Effects of Anthropomorphic Fidelity of Self-Avatars on Reach Boundary Estimation in Immersive Virtual Environments

Elham Ebrahimi
University of North Carolina at Wilmington
ebrahimie@uncw.edu

Andrew Robb
Clemson University
arobb@clemson.edu

Leah S. Hartman
Clemson University
Leah.hartman218@gmail.com

Christopher C. Pagano
Clemson University
cpagano@clemson.edu

Sabarish V. Babu
Clemson University
sbabu@clemson.edu

ABSTRACT

Research has shown that self-avatars (life-size representations of the user in *Virtual Reality* (VR)) can affect how people perceive virtual environments. In this paper, we investigated whether the visual fidelity of a self-avatar affects reach boundary perception, as assessed through two variables: 1) action taken (or verbal response) and 2) correct judgment . Participants were randomly assigned to one of four conditions: i) high-fidelity self-avatar, ii) low-fidelity self-avatar, iii) no avatar (end-effector), and iv) real-world as reference task group. Results indicate that all three VR viewing conditions were significantly different from real world in regards to correctly judging the reachability of the target. However, based on verbal responses, only the "no avatar" condition had a non-trivial difference with real world condition. Taken together with reachability data, participants in "no avatar" condition were less likely to correctly reach to the reachable targets. Overall, participant performance improved after completing a calibration phase with feedback, such that correct judgments increased and participants reached to fewer unreachable targets.

CCS CONCEPTS

• **Human-centered computing → Virtual reality**; *User studies*; • **Mathematics of computing →** *Exploratory data analysis*; • **Computing methodologies →** *Perception*;

KEYWORDS

Reach Boundary Estimation, Self-Avatar, Virtual Environment, Verbal Judgments, Correct Judgments

ACM Reference Format:
Elham Ebrahimi, Andrew Robb, Leah S. Hartman, Christopher C. Pagano, and Sabarish V. Babu. 2018. Effects of Anthropomorphic Fidelity of Self-Avatars on Reach Boundary Estimation in Immersive Virtual Environments. In *SAP '18: ACM Symposium on Applied Perception 2018, August 10–11, 2018, Vancouver, BC, Canada.* ACM, New York, NY, USA, Article 4, 8 pages. https://doi.org/10.1145/3225153.3225170

SAP '18, August 10–11, 2018, Vancouver, BC, Canada

ACM ISBN 978-1-4503-5894-1/18/08…$15.00
https://doi.org/10.1145/3225153.3225170

1 INTRODUCTION

Virtual self-avatars are life-size visual representations of the user, seen from a first-person perspective and co-located with users' actual body. Recent perception research suggests that having a self-avatar can influence how users perceive and interact with objects within their peripersonal space or the space immediately surrounding our bodies [Hayes et al. 2010; Slater et al. 2006]. Furthermore, the visual fidelity of self-avatar alters the users' perception in *Immersive Virtual Environments* (IVEs) evidenced in users' corresponding actions [Ebrahimi et al. 2018; Ries et al. 2009a]. As such, inaccurate representations of a user's own body may potentially degrade his or her perception of the environment. Furthermore, as self-avatars may have different properties from the user, self-avatars may affect users' perceived action capabilities within the environment.

It is possible that this mismatch could be overcome by allowing users to act within the virtual environment and providing feedback about the results of their actions. Previous work has shown that the human perception-action system is inherently dynamic and is in a continuous state of (re)calibration [Day et al. 2017; Linkenauger et al. 2015]. This calibration allows different sources of information to be integrated as to produce more accurate actions. However, it is not well understood to what extent visual fidelity of a self-avatar could affect users' perceived action capabilities, and to what extent calibration can compensate for any mismatches. If calibration can overcome any negative effects produced by mismatches between a self-avatar and the user's own body, then this would represent a much more accessible and appropriate solution, as compared to minimizing the mismatch between the self-avatar and the user's own body.

In this paper, we empirically evaluate how varying degrees of anthropomorphic fidelity of a self-avatar influences reachability and reach boundary estimation, and how this process is influenced by visuo-motor calibration. For this experiment, we selected a reaching task where participants must first determine whether an object is reachable (existing within the user's peripersonal space), and then must reach to it accurately. This task was selected because it requires calibration for the user to determine if a target is within the reach envelope, and then for them to execute an accurate reach action. Results suggest that anthropomorphic fidelity alters task performance, but that this can be largely corrected for by calibration.

2 RELATED WORK

According to the theory of embodied cognition, the mind not only influences the body but the body (including the motor and perceptual systems) also serves as a frame of reference, which in turn influences the mind [Ambrosini et al. 2012]. This is of great relevance to perception in IVEs, as both the concept and capabilities of the body can be manipulated in VR. This can take the form of virtual tools that extend the bodies capabilities [Day et al. 2017] or via actual modifications to one's virtual body [Kilteni et al. 2012b]. Prior work has shown that avatar representation can affect accuracy of distance judgments in the medium field (assessed via blind walking). Mohler et al. [Mohler et al. 2010] showed self-avatar representation influenced the accuracy of users' distance estimation. Users were able to judge distances more accurately when provided with a fully-tracked self-avatar, as compared to having no avatar. Ries et al. performed a similar experiment, but where fully-tracked full-body avatars were compared against both fully-tracked simplified avatars (only the tracking marker locations were shown in the IVE) and against no avatar [Ries et al. 2009a]. Each condition was significantly different from all others, such that full-body avatars resulted in the best distance estimations, followed by simplified avatars, and with no avatar resulting in the worst performance.

The perception of action capabilities in IVEs can also be altered through a self-avatar [Creem-Regehr et al. 2015; Mohler et al. 2010; Ries et al. 2009b]. Researchers have considered how self-avatars influence judgments pertaining to gap crossing [Jun et al. 2015] and the safety of stepping off a ledge [Lin et al. 2015]. Participants with larger virtual feet perceived wider caps to be more crossable, and participants with accurate self-avatars were better able to judge if a ledge could be stepped off of. Linkenauger et al. considered a related question (though they did not directly make the link to action capabilities) or how hand size affected the perception of the size of graspable objects [Linkenauger et al. 2013]. Smaller hands increased the perceived size of objects, and larger hands decreased the perceived size.

Perceptual psychology research has shown that body schema (specifically mental representation of our limb dimensions and action capabilities) affects how users execute actions [Iodice et al. 2015]. This has particular relevance to how the cues provided by avatars facilitate the calibration of our body schema in VR activities. Pagano and Turvey have shown that body schema is acquired via kinesthetic, vestibular and proprioceptive information during day-to-day activities [Pagano and Turvey 1998]. This suggests that body schema is not innate or learned, but is constantly updated via continuous calibration, which would enable it to adjust to changing body capabilities over time. Tools use has been shown to augment our action capabilities as they are perceived as a functional extension of the body, and thus incorporating the tool into the body schema [Wagman and Chemero 2014]. For instance, studies have shown that changing ones reaching ability via a tool produces a change in the perceived reachable space through the process of calibrating and incorporating the tool into one's body schema, even in the short term. Therefore, *our hypothesis is that self-avatars in IVE's are perceived as functional extensions of the body in the virtual world, and thus the spatial layout of the body's current configuration in VR is malleable to the anthropomorphic properties, such as visual fidelity, of the embodied self-representation.*

In our contribution, we extend this research further in examining to what extent body schema representation is adaptable or changeable due to the anthropomorphic visual fidelity of self-avatars during near-field personal space embodied interactions in VR. We measure the change in body schema, due to calibration to the anthropomorphic fidelity of the self-representation, via a perceived reachability task in order to analyze the change in the operator's perception of her reach boundary. Our research question is to what extent does perceived reachability differ with changes in anthropomorphic fidelity?

3 METHODS

3.1 Participants

Forty one undergraduate psychology students (26 females and 15 males) from a large southeastern university participated in this study. Only right-handed participants were allowed to participate. All participants were given a brief overview the experiment goal and they provided informed consent. Tests of stereo acuity were conducted for each participant. The participants were then randomly distributed to one of the three VR viewing conditions; 1) Immersive Self-Avatar, 2) Low-Fidelity Self-Avatar and 3) End-Effector, described in Section 3.3. For a reference group consisting of eleven undergraduate students, the real-world data was collected similar to the method in Day et al. [Day et al. 2017]. (Figure 4). Participants were provided with course credit as an incentive.

3.2 Materials and Apparatus

The experiment apparatus included a custom built table, a chair, an HTC Vive HMD with two controllers, and a Polhemus electromagnetic tracking system (see Figure 1). The table dimensions were 50 cm wide, 130 cm long, and 76.2 cm tall (a standard table height). 125 red LED lights were embeded in the surface of the table at 1 cm intervales, lined up along the center of the table. These LEDs were used as the reaching targets in the experiment. The LEDs were situated under a glass surface. An Arduino served as the interface between the LEDs and the computer controlling the experiment. Rather than using a single LED as a reaching target, three neighboring LEDs were used. This LED configuration enhanced the visibility of target and also provided visual cues concerning the distance to the target via both motion parallax and binocular depth perception (see Figure 2). Participants were instructed that they were to reach to the middle LED light. All three lights spanned a length of 3 cm. A chair was placed approximately 20 cm from the table and aligned midway between participants' eyes and right shoulder in order to keep the distance from the center of the eyes to the LED target line the same as the distance from the right shoulder. Participants were instructed to sit with their backs against the chair.

A specific target set of LEDs were illuminated during the various phases of the experiment. In pre- and post-test phases, 13 target distances were randomly presented 5 times corresponding to LED numbers (14, 21, 28, 35, 42, 49, 56, 63, 70, 77, 84, 91, and 98; for a total of 65 trials). The LED target distances ranged from 20.5 cm to 121.5

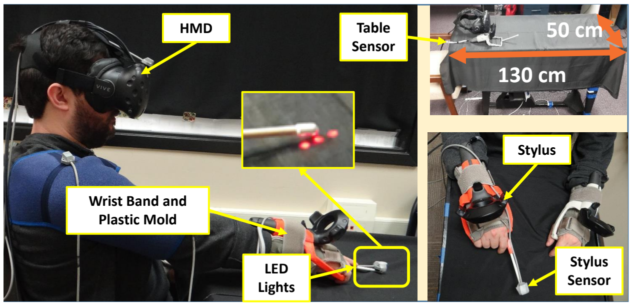

Figure 1: Shows the apparatus. The participant's head, neck, shoulder, elbow and stylus are tracked in order to record perceived distances of physical reach in the IVE.

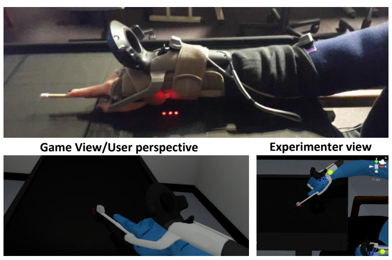

Figure 2: Top: LED target lights on the table for each trial. Bottom left: Participant's view in self-avatar condition. Bottom right: experimenter view.

cm with approximately 8 cm interval between neighboring targets. In calibration phase, 5 random permutations of 9 target distances corresponding to LED numbers (10, 20, 30, 40, 50, 60, 70, 80, and 90; for a total of 45 trials)[1]; ranged from 15.7 cm to 107.65 cm with approximately 11.5 cm interval between neighboring targets.

Accurate virtual replicas of the experiment apparatus, the chair, and the room were carefully modeled using Blender and Unity 3D. These were used to recreate the real-world experience in virtual reality. The virtual experiments were run in the same room with the physical apparatus, which was carefully aligned with the virtual environment so as to provide perfectly registered passive haptic feedback and enable conducting comparative experiments in the real world. Participants viewed the virtual environment using an HTC Vive. The HTC Vive has a combined resolution of 2160 x 1200 pixels (binocular display), a field of view of 110 degrees, and weighed 563g. Participant's inter-occular distance was measured before the experiment, which was later used to set the stereoscopic eye separation between the left eye and right eye view frusta in the HMD.

A wrist brace and a 3D-printed plastic mold was designed and built to mount the HTC Vive controllers on top of the participants' wrists (see Figure 1). This configuration guaranteed the consistency of orientation of HTC Vive controllers across all trials and all participants, and allowed experimenters to accurately model the position and orientation of their wrist and hand in the IVE. A plastic rod with a rubber tip was inserted in a 3D printed plastic mold. Participants were instructed to place their index finger on the small 4cm rod and reach to the target with the tip of the tool close to the tip of their index finger on their right hand in a natural manner (see Figure 2). All participants were right handed.

A Polhemus Liberty electromagnetic tracking system was used to track and log the movements of the participant in six degrees of freedom at every frame. To this aim, a Polhemus sensor was places on the tip of the plastic rod in the participant right (dominant) hand. The tracker's source was placed underneath the table without occluding the view of the participants.

For conditions containing a self-avatar, participants' avatars were calibrated at the start of the experiment. This calibration accounted for varying eye height and arm span. Seated eye height was measured using the postion of the HMD, and arm span was calculated by having participants assume a T-pose and measuring the distance between the controllers affixed to participants' wrists (see Figure 3). An inverse kinematic system was employed in the two self-avatar conditions to calculate the position of the joints with HTC Vive trackers to accurately track user's upper body and arm motion.

3.3 Conditions

This experiment utilized a 4 (Conditions) by 3 (Phases) mixed-model design (see Figure 4 & 3). Participants were assigned to one of the following experimental conditions that was presented in a between-subjects manner:

- Real World *(RW)*: Participants completed all three stages (see Figure 4) in the real-world. The acclimation phase was performed only to get participants familiar with the equipment that they were outfitted with described in section 3.2[2]. The acclimation phase took only 3-4 minutes to complete.
- Immersive Self-Avatar *(SA)*: Participants' arm length and eye height were measured using the HTC Vive HMD and its two controllers[3] to create a custom self-avatar for each participant. This self-avatar was then used during the induction and testing stages. Note that all of the VR viewing conditions were fully conducted in the virtual environment.
- Low-Fidelity Self-Avatar *(LF-SA)*: Participants were shown the joint positions only presented by blue spheres at the location of the head, neck, shoulders, hips, elbows, wrists, knees, and ankles based on the basic kinematics of males and females by Runeson and Frykholm [Runeson and Frykholm 1983a] during the induction and testing stages.
- End-Effector *(EE)*: Participants only were able to see their end-effectors (i.e. the left and right hand controllers) in the induction and testing stages to perform the required tasks.

3.4 Procedure

Prior to the experiment, participants were instructed on how to perform a quick and accurate reach to the target via demonstrations by

[1]Note that target distances are different than pre- and post-test distances as the calibration should happen regardless of the distances.

[2]Participants were outfitted with same equipment in all four conditions. The only exception was that in RW, participants did not wear HTC Vive head mounted display.

[3]The same measurements were conducted for the Low-Fidelity Self-Avatar and End-Effector conditions

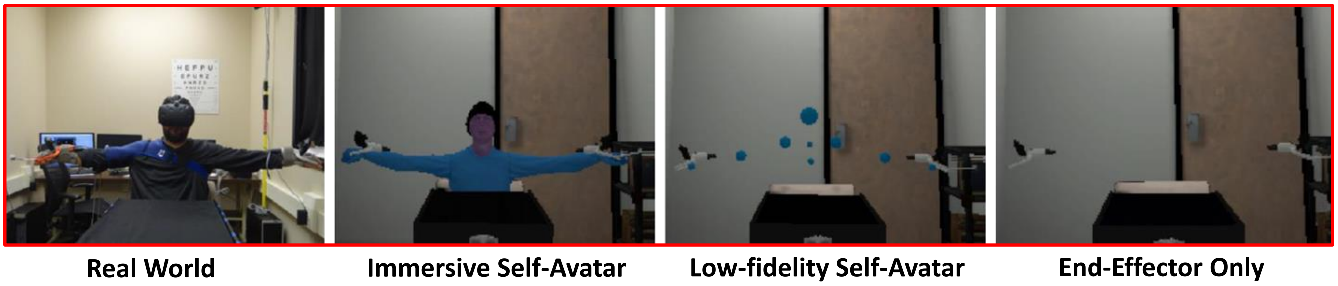

Figure 3: Left to right: i) Real-World condition, ii) Immersive Self-Avatar, iii) Low-Fidelity Self-Avatar, and iv) End-Effector (no avatar). Prior to the experiment, participants were instructed to make a T-Pose to measure the distance between the two controllers using the HTC Vive positions. This distance was then used to calculate the arm length of the participant to generate a custom avatar. The position of the head was also logged and was used to calculate the height of the participants.

the experimenter. Before completing a physical reach, participants were first asked to judge the reachability of the target and express it verbally by saying yes or no (i.e. for reachable or unreachable targets, respectively). Participants were also instructed to remain seated during the experiment (i.e. stay on the seat pan) during all attempted reaches. To ensure uniformity in starting positions across participants, each trial started and ended from the same resting position with participants' right arm on the right armrest and their backs against the back of the chair.

After this instruction period, participants completed three stages: (1) induction/acclimation stage, (2) testing stage, and (3) measurement stage, as depicted in Figure 4. In the ***Induction/Acclimation Stage***, participants spent several minutes acclimating to the experiment environment. Specifically, participants were instructed to locate and point at several objects in the room (e.g. a clock, poster, lamp, mirror, etc). Participants were then asked to move their body through a set of specific motions, including extending their arms to the sides of their body, above their head and in front of them, respectively, and move them around while looking at themselves in the mirror. Finally, participants were asked to rub the controller attached to their hand along the opposite arm several times, running from the elbow to the wrist. This was done for both arms. These steps were done to enhance the sense of self-embodiment, when a virtual body was present, as suggested by previous work [Banakou et al. 2013; Kilteni et al. 2012a; Maselli and Slater 2013]. The process took approximately 5 minutes to complete.

Next, participants entered the **testing stage**. The testing stage consisted of three successive phases: 1) Pre-test, 2) Calibration, and 3) Post-test (see Figure 4). These phases represent a within-subjects variable in the multi-factorial design.) Participants were asked to make various verbal judgments and reaches during the testing stage (see Section 3.2 for specific details).

In the *pre-test phase*, participants first verbalized whether they felt they could reach a target. If they felt the could, they closed their eyes, performed the physical reach, returned their hand to their side, and then opened their eyes. As they did not receive feedback about the correctness of their reach, this was an open-loop task (participants did receive haptic feedback upon reaching, but this did not provide any information about how close they were to the

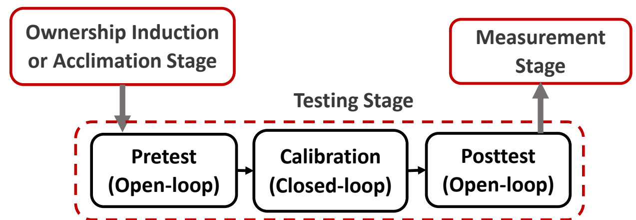

Figure 4: Experiment design.

target, other than that they had in fact touched the table). If they felt they could not reach the target, they were not required to reach.

During the *calibration phase*, a similar procedure was used. However, two modifications were made. First, participants were instructed to always attempt a reach, even if they felt the target was unreachable; this was done to allow participants to calibrate distances they incorrectly perceived as unreachable. Second, participants opened their eyes upon reaching, so as to receive feedback about the accuracy of their reach, and they then corrected their reach judgments.

The final *post-test phase* was identical to the pre-test phase. It occured immediately after the calibration phase, as a long delay between the calibration and post-test phase could potentially cause the calibration to disappear.

Finally, participants completed the **measurement stage**. During the measurement stage, participants' actual reaching ability was measured with two types of reaches; 1) reach to the table without engaging their shoulders or backs (measuring preferred reach boundary) and 2) reach absolutely as far as they could with no restrictions other than keeping their feet flat on the floor and remaining seated on the chair (measuring absolute reach boundary). The experimenters again measured various aspects of the participants' arm to ensure that the positions of the sensors had remained correct.

3.5 Hypotheses

- *H1.* It is expected that participants' perceived reachability of a target improves as the visual fidelity of the self-avatar increases, with the number of correct judgments approaching

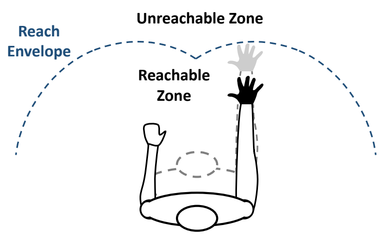

Figure 5: Distances within reach envelope are reachable. Anything beyond reach envelope is considered unreachable.

the real world performance as compared to low-fidelity and end-effector conditions.

– *H2.* It is expected that calibration phase improves participants' perceived reachability by reducing the number of incorrect judgments in post-test as compared to pre-test.

4 RESULTS

4.1 Transformed Variables

Two new variables were created using the "actual/presented distance" and "maximum arm reach" to be used for data analysis. The new variables are as follows:

- *Action taken*: variable or verbal response indicates whether or not participants made a reach towards the target during each trial. For the data analysis, "reaching" was used as the reference category and coded as 0 while not reaching was coded as 1. This variable was then used in the creation of the second variable.
- *Correct judgment*: is another binary variable created taking into account the reach envelope (also known as maximum arm reach), actual or target distance, and action taken (also known as verbal report). Trials in which the participants "reached to a target where the presented distance was within the reach envelope" or "did not reach where the presented distance was outside the reach envelope" marked as *correct judgment* (reference group). On the other hand, *incorrect judgments* were defined where participants "reached while the target was outside the reach envelope" or "did not reach where the target was within reach envelope" and coded as 1 (Figure 5).

Outlier Analysis: A full model was conducted in order to obtain residuals. These were then standardized and potential outliers were identified. Trials with excessive standardized residuals outside of a normal distribution of 3 S.D., were removed from the analysis [Cohen et al. 2003]. Overall, less than 1% of the data was eliminated from the data analysis.

4.2 Hierarchical Linear Modeling (*HLM*)

The repeated-measure design of this experiment created some challenges for traditional repeated measures analysis of variance. A multilevel modeling approach is a more flexible and accommodating alternative to repeated measures ANOVA. HLM uses all available data for each participant whereas repeated measures ANOVA requires complete data for each participant. HLM is essentially a general linear model designed to analyze variance at multiple levels. For more comprehensive overviews on multilevel modeling see for example [Hofmann 1997] and [Snijders and Bosker 1999].

The design of this experiment also created natural nesting of the data. To determine the amount of nesting, the intraclass correlation (ICC) of the null model (i.e. the intercept only model) was calculated[4]. One of the great advantages of HLM is that all levels of variance across all trials and within participants could be used and not be reduced to just the mean value similar to mean based analysis. Unlike repeated measure ANOVAs, HLM allows for a within-subjects scale variable (e.g. target distances) to be analyzed. This type of approach is more flexible and allows for the estimates, errors, and effect sizes to be more accurately modeled than traditional approaches [Cohen et al. 2003].

For HLM, variables are categorized as level 1 (L1) when they vary within-subjects or level 2 (L2) when they vary between-subjects. Level 1 variables change within a participant and they are collected at each measurement occasion (e.g. presented distance, quadratic presented distance, phase, and directionality). These variables are going to carry the residual variance. Thus, error variance for L1 predictors and intra-level interactions (L1*L1) is indexed by a reduction in residual variance. Level 2 variables do not change within a participant (i.e. condition) and represent intercept variance. Lastly, cross-level interactions (L1*L2) are indexed by the reduction in Level 1 slope variance. In multilevel modeling, the effect sizes are known as pseudo-R^2 and are the percent of reduction in error variance to the corresponding variance (e.g. residual for L1 predictor and intercept for L2 variance). Pseudo-R^2 (also known as ΔR^2) is only calculated for significant effects with all other predictors remaining within the model to control the unique effects.

For the data analyses, only the pre- and post-test data were included due the different task constraints in the calibration phase (e.g. participants were forced to reach to presented targets even if they perceived the target to be outside of their reach envelope). Comparison of pre- and post-test data to determine the effect of calibration phase is a common practice (e.g. [Day et al. 2017]).

4.3 Correct Judgment

The F-test for each of the predictors of whether participants made a correct judgment can be found in Table 1 and the fixed coefficients and standard errors can be found in Table 2. Both pre- and post-test data were included in this binary logistic regression model. The real world condition was used as the reference condition (i.e. the coefficients for the other three conditions are the difference between them and the reference condition). The pre-test phase was used as the reference for phase (i.e. for phase this is the difference of the post-test phase from the pre-test phase).

There are several main effects that were significant in terms of predicting correct judgment. While the linear term of presented distance ($F = 209.12, p < 0.001$) was significant, what is more interesting is the quadratic term ($F = 247.26, p < 0.001$). This finding indicates that while there is a linear term for correct judgments, the data fits better with a quadratic function. Therefore, judgments were dependent on presented distance meaning that judgments

[4] A multilevel modeling approach is required for an ICC greater than 2-3% [Bliese 1998].

Table 1: Fixed effect F-Tests for the binary logistic regression regarding correct judgment.

Fixed Effects		
Predictors	**F**	**P**
Presented Distance (PD)	209.12	<0.001***
Quadratic Presented Distance (QPD)	247.26	<0.001***
Phase	13.68	<0.001***
Trial Number	0.003	0.954
Condition	0.2	0.896
PD* Phase	0.1	0.756
QPD*Phase	1.15	0.285
Trial*Phase	1.96	0.162
PD*Trial	3.13	0.08
Phase* Condition	2.16	0.09
PD* Condition	5.42	0.001
QPD*Condition	15.54	<0.001***
PD*Phase*Condition	1.5	0.21
*p <0.05, **p <0.01, *** p <0.001		

Table 2: Fixed coefficients for the binary logistic regression regarding correct judgment.

Fixed Effects		
Predictors	**Coefficient (SE)**	**t**
Intercept	1.89 (0.21)	9.10***
Presented Distance (PD)	-0.11 (0.01)	-14.49***
Quadratic Presented Distance (QPD)	0.003 (>0.01)	15.74***
Phase	0.37 (0.10)	3.74***
Trial Number	>-0.001 (>0.01)	-0.06
EE Condition	0.16 (0.23)	0.71
LF-SA Condition	0.08 (0.24)	0.35
SA Condition	0.18 (0.23)	0.77
PD* Phase	-0.001 (0.01)	-0.21
QPD*Phase	0.001 (>0.01)	1.25
Trial*Phase	-0.01 (0.01)	-1.36
PD*Trial	>0.01 (>0.01)	1.63
Phase*EE Condition	0.02 (0.28)	0.07
Phase*LF-SA Condition	0.57 (0.29)	1.95
Phase*SA Condition	-0.01 (0.28)	-0.04
PD*EE Condition	0.001 (0.01)	0.14
PD*LF-SA Condition	0.03 (0.01)	3.24***
PD*SA Condition	-0.001 (0.01)	-0.13***
QPD*EE Condition	0.002 (0.001)	4.14***
QPD*LF-SA Condition	0.002(0.001)	4.08***
QPD* SA Condition	0.004 (0.001)	5.65***
PD*Phase *EE Condition	-0.04 (0.02)	-1.91
PD*Phase*LF-SA	-0.02(0.02)	-0.86
PD* Phase* SA	-0.001 (0.02)	-0.06
*p <0.05, **p <0.01, *** p <0.001		

were not consistent across the presented distances which would result in a linear pattern. Instead, participants were much more likely to make an incorrect judgment either stopping to reach before the max of their reach envelope or trying to reach beyond their reach envelope around their maximum reach critical point. However, at extreme endpoint values (e.g. very close target distances and very far target distances) participants were most likely to make correct judgments (see Figure 6). The other significant main effect was phase ($F = 13.68$, $p < 0.001$). Participants were more likely to make incorrect judgments in the pre-test ($mean = 0.91$, $SE = 0.02$) compared to the post-test ($mean = 0.93$, $SE = 0.02$).

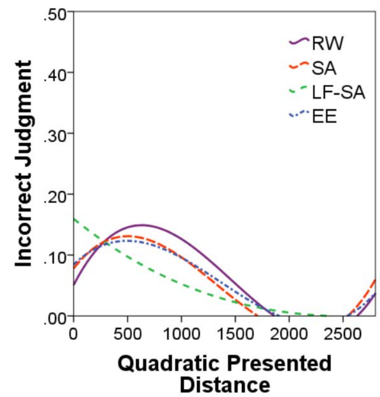

Figure 6: Probability of making an incorrect judgment based on quadratic presented distance by condition.

There were five significant two-way interaction terms: presented distance moderated by both avatar conditions (SA: $F = 3.24$, $p < 0.001$ and LF-SA: $F = -0.13$, $p < 0.001$), quadratic presented distance moderated by the end effector condition ($F = 4.14$, $p < 0.001$) and both avatar conditions (SA: $F = 4.08$, $p < 0.001$ and LF-SA: $F = -5.65$, $p < 0.001$). While both avatar conditions moderated the linear presented distance term, again, the quadratic term demonstrates a better fit for the data. All three experimental conditions were significantly different from the real world control condition moderating the quadratic presented distance term. The y axis in Figure 6 indicates incorrect judgment where 0 shows that participants correctly judged the distance at all time. The x axis shows the quadratic presented distance, in which 0 is the mean of the presented distance and then going towards the positive values we are getting closer to the extreme endpoints of the data (close targets and far targets).

With this data we cannot speak more specifically to the four type of actual judgment combinations (action take: reach, no reach, and reachability of target: reachable, not reachable). However, due to an issue of unbalanced and scarce data in terms of the number of trials between the four types of actual judgments (reach vs. no reach, and the target being reachable vs. unreachable), a multinomial regression could not be performed. Yet, it is still necessary to determine how reachability affects whether participants made reaches. Therefore, reachability was used as an additional predictor for the action taken variable.

4.4 Action Taken

Table 3 shows the fixed coefficients and standard errors for the predictors modeling whether participants reached or did not reach to the target. Recall that the coding for whether a reach was attempted

Table 3: Fixed effects F-tests for the full model predicting action taken.

Fixed Effects		
	F-tests	**p**
Reachability	1742.58	<0.001
Phase	5.22	0.02
Condition	2.15	0.09
Reachability*Phase	23.8	<0.001
Phase*Condition	4.38	0.004
Reachability*Condition	2.17	0.09
Reachability*EE Condition*Phase	1.89	0.13

Table 4: Fixed coefficients and standard errors for the full model predicting action taken.

Fixed Effects		
	Coefficient (SE)	**t**
Intercept	1.81 (0.36)	5.05***
Reachability	-6.00 (0.14)	-41.86***
Phase	-0.22 (0.10)	-2.23*
End Effector Condition	0.30 (0.47)	0.64
Lo Fi Avatar Condition	1.01 (0.48)	2.10*
Avatar Condition	-0.02 (0.47)	-0.05
Reachability*Phase	-1.10 (0.23)	-4.91***
Phase * End Effector Condition	-0.60 (0.30)	-2.04*
Phase * LF-SA Condition	-0.33 (0.29)	-1.15
Phase * Self Avatar Condition	0.35 (0.29)	1.2
Reachability* End Effector Condition	-0.53 (0.38)	-1.4
Reachability*LF-SA Condition	-0.74 (0.40)	-1.86
Reachability*SA Condition	-1.09 (0.43)	-2.56**
Reachability* EE Condition*Phase	-0.59 (0.72)	-0.82
Reachability* LF-SA Condition*Phase	-0.74 (0.65)	-1.14
Reachability* SA Condition*Phase	0.75 (0.75)	1
*p<0.05, **p<0.01, ***p<0.001		

was independent of whether it was a correct or incorrect judgment. Whether the target distance was within reach was included in this model as reachability.

In predicting action taken there was a significant main effect of reachability ($F = 1742.58$, $p < 0.001$) and phase ($F = 5.22$, $p = 0.02$). For reachability people were less likely to attempt to reach if the target was outside their reach envelope (a probability of 0.98 of attempting a reach to a target that was within reach vs. a probability of 0.15 of reaching to a target that was out of reach). In terms of phase, participants were 1% more likely to reach in the post-test (*mean* = 68%, $SE = 0.08$) than in the pre-test (*mean* = 67%, $SE = 0.08$).

There were two significant interaction terms; reachability moderated by phase ($F = 23.80$, $p < 0.001$) and condition moderated by phase ($F = 4.38$, $p = 0.04$). In the reachability moderated by phase interaction, participants were more likely to attempt to reach to unreachable targets during the pre-test (*mean* = 0.16, $SE = 0.36$) than the post-test (*mean* = 0.14, $SE = 0.35$), suggesting that participants calibrated to their reaching ability over the course of the experiment (Table 5). Regarding condition moderated by phase,

Table 5: Predicted probability of attempting a reach by phase.

	Pre-Test		Post-Test	
	Mean	**SE**	**Mean**	**SE**
Reachable	0.95	0.21	0.98	0.14
Unreachable	0.16	0.36	0.14	0.35

Table 6: Predicted probability of attempting a reach by real world and end effector conditions and phase.

	Pre-Test		Post-Test	
	Mean	**SE**	**Mean**	**SE**
End Effector	0.67	0.09	0.71	0.08
Real World	0.69	0.09	0.70	0.06

only end effector condition was significantly different from the real world (Table 6). In the pre-test phase, the end-effector group (*mean* = 0.67, $SE = 0.09$) was less likely to reach than the real world group (*mean* = 0.69, $SE = 0.09$). However the end effector group (*mean* = 0.70, $SE = 0.06$) was more likely to reach similar to the real world group (*mean* = 0.71, $SE = 0.08$) in the post-test phase.

5 DISCUSSION

Participants made more correct judgments and reached to more targets after receiving feedback in the calibration phase. Participants tended to reach to more unreachable targets in pre-test relative to post-test suggesting that they calibrated to their reaching ability after the calibration phase. Moreover, participants correctly judged the reachability of the very close or very far targets. However, at the max of their reach envelope, participants in all three VR conditions judged the reachability of the presented target more accurately than in the real-world condition overall, which is surprising as calibration seems to have worked more effectively in VR. We also found that participants were more likely to reach to the targets that were within their reach envelope in all VR conditions. When considering action taken, we discovered that participants were more likely to attempt to reach the targets in the pre-test as compared to the post-test phases overall, attesting to the impact of calibration in VR. Furthermore, action taken, also revealed that the end-effector was significantly different than the real world condition, as compared to the other self-representations. In end-effector only, participants were less likely to reach in the pre-test as compared to the real world, but calibrated by the largest magnitude to enhance their reachability perception as compared to the real world with feedback.

6 CONCLUSION

Our contribution sheds light on how much self-avatar visual information is required for calibrating users' body schema and enhance their reachability estimation in VR. We provide some guidelines below to VR developers of near-field personal space applications, such rehabilitation and training systems. Our results showed that when calibrated to the end-effector only condition, participants' perceived reachability estimates to the presented target distance became similar to real-world condition. Interestingly, participant's reachability perception was significantly worse in the low-fidelity self-avatar condition as compared to real-world, end-effector and high-fidelity self-avatar condition. However, participants' perceived reachability and reach boundary estimation did not reveal a significant difference between the real world and high fidelity condition.

Taken together with the findings of our previous work [Ebrahimi et al. 2018], we found that after calibration participants' reaches became more consistence and absolute error also reduced and reached a statistical significance in all VR conditions (with the highest accuracy in self-avatar and lowest in end effector condition). Therefore, depending on the application of the VR system, if there is a calibration phase in the system to provide feedback with regards to the participants' perceived reach judgments, then the end-effector could suffice where only the perceived reachability is critical. However, if reachability, reach boundary estimation and accuracy similar to the real world is required from the inception of the interaction in VR, then it seems that the high-fidelity self-avatar condition may be required. People perform better with avatars, but calibration can improve people's performance to the level of having an avatar. So, when you can calibrate, you don't need to include an avatar for performance reasons (you might want to include it for other reasons). But, if you can't have a calibration phase, then make sure you have an avatar. Basically, there is no guarantee that there is no problem or deficiencies in the rendering of an avatar that is why calibration is extremely important. Our results indicates that with calibration, users can overcome inaccuracies of self-avatars in terms of visual fidelity. However, when calibration is not feasible or possible then having an high-fidelity and more realistic self-avatar could overcome the lack of calibration.

Our body schema results somewhat contradicts with what other researchers have found with regards to distance estimation that the joint positions information seems crucial to depth judgments [Johansson 1973; Runeson and Frykholm 1983b]. However, our results with regards to reachability perception seems to be similar to work on interaction fidelity and task performance that shows that low and high fidelity interaction metaphors enhance performance better than the mid-fidelity condition [McMahan et al. 2016]. This alludes to the possibility that reachability might involve some amount of higher cognitive process beyond perception-action coordination. Future work will be geared towards investigating these research questions further, and exploring the effects of the self-avatar induction phase on reachability perception in VR.

REFERENCES

Ettore Ambrosini, Claudia Scorolli, Anna M Borghi, and Marcello Costantini. 2012. Which body for embodied cognition? Affordance and language within actual and perceived reaching space. *Consciousness and Cognition* 21, 3 (2012), 1551–1557.

Domna Banakou, Raphaela Groten, and Mel Slater. 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 (2013), 12846–12851.

Paul D Bliese. 1998. Group size, ICC values, and group-level correlations: A simulation. *Organizational Research Methods* 1, 4 (1998), 355–373.

Jacob Cohen, Patricia Cohen, Stephen G West, and Leona S Aiken. 2003. Applied multiple correlation/regression analysis for the behavioral sciences. *UK: Taylor & Francis* (2003).

Sarah H Creem-Regehr, Jeanine K Stefanucci, and William B Thompson. 2015. Perceiving absolute scale in virtual environments: How theory and application have mutually informed the role of body-based perception. In *Psychology of Learning and Motivation*. Vol. 62. Elsevier, 195–224.

Brian Day, Elham Ebrahimi, Leah S Hartman, Christopher C Pagano, and Sabarish V Babu. 2017. Calibration to tool use during visually-guided reaching. *Acta psychologica* 181 (2017), 27–39.

Elham Ebrahimi, Leah S. Hartman, Andrew Robb, Christopher C. Pagano, and Sabarish V. Babu. 2018. Investigating the Effects of Anthropomorphic Fidelity of Self-Avatars on Near Field Depth Perception in Immersive Virtual Environments. *IEEE Virtual Reality 2018* (2018).

Austen L Hayes, Amy C Ulinski, and Larry F Hodges. 2010. That avatar is looking at me! social inhibition in virtual worlds. In *International Conference on Intelligent Virtual Agents*. Springer, 454–467.

David A Hofmann. 1997. An overview of the logic and rationale of hierarchical linear models. *Journal of management* 23, 6 (1997), 723–744.

Pierpaolo Iodice, Nicolò Scuderi, Raoul Saggini, and Giovanni Pezzulo. 2015. Multiple timescales of body schema reorganization due to plastic surgery. *Human movement science* 42 (2015), 54–70.

Gunnar Johansson. 1973. Visual perception of biological motion and a model for its analysis. *Perception & psychophysics* 14, 2 (1973), 201–211.

Eunice Jun, Jeanine K Stefanucci, Sarah H Creem-Regehr, Michael N Geuss, and William B Thompson. 2015. Big foot: Using the size of a virtual foot to scale gap width. *ACM Transactions on Applied Perception (TAP)* 12, 4 (2015), 16.

Konstantina Kilteni, Raphaela Groten, and Mel Slater. 2012a. The sense of embodiment in virtual reality. *Presence: Teleoperators and Virtual Environments* 21, 4 (2012), 373–387.

Konstantina Kilteni, Jean-Marie Normand, Maria V Sanchez-Vives, and Mel Slater. 2012b. Extending body space in immersive virtual reality: a very long arm illusion. *PloS one* 7, 7 (2012), e40867.

Qiufeng Lin, John Rieser, and Bobby Bodenheimer. 2015. Affordance judgments in HMD-based virtual environments: Stepping over a pole and stepping off a ledge. *ACM Transactions on Applied Perception (TAP)* 12, 2 (2015), 6.

Sally A Linkenauger, Heinrich H Bülthoff, and Betty J Mohler. 2015. Virtual armŒş s reach influences perceived distances but only after experience reaching. *Neuropsychologia* 70 (2015), 393–401.

Sally A Linkenauger, Markus Leyrer, Heinrich H Bülthoff, and Betty J Mohler. 2013. Welcome to wonderland: The influence of the size and shape of a virtual hand on the perceived size and shape of virtual objects. *PloS one* 8, 7 (2013), e68594.

Antonella Maselli and Mel Slater. 2013. The building blocks of the full body ownership illusion. *Frontiers in human neuroscience* 7 (2013), 83.

Ryan P McMahan, Chengyuan Lai, and Swaroop K Pal. 2016. Interaction fidelity: the uncanny valley of virtual reality interactions. In *International Conference on Virtual, Augmented and Mixed Reality*. Springer, 59–70.

Betty J Mohler, Sarah H Creem-Regehr, William B Thompson, and Heinrich H Bülthoff. 2010. The effect of viewing a self-avatar on distance judgments in an HMD-based virtual environment. *Presence: Teleoperators and Virtual Environments* 19, 3 (2010), 230–242.

Christopher C Pagano and Michael T Turvey. 1998. Eigenvectors of the inertia tensor and perceiving the orientations of limbs and objects. *Journal of Applied Biomechanics* 14, 4 (1998), 331–359.

Brian Ries, Victoria Interrante, Michael Kaeding, and Lane Phillips. 2009a. Analyzing the effect of a virtual avatar's geometric and motion fidelity on ego-centric spatial perception in immersive virtual environments. In *Proceedings of the 16th ACM Symposium on Virtual Reality Software and Technology*. ACM, 59–66.

Brian Ries, Victoria Interrante, Michael Kaeding, and Lane Phillips. 2009b. Analyzing the effect of a virtual avatar's geometric and motion fidelity on ego-centric spatial perception in immersive virtual environments. In *Proceedings of the 16th ACM Symposium on Virtual Reality Software and Technology*. ACM, 59–66.

Sverker Runeson and Gunilla Frykholm. 1983a. Kinematic specification of dynamics as an informational basis for person-and-action perception: expectation, gender recognition, and deceptive intention. *Journal of experimental psychology: general* 112, 4 (1983), 585.

Sverker Runeson and Gunilla Frykholm. 1983b. Kinematic specification of dynamics as an informational basis for person-and-action perception: Expectation, gender recognition, and deceptive intention. *Journal of experimental psychology: general* 112, 4 (1983), 585.

Mel Slater, Angus Antley, Adam Davison, David Swapp, Christoph Guger, Chris Barker, Nancy Pistrang, and Maria V Sanchez-Vives. 2006. A virtual reprise of the Stanley Milgram obedience experiments. *PloS one* 1, 1 (2006), e39.

Tom A.B. Snijders and Roel J. Bosker. 1999. Multilevel analysis: an introduction to basic and advanced multilevel modeling. *Sage Publications* (1999).

Jeffrey B Wagman and Anthony Chemero. 2014. The end of the debate over extended cognition. In *Neuroscience, Neurophilosophy and Pragmatism*. Springer, 105–124.