Virtual Reality manuscript No.
(will be inserted by the editor)

Evaluation of the impact of different levels of self-representation and body tracking on the sense of presence and embodiment in immersive VR

Guilherme Gonçalves · Miguel Melo · Luís Barbosa · José Vasconcelos-Raposo · Maximino Bessa

Accepted Version. Changes were made to this version by the publisher prior to publication. The final version of record is available at https://doi.org/10.1007/s10055-021-00530-5

Abstract Despite the majority of Virtual Reality (VR) applications not showing the user's avatar nor its movements, this self-representation has been shown to make the difference in the user experience and, consequently, in the potential of VR applications. However, these solutions are often costly and complex. With the widespread of VR in the consumer market, low-cost approaches are beginning to appear, and it is of utmost importance to propose and evaluate these low-cost approaches to enable consumers to get the most out of VR.

The main goal of this paper is to investigate the effect of having different levels of control over the self-representation (defined by the number of tracking points available) and how the self-representation is done (full body avatar or floating members) on the sense of presence and embodiment using low-cost consumer technology. We compare different ways of self-representations through floating members (hands vs. hands+feet) and by using a body (floating members vs. full-body avatar), walking fidelity (static feet, simulated

G. Gonçalves
Universidade de Trás-os-Montes e Alto Douro, Vila Real Portugal
INESC TEC, Porto, Portugal
E-mail: guilhermeg@utad.pt

M. Melo
INESC TEC, Porto, Portugal

L. Barbosa
Universidade de Trás-os-Montes e Alto Douro, Vila Real Portugal
INESC TEC, Porto, Portugal

J. Vasoncelo-Raposo
Universidade de Trás-os-Montes e Alto Douro, Vila Real Portugal
INESC TEC, Porto, Portugal

M. Bessa
Universidade de Trás-os-Montes e Alto Douro, Vila Real Portugal
INESC TEC, Porto, Portugal

Accepted Version. Changes were made to this version by the publisher prior to publication.
The final version of record is available at https://doi.org/10.1007/s10055-021-00530-5

walking, real walking), and a number of tracking points used (head+hands, head+hands+feet, head+hands+feet+waist).

Results showed that: floating hands alone were better than the combination of hands and feet; simulated walking outperformed static feet and real walking; the addition of waist tracking to feet, hand and head tracking significantly improved embodiment scores; there were no differences between using floating members and full-body avatar.

Keywords Virtual Reality · Embodiment · Presence · Body Tracking · Body Representation

1 Introduction

The fast-paced evolution of Virtual Reality (VR) technologies has contributed for its affordability, and consequently, to the widespread of immersive VR systems based on head-mounted displays (HMD) which bring the first-person point of view experience to a higher level. The main goal of a VR application can be summed up to the transition from the real world to a virtual one where users can interact as if they were there. This sense of "being there" is often referred to as "presence" and it has been widely discussed in the literature (please refer to Skarbez et al. (2017) for a comprehensive overview of presence theorising). In this paper, we adopt Slater and Wilbur (1997) definition, which states that presence depends on the capacity of the VR application to create an illusion on the user to a level that makes them feel like actors in a VE rather than operating on it from outside.

Following the rationale of developing a sense of presence when experiencing VR, the cognitive process associated with "being there" triggers a mental model of the virtual space where the location of the user's own body is inferred as also being in that same virtual space Biocca (1997). This sense of having a virtual body and be able to control it inside the VR context is defined in the literature as the sense of embodiment (hereinafter referred to as embodiment) Kilteni et al. (2012). In this work we consider a self-representation through avatar when the user is fully represented by a virtual humanoid body. Based on a review over embodiment studies, Gonzalez-Franco and Peck (2018) identified six main components that elicit the embodiment illusion: body ownership (the extent to which users feel the avatar as their own body), agency (the degree to which users can move the parts or all of the virtual body), tactile sensations (the extent that tactile/haptic stimulation enhance the embodiment illusion), location of the body (the extent the virtual body is perceived as being at the same location as the real body), external appearance (how much the avatar ressembles the individual), and response to external stimuli (the extent the individual feels their own body is threatened when facing dangerous situations).

Self-representation triggers a feeling of self-identification with the virtual body on users Slater and Wilbur (1997); Biocca (2014). The literature discusses that there is a relationship between embodiment and presence in a VR context, suggesting that having a self-representation in a VE is a key contributor to developing the sense of presence Slater et al. (2010). A well-known experimental study that shows the relationship between self-representation and presence is Slater et al. (1995) where authors have evaluated the walking metaphor (pointing vs. walk-in-place) using presence and, as an explanatory variable, the association with the virtual body. The pointing metaphor was achieved using a 3D mouse. The experiment consisted of having participants walking on a virtual room to pick various objects and place them on a open travel bag. Results have revealed that presence was increased with the more natural walking metaphor (walk-in-place). There was also a correlation between the participant's association with the virtual body and the reported sense of presence. Usoh et al. (1999) have further extended this study by adding real-walking using optical tracking, improving the visual fidelity of the models, and by complementing the presence measurement with oral debriefing, thus gathering feedback about the highlights of the experience, and breaks in presence (the moment when users "wake up" from the illusion of being in the virtual world (Slater and Steed, 2000)). The study corroborated Slater's Slater and Wilbur (1997) findings that showed a strong correlation between presence and the participant's degree of association with the virtual body. Another observation highlighted by the authors was that real-walking has revealed to be a compelling virtual experience where participants have expressed themselves with positive reactions.

The impact of using self-representation in VE is not limited to presence, as shown by Steed et al. (2016) which studied how having an avatar would impact the cognitive load in immersive VR. In that study, the authors asked the participants to memorise a pair of letters, perform a spatial rotation exercise (users had to rotate a 3D object mentally), and then recall the pair of letters. The condition with self-avatar had two additional sub-conditions that consisted of representing hand gestures or not. The self-avatar hand gestures were achieved by using hand-tracking. Moreover, inverse kinematics (i. e. estimating the movements of the body parts between the tracking points) based on the hands' positions was used to represent the movements of the arms. Results have revealed that participants who had an avatar and were also allowed to move their hands had significantly higher letter pair recall. Thus, the authors concluded that an active self-representation by avatar could alleviate the mental load of doing spatial rotation exercises, improving letter recall.

Having a self-representation can also bring benefits in other mental processes such as distance perception, which that can influence decision-making. In Ries et al. (2008), authors have investigated the effect of self-representation on distance perception under two conditions: with avatar and without avatar. The avatar was realistically represented by using a full-body optical tracking apparatus where participants could view their movements properly replicated in the VE. The study consisted of placing participants in a virtual hallway and asking them to perform twenty blind walks between two virtual markers. Results have revealed that the distance estimations were significantly better in the condition with an avatar. McManus McManus et al. (2011) went beyond

Accepted Version. Changes were made to this version by the publisher prior to publication.
The final version of record is available at https://doi.org/10.1007/s10055-021-00530-5

the effect of self-embodiment on distance perception by also considering object interaction and locomotion. In addition to the avatar representation, the study added to the equation an avatar's pre-recorded animation. Results have revealed that there were benefits on task performance when adopting both an avatar and character animations.

More than a simple virtual body, self-representation realism can stimulate both presence and embodiment and take immersive VR applications to another level as it can lead to improvements on different variables such as task performance, distance perception, decision-making or virtual social interaction as described next. Lugrin Lugrin et al. (2015b) studied the impact of self-representation realism (iconic *vs.* realistic representation) on presence and user experience and the performance under different VEs (stressful *vs.* stress-free) and different levels of visual realism of the scene (low polygon models and plain colours as materials *vs.* high polygon models and photo-realistic texturing). Results have revealed that presence was higher under the realistic representation condition. In another study, LugrinLugrin et al. (2015a) evaluated the impact of avatar realism in virtual body ownership in a virtual fitness training context by using no avatar (two spheres), non-realistic avatar (a giant warrior, a humanoid robot and a simplified humanoid) and realistic avatar (male and female human adults). Participants were asked to touch the maximum number of cubes that appeared consecutively with their virtual hands in front of a virtual mirror that replicated their movements through Microsoft Kinect. Results have revealed that task performance increased towards the avatars with a higher degree of resemblance with participants.

The impact of self-representation realism on embodiment and task performance was further addressed by Argelaguet et al. (2016). In their study, they investigated the impact of virtual hand representation (abstract - represented by spheres vs iconic - robotic hand vs realistic) on embodiment (namely agency and body ownership) and task performance. The experimental study consisted of participants having to perform a pick-and-place task while avoiding obstacles as well as to perform a potentially dangerous operation by placing their virtual hands close to a virtual spinning saw. After each of the 33 participants performed three trials, authors verified that the reported agency was similar between the three conditions while for ownership the realistic hand condition revealed higher ratings. Despite the results, the authors recognised that the agency results might have been affected by tracking issues since participants reported that when performing the tasks their hands were not precisely tracked. As for task performance, results have revealed that, in the first trial, the task performance was better on the realistic hand condition but at the end of the third trial the performance was similar.

Lin et al. (2013) conducted a study to evaluate the impact of having a matched avatar (gender and height) or no avatar in the decision-making process based on distance judgement. In the experimental study, participants were guided into an edge of a ledge and asked if they were able to step off the ledge comfortably and without losing balance. Results have revealed that using a matched avatar has played a positive role in the judgement of the situation.

Accepted Version. Changes were made to this version by the publisher prior to publication.
The final version of record is available at https://doi.org/10.1007/s10055-021-00530-5

The impact of avatar realism was also studied by Latoschik et al. (2017). An experimental study was conducted where participants had to go through two phases: in the first phase participants were asked to step in front of a virtual mirror and asked to perform a set of movements that were depicted in the self-representation and reflected in the virtual mirror to induce the illusion of body ownership. In the second phase, participants were asked to go to a specific spot in the VE where they could see a virtual counterpart waving. Participants were asked to repeat that same behaviour. Full-body tracking was achieved using optical body tracking. The independent variable was composed of four levels resulting from the combination of the gender-matched self-representation and the realism of the virtual counterpart (wooden mannequin or realistic avatar). Realistic self-representation increased the sense of virtual body ownership and the appearance of a realistic counterpart had a positive effect on the self-perception of the virtual body.

Due to the high plasticity of the human brain, the use of self-representation can not only increase the self-identification but also shape how the user would drive his behaviour and promote attitudinal or and behavioural changes (Yee and Bailenson, 2007). An example of such is the study by Van Der Hoort et al. (2011) that embodied users in a small virtual body and where it was verified that they overestimate the size of objects. Another example is when embodying users in a child's body, users have revealed to have more childish attitudes (Banakou et al., 2013; Tajadura-Jiménez et al., 2017).

More than attitudes, self-representation can also have an impact on cognitive performance: Banakou et al. (2018) conducted a study where participants were embodied with a virtual body strongly associated with high performing cognitive abilities (depicted by representing an Einstein-alike avatar) or a normal virtual body. Results have revealed that when the embodied virtual body was the Einstein body, participants performed better on cognitive tasks. However, there are also studies which shows, depending on the context, how the use of self-representation can bring no significant benefits. For instance, in a gaming context, Lugrin et al. (2018) has studied the impact of different levels of self-representation on virtual body ownership, game experience, game performance, and physical engagement. Using an action-based game, the self-representation was divided into three levels: none (no avatar body parts), low (only hands and forearms), and medium (head, neck, trunk, forearms, and hands). The self-representation was based on a genius figure (the lower body was represented as a tail) and the tracking of head and hands was achieved with the remotes of the HTC VIVE VR setup. Results have revealed that, unlike in other VR contexts, increasing the level of self-representation did not have a positive impact on the studied dependent variables. The authors point this result to the fact that in action-based games the enemy awareness and control efficiency creates an intense flow state (refer to Nakamura and Csikszentmihalyi (2009) for a description of flow theory) that reduces self-awareness, and consequently, the impact of self-representation details. Nevertheless, authors also point that lower body was not included in the study and further investigation is required to extend these findings.

Accepted Version. Changes were made to this version by the publisher prior to publication.
The final version of record is available at https://doi.org/10.1007/s10055-021-00530-5

Although a lot of studies have addressed the impact of different variables associated with self-representation on user experience, the technological factor is still a bit disregarded, especially when the goal is to represent full-body movement in the VE. For achieving full-body tracking, as verified above, most of the systems rely on optical tracking which implies costs that are not reasonable for consumers. However, with the widespread of immersive VR systems, it becomes important to ensure that there are alternative solutions that could allow taking the most of VR technology at a consumer level on a variety of application fields.

One proposal in this direction was put forward by Park and Jang (2019) which has worked on the problems associated with the walk-in-place technique, where walking in the VE is controlled by the user's legs but that there is no translation of those movements to their self-representation on the VE. This lack of information originates a contradiction between the sensory data and proprioception as there is no self-representation or the self-representation is not synced with the user's movements in the real world (e.g. left leg is moved in the real world while in the VE this movement is associated to the right leg). To overcome such, authors have proposed a method that makes use of inertial sensors to enable full-body tracking and translate the real movements to the avatar representation in the VE. A preliminary user study has revealed that the proposed method increased the sense of presence.

Another contribution for achieving full-body tracking accessible to consumers was made by Kilimann et al. (2018) that proposed a method that allows the full-body tracking in VR applications using consumer equipment. The method requires 6-DOF position and orientation trackers and the implementation example makes use of an HTC VIVE system where the HMD is used for the head tracking, the two controllers are used to track the hands and three additional VIVE trackers used to track each ankle and the lower back. The method processes the tracking data jointly with inverse kinematics to represent full-body movements. The method also encompasses a calibration step that allows correcting the lengths of the joints as the humanoid rig has different proportions from the user.

Although there are already low-cost proposals for full-body mimic in VR, the evaluation work is lacking. The main goal of this paper is to investigate the effect of having different levels of self-representation on the sense of presence and embodiment using low-cost consumer technology by conducting a user study. The different levels of self-representation vary in terms of visual self-representation and movements replication. An additional contribution to the literature is the proposal and evaluation of a method that, by taking advantage of inverse kinematics, allows to deliver a self-representation with full-body movement representation to the VE using only the typical VR setup based on the HMD and a pair of controllers. The outcomes of this work will not only provide a more detailed knowledge over the impact different low-cost self-representation methods that allow body movement simulation, but also a new method that simulates virtual body movement by simplifying the apparatus. Such contribution to the literature will enable to make more informed decisions

Accepted Version. Changes were made to this version by the publisher prior to publication.
The final version of record is available at https://doi.org/10.1007/s10055-021-00530-5

when designing or using VR applications where the sense of presence and embodiment play a role. Consequently, it will allow developing more effective VR applications for a wide area of application fields where VR is successfully being adopted, such as motor rehabilitation, exposure therapy, training or education.

2 Methodologies

This paper has a two-fold contribution: evaluate different sets of already existing self-representation approaches (varying both visual representation and movements replication) as well as to present and include into the evaluation, a new approach for movement replication in VEs. The evaluation is based on an experimental cross-sectional study of comparative character adopting a between-group design.

2.1 Variables

A total of six groups were created to conduct the experiment using a between-group design. Each group is as follows:

- **No avatar - Hands (H)**: Users can see their hands floating.
- **No avatar - Hands and feet (HF)**: Same as H but users can also see their feet, represented using floating boots.
- **Three points (3P)**: Users are self-represented trough a full-body avatar. Head and hands are controlled by the users, feet are static, and knees can bend when users crouch.
- **Three points with simulated walking (3PS)**: Same as 3P, but the feet movement is done through animations to simulate walking.
- **Five points (5P)**: Same as 3P, but users also have control over the avatar's feet.
- **Six points (6P)**: Same as 5P, but users can also control the avatar's waist based on the torso movement.

The study was split into several studies, each one focusing on an individual independent variable (IV). A description follows for each IV:

1. **Floating members**: Number of floating members visually represented, with two levels: hands (H) *vs.* hands and feet (HF).
2. **Walking fidelity**: The fidelity of the walking animation while using an avatar with three levels: no walking (3P) *vs.* simulated walking (3PS) *vs.* real walking (5P).
3. **Tracking points**: The number of tracking points used to control the avatar, with four levels: three points (3P), five points (5P) and six points (6P).

Accepted Version. Changes were made to this version by the publisher prior to publication.
The final version of record is available at https://doi.org/10.1007/s10055-021-00530-5

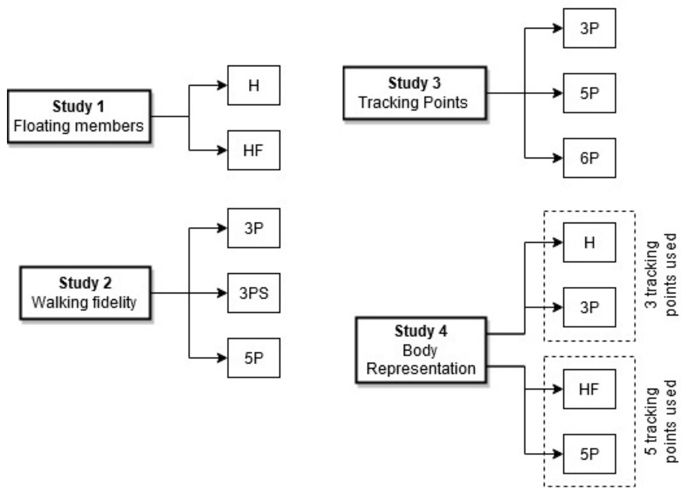

Fig. 1 Study layout depicting which conditions are analysed in each of the four studies.

4. **Body representation**: How the user is self-represented in the virtual world, by using the same number of tracking points. There are two comparisons: floating members (HF) *vs.* five point tracking with avatar (5P), both using 5 point tracking, and floating hands (H) *vs.* three point tracking avatar (3P), both using 3 point tracking.

Please refer to Figure 1 for a better understanding of how the studies are organized.

In the first study (1), we investigate the user representation using floating members. We analyse whether the addition of feet alongside hands can represent a benefit for presence and embodiment over hands alone.

In the second study (2), we focus on the walking fidelity. Following a conventional VR setup, normally only the hands and head are tracked (3P condition). Because feet are not tracked, both stay static, sliding over the floor when users walk. In this study, we included a condition which simulates walking through pre-recorded animations while only using hands and head tracking (3PS). In order to also compare the simulated walking with real walking, we also included a condition that features tracking of the users' feet in addition to hands and head (5P). This study will allow us to understand if simulated walking (3PS) is better than no walking (3P), and if it is worst than using real feet movement (5P). In order to better isolate the variable being studied, 6P condition was not included in this study because it also adds waist control.

The third study (3) aims to analyse the level of control the user has on the full-body avatar. The level of control is represented through the number of tracking points. The more the tracking points, the more control the users have over their avatar, resulting in more realistic movements. Therefore we compared three (3P), five (5P), and six (6P) tracking points. These conditions

Accepted Version. Changes were made to this version by the publisher prior to publication.
The final version of record is available at https://doi.org/10.1007/s10055-021-00530-5

represent the control of *head + hands*, *head + hands + feet*, and *head + hands + feet + waist*, respectively.

The fourth study (4) investigates weather is better to have an avatar over floating members. Such is done by comparing two pairs of conditions. In each pair of conditions, the number of trackers is equal (same control over the self-representation), changing only how the users are represented. For a situation where only three trackers are available, we compare the condition H against its equivalent 3P. In situations where five trackers are being used, we compare the HF against its equivalent 5P.

A more detailed description of how the avatar is animated through the use of inverse kinematics can be found in section 2.4. The considered dependent variables are:

– **Sense of presence** and its subscales (spatial presence, involvement, experienced realism).
– **Embodiment** and its subscales (ownership, agency, tactile sensations, location, appearance and response).

2.2 Sample

The sampling technique was based on a non-probabilistic method, being the sample composed by 98 volunteers (53 male and 39 female) aged between 16 and 49 years ($M = 22.970$, $S.D. = 4.855$). All participants reported normal or corrected-to-normal vision. Participants were randomly distributed between the different experimental conditions (Table 1).

Table 1 Distribution of the participants per condition

	H	HF	3P	3PS	5P	6P
Participants	15	16	18	18	16	16

2.3 Materials

A computer with NVIDIA GeForce GTX 1080 graphics card and Intel Core i7-7700K processor was used to run the experiences. A good application performance (90fps) was achieved with the described hardware. Besides, this computer was also responsible for managing all virtual reality equipment such as helmet, controls, cameras, headphones and sensors. The HTC VIVE HMD was used to allow the user to see the virtual environment and its controllers allowed the user to interact with virtual objects. Also, both of them had built-in sensors that were used to track the user's head and hands, respectively. Also, the HTC VIVE trackers were used, depending on the experimental condition, to track the user's waist and feet using the built-in sensors. Active noise cancelling headphones were also used to remove additional noises and to deliver

Accepted Version. Changes were made to this version by the publisher prior to publication.
The final version of record is available at https://doi.org/10.1007/s10055-021-00530-5

Fig. 2 The HTC vive HMD, controllers and trackers

the sound of the virtual environment. In Figure 2, it is possible to observe the equipment the participants used.

Two virtual environments were developed using the Unity® game engine. The purpose of the first environment was to calibrate the trackers and the IK system. Right after, the users would enter the second environment where they had to complete a task. As can be seen in Figure 3, the two scenarios are similar. The first environment lacks some objects like the bed and suitcase to make space for the calibration process. The environments were created to resemble a wood cabin. Some elements were introduced that may cause danger to the player, such as the fireplace and objects that fall from the shelves when they are opened. There are several cabinets and closets scattered around the cabin in which objects are hidden. A mirror is also present in both environments so users can see themselves self-represented.

2.4 Tracking System / IK System

The IK algorithm used was the "SAFullBodyIK". The algorithm parameters were fine-tuned to provide smoother and more correct movements.

Because there are participants with different heights, a resize would be done to the avatar depending on the HMD height. Two avatars were created,

Accepted Version. Changes were made to this version by the publisher prior to publication.
The final version of record is available at https://doi.org/10.1007/s10055-021-00530-5

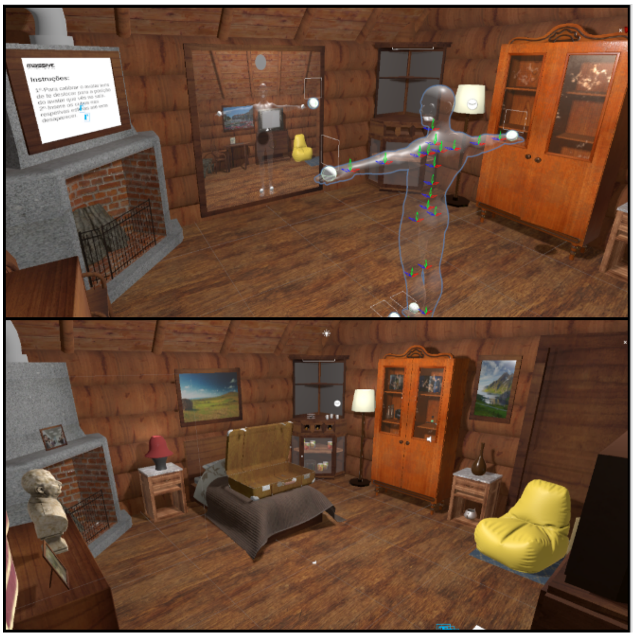

Fig. 3 Top: Calibration environment. Bottom: Task environment

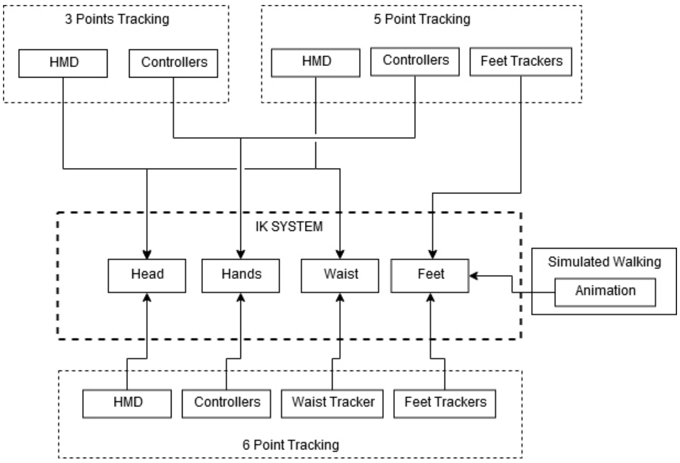

Fig. 4 Diagram describing the IK System input.

one male and one female (Figure 5) to allow better compatibility between male and female participants.

A diagram describing the origin of the data fed to the IK system can be found in Figure 4.

Accepted Version. Changes were made to this version by the publisher prior to publication.
The final version of record is available at https://doi.org/10.1007/s10055-021-00530-5

Fig. 5 Male and Female avatars used depending on the gender of the user.

2.4.1 6 point tracking

The 6P is considered the best condition. With the tracking data of the head, hands, waist and feet as inputs of the IK system, avatar movements are more realistic and lifelike. Because there is a waist tracker, the avatar orientation, as well as position, is calculated through that same tracker.

2.4.2 5 point tracking

In the condition 5P, the head, hands, and feet are tracked, but the waist is not. The avatar waist is the "center of mass" of the avatar, and controls the avatar's position and orientation. As such, the head orientation and position were used to guide the avatar's waist. However, some issues arise due to the lack of waist tracking. The body yaw rotation would always be the same as the head, so participants could not look around without their body also rotating. When crouching, the body would not bend correctly.

2.4.3 3 point tracking

The 3 point tracking is considered the most basic tracking done with the "standard" equipment of the HTC Vive (HMD and controllers). Similar to 5P, the waist position and orientation was controlled by head tracking. However, because there is no feet tracking, the avatar legs are static and "slide" throughout the floor when participants walk. However, legs do still bend when participants crouch.

2.4.4 3 point tracking with simulated walking

The simulated walking condition (3PS), tries to overcome the issue of having static legs of the 3P condition. In order to move the feet without trackers, four feet animations were used: walk forward, backwards, strafe left and right. Depending on the velocity and direction of the HMD, the correct type of animation and it's intensity would play. If the participant walked forwards and

Accepted Version. Changes were made to this version by the publisher prior to publication.
The final version of record is available at https://doi.org/10.1007/s10055-021-00530-5

left at the same time (diagonally) a blend of the walk forward and strafe left would play. If the participant would walk forward but slowly, the walk animation speed would scale down to the speed of the participant. The movement from the feet animations would serve as input for the IK system. This way, if the participants crouch while still walking, the knees would bend properly.

2.5 Instruments

The following questionnaires were used to collect data from the participant:

- **Generic socio-demographic questionnaire**: age and gender were asked for purposes of sample characterisation;
- **Igroup Presence Questionnaire - Portuguese Version (IPQp)** Vasconcelos-Raposo et al. (2016): presence questionnaire constituted by 14 questions using a 5-point Likert scale. The questionnaire includes the subscales presence (overall feeling of being "there"), spatial presence (feeling of being physically present in the VE), involvement (the attention given to the VE), and experienced realism (subjective experience of realism in the VE);
- **Embodiment questionnaire**: a twenty-item Portuguese version of the questionnaire proposed by Gonzalez-Franco and Peck (2018) achieved by adopting the back-translation method Hambleton and Zenisky (2011) with two Computer Science doctorates specialised in VR to ensure its psychometric properties. In this work, we considered the following subscales: body ownership (the extent users sense the self-representation as being theirs); agency (the extent users perceive the self-representation movements as being theirs and the level of control over it); tactile sensations (the extent users perceive the touch of the self-representation as real); external appearance (how the self-representation resembles the user's appearance); response to external stimuli (how users respond to threats to their self-representation).

2.6 Procedures

All experiments were conducted at MASSIVE Virtual Reality Laboratory, located at the University of Trás-os-Montes e Alto Douro. The experimental room has soundproof walls and doors, temperature control, light control as well as air extraction allowing full control of external variables that could influence results. Before starting, participants completed the socio-demographic questionnaire. Then the investigator explained the task to be performed during the experiment, as well as the calibration process. A sheet containing this information accompanied by an illustration was also given for the participant to read.

Depending on the condition, participants could wear all the trackers (Figure 6) or only the HMD and controllers. Participants were equipped with the help of a researcher.

Accepted Version. Changes were made to this version by the publisher prior to publication.
The final version of record is available at https://doi.org/10.1007/s10055-021-00530-5

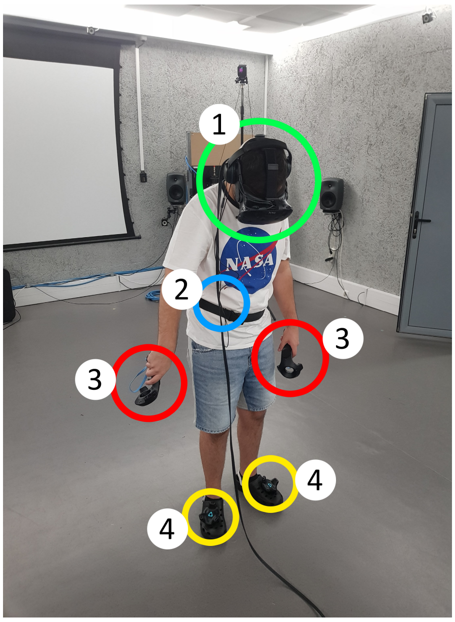

Fig. 6 Participant wearing all the tracking devices. 1-HMD for head tracking; 2- HTC Vive tracker to track waist (situated in the participant's back); 3-Controllers for interaction and hand tracking; 4-HTC Vive trackers to track both feet.

After everything was up and running, the participant began the experience in a dark virtual room with a message saying "Please Wait". A virtual arrow is marked in the ground, indicating where the participants should position and orient themselves. In this pre-calibration stage, the researcher selects the condition and introduces the participant gender (so in the avatar conditions, the gender would match the participants one). The participant is then virtually transported to the calibration scene. Here the participant can see a transparent avatar, in T-pose, in front of him, facing the room's mirror. The user has to step "inside" this transparent avatar and put his/her head, hands and feet in the same place as the avatars ones. At the extremities of the avatars members, the user will see white spheres, serving as helpers to correctly position the trackers. For the head calibration, the sphere will be at a certain distance in front of the transparent avatar (in order for the user to be able to see the sphere).

The calibration process is done by positioning the trackers (represented by red cubes) inside white spheres (located at the extremities of the avatar members) (Fig. 7). When in the right place, the sphere will turn green and users must hold the tracker position for 3 seconds. This process can only be done with one tracker at a time. After one tracker is calibrated, there is no need to continue holding its position.

After all the points are calibrated, the transparent avatar is replaced with the users' self-representation. The participants can inspect their representa-

Accepted Version. Changes were made to this version by the publisher prior to publication.
The final version of record is available at https://doi.org/10.1007/s10055-021-00530-5

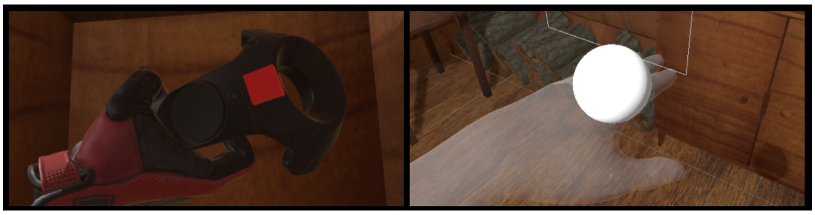

Fig. 7 Calibration process. The user has to position the tracker represented with a red cube (leftt image) inside the white sphere (right image) and hold it's position for 3 seconds.

tion in the mirror. After 15 seconds, the participant is transported to the second scenario. Here, the participant must complete the task of finding a series of objects scattered around the environment and put them in a travel bag. These objects were scattered throughout the environment, inside drawers and cabinets, positioned at different heights. There was no time limit, and the participants could take the time they wanted to explore and find the objects (even though time was being recorded without their knowledge). Beyond the predefined objects, the environment was made up of more interactable objects that were not necessary to complete the task.

All the participants did the experiment without any information about what was being evaluated.

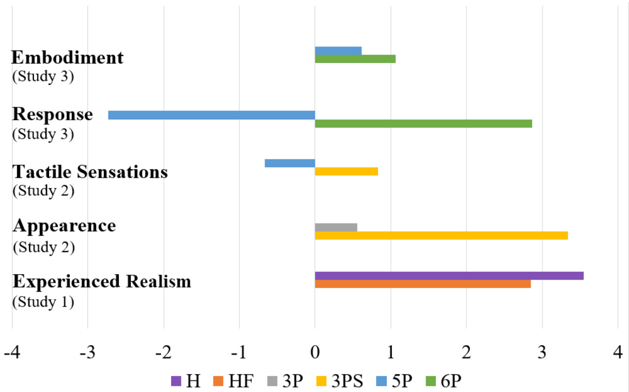

Fig. 8 Means for each significantly different dependent variable between groups

3 Results

A preliminary analysis of the data was conducted to verify its normal distribution. From the analysis of the Presence and Embodiment data, it was verified that the distribution was normal ($|Skewness| < 2$ and $|Kurtosis| < 2$). We identified seven outliers through visual inspection of boxplots, which were re-

Accepted Version. Changes were made to this version by the publisher prior to publication.
The final version of record is available at https://doi.org/10.1007/s10055-021-00530-5

moved from the analysis (four removed in H, one in HF, one in 5P and one in 6P).

Descriptive data are express as mean ± standard deviation, unless stated otherwise. Homogeneity of variances was assessed by Levene's test. In ANOVA analysis, if the assumption of homogeneity was violated, then a Welch's ANOVA would be performed instead. Descriptive statistics for every subscale and condition can be visualised in Table 2 and 3. The means for each significantly different dependent variable can be visualised in Figure 8.

3.1 Study 1 - Floating members

An independent-samples t-test was run to determine if there were differences in presence and embodiment scores between the self-representation of the users through hands alone (H) against hands and feet (HF).

The score of the subscale experienced realism was higher in the condition H (3.546 ± 0.835) than HF (2.850 ± 0.778), with a statistically significant difference of 0.695 (95% CI, 0.038 to 1.353), t(24) = 2.183, p = 0.039. For all other dependent variables, no statistically significant differences were found.

3.2 Study 2 - Walking fidelity

An independent-samples t-test was run to determine if there were differences in presence and embodiment scores when comparing walking simulation (3PS) with no walking (3P) and walking simulation (3PS) with real walking (5P). Bellow follows the significant differences found.

3.2.1 Simulated vs no walking

The score of the subscale appearance was higher in the condition 3PS (3.333 ± 2.870) than 3P (0.556 ± 3.884), with a statistically significant difference of 2.778 ($95\% CI, 0.465$ to 5.090), $t(34) = 2.440$, $p = 0.020$.

3.2.2 Simulated vs real walking

The score of the subscale tactile sensations was higher in the condition 3PS (0.833 ± 1.948) than 5P (-0.667 ± 1.952), with a statistically significant difference of 1.500 ($95\% CI, 0.110$ to 2.890), $t(31) = 2.201$, $p = 0.035$.

3.3 Study 3 - Tracking points

An ANOVA was run to determine if there were differences in presence and embodiment scores between a different number of tracking points (3P, 5P, and 6P). Response and Embodiment subscales violated the homogeneity of variances. Thus a Welch's ANOVA was performed.

Accepted Version. Changes were made to this version by the publisher prior to publication.
The final version of record is available at https://doi.org/10.1007/s10055-021-00530-5

Significant differences were found in Response ($F(2, 27.174) = 7.777, p = 0.002$) and Embodiment ($F(2, 29.789) = 4.338, p = 0.022$) between groups.

Games-Howell post-hoc analysis revealed that, for Response, an increase from 5P (-2.733 ± 2.344) to 6P (2.867 ± 5.027) of 5.600 ($95\% CI(1.974$ to $9.226)$) was statistically significant ($p = 0.002$). For Embodiment, it also revealed a increase from 5P (0.616 ± 0.406) to 6P (1.060 ± 0.454) of 0.444 ($95\% CI(0.0546$ to $0.833)$).

3.4 Study 4 - Body representation

An independent-samples t-test was run to determine if there were differences in presence and embodiment scores when comparing the use of floating members (HF) against the use of avatar (5P) and also the use of hands alone (H) against its avatar counterpart (3P). The analysis revealed no significant differences for HF *vs.* 5P and H *vs* 3P.

Table 2 Mean, Standard Deviation of all subscales for H, HF and 3PS conditions.

	H	HF	3PS
	M ± SD	M ± SD	M ± SD
Spatial Presence	4.182 ± 0.565	4.122 ± 0.353	4.352 ± 0.383
Involvement	3.591 ± 0.889	3.600 ± 0.855	3.667 ± 0.818
Experienced Realism	3.546 ± 0.835	2.850 ± 0.778	2.833 ± 0.795
Presence	3.777 ± 0.399	3.524 ± 0.373	3.617 ± 0.366
Ownership	7.182 ± 4.355	7.467 ± 3.357	2.389 ± 5.772
Agency	7.000 ± 2.683	6.133 ± 1.885	6.000 ± 1.879
Tactile Sensations	-0.455 ± 1.753	-0.333 ± 2.193	0.833 ± 1.948
Appearance	-0.818 ± 2.183	-1.467 ± 4.172	3.333 ± 2.870
Response	0.818 ± 4.045	-5.018 ± 5.208	-0.333 ± 4.446
Embodiment	0.887 ± 0.115	0.720 ± 0.403	0.712 ± 0.426
Time (seconds)	300 ± 101	294 ± 89	275 ± 57

Table 3 Mean, Standard Deviation of all subscales for 3P, 5p and 6P conditions.

	3P	5P	6P
	M ± SD	M ± SD	M ± SD
Spatial Presence	4.397 ± 0.473	4.222 ± 0.457	4.533 ± 0.389
Involvement	3.529 ± 0.838	3.750 ± 0.720	3.683 ± 0.691
Experienced Realism	3.015 ± 0.773	3.217 ± 0.597	3.433 ± 0.594
Presence	3.647 ± 0.494	3.730 ± 0.383	3.883 ± 0.394
Ownership	4.235 ± 5.262	5.13 ± 4.533	7.667 ± 4.287
Agency	5.882 ± 2.619	5.733 ± 1.534	6.200 ± 2.981
Tactile Sensations	0.647 ± 1.967	-0.667 ± 1.952	-1.089 ± 2.086
Appearance	0.530 ± 4.002	1.200 ± 4.057	2.067 ± 2.251
Response	-0.824 ± 4.613	-2.733 ± 2.344	2.867 ± 5.027
Embodiment	0.683 ± 0.693	0.616 ± 0.406	1.060 ± 0.454
Time (seconds)	253 ± 65	244 ± 68	281 ± 141

Accepted Version. Changes were made to this version by the publisher prior to publication.
The final version of record is available at https://doi.org/10.1007/s10055-021-00530-5

4 Discussion

One of the most used forms of self-representation in virtual environments is through floating members, more concretely, through floating hands. We compared the groups H and HF to understand if, in a body-less self-representation, seeing and controlling more members would improve presence and embodiment. Results indicated that the addition of floating feet worsened the Experienced Realism scores. This result was unexpected, as the more control we have over our self-representation, the better should be our experience. One possible justification could be that even though the hands were also not connected to a body, the addition of feet may have strengthened the sense of inconsistency of not having a virtual body connecting the members, thus lowering the experienced realism.

In a standard three-point tracking (HMD and controllers), the number of tracking points is not enough to track feet. In result, if the user is self-represented with an avatar, this could lead to static feet. A simulation of the feet animation could solve the issue by analysing the user movement and producing a walking animation. To analyse how the walking simulation would perform against a static feet configuration and also against using real feet tracking data to drive the avatar's feet, we compared 3PS *vs* 3P and 3PS *vs* 5P. Results showed that the simulated walking provided better appearance scores than static feet and better tactile sensations than real walking. The first difference could be explained by the fact that in 3P users see a "floating" body sliding throughout the environment when walking. By providing a simulated walking, even though it won't correspond to the real user movement, the sense of the avatar's appearance to the user will increase. Although no difference was found, the agency would have been expected to increase, as it evaluates the extent to which the user can control his avatar. Regarding the tactile sensations being better in the simulated walking than real walking, more studies on this topic are needed to clarify this phenomenon.

A standard VR system usually comes with 3 trackers, head and hands. However, more trackers could be bought to provide better control over the user's avatar. In this study, we analysed three conditions representing three points, five points, and six points in order to understand if more tracking points represent a benefit for presence and embodiment. The results indicated a significant increase in response and overall embodiment scores from 5P to 6P. The difference between conditions is the addition of a waist tracker in 6P. This waist tracker further increases the fidelity of the avatar movement. In 5P the waist positions are estimated using the head position. One of the major disadvantages is that when the user crouches, the waist will still have the same relative position of the head, resulting in a strange avatar position. The 6P corrects this issue by tracking the waist, the only major avatar movement estimations left to calculate by the IK would be the knees and elbows. By adding a waist tracker, to the already tracked head, hands and feet, users felt more embodied and responded to stimuli inside the virtual world more realistically, as if their avatar was in danger. On the other hand, there was no

Accepted Version. Changes were made to this version by the publisher prior to publication.
The final version of record is available at https://doi.org/10.1007/s10055-021-00530-5

statistically difference between 3P and 5P, suggesting that waist tracking in 6P is an important element.

In this study, we analysed whether the use of a full-body avatar, animated through trackers and IK, would outperform a self-representation by floating members. Therefore we compared the pair of conditions that shared the same number of trackers: H vs 3P and HF vs 5P. The results indicated no differences between both pairs of conditions for any dependent variable. This result contradicts state of art, where self-representation through an avatar would improve the level of presence and embodiment. However, in this study conditions, as long as there is the same level of control, the self-representation could be done through floating members or avatars without influencing presence and embodiment.

5 Conclusion

In most VR applications, usually, there is no avatar to represent users. Full-body tracking solutions are expensive for the average consumer. Although consumer level, low-cost tracking equipment can be used to enhance body tracking, there are no studies that thoroughly evaluate how they perform. Therefore, this work intended to study how low-cost solutions performed regarding self-representation and movement replication fidelity by evaluating the participant's presence and embodiment. As a secondary objective, we also evaluated how our method to simulate avatar walking, without feet tracking, would perform against more expensive solutions. This walking simulation method was intended to work with the "standard" VR equipment available to consumers, with no need to buy expensive extra equipment.

Our results indicate that the use of floating hands alone outperforms hands and feet in experienced realism. As such, more tracker (to track the feet) will not produce better results than having only floating hands. Simulated walking presents a good solution to solve the lack of feet animation when using three tracking points and a full-body avatar. When using a full-body avatar, six points tracking (tracking every extremity of the human body) led users to have a more realistic response to stimuli and have a higher overall feeling of embodiment than five points tracking. The results also led us to conclude that, as long as there is the same level of control over the self-representation, both floating members or avatar could be used without affecting presence and embodiment.

Some results contradict the literature, as such more studies are needed to confirm the reason. Future lines of work could be to add floating heads in body-less self-representations, different ways to represent floating members, add more avatar variety, and analyse the user performance in different types of tasks.

Accepted Version. Changes were made to this version by the publisher prior to publication.
The final version of record is available at https://doi.org/10.1007/s10055-021-00530-5

Acknowledgements

This work is financed by the ERDF – European Regional Development Fund through the Operational Programme for Competitiveness and Internationalisation - COMPETE 2020 Programme and by National Funds through the Portuguese funding agency, FCT - Fundação para a Ciência e a Tecnologia within project POCI-01-0145-FEDER-028618 entitled PERFECT - Perceptual Equivalence in virtual Reality For authEntiC Training.

References

Argelaguet F, Hoyet L, Trico M, Lecuyer A (2016) The role of interaction in virtual embodiment: Effects of the virtual hand representation. In: 2016 IEEE Virtual Reality (VR), pp 3–10, DOI 10.1109/VR.2016.7504682

Banakou D, Groten R, Slater M (2013) Illusory ownership of a virtual child body causes overestimation of object sizes and implicit attitude changes. Proceedings of the National Academy of Sciences 110(31):12846–12851

Banakou D, Kishore S, Slater M (2018) Virtually being einstein results in an improvement in cognitive task performance and a decrease in age bias. Frontiers in Psychology 9:917, DOI 10.3389/fpsyg.2018.00917, URL https://www.frontiersin.org/article/10.3389/fpsyg.2018.00917

Biocca F (1997) The cyborg's dilemma: Progressive embodiment in virtual environments. Journal of computer-mediated communication 3(2):JCMC324

Biocca F (2014) Connected to my avatar. In: International Conference on Social Computing and Social Media, Springer, pp 421–429

Gonzalez-Franco M, Peck TC (2018) Avatar embodiment. towards a standardized questionnaire. Frontiers in Robotics and AI 5:74, DOI 10.3389/frobt.2018.00074, URL https://www.frontiersin.org/article/10.3389/frobt.2018.00074

Hambleton R, Zenisky AL (2011) Translating and adapting tests for cross-cultural assessments. Cross-cultural research methods in psychology pp 46–70

Kilimann J, Friesen J, Pott C, Lensing P (2018) A fast avatar calibration method for virtual environments using six 6-dof trackers. In: 2018 IEEE 1st Workshop on Animation in Virtual and Augmented Environments (ANIVAE), pp 1–4, DOI 10.1109/ANIVAE.2018.8587272

Kilteni K, Groten R, Slater M (2012) The sense of embodiment in virtual reality. Presence: Teleoperators and Virtual Environments 21(4):373–387

Latoschik ME, Roth D, Gall D, Achenbach J, Waltemate T, Botsch M (2017) The effect of avatar realism in immersive social virtual realities. In: Proceedings of the 23rd ACM Symposium on Virtual Reality Software and Technology, ACM, p 39

Lin Q, Rieser JJ, Bodenheimer B (2013) Stepping off a ledge in an hmd-based immersive virtual environment. In: Proceedings of the ACM Symposium on Applied Perception, ACM, New York, NY, USA, SAP '13, pp 107–110, DOI 10.1145/2492494.2492511, URL http://doi.acm.org/10.1145/2492494.2492511

Lugrin J, Landeck M, Latoschik ME (2015a) Avatar embodiment realism and virtual fitness training. In: 2015 IEEE Virtual Reality (VR), pp 225–226, DOI 10.1109/VR.2015.7223377

Lugrin J, Wiedemann M, Bieberstein D, Latoschik ME (2015b) Influence of avatar realism on stressful situation in vr. In: 2015 IEEE Virtual Reality (VR), pp 227–228, DOI 10.1109/VR.2015.7223378

Lugrin J, Ertl M, Krop P, Klüpfel R, Stierstorfer S, Weisz B, Rück M, Schmitt J, Schmidt N, Latoschik ME (2018) Any body there? avatar visibility effects in a virtual reality game. In: 2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR), pp 17–24, DOI 10.1109/VR.2018.8446229

McManus EA, Bodenheimer B, Streuber S, de la Rosa S, Bülthoff HH, Mohler BJ (2011) The influence of avatar (self and character) animations on distance estimation, object interaction and locomotion in immersive virtual environments. In: Proceedings of the ACM SIGGRAPH Symposium on Applied Perception in Graphics and Visualization, ACM, New York, NY, USA, APGV '11, pp 37–44, DOI 10.1145/2077451.2077458, URL http://doi.acm.org/10.1145/2077451.2077458

Nakamura J, Csikszentmihalyi M (2009) Flow theory and research. Handbook of positive psychology pp 195–206

Park C, Jang K (2019) Investigation of visual self-representation for a walking-in-place navigation system in virtual reality. In: 2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR), pp 1114–1115, DOI 10.1109/VR.2019.8798345

Ries B, Interrante V, Kaeding M, Anderson L (2008) The effect of self-embodiment on distance perception in immersive virtual environments. In: Proceedings of the 2008 ACM Symposium on Virtual Reality Software and Technology, ACM, New York, NY, USA, VRST '08, pp 167–170, DOI 10.1145/1450579.1450614, URL http://doi.acm.org/10.1145/1450579.1450614

Skarbez R, Brooks FP Jr, Whitton MC (2017) A survey of presence and related concepts. ACM Comput Surv 50(6):96:1–96:39, DOI 10.1145/3134301, URL http://doi.acm.org/10.1145/3134301

Slater M, Steed A (2000) A virtual presence counter. Presence: Teleoperators & Virtual Environments 9(5):413–434

Slater M, Wilbur S (1997) A framework for immersive virtual environments five: Speculations on the role of presence in virtual environments. Presence: Teleoper Virtual Environ 6(6):603–616, DOI 10.1162/pres.1997.6.6.603, URL http://dx.doi.org/10.1162/pres.1997.6.6.603

Slater M, Usoh M, Steed A (1995) Taking steps: The influence of a walking technique on presence in virtual reality. ACM Trans Comput-Hum Interact 2(3):201–219, DOI 10.1145/210079.210084, URL http://doi.acm.org/10.1145/210079.210084

Slater M, Spanlang B, Corominas D (2010) Simulating virtual environments within virtual environments as the basis for a psychophysics of presence. In:

Accepted Version. Changes were made to this version by the publisher prior to publication.
The final version of record is available at https://doi.org/10.1007/s10055-021-00530-5

ACM Transactions on Graphics (TOG), ACM, vol 29, p 92

Steed A, Pan Y, Zisch F, Steptoe W (2016) The impact of a self-avatar on cognitive load in immersive virtual reality. In: 2016 IEEE Virtual Reality (VR), IEEE, pp 67–76

Tajadura-Jiménez A, Banakou D, Bianchi-Berthouze N, Slater M (2017) Embodiment in a child-like talking virtual body influences object size perception, self-identification, and subsequent real speaking. Scientific reports 7(1):9637

Usoh M, Arthur K, Whitton MC, Bastos R, Steed A, Slater M, Brooks Jr FP (1999) Walking¿ walking-in-place¿ flying, in virtual environments. In: Proceedings of the 26th annual conference on Computer graphics and interactive techniques, ACM Press/Addison-Wesley Publishing Co., pp 359–364

Van Der Hoort B, Guterstam A, Ehrsson HH (2011) Being barbie: the size of one's own body determines the perceived size of the world. PloS one 6(5):e20195

Vasconcelos-Raposo J, Bessa M, Melo M, Barbosa L, Rodrigues R, Teixeira CM, Cabral L, Augusto Sousa A (2016) *ipq* adaptation and validation of the igroup presence questionnaire in a portuguese sample. Presence: Teleoperators & Virtual Environments 25(3)

Yee N, Bailenson J (2007) The proteus effect: The effect of transformed self-representation on behavior. Human communication research 33(3):271–290