new media & society
2021, Vol. 23(8) 2189–2209

DOI: 10.1177/1461444821993121
journals.sagepub.com/home/nms
SAGE

Virtual reality perspective-taking at scale: Effect of avatar representation, choice, and head movement on prosocial behaviors

Fernanda Herrera
and Jeremy N Bailenson
Stanford University, USA

Abstract

The present investigation examined the effect of avatar representation, choice, and head movement on prosocial behaviors and measures of presence after a virtual reality perspective-taking (VRPT) task. Participants were either represented by a set of virtual hands or had no representation during a VRPT task. Of those with hands, only half were able to choose their skin tone. Results showed that there was no significant advantage to having an avatar representation. However, if participants had an avatar and were able to choose their own skin tone, a higher proportion of participants performed prosocial behaviors and reported higher social presence scores compared with participants who had no choice. Regardless of condition, head rotations significantly predicted petition signatures such that the more participants rotated their heads side to side, the more likely they were to sign the petition. Moreover, when participants do not consistently rotate their head side to side, the proportion of petitions signed is on par with individuals who do not complete a VRPT task at all.

Keywords

Avatar representation, head movement, perspective-taking, prosocial behaviors, virtual reality

Corresponding author:
Fernanda Herrera, Department of Communication, Stanford University, 450 Serra Mall, Stanford, CA 94305, USA.
Email: fernhc@stanford.edu

An extensive line of research in psychology has demonstrated that perspective-taking (i.e. imagining what it would be like to be someone else under specific circumstances) is an effective way of increasing empathy and helping behaviors toward members of stigmatized groups (Batson et al., 1997a). Moreover, perspective-taking tasks have been shown to help reduce prejudice toward out-group members (Todd and Galinsky, 2014) as well as foster positive social bonds between individuals (Galinsky et al., 2005). More recently, researchers have found that virtual reality perspective-taking (VRPT; i.e. experiencing what it is like to be somebody else in an immersive virtual environment [IVE]) can reduce prejudice and bias (Oh et al., 2016a; Peck et al., 2013), increase understanding (Ahn et al., 2013), lead to prosocial behaviors (Rosenberg et al., 2013), and improve attitudes toward specific social targets up to 2 months after the VRPT task (Herrera et al., 2018).

IVEs are fully immersive, digital environments that allow users to interact with virtual content in real time and often lead to feelings of *spatial presence*. Spatial presence refers to the feeling of being inside a virtual environment (Steuer, 1992). Given any scenario can be coded and presented to the user from any perspective, IVEs help users viscerally experience specific situations and understand what it would be like to be someone else without having to rely on their imagination or preconceived biases. Past research in this area has provided evidence that VRPT tasks can increase empathy for many different social targets, namely, the elderly (Yee and Bailenson, 2006), schizophrenics (Kalyanaraman et al., 2010), and the homeless (Herrera et al., 2018).

Research has also shown that VRPT can reduce implicit racial biases for specific social targets when participants are able to control an avatar that looks like the social target with their own body (Peck et al., 2013). This is particularly prescient given past research has found that avatar representation (i.e. digital representations of users that are controlled in real time) can affect the way that users interact with others and influence behavior inside and outside IVEs (Yee et al., 2009). However, not all VRPT tasks afford an avatar representation. For example, Herrera et al. (2018) showed that a VRPT task where users did not have an avatar increased empathy and improved attitudes toward the homeless. As such, the role, if any, that avatars play during VRPT tasks remains unclear. Moreover, past research has shown that having the ability to choose and customize an avatar can lead to increased arousal (Lim and Reeves, 2009) and identification with the avatar being used (Klimmt et al., 2009). Yet, despite growing literature supporting VR as an effective empathy tool for multiple social targets, no study to date has assessed the effect that the presence or absence of an avatar representation or avatar choice (i.e. user's ability to choose specific aspects of their avatar) during VRPT tasks have on prosocial behaviors.

In addition, there is a growing line of research examining the effect that user movement during VR experiences has on affect (Li et al., 2017). However, most of this work has small sample sizes and has focused on examining the relationship between movement and emotional states (Won et al., 2016), rather than movement and subsequent behavior outside of VR. Thus, the current investigation focuses on (1) examining the effect that avatar representation and choice have on prosocial behaviors as well as social presence, spatial presence, and body transfer and (2) exploring the relationship between head movement and prosocial behaviors with a large, demographically diverse sample.

Avatar representation and choice

In non-immersive virtual worlds, such as online games and video games, research has shown that users tend to give preference to avatars that resemble them in some way, and that this resemblance leads to more in-game enjoyment (Trepte and Reinecke, 2010) and identification with their avatar (Klimmt et al., 2009). A study by Lim and Reeves (2009) compared the effect of being able to choose your own avatar on subjective evaluations of game experiences. Results demonstrated that being able to choose the avatar that will represent you during game-play led to higher arousal and more positive feelings about the game experience, especially for male players.

Within IVEs, past research has demonstrated that users take ownership of their avatars and exhibit similar neurological responses when their virtual body is threatened in comparison with when their own body is threatened (González-Franco et al., 2014). Studies have shown that participants take ownership of virtual bodies that are completely different from their own (e.g. adults embodying avatars that look like children) as long as there is *visuomotor synchrony* (i.e. participants see their avatar move exactly as they are moving their physical body in real time; Banakou et al., 2013). In addition, altering users' avatar representations can lead to behavioral changes. For example, Yee and Bailenson (2007) demonstrated that participants who embodied avatars that were shorter in a negotiation task were more willing to accept unfair offers than their counterparts who embodied taller avatars. An extension of this line of work showed that when participants embodied taller avatars inside a collaborative virutal environment (CVE), the behaviors exhibited inside the virtual environment carried over to the physical world as well (Yee et al., 2009).

Waltemate et al. (2018) examined the effect that avatar personalization from using photorealistic scans of users had on feelings of *body transfer* (i.e. the extent users feel their avatar's body is an extension of their own) as well as spatial presence. Results showed that participants reported higher body transfer and spatial presence when they embodied personalized avatars compared with generic ones even when both the personalized and generic avatars were similar in terms of realism and image quality. Even though a personalized body scan is different than choosing among customized options, these findings provide support for the current study because participants can choose hands that look more like their own. Allowing users to customize their avatar may increase spatial presence which could help users better understand what it is like to be in someone else's situation. Increased body transfer could further enhance these effects by making users feel like their body is part of the experience as well.

Moreover, it is important to note that some VRPT tasks allow users to interact with other virtual humans inside the IVE. These interactions are specifically designed to help mitigate stereotypes or to show users what it would be like to be treated as someone else (Herrera et al., 2018; Oh et al., 2016a). As such, *social presence*, the feeling users experience when they believe virtual humans are real people that are in the same virtual environment as them, may play an important role on the effectiveness of VRPT tasks. In social psychology, Allport (1954) proposed the contact hypothesis, which states that intergroup contact under certain conditions can lead to a reduction in prejudice and an increase in understanding. Given virtual contact could also produce these results, social presence could play a key role in the relationship between VRPT tasks and prosocial behaviors.

Head movements predict mental states

Most modern VR systems work by tracking the position and rotation of users' head and hands, rendering the appropriate perceptual input from the virtual world based on the user's current position, and displaying the scene to the user through the VR headset. Because most current consumer VR systems need to track users' head and hand movements 90 times per second to render and display VR content appropriately, researchers are now able to collect large amounts of nonverbal data. Although researchers have studied nonverbal behavior for decades, VR technology now allows researchers to gather and analyze large data sets of nonverbal behavior without having to rely on hand-coding from recoded videos (Bailenson, 2018).

Using these large data sets, past researchers have found that users' head movements, particularly their head rotations, are diagnostic. For example, they can reveal emotional states (Won et al., 2016), confirm mental health diagnoses (Rizzo et al., 2004), and be used to predict performance (Won et al., 2014a) and interpersonal attraction between two individuals (Herrera et al., 2020). In one study, Won et al. (2014b) asked participants in dyads to collaborate with each other and measured creativity while participant's nonverbal behavior was tracked. Results showed that higher synchrony between participants' head movements predicted higher creativity scores. In a different study by Won et al. (2016), students' head and hand movements were tracked while they attended a virtual class. Results showed that participants' head rotation, particularly side to side movements, predicted participants' self-reported anxiety.

Past research has also demonstrated that participants' head movements are positively correlated with valence and arousal regardless of content when it comes to 360 degree videos (Li et al., 2017). More specifically, the more participants scanned the virtual environment from side to side, the more pleasant and positive they found a 360 degree video, and the more participants moved their head up and down, the more exciting they found a 360 degree video. A different study assessed the effect of behavioral realism (i.e. the degree to which human avatars are able to exhibit naturalistic behaviors in real time the way humans are expected to) and nonverbal behaviors (e.g. head/hand rotations) inside IVEs found that when participants are not able to control their avatar's body in real time, they tend to move less during the interaction compared with participants who are able to control their avatar (Herrera et al., 2020). Even though head movements seem to be predicting different emotional states, the overall results of these studies offer cogent evidence suggesting nonverbal behavior is both automatic and context dependent, and could potentially be used to predict how participants feel about a particular task and how they behave after the task is done. However, to date, no extant study has empirically assessed the extent to which avatar representation, or lack thereof, affects head movement.

Present study

The current investigation sought to (1) examine the effect of avatar representation and choice on body transfer, presence (spatial and social), and whether or not participants signed a petition after a VRPT task, and (2) explore the relationship between head movement and prosocial behavior. As mentioned above, past research has demonstrated that VRPT can increase empathy and helping behaviors toward specific social targets (Ahn et al., 2013; Rosenberg et al., 2013), that avatar representation affects behavior both inside and outside the virtual environment (Yee et al., 2009), and demonstrated the benefits of being able to choose your own avatar (Klimmt et al., 2009; Lim and Reeves, 2009). Past research has also found a relationship between head movement patterns and multiple mental states (Won et al., 2014a, 2014b, 2016). However, less is known about the relationship between head movement and behavior. Thus, our hypotheses and research questions are as follows:

Hypothesis 1: Participants that go through a VRPT task will sign a petition in support of the social target they take the perspective of at higher rates than participants that do not go through a VRPT task at all.

Hypothesis 2: Participants who have an avatar representation will (a) sign the petition at higher rates and (b) report higher social presence, spatial presence, and body transfer scores compared with participants that do not have an avatar representation.

Hypothesis 3: Participants who are able to choose their avatar representation's skin tone will (a) sign the petition at higher rates and (b) report higher social presence, spatial presence, and body transfer scores compared with participants who are not given a choice.

Research Question 1: Does avatar representation and choice affect the amount of head movement made by participants throughout the VRPT task?

Research Question 2: Is participants' head movement a significant predictor of future prosocial behavior?

It is also important to note that despite the results mentioned above, most psychological studies using VR systems have been conducted in laboratories where sample sizes are small and mostly composed of college students with little demographic variance (Oh et al., 2016b). To address these methodological issues, we gathered a large sample of data from a more diverse population by taking a mobile lab unit outside of the lab and into more naturalistic settings.

Method

Participants

A total of 937 participants were recruited at a museum in the San Francisco Bay Area ($n=562$) and a film festival in New York City ($n=375$). Our sample consisted of 503 males, 422 females, and nine participants who identified as other. Of these participants, 95 (10.14%) were 15–18 years old, 423 (45.14%) were 19–35, 377 (40.24%) were 36–65, and 42 (4.48%) were above 65 years of age. Participation in this study was voluntary and participants did not receive monetary compensation. All recruitment and experimental procedures were approved by (*Institution omitted for review*)'s Institutional Review Board.

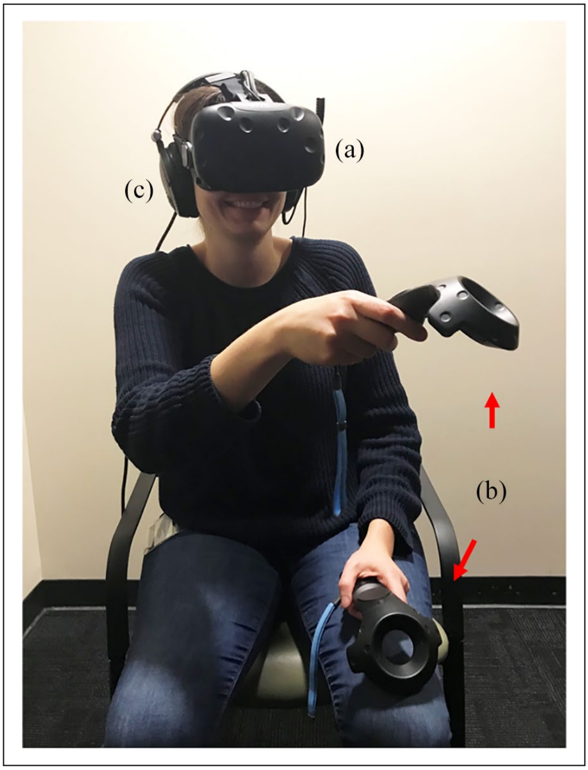

Figure 1. VR equipment: (a) HTC Vive (b) Vive controllers, and (c) headphones.

Materials and apparatus

Participants used the HTC Vive head-mounted display (HMD), hand controllers, and headphones. The HMD had a resolution of 1080 × 1200 and an update rate of 90 Hz. An optical tracking system (Valve Lighthouse, update rate of 120 Hz) and three 6 degree of freedom sensors were used to track participants' head position (x, y, and z) and rotation (pitch, yaw, and roll). Figure 1 shows the equipment setup.

The VRPT task used in this study was *Becoming Homeless* (Asher et al., 2018). The VRPT task consists of four slightly different scenes. In the first scene (apartment scene), participants are sitting in their apartment when they learn that they have lost their job and have to sell their belongings to raise enough money for rent. However, they are unable to raise enough money and are evicted. In the second scene (car scene), participants are living out of their car when they are cited for living in their car on public property. In the third scene (bus scene), participants are now traveling on a bus at night for shelter when they learn that there two men who are known to harass homeless people. One man tries to steal their backpack, while the other tries to invade their personal space. In the last scene (virtual human scene), participants are able to interact with other homeless people in the same bus and hear real stories about how the other passengers became homeless. Figure 2 shows the scene progression of the VRPT experience

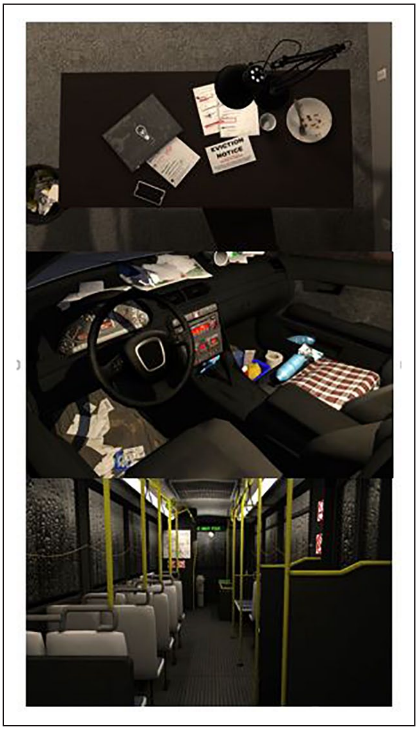

Figure 2. Scene progression of VRPT experience. Top: apartment scene; middle: car scene; bottom: bus scene and virtual human scene.

Design and procedure

This study adopted a between-subjects design with four conditions: (1) Control, (2) No Hands (3) Chosen Hands, or (4) Random Hands. After agreeing to participate, participants were fitted with the HMD and followed a quick tutorial scene where they learned how to select answer choices and interact with the IVE. After participants completed the training scene, they were asked to answer a pre-intervention questionnaire which included demographic questions as well as a question asking about participants' attitudes toward the homeless. The entirety of the study, including consent, pre and post-questionnaires, and the VRPT task, took place inside the HMD, making this study the first ever self-contained VR study.

Once participants completed the pre-questionnaire, they were randomly selected into one of the four conditions. Participants in the No Hands, Chosen Hands, and Random Hands conditions went through the exact same VRPT task where they experienced *Becoming Homeless*. These conditions differed in the type of avatar representation afforded to participants. Participants in the No Hands condition did not have any type of avatar representation, while participants in the Chosen and Random Hands conditions were represented by a set of hands that mimicked their physical hand movements in real

Figure 3. Sampling of skin tones.

time. However, participants in the Chosen Hands condition were able to choose their own skin tone from a sampling of eight tones (Figure 3), while a random skin tone from this sample was assigned to participants in the Random Hands condition. Participants were only represented by a set of virtual hands rather than a full-bodied representation given past research has demonstrated that being able to control the entirety of your avatar representation leads to increased body transfer in comparison with full-bodied avatars that cannot be fully controlled (Herrera et al., 2020).

Participants followed a virtual narrative that allowed them to interact with the IVE and guided them through the experience. Regardless of condition, participants could see a pointer in the form of a white line extending out onto the virtual environment from their physical hands. Participants used their physical hand movements to control this line and to point at objects or virtual humans they wanted to interact with (Figure 4). Participants selected answer choices, objects, or virtual humans by first pointing at them and then pulling the trigger found on the back of the controller. The line would turn green whenever participants hovered over an object or virtual human they could interact with. After the experience, participants completed a post-questionnaire, took the HMD off, and were debriefed.

Unlike the three conditions mentioned above, participants in the Control condition proceeded to the post-questionnaire immediately after they completed the pre-questionnaire. However, since participants in this condition did not go through the VRPT experience prior to answering the post-questionnaire, some questions directly related to the experience itself were omitted. After participants completed all questionnaires, they were able to go through the VRPT experience. Participants in the Control condition

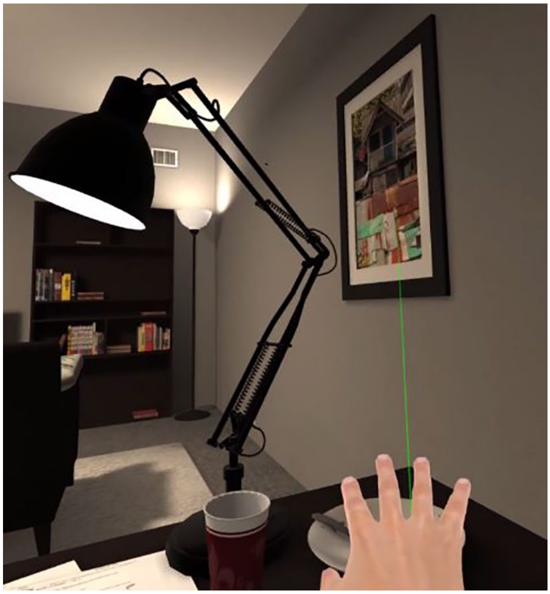

Figure 4. Pointer used for the Hands conditions. Participants in the No Hands condition used the same pointer, but were not represented by virtual hands.

completed the same VRPT experience as participants in the Chosen Hands condition. Their head movements tracked just like the other conditions. Once they were done with the VRPT experience they took off the HMD and were debriefed.

Measures

Attitudes. Adapted from the work of Batson et al. (1997b), participants were asked the extent to which they cared about the plight of the homeless prior to the VRPT task to gauge attitudes toward the homeless (M=3.55, SD=.83) using a 5-point Likert-type scale (1=*not at all*, 5=*very strongly*).

Petition signing. Participants were presented with an excerpt of a proposition supporting affordable housing for vulnerable populations adapted from the work of Herrera et al. (2018). After participants read the excerpt, they were asked whether or not they were willing to sign the petition. Participants who decided to sign the petition used the Vive controller to sign the petition inside the headset. Participants who decided not to sign simply clicked on a button that read “I do not want to sign.”

Head rotation. To determine whether avatar representation or choice influenced participant’s head movements in the form of head rotations, participant’s yaw, pitch, and roll rotations were recorded four times per second while participants completed the VRPT task. Yaw represents the participant's head movements around the Y-axis (e.g. moving their head side to side as if to say "no" or look over one's shoulder). Even though yaw can represent negation, it can also represent user's head rotations to explore the virtual environment surrounding them. Pitch represents the participant's head movements around the X-axis (e.g. nodding "yes"). Similar to yaw, pitch could be considered a sign of approval, however, it is also a measure of how much participants explore the virtual environment above and below them. Finally, roll represents the participant's head tilting around the Z-axis (e.g. touching the left ear to the left shoulder or the right ear to the right shoulder). Figure 5 shows these head rotations. Following Won et al. (2016), the extent to which the participant moved their head was calculated by averaging the standard deviations of the rotational data for the duration of the VRPT task. Thus, the standard deviation of yaw reflects the tendency of a participant to rotate their head from left to right during the task.

Social presence. Participants were asked to rate how strongly they felt like they were with other people inside the virtual environment ($M = 3.37$, $SD = 1.00$) using a 5-point Likert-type scale (1 = *not at all*, 5 = *very strongly*). This single item was adapted from the work of Nowak and Biocca (2003).

Spatial presence. Participants were asked to rate how strongly they felt like they were actually inside the virtual environment ($M = 3.25$, $SD = .96$) using a 5-point Likert-type scale (1 = *not at all*, 5 = *very strongly*). This single item was adapted from the work of Bailenson et al. (2005).

Body transfer. Participants were asked to rate the extent to which they felt the hands they saw in the virtual environment were their own ($M = 3.49$, $SD = .95$). This measure was adapted from the work of Slater et al. (2010).

Results

An aligned-rank transform analysis of variance test was used in lieu of an analysis of variance (ANOVA) given the self-report data collected was ordinal (Wobbrock et al., 2011). All significant effects of condition on the outcome variables were followed up with three planned orthogonal contrasts that specifically tested our hypotheses. We predicted (1) that going through a VRPT task will be more effective at eliciting prosocial behaviors than not going through a VRPT task at all (VRPT conditions vs Control), (2) that having virtual hands would be more effective than having no hands (i.e. Chosen and Random Hands conditions vs No Hands condition), and (3) that being able to choose your own skin tone would be more effective than being assigned a random skin tone (i.e. Chosen Hands vs Random Hands). The means and standard deviations of the outcome variables and the proportion of petitions signed by condition are summarized in Table 1.

Attitudes

Results showed that there was no significant difference on attitudes toward the homeless between conditions before the study, $F(3, 933) = .81$, $p = .486$. These results show

Figure 5. Representations of head rotations. Top: pitch; middle: yaw; bottom: roll.

that random assignment was successful. In addition, there were no differences across conditions in terms of age, $F(3, 933) = 1.18$, $p = .313$, or gender, $F(3, 933) = 1.13$, $p = .338$.

Table 1. Means and standard deviations for all outcome variables by condition.

	Control		No Hands		Chosen Hands		Random Hands	
	(n=238)		(n=236)		(n=234)		(n=229)	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Attitudes toward the homeless	3.54	.79	3.47	.82	3.61	.84	3.56	.84
Social presence	N/A	N/A	3.36	1.03	3.48	.99	3.26	.97
Spatial presence	N/A	N/A	3.29	.98	3.33	.96	3.14	.93
Body transfer	N/A	N/A	3.49	1.02	3.55	.92	3.43	.92
Yaw	79.82	15.95	80.24	14.85	81.32	15.97	80.87	14.08
Proportion of petitions signed	166 of 238 (69.7%)		173 of 236 (73.3%)		190 of 234 (81.2%)		164 of 229 (71.6%)	

SD: standard deviation.

Petition signing

There was a significant effect of condition on the proportion of petitions signed (Fisher's exact test: p=.022). The first planned contrast (Control vs any VRPT condition) showed that there was a marginally significant effect of VRPT on petitions signed (z=1.78, p=.074) such that a higher proportion of participants who experienced any kind of VRPT signed the petition, supporting Hypothesis 1. However, there was no significant difference between the No Hands and the Hands conditions (z=1.78, p=.318), rejecting Hypotheses 2a. The last planned contrast between the Hands conditions showed that a significantly higher proportion of participants in the Chosen Hands condition signed the petition in comparison with the Random Hands condition (z=2.46, p=.016), supporting Hypotheses 3a.

Social presence

There was a marginally significant difference between conditions on social presence scores, F(2, 696)=2.90, p=.056. Results from planned contrasts demonstrated that there was no significant difference in social presence scores between participants in the No Hands and the Hands conditions (i.e. Chosen and Random Hands; b=.542, t(696)=.016, p=.916) rejecting Hypothesis 2b. However, there was a significant difference between Chosen and Random Hands, b=21.49, t(936)=2.41, p=.016, with participants in the Chosen Hands condition reporting higher social presence scores, partially supporting Hypothesis 3b. Social presence was not significantly correlated with yaw head movements (r=.06, p=.092).

Spatial presence and body transfer

There was no significant difference between conditions on spatial presence, F(2, 696)=2.25, p=.106, or body transfer scores, F(2, 696)=1.33, p=.266, rejecting Hypothesis 2b and

Table 2. Proportion of petitions signed between high movers, low movers, and the Control condition.

	Control	Low movers	High movers
Proportion of petitions signed	166 of 238 (69.7%)	247 of 352 (70.2%)	280 of 347 (80.7%)

partially rejecting Hypothesis 3b. Given there was no significant effect of condition on these measures, the three planned contrasts were not conducted. Moreover, spatial presence ($r=.02$, $p=.503$) and body transfer ($r=.04$, $p=.200$) were not significantly correlated with yaw head movements, indicating that head movements were not a proxy for measures of presence or body transfer.

The effect of condition on petition signing was not mediated by body transfer, ACME = .001, 95% CI = [–.004, .01] $p=.68$; social presence, ACME = .003, 95% CI = [–.001, .01], $p=.20$; or spatial presence, ACME = .0004, 95% CI = [−.003, .01], $p=.69$.

Head rotation

There was no significant effect of condition on the amount of general head rotation exhibited by participants across conditions, $F(3, 933)=.447$, $p=.72$, indicating that our experimental manipulations did not affect the amount of movement exhibited by participants across conditions. However, whether or not participants signed the petition regardless of condition was significantly predicted by participants' yaw head movements, $b=.024$, $z(936)=4.99$, $p<.001$, such that the more participants moved their heads from side to side during the VRPT task, the more likely it was that they would sign the petition. Pitch, $b=-.039$, $z(936)=-1.54$, $p=.123$, and roll, $b=-.007$, $z(936)=-.423$, $p=.674$, were not significant predictors of petition signing.

Analysis of participants' head movements by scene revealed that it was participants' yaw rotations in the last two scenes of the VRPT experience, that is, bus scene, $b=.008$, $z(936)=3.52$, $p<.001$, and virtual human scene, $b=.009$, $z(936)=3.88$, $p<.001$, that significantly predicted signing. The first two scenes did not predict petition signing, apartment scene, $b=.006$, $z(936)=.46$, $p=.644$, and car scene, $b=.004$, $z(936)=1.23$, $p=.219$.

High versus low movers

Given participants' head movements predicted petition signing, and that a similar proportion of participants signed the petition between the Control (69.7%), No Hands (73.3%), and Random Hands (71.6%) conditions, an exploratory analysis was conducted to understand how head movement influenced behavior. Participants across the three VRPT conditions were divided into two levels of movement, high and low based on their yaw rotations, and compared with the Control condition using a median split (Iacobucci et al., 2015). Table 2 shows the proportion of petitions signed based on high and low movers compared with the Control condition. A chi-square test of proportions revealed that there was a significant difference between these three groups, $X^2(2)=12.98$, $p=.002$,

Table 3. Proportion of petitions signed between high and low movers in the Control condition.

	Control condition	
	Low movers	High movers
Proportion of petitions signed	77 of 118 (65.3%)	89 of 120 (74.2%)

with high movers signing a significantly higher proportion of petitions (80.7%) compared with low movers (70.2%; Fisher's exact test: $p=.002$) and the Control condition (69.7%; Fisher's exact test: $p=.003$). However, there was no difference between the proportion of participants who were low movers and participants in the Control condition (Fisher's exact test: $p=.927$). These results indicate that when participants do not move as part of their VRPT experience (regardless of avatar representation and choice), there is no difference in prosocial behavior between low movers and participants in the Control condition.

Even though participants in the Control condition answered the petition question before they completed the VRPT task, their tracking data during this task were still collected. Post hoc analyses show that there was no significant difference in the proportion of petitions signed between the high and low movers of the Control condition (Fisher's exact test: $p=.158$). Table 3 shows the proportion of petitions signed within the Control condition between high and low movers.

General discussion

This study extends previous literature by being the first to examine the effect that avatar representation, choice, and head movements had on prosocial behaviors, measures of presence (social and spatial), and body transfer after completing a VRPT task. Even though participants in the VRPT conditions (No Hands, Chosen Hands, and Random Hands) signed the petition at higher rates than participants in the Control condition, the result was marginal, providing only partial support for Hypothesis 1. In terms of avatar representation, there was no significant difference between the No Hands and the Hands conditions (Chosen Hands and Random Hands) in the proportion of petitions signed. However, when it came to choice, as predicted, a significantly higher proportion of participants in the Chosen Hands condition signed the petition supporting affordable housing toward the homeless in comparison with participants in the Random Hands condition. These results suggest that there may not be a significant advantage to being virtually represented during VRPT tasks when it comes to petition signing. However, if users are going to be virtually represented, as is the case with most of the virtual experiences that are publicly available, giving them the ability to choose an aspect of their avatar increases prosocial behaviors.

Results also showed that despite our experimental manipulations (i.e. avatar representation and choice), there was no significant difference in the amount of head movements performed during the VRPT task. However, analysis of the tracking data revealed that yaw significantly predicted petition signing such that the more participants rotated their heads from side to side, the more likely they were to sign the petition. Past research has demonstrated that yaw specifically can predict mental states (Li et al., 2017; Won et al., 2016). However, this study, to our knowledge, is the first to show that those same head rotations can be used to predict subsequent behaviors.

Embodied cognition theory posits that cognition is an interaction of the body and mind (Barsalou, 2010). In other words, when an event occurs, the underlying sensory, somatic, and motor states are stored, and when the event is remembered, the original states are partially simulated. Thus, the interpretation and memory of a particular event is rooted in the interaction of both cognitive and bodily states. VRPT tasks allow users to viscerally experience other people's circumstances as if they were happening to them by allowing users to physically react to the virtual experience using naturalistic movements (e.g. head rotations). VRPT tasks thus activate the same sensory, somatic, and motor states that would be activated if the experience was happening to the user in real life. Compared with traditional perspective taking tasks, VRPT tasks allow for the interaction of both cognitive and bodily states, which could lead to a more accurate interpretation of the virtual experience, and thus, a better understanding of other people's perspectives and circumstances. As such, it is possible that by rotating their heads at higher rates, participants may be gathering more accurate information about what it would be like to be homeless and how their body would feel like if they were actually in that situation thereby enhancing the perspective taking task.

Moreover, scene-by-scene analysis showed that it was particularly participants' yaw head movements in the last two scenes of the *Becoming Homeless* experience that drove these results. The last two scenes of the task are unique in that they are the only scenes where participants get to interact with other virtual humans and where yaw rotations were encouraged by making participants aware of their surroundings. For example, in the bus scene, participants are warned that the man to their left may try to steal their backpack while the man to their right has been known to harass homeless people. To keep the man to their right away, they have to look at him. However, to keep the man to their left from stealing their belongings, they have to keep an eye on him as well. Even though all participants heard the same instructions and we observed variance across participants in the amount of head rotations regardless of condition, it is possible that task compliance (e.g. participants looking right and left to keep the men away from them) and social interactions with the virtual humans improved the level of perspective taking and led to subsequent prosocial behaviors. However, it is also possible that head rotations act as a proxy for specific psychological experiences, for example vigilance (i.e. the state of keeping careful watch for possible danger). Given participants were placed in virtual environment depicting eviction, a police stop, thievery, and harassment, it is possible that participants felt like they were in potential danger and constantly scanned the virtual environment for threats. These scans, in turn, helped participants further understand not only what it would be like to be homeless but what it would feel like to be in that situation as well. Future studies should assess whether or not head movements are a proxy for vigilance or other possible explanations that tie movements to mental states.

Participants in the Random Hands, No Hands, and Control signed the petition at similar rates. To further understand these results, an exploratory analysis of the relationship between head movement and behavior was conducted. Results showed that high movers signed the petition at significantly higher rates than low movers and participants in the Control condition. In addition, participants in the Control condition were then divided into high and low movers to examine the possible causal relationship between movement and signatures. Given participants in the Control condition answered the petition question prior to completing the VRPT task, if participants who signed the petition also rotated their head more, this would be indicative that, in general, participants who sign the petition tend to move more but movement itself does not lead to more signatures. However, if there was no difference in the proportion of petitions signed between high and low movers in the Control condition, this result would indicate that movement during the VRPT task plays a significant role leading to more petitions being signed. Analysis of the tracking data for participants in the Control condition revealed that there was no difference between high and low movers, suggesting that increased yaw movements during VRPT task led to more signatures. These results suggest that when participants stay facing forward or forget that VR surrounds them, the VRPT experience yields less prosocial behaviors that are on par with the prosocial behaviors exhibited by those who did not complete a VRPT task. Although these results shed light on the possible causal relationship between movement and prosocial behaviors, future studies should examine this relationship more closely.

When it comes to social presence scores, planned contrasts demonstrated that there was no significant difference between the Hands (Chosen and Random Hands) and the No Hands condition. However, if participants have an avatar representation, which tends to be the case for most VRPT tasks, being able to choose their avatar's skin tone led to significantly higher social presence scores in comparison with participants who were randomly assigned a skin tone. Past research has shown that when users are able to choose some aspect of their avatar, they are more likely to identify with their avatar (Klimmt et al., 2009) and report increased arousal (Lim and Reeves, 2009). It is possible that by being able to choose their skin tone, participants in the Chosen Hands condition identified with their avatar and felt they were better represented in the virtual environment. This increase in self-identification and perception of being accurately portrayed inside the virtual environment may have translated to participants feeling like the other virtual humans were also accurately depicting real people, and increased participant's feeling of being inside the virtual environment with others (even if the virtual humans are programmed to behave a certain way).

There was no significant effect of condition on spatial presence, indicating that regardless of avatar representation and choice, participants felt they were inside the virtual environment at similar rates. There was also no significant effect of condition on body transfer scores. These results indicate that participants felt their avatar was their own at similar rates regardless of whether or not they got to choose its skin tone. Interestingly, participants in the No Hands condition report similar body transfer scores even though they did not embody an avatar.

Limitations and future directions

There are a number of limitations to this study. First, even though participants in the No Hands condition were not represented by a set of hands, they could still see a pointer in the form of a line extending from their physical hands. Given participants were still able to control the pointer with naturalistic hand movements, it is possible that participants felt that the pointer was an extension of themselves. Future studies should have a condition where participants are not virtually represented at all to be able to more accurately assess the difference between having and not having a virtual representation during VRPT tasks.

Participants in the Random Hands condition in the California locations answered whether or not the skin tone that was randomly assigned to them matched their own. Two participants were excluded from the analysis because the skin tone randomly assigned matched the participants' own. However, due to technical difficulties, these data were not collected for the participants in the New York location, so it is possible that a small percentage of participants in the Random Hands condition had hands that matched their own. While our goal was to examine the effect of being able to choose your own skin tone rather than the effect of embodying an avatar that matches your skin tone, future studies should ask participants in the Random Hands condition whether or not the skin tone they were assigned matched their own to avoid this confound.

In this study, participants' head rotations were positively associated with prosocial behaviors and results indicated that this was driven by the VRPT task's last two scenes which encouraged interactions between participants and both the virtual environment and virtual humans. However, the actual number of interactions with the virtual environment and the duration of those interactions were not recorded. Future studies should count and measure the length of these interactions and examine whether or not these results are driven by concrete interactions rather than exploration alone.

Participants in this study participated in the first ever self-contained study using VR, meaning they completed the entirety of the study while wearing the HMD. Even though this format allowed us to run participants at different locations around the country and gather a large, heterogeneous sample, we employed single-item scales for our dependent variables to prevent participant fatigue. Future studies should attempt to replicate this study's findings using full-item scales measuring social presence, spatial presence, and body transfer to address this limitation.

In terms of task compliance and movement, future studies should examine whether VRPT experiences that are designed to encourage more head movements are more effective in comparison with VRPT experiences that restrict movement. Researchers should also assess whether the amount of movement is dependent on the type of VR experience the user is going through or if movement is an individual trait (i.e. users will tend to move their head at similar rates regardless of VR content). In addition, in terms of social interaction, future studies should examine whether or not exposure to and interaction with virtual humans depicting the social target of the VRPT task offer similar results to Allport's (1954) intergroup contact hypothesis, which suggests that contact under appropriate conditions between different groups of people can result in reduced prejudice and potentially increased empathy.

It is also important to note that some of our hypothesis only received partial support. Results of this study suggest that the experience *Becoming Homeless* only marginally increased the number of petitions signed compared with the Control condition. Future studies should examine the relationship between VRPT tasks and prosocial behaviors using different experiences to be able to generalize this finding across VRPT tasks. Past research has demonstrated that VRPT experiences can increase empathy for a myriad of social targets (Herrera et al., 2018); however, whether or not the increase in empathy mediates the relationship between VRPT tasks and prosocial behaviors is still unknown. Future studies should examine this relationship further by examining possible mediators and moderators.

Conclusion

This investigation sought to examine the effect that avatar representation, choice, and head movements had on prosocial behaviors and measures of presence after completing a VRPT task. Results showed that there is no significant advantage of having an avatar representation when it comes to prosocial behaviors or spatial presence. However, when participants were represented by an avatar, being able to choose their avatar's skin tone led to more petition signatures and higher social presence scores in comparison to being randomly assigned a skin tone. This investigation also found that participant's side to side head rotations significantly predicted whether or not participants signed the petition. Further analyses revealed that participants that have a tendency to continuously rotate their heads during a VRPT task signed more petitions than those who did not rotate their head as much. Overall, these results suggest that allowing users to embody and customize their avatar can increase prosocial behaviors. In addition, this investigation found that when participants fail to rotate their heads from side to side during a VRPT task, the amount of prosocial behaviors exhibited is on par with individuals who do not complete a VRPT task at all.

Authors' note

The work enclosed has not been published in whole or in part nor is it being considered for publication elsewhere. The data have not been presented at any scientific meeting, and there are no conflicts of interest to disclose. All figures and tables are original and were created by the authors. All authors have agreed to the submission of the manuscript in its current form and to the order of authorship.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD

Fernanda Herrera https://orcid.org/0000-0001-5616-0706

References

Ahn SJ, Le AMT and Bailenson J (2013) The effect of embodied experiences on self-other merging, attitude, and helping behavior. *Media Psychology* 16(1): 7–38.

Allport GW (1954) *The Nature of Prejudice*. Cambridge, MA: Addison-Wesley.

Asher T, Ogle E, Bailenson J, et al. (2018). Becoming homeless: a human experience. In: *ACM SIGGRAPH 2018 Virtual, Augmented, and Mixed Reality* (p. 7). ACM. Available at: https://vhil.stanford.edu/pubs/2018/becoming-homeless-a-human-experience/

Bailenson J (2018) Protecting nonverbal data tracked in virtual reality. *JAMA Pediatrics* 172(10): 905–906.

Bailenson JN, Swinth K, Hoyt C, et al. (2005) The independent and interactive effects of embodied-agent appearance and behavior on self-report, cognitive, and behavioral markers of copresence in immersive virtual environments. *Presence: Teleoperators & Virtual Environments* 14(4): 379–393.

Banakou D, Groten R and 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.

Barsalou LW (2010) Grounded cognition: past, present, and future. *Topics in Cognitive Science* 2(4): 716–724.

Batson CD, Early S and Salvarani G (1997a) Perspective taking: imagining how another feels versus imaging how you would feel. *Personality and Social Psychology Bulletin* 23(7): 751–758.

Batson CD, Polycarpou MP, Harmon-Jones E, et al. (1997b) Empathy and attitudes: can feeling for a member of a stigmatized group improve feelings toward the group? *Journal of Personality and Social Psychology* 72(1): 105.

Galinsky AD, Ku G and Wang CS (2005) Perspective-taking and self-other overlap: fostering social bonds and facilitating social coordination. *Group Processes & Intergroup Relations* 8(2): 109–124.

González-Franco M, Peck TC, Rodríguez- Fornells A, et al. (2014) A threat to a virtual hand elicits motor cortex activation. *Experimental Brain Research* 232(3): 875–887.

Herrera F, Bailenson JN, Weisz E, et al. (2018) Building long-term empathy: a large-scale comparison of traditional and virtual reality perspective-taking. *PLoS ONE* 13(10): e0204494.

Herrera F, Oh SY and Bailenson JN (2020) Effect of behavioral realism on social interactions inside collaborative virtual environments. *PRESENCE: Virtual and Augmented Reality* 27(2): 365–372.

Iacobucci D, Posavac SS, Kardes FR, et al. (2015) Toward a more nuanced understanding of the statistical properties of a median split. *Journal of Consumer Psychology* 25(4): 652–665.

Kalyanaraman S, Penn DL, Ivory JD, et al. (2010) The virtual doppelganger: effects of a virtual reality simulator on perceptions of schizophrenia. *The Journal of Nervous and Mental Disease* 198(6): 437–443.

Klimmt C, Blake C, Hefner D, et al. (2009) Player performance, satisfaction, and video game enjoyment. In: *International conference on entertainment computing* (pp. 1–12). Berlin: Springer. Available at: https://link.springer.com/chapter/10.1007/978-3-642-04052-8_1

Li BJ, Bailenson JN, Pines A, et al. (2017) A public database of immersive VR videos with corresponding ratings of arousal, valence, and correlations between head movements and self report measures. *Frontiers in Psychology* 8: 2116.

Lim S and Reeves B (2009) Being in the game: effects of avatar choice and point of view on psychophysiological responses during play. *Media Psychology* 12(4): 348–370.

Nowak KL and Biocca F (2003) The effect of the agency and anthropomorphism on users' sense of telepresence, copresence, and social presence in virtual environments. *Presence: Teleoperators & Virtual Environments* 12(5): 481–494.

Oh SY, Bailenson J, Weisz E, et al. (2016a) Virtually old: embodied perspective taking and the reduction of ageism under threat. *Computers in Human Behavior* 60: 398–410.

Oh SY, Shriram K, Laha B, et al. (2016b) Immersion at scale: researcher's guide to ecologically valid mobile experiments. In: *IEEE virtual reality (VR)* (pp. 249–250). IEEE. Available at:

https://vhil.stanford.edu/pubs/2016/immersion-at-scale-researchers-guide-to-ecologically-valid-mobile-experiments/

Peck TC, Seinfeld S, Aglioti SM, et al. (2013) Putting yourself in the skin of a black avatar reduces implicit racial bias. *Consciousness and Cognition* 22(3): 779–787.

Rizzo A, Klimchuk D and Mitura R (2004) The virtual classroom: a virtual environment for the assessment of attention processes in children with attention deficit hyperactivity disorder. In: *Virtual reality for public consumption, IEEE virtual reality 2004 workshop*, Chicago, IL, vol. 27. Available at: https://www.researchgate.net/publication/228446830_The_Virtual_Classroom_A_Virtual_Environment_for_the_Assessment_of_Attention_Processes_in_Children_with_Attention_Deficit_Hyperactivity_Disorder

Rosenberg RS, Baughman SL and Bailenson JN (2013) Virtual superheroes: using superpowers in virtual reality to encourage prosocial behavior. *PLoS ONE* 8(1): e55003.

Slater M, Spanlang B, Sanchez-Vives MV, et al. (2010) First person experience of body transfer in virtual reality. *PLoS ONE* 5(5): e10564.

Steuer J (1992) Defining virtual reality: dimensions determining telepresence. *Journal of Communication* 42(4): 73–93.

Todd AR and Galinsky AD (2014) Perspective-taking as a strategy for improving intergroup relations: evidence, mechanisms, and qualifications. *Social and Personality Psychology Compass* 8(7): 374–387.

Trepte S and Reinecke L (2010) Avatar creation and video game enjoyment: effects of life-satisfaction, game competitiveness, and identification with the avatar. *Journal of Media Psychology: Theories, Methods, and Applications* 22(4): 171–184.

Waltemate T, Gall D, Roth D, et al. (2018) The impact of avatar personalization and immersion on virtual body ownership, presence, and emotional response. *IEEE Transactions on Visualization and Computer Graphics* 24(4): 1643–1652.

Wobbrock JO, Findlater L, Gergle D, et al. (2011). The aligned rank transform for nonparametric factorial analyses using only ANOVA procedures. In: *Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 143–146). ACM. Available at: https://faculty.washington.edu/wobbrock/pubs/chi-11.06.pdf

Won AS, Bailenson JN and Janssen JH (2014a) Automatic detection of nonverbal behavior predicts learning in dyadic interactions. *IEEE Transactions on Affective Computing* 5(2): 112–125.

Won AS, Bailenson JN and Stathatos SC (2014b) Automatically detected nonverbal behavior predicts creativity in collaborating dyads. *Journal of Nonverbal Behavior* 38(3): 389–408.

Won AS, Perone B and Friend M (2016) Identifying anxiety through tracked head movements in a virtual classroom. *Cyberpsychology, Behavior, and Social Networking* 19(6): 380–387.

Yee N and Bailenson JN (2006) Walk a mile in digital shoes: the impact of embodied perspective-taking on the reduction of negative stereotyping in immersive virtual environments. In: *Proceedings of presence 2006: The 9th annual international workshop on presence*, August 24–26, Cleveland, OH.

Yee N and Bailenson JN (2007) The Proteus effect: the effect of transformed self-representation on behavior. *Human Communication Research* 33(3): 271–290.

Yee N, Bailenson JN and Ducheneaut N (2009) The Proteus effect: implications of transformed digital self-representation on online and offline behavior. *Communication Research* 36(2): 285–312.

Author biographies

Fernanda Herrera received her B.S. in Psychology from the University of Texas at Austin and her M.A. and Ph.D. in Communication from Stanford University. Her research focuses on examining the psychological and social effects of Augmented Reality and Virtual Reality experiences. More specifically, her research examines the effect of VR experiences on empathy and prosocial behaviors, assesses the effect of avatar representation and system affordances on social interactions inside collaborative virtual environments, and studies how face-to-face social interactions are impacted by the use of AR technology.

Jeremy N Bailenson is founding director of Stanford University's Virtual Human Interaction Lab, Thomas More Storke Professor in the Department of Communication, Professor (by courtesy) of Education, Professor (by courtesy) Program in Symbolic Systems, a Senior Fellow at the Woods Institute for the Environment, and a Faculty Leader at Stanford's Center for Longevity. He earned a B.A. cum laude from the University of Michigan in 1994 and a Ph.D. in cognitive psychology from Northwestern University in 1999. He spent four years at the University of California, Santa Barbara as a Post-Doctoral Fellow and then an Assistant Research Professor.