Germany) (see Figure 3 (d and e)) and the fourth in Athens (Greece) (see Figure 3 (f and h)). For the purpose of these focus groups, all prototypes and questionnaires were localized (i.e. prepared in both German and Greek languages). Since the participants of the focus groups in London, wished for a natural human voice in the prototypes, in the second iteration, the voices of a native German and a native Greek speaker were recorded and integrated in the prototypes. All participants used the system for about 15 min supervised by at least a clinical and a technical expert before the interviews started.

**Focus Group with Older Adults III & IV**

The participants of both focus groups wore Docooler AR HMD and used BPH and CTEG prototypes, which were running on the Pixel 3 XL smartphone. ATT and APA do not require the headset, and therefore were displayed on the smartphone alone. The participants evaluated these prototypes using several standard questionnaires and a few open-ended questions in the end. We also collected their feedback during the focus group.

We interviewed in total 24 older adults (21 female) between 60 and 84 years old (M=71.42, Mdn=72, SD=7.99). Eight of them were interviewed in Freiburg and the rest in Athens. Half of the participants had balance disorder and reported an average of 0.56 (SD=0.82) falls and 0.85 (SD=1.32) near falls during the past 6 months. Seven responses to the near fall question were not specific, and therefore were not included in the analysis. These responses were: “Daily”, “2-3 per week”(n=1), “several”(n=1), “only dizziness”(n=1), “a few”(n=1), and 1 empty response.

Our participants were also physically active since more than 70% of them reported daily (37.5%) or weekly (33.33%) performance of exercises. The rest (25%) stated that they never do exercises. To perform exercises, about 70% of the participants indicated that they have 3m² or more free space at home. 13.04% of them had 2m² and the rest (17.39%) had 1m² available space at home. Moreover, about 2/3 of these seniors had broadband internet at home and used computer (including smart phones, tablet PCs, etc.) on a daily (58.33%) or weekly (4.16%) basis. On a scale of 1 (very bad) to 5 (very good) they rated their confidence in using technology on average 2.33 (SD=1.34, Mdn=2.5). Furthermore, the majority of

---

**Table 1. Means (standard deviations), medians, and percentage of the responses to a modified version of the SUS questionnaire during the Focus Group II. Its items are: Frequency of use: “I would like to use this system frequently if it meant that it would reduce my risk of falling”, Ease of use: “I think this system would be easy to use”, Learn to use: “I would imagine that most people would learn to use this system very quickly”, Confidence to use: “I would feel very confident using this system”, Need for Assistance: “I think that I would need assistance to be able to use this system. If yes, what type?”. All items (with an exception of the Need for Assistance) were answered on a scale of 1 (strongly disagree) to 5 (strongly agree). The Need for Assistance item was a multiple choice question.**

<table>
<thead>
<tr>
<th>Item</th>
<th>BPH (M, SD)</th>
<th>Mdn</th>
<th>MCWS (M, SD)</th>
<th>Mdn</th>
<th>CTEG (M, SD)</th>
<th>Mdn</th>
<th>AAT (M, SD)</th>
<th>Mdn</th>
<th>APA (M, SD)</th>
<th>Mdn</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequency of use</td>
<td>4.15 (.102)</td>
<td>4</td>
<td>3.81 (.31)</td>
<td>4</td>
<td>3.57 (.13)</td>
<td>4</td>
<td>3.74 (.115)</td>
<td>4</td>
<td>3.75 (.128)</td>
<td>4</td>
</tr>
<tr>
<td>Ease of use</td>
<td>3.69 (.99)</td>
<td>4</td>
<td>3.6 (.07)</td>
<td>4</td>
<td>3.64 (.19)</td>
<td>4</td>
<td>3.64 (.16)</td>
<td>4</td>
<td>3.46 (.32)</td>
<td>3</td>
</tr>
<tr>
<td>Learn to use</td>
<td>3.04 (.03)</td>
<td>3</td>
<td>3.05 (.11)</td>
<td>3</td>
<td>3.32 (.22)</td>
<td>3</td>
<td>3.38 (.08)</td>
<td>3</td>
<td>3.53 (.98)</td>
<td>3</td>
</tr>
<tr>
<td>Confidence to use</td>
<td>3.67 (.87)</td>
<td>4</td>
<td>3.57 (.21)</td>
<td>4</td>
<td>3.77 (.17)</td>
<td>4</td>
<td>3.6 (.17)</td>
<td>4</td>
<td>3.73 (.12)</td>
<td>4</td>
</tr>
<tr>
<td>Need for Assistance?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>What type?</td>
<td>%</td>
<td>%</td>
<td>%</td>
<td>%</td>
<td>%</td>
<td>%</td>
<td>%</td>
<td>%</td>
<td>%</td>
<td>%</td>
</tr>
<tr>
<td>No</td>
<td>41.3</td>
<td>34.9</td>
<td>48.8</td>
<td>43.9</td>
<td>30.6</td>
<td>5.6</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Physical</td>
<td>10.9</td>
<td>9.3</td>
<td>14</td>
<td>9.8</td>
<td>5.6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Technical</td>
<td>39.1</td>
<td>44.2</td>
<td>25.6</td>
<td>34.1</td>
<td>50</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Educational</td>
<td>17.4</td>
<td>18.6</td>
<td>11.6</td>
<td>7.3</td>
<td>16.7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medical</td>
<td>6.7</td>
<td>7</td>
<td>4.7</td>
<td>7.3</td>
<td>2.8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
the participants (91.66%) did not have prior experience with AR or VR.

In order to calculate the SUS score we had to remove four responses, which were not fully available. The mean SUS score for the remaining 20 responses was calculated as 57 (min=10, max=92, Mdn=60, SD=24.97). Based on the Shapiro-Wilk test, the SUS scores were normally distributed. Therefore, we ran Pearson correlation, which suggested a significant negative correlation between the subject’s age and their SUS score (cor=-0.628, df = 18, p<.01). In other words, as the age of the participants increased, their evaluation of the usability of the system decreased. Figure 4 depicts the SUS scores per age groups. The results of the ANOVA suggested a significant difference between the age groups (F(4,15)=5.57, p<.01, $\eta^2_G=.6$). The pairwise comparisons with Bonferroni corrections showed that the SUS score of the participants, which were over 80 years old, was significantly lower than the participants aged 60-64 (p<.01) and 70-74 (p=.05) years.

As it can be seen in Figure 5, the results of the NASA-Task Load Index (NASA-TLX) indicates that despite wearing an HMD, the system was not rated as physically demanding. The effort and frustration levels were also rated low. Although the mental and temporal demand was rated higher than other factors, the values are still around the medium score. Therefore, the overall workload during the testing can be considered as low to medium.

The results of the User Experience Questionnaire (UEQ) are shown in Figure 7. All scales have a mean value above 0.8, which corresponds to a positive evaluation. In particular, the system was perceived as novel, stimulating, and attractive.

The Godspeed questionnaire [7] was employed to evaluate the representation of the virtual coach in the BPH module. As it can be seen in Figure 6, different aspects of the virtual coach are rated quite high. Therefore, the participants have perceived the virtual coach like a human (Anthropomorphism), who is alive (Animacy) and calm (Perceived Safety), has intelligence (Perceived Intelligence), and has left a positive impression on them (Likeability).

In addition, the mean Social Presence [5] score (on a scale of -3 to +3) was 0.26 (SD=1.19). The positive value suggests that the participants perceived the virtual coach as conscious and aware.

**DISCUSSION**

In order to have a successful technology-based intervention for fall prevention, it is essential to understand the older adults’ view on such programs. Through several focus groups across three countries, a large number of requirements were defined, which helped us to design and develop working prototypes for multiple technical modules of our concept.

First, it is essential to be aware of the individual differences between the older adults. The 20 years difference between the age of 60 and 80 makes a significant effect on older adults’ use of technology. Our results confirms that the use of our system for participants in their 60s was significantly easier than those over 80 years of age. Indeed, we observed that some 80+ participants did neither have internet at home, nor mobile phones or computers, and as a result had difficulty interacting via the smartphone’s touch screen.

Next, one needs to acknowledge the older adults’ interest in technology and its application in their everyday life. For instance, the older adults living in London we had the chance to introduce our concept to, were not completely unfamiliar with 3D holographic displays, AR/VR HMDs, or social media. And as a result they were very positive and supportive about the project.

Our findings from the focus groups in Freiburg and Athens suggest that the overall impression of the product was positively evaluated by the older adults. They also showed that the given tasks of the prototypes could be solved without unnecessary effort and they felt in control of the interaction. The system was as well very exciting, creative, and motivating for them.

Furthermore, the results of the social presence and Godspeed questionnaire showed that the older adults perceived the virtual coach not as a computer-generated image, but rather a conscious and alive person, who is present in the room and observes them. Moreover, they found the virtual coach intelligent, friendly, and safe to interact with.
CONCLUSION

In this paper, we presented a first investigation of the acceptance of an individualized virtual coaching system based on AR for balance training.

The proposed system is intended to support balance training at home without the requirement of the physical presence of human physiotherapeutic or medical experts. In a human-centered design approach we developed and evaluated several mock-ups and prototypes with more than 70 older adults across three countries.

In general, the results suggest that the participants found the system encouraging and stimulating. Furthermore, the virtual coach was perceived like an alive, calm, intelligent, and friendly human and has therefore enormous potential to enhance social facilitation, which is important for motivation. However, the usability of the AR system showed a significant negative correlation with the participants’ age. To summarize, the results give important implications for the use of AR in the overall aging population.

The overall positive feedback of the participants could be caused by a novelty effect, which may have led to the tendency for performance to initially improve when new technology is instituted. Hence, feedback of the participants should be collected during long-term usage of our system. In this context, we plan to validate the effectiveness of this novel rehabilitation and training concept through a randomized control trial on each three locations, which hosted the above described focus groups (i.e. London, Freiburg, and Athens).

ACKNOWLEDGMENTS

We wish to thank all our participants and project partners. This work was funded by the European Union’s Horizon 2020 research and innovation program under grant agreement No 769574 as well as German Research Foundation (DFG), the German Federal Ministry of Education and Research (BMBF) and Federal Ministry for Economic Affairs and Energy (BMWi).

REFERENCES


[10] Timothy W Bickmore, Rebecca A Silliman, Kerrie Nelson, Debbie M Cheng, Michael Winter, Lori Henault,


