The Effect of Viewing a Self-Avatar on Distance Judgments in an HMD-Based Virtual Environment

Abstract

Few HMD-based virtual environment systems display a rendering of the user's own body. Subjectively, this often leads to a sense of disembodiment in the virtual world. We explore the effect of being able to see one's own body in such systems on an objective measure of the accuracy of one form of space perception. Using an action-based response measure, we found that participants who explored near space while seeing a fully-articulated and tracked visual representation of themselves subsequently made more accurate judgments of absolute egocentric distance to locations ranging from 4 m to 6 m away from where they were standing than did participants who saw no avatar. A nonanimated avatar also improved distance judgments, but by a lesser amount. Participants who viewed either animated or static avatars positioned 3 m in front of their own position made subsequent distance judgments with similar accuracy to the participants who viewed the equivalent animated or static avatar positioned at their own location. We discuss the implications of these results on theories of embodied perception in virtual environments.

1 Introduction

The classic view of visual space perception involves what is sometimes called inverse optics, in which geometric analysis is used to infer the structure of the world that is likely to have generated the sensed view of the world. An alternate approach, arising from converging research in psychology and neuroscience, considers the viewer's body as central to the act of perceiving (Wilson, 2002; Barsalou, 2008; Proffitt, 2006). This body-based approach to perception is sometimes referred to as embodied perception. The importance of body-based perception in immersive virtual environments (IVEs) has been recognized for some time (Slater & Usoh, 1994; Biocca, 1997a, 1997b; Hillis, 1999). How the body is represented in IVEs has significance for how a user may perceive, think, and interact within the environment. However, only recently has IVE technology advanced sufficiently to easily allow incorporation of a user's own body into a high fidelity virtual simulation.

Explicit awareness of body-based information can potentially provide two types of information that might be useful in spatial perception within the context of locations beyond the body itself.
Awareness of the body may serve to ground or anchor the body’s position in space. Visual information about body position might serve to establish a frame of reference, particularly in situations in which the body’s position is ambiguous or when cues about location conflict. A second function of body awareness is to provide a metric for scaling of absolute dimensions of space. The body may provide metric scaling information through cues such as familiar size or through visual-motor feedback obtained when moving the body.

In this paper, we examine how prior experience viewing a realistic human avatar affects subsequent spatial judgments. Since the effect of viewing such an avatar on space perception might be affected by the sense a user has of body ownership or by other aspects of the simulation tying the virtual body to the user, we considered both static avatars and animated avatars coupled to the user’s own body motions, along with avatars collocated with the viewer’s virtual world position and avatars positioned in front of the viewing location in the virtual world. Two control conditions were included, a simple marking on the floor collocated with the viewer’s location but otherwise having no visual similarity to a person, and lines indicating the location and height corresponding to the displaced avatars, but otherwise having no human characteristics.

The spatial judgment evaluated was absolute egocentric distance to floor locations several meters away from the observation point. Egocentric distance refers to the interval between the viewpoint and an environmental location, which differs from exocentric distances which are intervals between two environmental locations. Absolute distance refers to distances represented in some absolute scale, rather than being relative to other distances. Multiple studies of the accuracy of absolute egocentric distance judgments in HMD-based virtual environments have reported that distances appeared to be compressed in such environments, at least for locations beyond a few meters (e.g., Loomis & Knapp, 2003; Thompson et al., 2004). One motivation for the study presented in this paper was to determine whether the presence of realistic avatars might reduce or eliminate this perceptual distortion.

2 Background

Avatars are the digital representation of humans in online or virtual environments (Bailenson & Blascovich, 2004). Avatars are now in common usage in applications ranging from online games to IVEs. The majority of avatars are used in third person perspective contexts, representing views of either the user at a distance or other actors in a simulation. A substantial body of research shows that users interact with avatars representing other animate entities in a simulation in a manner similar in important ways to their interactions with real people (e.g., Durlach & Slater, 2000; Slater et al., 2006; Zhang, Yu, & Smith, 2006), as long as the avatars exhibit behavior that appears to respond to the user’s actions appropriately, with gaze direction being particularly important (Bailenson, Beall, & Blascovich, 2002).

Less common is the use of first person perspective, first person avatars (sometimes called self-avatars), which allow the user to see her or his own body. Slater, Usoh, and Steed (1995), while investigating alternative ways to control locomotion in a virtual environment, found that the subjective rating of presence was enhanced in some circumstances if participants also reported a subjective association with a virtual body that was rendered as part of the simulation. In a similar experiment but with higher quality graphics and a more complete coupling of user movement to avatar animation, Usoh et al. (1999) concluded that “presence correlates highly with the degree of association with the virtual body,” and argued that “presence gains can be had from tracking all limbs.” Lok, Naik, Whitton, and Brooks (2003) used a task that involved manipulating real objects while viewing a virtual simulation of the same objects and either generic or faithful self-avatars of the user’s hands. They evaluated both task performance and subjective presence and concluded that the fidelity of the motion of the avatars was more important for a “believable” self-avatar than the visual fidelity of the avatar, though users indicated a preference for the more self-accurate avatar. Finally, immersive virtual environments can be manipulated in ways that generate a conflict between first person and third person...
perspectives for a user’s avatar and can be used to investigate “out-of-body” experiences (Lenggenhager, Tadi, Metzinger, & Blanke, 2007) in which the observer feels and acts as if they are in the visual location of the avatar.

Only a few studies have explored whether rendering parts of a user’s body from a first person perspective in a visually immersive environment affects space perception. Draper (1995) investigated the effect of a self-avatar on spatial orientation and distance estimation tasks, with equivocal results. There was no effect of the presence of the self-avatar on a search and replace task, possibly due to overall high performance, and a complex interaction between the avatar and height of targets in a perceived reachability task. Mohler, Bülthoff, Thompson, and Creem-Regehr (2008), in a preliminary study to the work presented here, showed that prior experience with a tracked self-avatar improved the accuracy of distance judgments in a virtual environment and that the effect was not due to visually attending to the ground on which the user was standing when an avatar was present. Ries, Interrante, Kaeding, and Anderson (2008) found similar results, though they identified two possible confounds with their methodology, one due to participants in their control condition not wearing a motion capture suit and the other due to participants being allowed to walk through the virtual hallway before giving distance judgments, allowing feedback which likely resulted in calibration of responses. In addition, Williams, Johnson, Shores, and Narasimham (2008) found that viewing a rendering of one’s static feet decreased the foreshortening of a bisection task within an HMD-based virtual environment.

There is evidence from the perception community to support the important role of visual representations of body parts on judgments of the spatial position. The dominance of vision in body representation was demonstrated over 40 years ago with the finding of visual capture of felt body position using distorting prisms (Hay, Pick, & Ikeda, 1965). When presented with a conflict between the visual and proprioceptive position of one’s arm, we have a strong tendency to resolve the conflict in favor of the visual position, feeling the arm to be where it is seen. Similar visual capture effects have also been induced with artificial limbs such as the rubber hand illusion in which a visible fake hand is stroked simultaneously with an unseen real hand, and a person often feels and acts as if the stroking is occurring at the location of the fake hand (Botvinick & Cohen, 1998). This effect has also been demonstrated recently in virtual environments (Slater et al., 2007; Slater, Perez-Marcos, Ehrsson, & Sanchez-Vives, 2008).

There is, as yet, little perceptual research on the effects of the visual presence of a person’s body on spatial judgments beyond the reaching space. Creem-Regehr, Willemesen, Gooch, and Thompson (2005) investigated distance perception in real environments and showed that viewing one’s feet and the immediately surrounding areas on the ground had no effect on the accuracy of distance estimates. This experiment was done in a real world setting rich in depth cues, and the differences in perceptual uncertainty between real and virtual environments may well make observers rely on different environmental and body-based information in the two situations. Furthermore, factors such as the relative importance of rendering different body parts, body tracking, and realism of rendering may also have differential effects on performance within the IVE.

Rigorously evaluating the accuracy of distance perception is quite difficult, since there is no direct way to measure what someone “sees.” In the real world, visually directed actions are often used as indicators of the accuracy of space perception (Rieser, Ashmead, Taylor, & Youngquist, 1990; Loomis, Silva, Fujita, & Fukusima, 1992). Participants are presented with an appropriately controlled visual stimulus and then asked to perform tasks based on the visual information they have been given. Visually directed actions are open-loop tasks in which visual feedback is not provided. Thus it is argued that the accuracy of the resulting actions reflects the accuracy of the underlying perception. Visually directed action tasks used to probe distance perception include eyes-closed walking (Rieser et al.; Loomis et al.; Fukusima, Loomis, & Silva, 1997), pointing (Loomis et al.), and throwing (Eby & Loomis, 1987; Sahm, Creem-Regehr, Thompson, & Willemesen, 2005) to or toward a previously seen target.

Absolute distance perception over action space, as indicated by performance on visually-directed action...
tasks, is quite accurate in the real world (Rieser et al., 1990; Loomis et al., 1992; Loomis, Da Silva, Philbeck, & Fukusima, 1996). This is not true for absolute distance perception in HMD-based virtual environments, where multiple studies have now reported that actions are performed as if distances were perceived as 20%–50% smaller than intended (e.g., Henry & Furness, 1993; Loomis & Knapp, 2003; Thompson et al., 2004; Sahm et al., 2005; Richardson & Waller, 2007; Waller & Richardson, 2008). Much speculation and research has been directed at the phenomenon of distance compression in IVEs. There is evidence that at least in isolation, compression of distance judgments is not due to issues involving binocular stereo (Willemsen, Gooch, Thompson, & Creem-Regehr, 2008), or a variety of other effects including restricted field of view (Knapp & Loomis, 2004; Creem-Regehr et al., 2005), motion parallax (Beall, Loomis, Philbeck, & Fikes, 1995), or image quality (Thompson et al., 2004). Physical properties of an HMD may influence the effective scale of virtual space, but this only partially accounts for the results that have been observed (Willemsen, Colton, Creem-Regehr, & Thompson, 2004). Other suggestions have been made that cognitive effects such as expectations about room size or explicit feedback about responses may affect the scaling of actions in IVEs (Interrante, Anderson, & Ries, 2006a, 2006b; Interrante, Ries, Lindquist, & Anderson, 2007; Foley, 2007; Richardson & Waller, 2005).

3 Experiment

The current work expands beyond Mohler et al. (2008) and Ries et al. (2008) by exploring two important questions relevant to the potential effect of self-avatars on space perception in virtual environments: (1) Is it important that the self-avatar accurately reflect the user’s own body motions? and (2) Is it important that the self-avatar be colocated with the user’s position in the virtual environment?

The experiment was divided into two phases. During the initial exploration phase, participants visually explored a virtual environment space immediately around themselves. During the subsequent distance judgment phase, participants performed eyes-closed walking to previously seen targets, with a comparison between the actual target distance and the walked distance used as a measure of the accuracy of their egocentric distance judgments. Six different conditions were explored, involving variations in the exploration phase but with all conditions using an identical distance judgment phase. The exploration phases varied in terms of (1) whether the avatar was static versus animated in correspondence with the user’s movements compared to a nonavatar location marker present, and (2) whether this avatar or nonavatar location marker was located at the participant’s body location in the virtual environment or was displaced forward from that location.

3.1 Method

3.1.1 Participants. Forty-eight paid volunteers participated in this experiment, eight in each condition. Conditions were roughly balanced for gender. Participants were drawn from the university community in Tübingen, Germany, and compensated for their time at the rate of 8 €/hour. Participants ranged in age from 19 to 43 years (mean 29.9). None had prior experience with head mounted displays. All had normal or corrected to normal vision and were screened for the ability to fuse stereo displays (using the stereo fly test, Stereo Optical Co., Inc.).

3.1.2 Stimuli and Apparatus. The study was carried out in a fully tracked free-walking space, 11.9 m × 11.7 m in size and 8 m high. Users’ full-body position and orientations were tracked using a 16 camera Vicon MX13 optical tracking system and reflective markers on the HMD and the user’s body. Each Vicon camera had a resolution of 1280 × 1024 and the tracking system had a frame rate of 120 Hz, with the tracking latency ranging from 26 ms to 60 ms for this experiment. In addition to updating the visual environment as a function of users’ head movements, the tracking system also allowed for the capturing of full body motion data. While head motion data could be captured over the whole of the area of the tracked space, technical
issues limited full body motion capture to a smaller area. For this study, head, hands, torso, and feet were tracked and inverse kinematics were used to fully articulate the avatar. The virtual model/avatar was rendered using Virtuools from Dassault Systems. The 3D models of the two different virtual environments used for the exploration and distance judgment phases were developed using 3DStudio Max. An animation-enabled avatar purchased from Rocketbox Studios was used. Height, arm span, and leg height of the avatar were scaled to match the physical dimensions of each individual participant. The visual display was an NVIS nVisor SX HMD, with a 47°(h) by 37°(v) field of view and 1280 by 1024 resolution in each eye, which yields a spatial resolution of approximately 2.2 arc-minutes/pixel.

3.1.3 Design. In the exploration phase, participants saw a static avatar matched to their own body dimensions, an animated avatar which moved consistent with their own body motions and which matched their own body dimensions, or a nonavatar location marker made up of one or two lines. The avatars or markers were placed either at the location of the participant in the virtual world (colocated conditions), so that they were visible when the participant looked down, or were placed 3 m in front of the virtual world position of the participant (displaced conditions). The colocated nonavatar marker was a 0.5 m line on the floor indicating where the participant was standing. The displaced nonavatar marker was a 0.5 m line on the floor indicating a location 3 m in front of where the participant was standing, together with a vertical line matching the height of the participant. For the displaced avatar conditions, the avatar was facing away from the participant, so that only the back of the avatar was visible. No avatar or nonavatar location marker was present for the distance judgment phase. A between-subject, six condition design was utilized: 3 (avatar/marker type) × 2 (colocated or displaced), as shown in Figure 1.

3.1.4 Procedure. In all conditions, participants began by putting on gloves, shoes, and a light backpack which provided targets for the motion tracker. The backpack with the laptop was carried by the experimenter. The participants were then guided into the tracking space with their eyes closed, at which point they were fitted with the head mounted display. With eyes still closed, they spent approximately 30 s in a T pose looking straight ahead toward one of the tracking space walls, while a calibration procedure registered the location of the virtual avatar to the real person and scaled the avatar body segments to that of the actual person (see Figures 2–3). During the T pose, the height, armspan, and pelvis height were recorded. Given the overall and pelvis height of the participant, a version of the avatar with the appropriate overall pelvis height ratio was chosen. This ensured that the leg length would be accurate after scaling the avatar’s height. Next, this avatar was scaled in height and in width so that the overall height and arm length was accurate. Finally, since inverse kinematics were used to animate the character, the virtual hands and feet were always in the same absolute position as the participant’s.

Once the setup was complete, participants opened their eyes and were exposed to one of the six exploration phase conditions, which lasted 5 min. All six exploration phase conditions were done in an identical virtual room that was 6 m × 7 m in size and 2.8 m high. Participants were instructed to look down or forward and explore the space immediately around where they were standing, but were not permitted to move from this location. In all conditions, including those involving nonavatar location markers, full body motion tracking was utilized. These data were recorded to support post-experiment analysis of gaze direction and body movement. For the animated avatar conditions, real time motion tracking was used to move the visually rendered avatar arms and limbs in a manner consistent with the participant’s own movements.

After the exploration phase, participants were presented with a different virtual world, consisting of a simulated hallway 3.5 m × 15 m in size and 2.8 m high (see Figure 4). In this new virtual world, they performed a series of direct-blind walking tasks to targets placed randomly at 4 m, 5 m, or 6 m from the participant in a blocked random order where each distance was repeated five times. On each trial, participants viewed the target on the floor and were instructed to create a “good”
Figure 1. The six experimental conditions, varying in avatar/marker type (animated, static, line) and location (colocated or displaced). There were eight participants in each condition.
image of the target and the surrounding environment. They then were instructed to close their eyes, the HMD screens were blanked, and the participants walked without visual feedback to the target location. Because the distance walked in some of these trials exceeded the range over which full body tracking was possible, the distance judgment phase was done without a view of the avatar or nonavatar location marker. (The implications of this are explored in Section 3.3.)

3.2 Results

Distance estimations increased after the static avatar experience relative to the line on the floor control, with an even larger effect after the animated avatar experience, regardless of the location of the avatar (as shown in Figure 5). A 2 (location: colocated vs. displaced) × 3 (avatar condition: animated, static, line) univariate ANOVA performed on the ratio of walked
Table 1. *Ratio of Walked Distance to Actual Distance, Reported for Three Different Distances for Six Different Conditions*

<table>
<thead>
<tr>
<th>Condition</th>
<th>4 m</th>
<th>5 m</th>
<th>6 m</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>1: Colocated animated avatar</td>
<td>0.93</td>
<td>0.98</td>
<td>0.98</td>
<td>0.96</td>
</tr>
<tr>
<td>2: Colocated static avatar</td>
<td>0.86</td>
<td>0.88</td>
<td>0.90</td>
<td>0.88</td>
</tr>
<tr>
<td>3: Colocated line on floor</td>
<td>0.79</td>
<td>0.80</td>
<td>0.82</td>
<td>0.80</td>
</tr>
<tr>
<td>4: Displaced animated avatar</td>
<td>0.99</td>
<td>1.01</td>
<td>0.99</td>
<td>1.00</td>
</tr>
<tr>
<td>5: Displaced static avatar</td>
<td>0.89</td>
<td>0.86</td>
<td>0.91</td>
<td>0.89</td>
</tr>
<tr>
<td>6: Displaced line on floor</td>
<td>0.80</td>
<td>0.83</td>
<td>0.85</td>
<td>0.83</td>
</tr>
</tbody>
</table>

distance to actual distance (averaged across 15 trials for each subject) confirmed a main effect of avatar condition, \(F(2,48) = 74.74, p < .01, \eta_p^2 = .78\), a marginal effect of location \((p = .06, \eta_p^2 = .08)\), and no avatar \(\times\) location interaction \((p = .49, \eta_p^2 = 0.3)\). Partial eta squared \(\eta_p^2\) is used as an indication of effect size reflecting the amount of variance accounted for by the independent variable, and shows a large effect for avatar condition and small effects for location and the avatar \(\times\) location interaction. Scheffe post hoc tests confirmed significant differences \((p < .01\) for all comparisons) between the ratio of distance walked/actual distance in the colocated conditions with the animated avatar \(\text{mean} = 0.964\), static avatar \(\text{mean} = 0.884\), and line on the floor \(\text{mean} = 0.804\). In the displaced condition, the effects were similar \((p < .05\) for all comparisons) with animated avatar \(\text{mean} = 0.999\), static avatar \(\text{mean} = 0.887\), and line on the floor \(\text{mean} = 0.829\). Table 1 reports the ratio of walked distance to actual distance.

Statistical analyses were also performed on the gaze direction data to assess whether there were differences in time directed looking at locations near to the body during the six exploration conditions or during the interval preceding walking in the distance estimation task (see Table 2). Separate 2 (location: colocated vs. displaced) \(\times\) 3 (avatar condition: animated, static, line) univariate ANOVAs were performed on time spent “looking down” for the exploration and prewalking intervals. For the exploration phase, there was an effect of location, in which more time was spent looking down in the colocated versus displaced condition, \(F(1,42) = 275.94, p < .01, \eta_p^2 = .87\). More importantly, there was no effect of avatar condition \((p = .62, \eta_p^2 = .02)\), demonstrating that observers’ looking patterns were similar regardless of the presence of the animated avatar, static avatar, or line during exploration. For the prewalking interval during the distance estimation task, there was no effect of location \((p = .36, \eta_p^2 = .02)\) or avatar condition \((p = .64, \eta_p^2 = .02)\).

Statistical analyses were also performed on the body movement data to access whether there were differences in the amount of body movement as a function of location or avatar condition (see Table 3 for means and standard deviations). A 2 (location: colocated vs. displaced) \(\times\) 3 (avatar condition: animated, static, line) univariate ANOVA was performed on the average speed of limb movements \((m/s)\) during the exploration interval, separately for arm and leg movements. For the arm movements, there was an effect of the avatar condition, \(F(1,42) = 1023.22, p < .01, \eta_p^2 = .98\). Planned contrasts revealed that more arm movements were made when the avatar was animated versus static \((p < .01\) and there was no difference in arm movements in the static versus line condition \((p = .85\). Importantly, there was no effect of location \((p = .35, \eta_p^2 = .021)\), demonstrating that arm movements were similar regardless of location of the avatar. We found the same effects for leg movement, demonstrating an effect of avatar condition \(F(1,42) = 137.08, p < .01, \eta_p^2 = .87\) (animated versus static, \(p < .01\); static versus line, \(p = .27\)) and no effect of location \((p = .114, \eta_p^2 = .06)\). Together, the body movement data provide additional information about the tendency to move one’s body when provided with visual feedback (or lack of feedback) from an avatar. Significantly greater movement was seen in both animated conditions, with notably greater movement in the arms compared to the legs. Also, it is important to note that there was no difference in body movements from the static to the line condition, although there was a statistical difference in spatial judgments. This indicates that body movements alone were not responsible for the changes in spatial judgments.
Table 2. Gaze Direction Data for the Six Different Conditions*

<table>
<thead>
<tr>
<th>Condition</th>
<th>Exploration phase</th>
<th>Prewalking phase</th>
</tr>
</thead>
<tbody>
<tr>
<td>1: Colocated animated avatar</td>
<td>69.54% (5.70)</td>
<td>17.11% (7.90)</td>
</tr>
<tr>
<td>2: Colocated static avatar</td>
<td>63.37% (9.09)</td>
<td>14.05% (4.45)</td>
</tr>
<tr>
<td>3: Colocated line on floor</td>
<td>64.00% (4.99)</td>
<td>11.97% (4.79)</td>
</tr>
<tr>
<td>4: Displaced animated avatar</td>
<td>28.25% (6.77)</td>
<td>14.03% (5.09)</td>
</tr>
<tr>
<td>5: Displaced static avatar</td>
<td>29.17% (9.27)</td>
<td>17.21% (3.02)</td>
</tr>
<tr>
<td>6: Displaced line on floor</td>
<td>31.13% (8.33)</td>
<td>16.12% (4.73)</td>
</tr>
</tbody>
</table>

*Average percentage of time that the participants looked down (defined as when their gaze direction intersected with the floor with a radius from their standing point of 2.5 m). Standard deviations are in parentheses.

Table 3. Body Movement Data for the Six Different Conditions Averaged over All Participants*

<table>
<thead>
<tr>
<th>Condition</th>
<th>Hands</th>
<th>Feet</th>
</tr>
</thead>
<tbody>
<tr>
<td>1: Colocated animated avatar</td>
<td>0.310</td>
<td>0.090</td>
</tr>
<tr>
<td>2: Colocated static avatar</td>
<td>0.040</td>
<td>0.023</td>
</tr>
<tr>
<td>3: Colocated line on floor</td>
<td>0.036</td>
<td>0.014</td>
</tr>
<tr>
<td>4: Displaced animated avatar</td>
<td>0.290</td>
<td>0.120</td>
</tr>
<tr>
<td>5: Displaced static avatar</td>
<td>0.036</td>
<td>0.020</td>
</tr>
<tr>
<td>6: Displaced line on floor</td>
<td>0.034</td>
<td>0.016</td>
</tr>
</tbody>
</table>

*Body movement is the average movement (m/s) for the 5 min exploration phase and is reported for hands and feet. Standard deviations are in parentheses.

3.3 Discussion

Consistent with Mohler et al. (2008) and Ries et al. (2008), we have demonstrated that experience with a self-avatar within an HMD virtual environment has large effects on subsequent distance judgments. Furthermore, we examined the importance of self-body-motion and colocation with the user’s physical position in the VE. Our results suggest that although significant increases in distance judgments occurred when the avatar did not move with the user’s movements, greater effects were found with accurate avatar movement. Additional analyses of gaze behavior suggest that these effects are not a result of differential attentional or viewing behavior among the avatar/marker conditions. Interestingly, the effects of static and animated avatars were essentially the same regardless of the visual location of the avatar.

There are several possible accounts for changes in spatial estimates as a function of the presence and movement of a visual body including (1) the visual presence of the body, (2) a frame of reference that grounds the observer in the environment, and (3) perceptual-motor feedback that informs absolute scale or perceived body ownership. Each of these accounts is informed by the present results. First, the comparison between the line and avatar conditions suggests that the presence of the body itself, regardless of its fidelity of motion, had an influence on distance estimations. One explanation for this outcome is that the human body serves as a familiar size cue which may provide metric scaling information for the virtual environment. Related to this, we found generalized effects of the avatars regardless of whether the avatar was colocated with the physical body. This effect suggests that the changes in spatial estimates may...
be more of a result of scaling by information of the body rather than grounding the body in a location in the environment. However, further work to discriminate between effects associated with establishing the body as a frame of reference versus scaling of spatial dimensions is necessary. One approach would be to determine whether judgments change as a result of the body in an additive or multiplicative way as a function of distance. An additive change would support the account of the body serving as a frame of reference, grounding the observer in a location within the environment and leading to a systematic shift in responses. In contrast, a change in scaling should be manifested in a multiplicative effect. This type of analysis relies on regression models requiring more data points at more distances and is an avenue for future research.

Notably the largest difference in distance estimates was found for the animated/tracked body part avatars, suggesting that visual-motor feedback does contribute in some way, either to a greater sense of body ownership, increased information for scaling of the virtual space, or both. Understanding the mechanisms underlying perceived body ownership with avatars could have important implications for increasing the fidelity of space perception in virtual environments. Several studies with and without VE technology have begun to test the parameters involved in self-identification of bodies and body parts. With artificial hands such as in the rubber hand illusion, it has been shown that parameters such as the orientation of the hands and visual similarity to the participant’s hands modulate the effects of the illusion (Pavani, Spence, & Driver, 2000). Lenggenhager et al. (2007) found that an “out of body” experience created with a virtual body was not effective when that body was replaced by a size-matched simple rectangle volume. Slater et al. (2008) recently demonstrated self-ownership of a virtual arm after stimulation with a virtual ball, following the rubber hand illusion paradigm.

In this context of body ownership, the effects of the displaced avatar are both intriguing and puzzling. The introduction of a self-avatar does not necessarily imply embodiment—that the user and avatar are experienced as the same self. In the present study, the displaced avatar is viewed as in front of the user with movements that correspond to the user’s own movements. One line of reasoning would suggest that if the user did experience himself or herself in the location of the avatar (3 m in front of his or her physical position) then we might expect the distance walked by the observer to decrease rather than increase in the displaced avatar condition, as the avatar was moved forward in space. However, our results show an increase in distance estimations, consistent with the effects of the colocated animated condition. On the other hand, we do not rule out the possibility of embodiment of the avatar. One important factor may be the extent to which agency is experienced in the avatar (Short & Ward, 2009). Agency is the sense of control of one’s own body and events in the external environment. Our displaced, animated condition directly linked the control of one’s own body movement with the visual movement of the avatar. These circumstances result in efferent feedback from voluntary motor commands, afferent feedback from signals such as proprioception indicating the state of body position, as well as reafferent sensory feedback dependent on one’s actions (Wexler & van Boxtel, 2005) from the visual avatar movement. The control over self-movement has been shown to increase self-recognition (Tskaris, Haggard, Franck, Mainy, & Sirigu, 2005) and may have contributed to the similarity of the effect to the colocated condition. More research is needed with self-avatars viewed from first-person or third-person perspectives to examine the roles of agency and dynamic movement on one’s subjective embodiment of an avatar as well as on spatial behavior.

Furthermore, the mechanisms underlying perceptual-motor calibration with avatars and the subsequent effects on space perception are unknown. Our previous work has demonstrated that observers readily calibrate locomotion after short adaptation periods of walking (without avatars) which manipulate visual-motor pairings of information for self-motion (Mohler et al., 2007). Perceptual-motor calibration of body part movement may be further tested using a similar approach in which visual information for self-motion is mismatched to actual limb motion. While the present results cannot distinguish whether the effects of the animated avatar are due to the presence of more information for scaling
versus a more salient sense of ownership of the avatar, they do emphasize the overall importance of body-motion. Future work which allows for the viewing of animated avatars whose movements are dissociated from the viewer’s own movements may differentiate between these two explanations.

Finally, it is important to consider that our laboratory tracking constraints led to an experimental design in which the avatar was experienced only during an exploration phase which occurred in a different environment from the distance judgment phase. Thus, we have shown that avatar effects on distance estimations do not require the presence of the avatar during the distance estimations and that avatar effects generalize from one environment to another.

4 Conclusions

The utility of virtual environments for applications involving spatial behavior will likely increase when perceptual experiences within virtual environments mirror those of the real world. Our goal was to investigate the influence of self-avatars on one type of spatial judgment requiring absolute distance perception. We found that the presence of an avatar changed the typical pattern of distance underestimation seen in many HMD-based virtual environment studies. Users showed a remarkable increase in distance estimations with avatar experience, especially when the avatar was animated in correspondence with their own real body movements. These results are an important advance in our understanding of the role of embodied perception in virtual environments. At the same time, our results introduce several new questions about the nature of self-representation in virtual environments and its effects on spatial perception.

Acknowledgments

This work was supported in part by NSF grant IIS-0745131 and by the World Class University program through the National Research Foundation of Korea funded by the Ministry of Education, Science and Technology (R31-2008-000-10008-0). The authors wish to thank Michael Weyel, Martin Breidt, Naima Laharnar, Stephan Streuber, and Jennifer Campos for discussions and experimental support.

References


