A Thousand Cuts

Virtual Reality and Inclusion

By Catherine McGregor

WHAT WOULD IT TAKE FOR YOU TO UNDERSTAND THAT RACISM IS MORE COMPLEX THAN BEING CALLED A NAME OR EVEN VIOLENCE AT THE HANDS OF POLICE?

What would it take for you to accept that we still live in a world where the color of someone's skin can get them killed, passed over for a job, denied a loan, denied housing, denied a fair shot?

What would it take for the reports of black bodies being gunned down by police to register as a pattern of racism, not merely insubordination or criminality?

What would it take for the outcries of racial injustice to register not as a sensitivity or a card to be played but as an unacceptable social reality?

What would it take for the data you see and that you say you "get"; what would it take for you to feel the reality behind that data?

– Dr. Courtney Cogburn

What would it take for you not just to feel bad and to empathize but to act and think differently?"

In the second of our series on the positive impact and potential future positive impact of technology, we take a deep dive into virtual reality through examining an ongoing project which aims to use VR technology to help foster greater understanding of the impact of the experience of racism. We spoke to the architects of the project Dr Courtney D. Cogburn of Columbia University and Professor Jeremy Bailenson of the Stanford Virtual Human Interaction Lab to get behind the platitudes.

IN SOMEONE ELSE'S SHOES

While there's obviously been great interest and take up of VR technology in the worlds of pornog-

raphy and gaming (neither of which are generally very inclusive) the fact that the technology and the industry is still so nascent does mean its potential is still wide open.

The fact that VR literally can allow you to walk in somebody else's shoes means it has huge potential to encourage empathy and to be used very powerfully in training situations for inclusion.

The quest for a fully sensory experience outside of one's own reality has a long genesis and could encompass a wide range of art forms and experiences. What is traditionally understood as the beginning of today's virtual reality journey happened in 1987. Jaron Larnier founded the Visual Programming Lab (VPL), and either coined or popularized the term "virtual reality". VPL developed a range of virtual reality gear including head-mounted goggles and gloves. The haptic experience (being able to touch and pick things up) of the new technology started to increase its potential and also the media's interest in it. During the 1990's, films such as The Lawnmower Man and The Matrix trilogies explored the concepts of virtual reality but real-life uptake was still some way off. Both Sega and Nintendo introduced VR gaming headsets that were flops.

During the 21st century VR has had an upturn, partly due to greater technological capacity and the development of smartphones and tablets. Recent years have seen the movement of many of the technological giants such as Google, Microsoft and Facebook invest and research into the field.

EMPATHY MACHINE

The immersive aspect of VR inevitably can lead to dangers which those working and researching in this field have highlighted. Exposure to violent video games has been shown to increase violent impulses in users and this is likely to be heightened in the use of immersive technology like VR. There's also the danger of reality drift which can occur through lengthened exposure to VR. Users become disorientated as to what is real and what isn't. I will return to consider the implications of this for inclusion.

In 2015 Chris Milk created a VR documentary entitled Clouds Over Sidra. In this film, viewers get to experience life in the Za'atari refuge camp in Jordan, home to over 80,000 displaced Syrian refuges. The film is narrated by Sidra, an eight year old girl. It does not show anything highly dramatic, rather the realities of day-to day-life in the camp. Describing seeing the film at the Tribeca Film Festival, Jeremy Bailensen, head of the Stanford Virtual Human Interaction Lab, in his forthcoming book 'Experience on Demand' explains how the experience does not employ any of the traditional tools used by film-makers to heighten empathy and engagement but rather simply uses the experience of having walked in the shoes of the refugees.

Chris Milk has spoken of his film, which is only eight and half minutes long, as being able to truly connect the experiences of the refugees to those who can help change that reality. Chris has deemed VR 'the ultimate empathy machine', highlighting the fact that it can connect people to other's in a way no other media can; and to that end he believes that it can change the world. Indeed, Clouds Over Sidra was created with the help of the United Nations and Samsung and the UN has reported that seeing it doubles the amount of people who donate to refugees.

Experiences like this mirror Jeremy Bailenson's own work in VR, which over the last twenty years has broadened the experience of walking in another's shoes - or even hooves. In his book "Experience on Demand" Jeremy describes a project spearheaded by a student of his, Joshua Bostick, which focused on giving participants the same experience as cows who were being reared for meat consumption. In this study volunteers would crawl

on all fours, wearing a special vest which allowed them to mirror the gait of a cow. Participants saw the cow they were controlling via a head-mounted monitor but could also 'feel' how the cow felt, by seeing their cow avatar get prodded, simultaneously getting prodded in the side themselves and also feel intense sound vibrations through the floor to mimic the sensation of getting an electric shock.

While running the VR experience, Jeremy and his researchers also ran a control group where participants viewed a video of cows going through the experiences in the VR simulation but without experiencing the sensations themselves. At the end of the experiments, those who had been through the VR simulation reported much greater empathy with the cows. In this case, the aims of Jeremy Bailenson and Bostick's study was not to promote whole-scale veganism but to make participants think about their consumption of meat.

In recent years the Stanford Virtual Human Interaction Lab has worked on a range of projects which explore using VR to enhance empathy. These have included projects where participants experience aging and homelessness. The lab also works with different organizations, using VR experience in training to increase empathy which certainly seems to have tangible effects in the real world. Jeremy explained:

"Overall we are very pleased with the level of user engagement and initial outcomes of using VR for empathy-related training within actual organizations. Fidelity saw a ten percent improvement in customer satisfaction scores after employing VR for empathy training, using VR to help employees form a better understanding of their customers. Recent work with Stanford Children's Hospital points to VR helping doctors trained using a perspective-taking simulation to feel more prepared to have tough conversations with patients. VR helps people to better manage their biases, by giving them practice on how to behave in intense situations which are rare but have a high impact in the real world."

> Chris has deemed VR 'the ultimate empathy machine' and has spoken of the fact that it can connect people to other people in a way no other media can; and to that end he believes that it can change the world.

But how easy is it to actually measure the effectiveness of VR experiences? It's all about measuring behavior, according to Jeremy. "When one is trying to change 'entrenched' behaviors - ones that are built up over decades and hard to change - it is not enough to use questionnaires. Therefore, we look at how people actually behave later on. For example, in a large-scale, longitudinal study where thousands of people have become 'homeless' in VR, our measure for effectiveness is whether or not they will actually sign a real petition that will increase their personal taxes to support affordable housing."

What's interesting though is that the work that Jeremy and his researchers have been doing suggests that VR can help create more empathetic responses in those who might traditionally not

be as able to show empathy.

"We have published a number of papers now that demonstrate VR empathy-training works best for people who have a hard time, in general, at taking the perspectives of others, for example, the first study in this paper: https://vhil.stanford.edu/pubs/2013/the-effect-of-embodied-experiences-on-self-other-merging-attitude-and-helping-behavior/.

The "Inter-Reactivity Index" measures the ability for people to engage in perspective taking as an individual difference, and VR has shown uniquely effective for those who score low on this measure."

It might then also prove very useful in training for diversity and inclusion, as a key aspect of inclusivity is often being able to understand the perspectives of others.

In recent years the Stanford Virtual Human Interaction Lab has worked to confront the issues of systematic racsim through their ongoing work with Professor Courtney D. Cogburn of Columbia University in the 1000 Cut Journey project.

1000 CUT JOURNEY

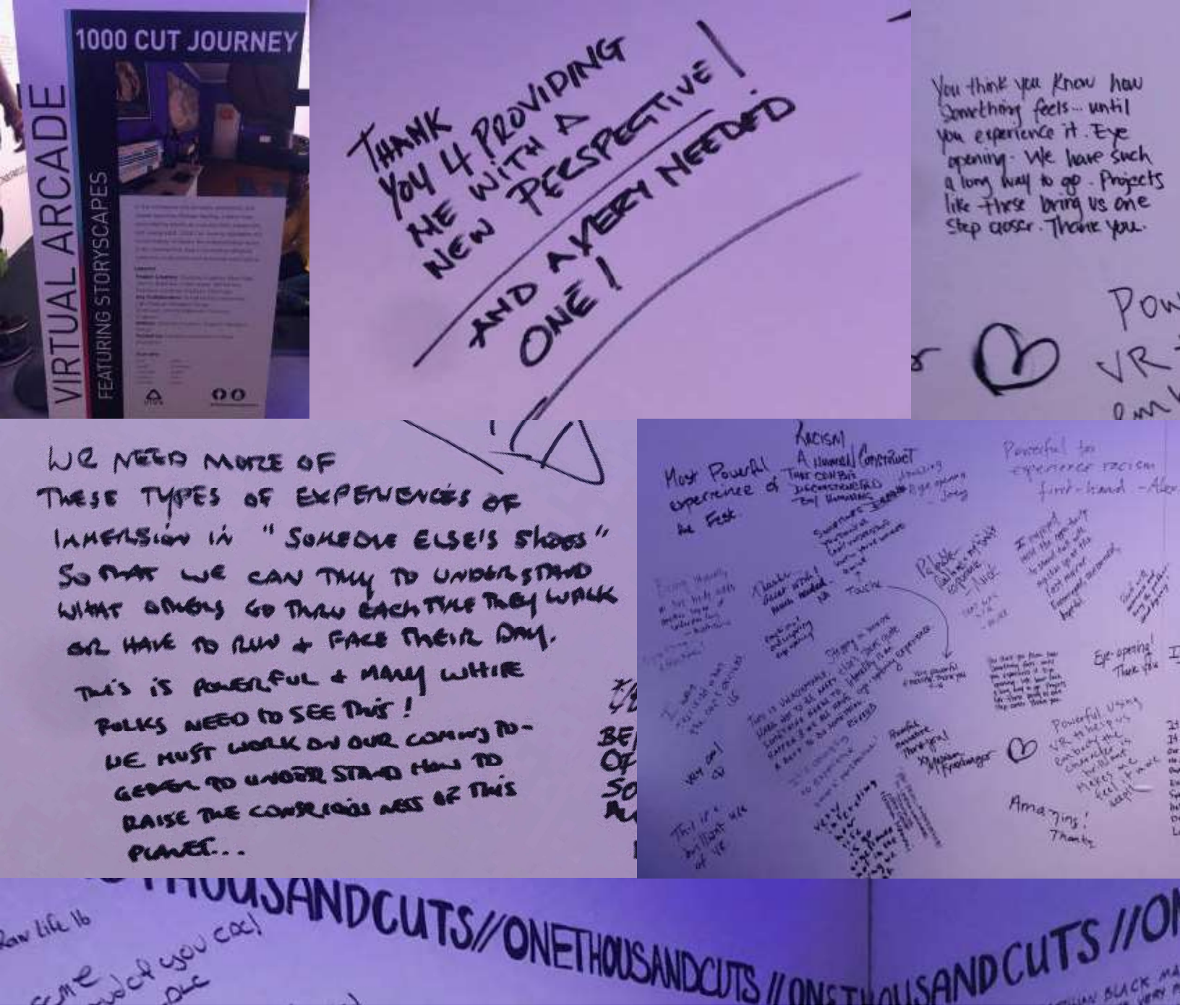
This project aims to create an immersive experience of racism from the perspective of a black male, allowing participants to become Michael Sterling who encounters racism as a young child, an adolescent, and as an adult. The genesis of the project is not just to stimulate empathy but to help others understand the social realities of racism, which is critical to promote effective and collective social action. Courtney Cogburn explains what led to the project:

"We communicate racism and racial inequality through language and symbols, such as through media representations that have an impact on health; it led me to think about how to leverage the media and engage a more complex narrative around racism. That eventually led me to VR - drawing on the adage of walking in someone else's shoes. My feeling was that many white people do not really understand racism or its impact: could I create something that gives them some first-hand insight into experiences of racism? Could we improve people's competence around racial inequality? My position is that emotional empathy is insufficient; we need people to act and think differently but what experience could we create that could push people to think in a different way?"

The target audience for 1000 Cut Journey is self-identified white liberals who are most likely to espouse beliefs of racial equality but may lack understanding of what that really means in terms of lived experiences of racism. There is an increasing body of research which suggests that true systemic change cannot be achieved unless the real pain of oppression is addressed, not just in its extreme forms but its everyday forms. It's those ideas that informed Remi Eddo Lodge's "Why I am No Longer Talking to White People About Race", where Eddo Lodge suggests that too often the burden of eliminating racism is falling to black people and that what is needed is greater consciousness and action from white people. It's this consciousness and action that 1000 Cut Journey seeks to stimulate.

The experience of the 1000 Cut Journey makes the participants experience three moments in the life of Michael Sterling a black man; first as a young child in a classroom, as an adolescent in an NYC neighborhood and then as young adult in a professional work environment. I asked Courtney why she chose the experiences of a black man over a black woman?

"At the moment, I think that our public discourse is heavily focused on black men and thus it was a salient point to begin engagement. I have written about and am personally quite aware, however, of the ways in which black women are made invisible in this discourse. Subconsciously I think that, for me, creating an experience about a black



woman would be more personal, complex and challenging. Trying to talk about a more deeply personal experience and the intersection of sexism and racism would be more complicated for me but it is definitely something I hope to explore in the future."

For Courtney, the key aspect of the experience is the accumulating effect of the different experiences throughout various phases of the life cycle. "The cumulative effect is important; we want participants to go through the three experiences. We wanted to give people the sense that it's not just a one time encounter and also that it starts very young. Similarly that it's not going to affect school or work because racism permeates every space and facet of society."

This amalgamation of experiences is derived from a range of narratives shared with her research team in the School of Social Work at Columbia University and also from empirical data drawn from multiple disciplines.

In the first segment of 1000 Cut Journey you are the young child Michael who is singled out and put on a 'time out' by his teacher for playing more roughly than his non-black peers, even though this isn't the case.The racism is arguably more subtle and

certainly less dramatic than the second segment but demonstrates how the narrative of the 'violent black male' starts incredibly young and how this trope is being imposed unfairly on Michael.

In the second segment you are a teenage Michael going out to play basketball with a friend; reports come in of police searching for a young black man who is implicated in a crime. The suspect is wearing something similar to you and your mother urges you to change. The segment concludes with you (the user) being stopped and searched by the police, including being forced onto your knees with your hands in the air.

In the final segment you're a 30 year old Michael at a job interview. You're waiting with a white candidate to be called in. The interviewer, ignoring you, walks over to the white candidate and cheerfully states "You must be our applicant from Yale." But he's not, Michael (you) is the candidate from Yale; the receptionist points out his mistake and you are ushered in for interview but you don't get the job.

The experience premiered at the Tribeca Film Festival in April, and afterwards, participants were asked to write their responses on a wall.

Responses were expressed differently but generally focussed on the fact that it's an incredibly powerful experience for the participants. In a number of cases the responses stated that this behavior is unacceptable and that something has to change.

These reactions to the experience heartened Courtney. "Based on the reactions throughout the festival I became more encouraged that we were doing something meaningful and impactful. All of the comments were focused on ideas like, 'I thought I understood this but I didn't', and that was really important to us, as were the delegates' emotional responses - people were upset and crying; really churned up."

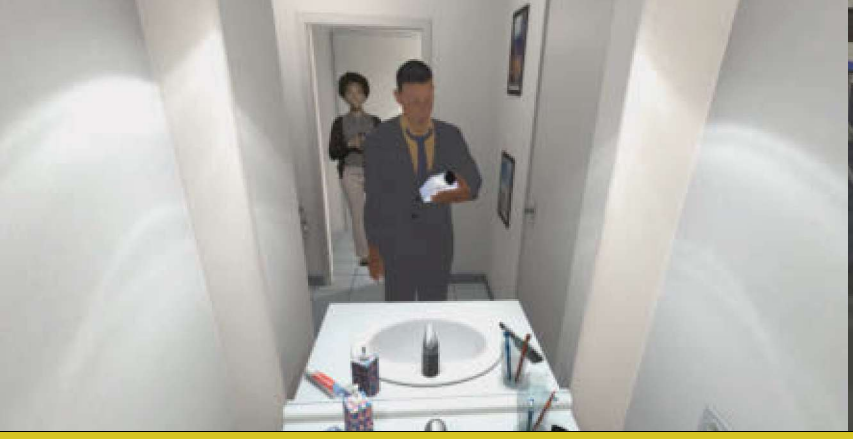
Jeremy Bailenson also sees the potential for greater application after the experience of premiering the piece at Tribeca. "Since then, we have been in conversations with dozens of organizations to install it as a part of their diversity and inclusion training. The idea is not to replace their existing curriculum, but instead to use it to augment the substantive information with an intense experience, one that is only possible in VR."

THE LIMITS OF THE VIRTUAL

One caveat that might be raised by the use of VR in training for inclusion is that it's not a magic bullet despite its effect often being very potent and it cannot replace the need for other aspects of training; a point highlighted by both Courtney and Jeremy in their conversations with me. Uncontrolled and unlimited access to VR experiences could lead to 'reality drift'; in inclusion terms, this could potentially lead to non- diverse participants imagining their diversity VR experience was the real thing and that they themselves were diverse. Without thoughtful application this could become an aspect of diversity tourism.

But the immediacy of walking in someone else's shoes is certainly a truly valuable part of training. Would the white manager in the third segment of 1000 Cuts Journey blithely assume a black candidate couldn't be the Yale candidate if he had been in Michael's shoes? Hopefully the emotional memory of that moment would stop future assumptions in their tracks.

1000 Cut Journey is a project which is still ongoing and Courtney and Jeremy and their teams are looking at further ways to apply this in real life. If your organization would like to find out more please email either Courtney at cc3803@columbia.edu or Jeremy at vhil@stanfordvr.com.



Experiencing the 1000 Cut Journey

I feel like a little child with a big world pitted against me.

I want to curl up and just cry.

I realize I am crying uncontrollably. At that point the feeling of the VR headset feels like some sort of shield or protection.

These are my immediate reactions at taking part in the 1000 Cut Journey developed by Professor Jeremy Bailenson and Dr Courtney D. Cogburn. I had wanted to take part in the experience for writing this article but had worried that knowing what it was through research for this story might lessen its effects. I was wrong.

During the first segment when Michael is a child in kindergarten and wrongly accused of playing roughly, I just felt confused. But during segments two and three, when Michael is stopped by the police as an adolescent and then faces discrimination in the workplace as an adult, I felt a very powerful throwback to feelings of childhood - feeling completely vulnerable, powerless, feeling threatened and unfairly treated, but there was nothing I could do about it.

I was also amazed at how attached I got to Michael's body in a short space of time. It felt quite sad to say goodbye to my 'Michael self', maybe because of what we had experienced together.

Did it change my perspective? Well, I am certainly predisposed to be empathetic but the VR experience takes this to another level. Afterwards I retreated to a nearby coffee shop to reflect. While there I read a newspaper report about the British rapper, Giggs, criticizing *The Sun* newspaper for their coverage of knife crime in London. It referenced an earlier Instagram comment from Giggs, which took to task Piers Morgan for his remarks on stop and search.

"Have you yourself ever experienced the traumas or violations, and a lot of the time abuse of power of being 'stopped and search' because you look like …"

Whilst I previously would have sympathized with this remark, it would have been academic. I did feel a subtle shift in perspective coming off of my virtual experience as Michael being stopped and searched.

PLOS ONE |

RESEARCH ARTICLE

Building long-term empathy: A large-scale comparison of traditional and virtual reality perspective-taking

Fernanda Herrera[1]*, **Jeremy Bailenson**[1], **Erika Weisz**[2], **Elise Ogle**[1], **Jamil Zaki**[2]

1 Department of Communication, Stanford University, Stanford, California, United States of America,
2 Department of Psychology, Stanford University, Stanford, California, United States of America

* fernhc@stanford.edu

OPEN ACCESS

Citation: Herrera F, Bailenson J, Weisz E, Ogle E, Zaki J (2018) Building long-term empathy: A large-scale comparison of traditional and virtual reality perspective-taking. PLoS ONE 13(10): e0204494. https://doi.org/10.1371/journal.pone.0204494

Editor: Brock Bastian, University of Melbourne, AUSTRALIA

Received: February 13, 2018

Accepted: September 10, 2018

Published: October 17, 2018

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: This research was funded by a grant from the Robert Wood Johnson Foundation; Grant ID# 72394. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

Abstract

Virtual Reality (VR) has been increasingly referred to as the "ultimate empathy machine" since it allows users to experience any situation from any point of view. However, empirical evidence supporting the claim that VR is a more effective method of eliciting empathy than traditional perspective-taking is limited. Two experiments were conducted in order to compare the short and long-term effects of a traditional perspective-taking task and a VR perspective-taking task (Study 1), and to explore the role of technological immersion when it comes to different types of mediated perspective-taking tasks (Study 2). Results of Study 1 show that over the course of eight weeks participants in both conditions reported feeling empathetic and connected to the homeless at similar rates, however, participants who became homeless in VR had more positive, longer-lasting attitudes toward the homeless and signed a petition supporting the homeless at a significantly higher rate than participants who performed a traditional perspective-taking task. Study 2 compared three different types of perspective-taking tasks with different levels of immersion (traditional vs. desktop computer vs. VR) and a control condition (where participants received fact-driven information about the homeless). Results show that participants who performed any type of perspective-taking task reported feeling more empathetic and connected to the homeless than the participants who only received information. Replicating the results from Study 1, there was no difference in self-report measures for any of the perspective-taking conditions, however, a significantly higher number of participants in the VR condition signed a petition supporting affordable housing for the homeless compared to the traditional and less immersive conditions. We discuss the theoretical and practical implications of these findings.

Introduction

Empathy, the ability to share and understand someone else's emotions, is an imperative component of successful social interactions [1]. Empathy has been shown to increase understanding and motivate prosocial behaviors [2–5]. Given its importance and positive effects on

intergroup social interactions, researchers, artists, and tech companies have tried to find novel ways to increase empathy via different types of empathy interventions.

In 2015, virtual reality (VR) was described as the "ultimate empathy machine" since it allows people to viscerally experience anything from another person's point of view [6]. VR experiences, or immersive virtual environments (IVEs), are computer generated, 3D environments where people can move around freely and interact with their surroundings. IVEs replace the user's perceptual input of the real world with perceptual input from a virtual world, and make users feel like they are actually inside a virtual environment. De la Peña and colleagues [7] argue that VR's ability to elicit *presence*, the user's subjective feeling of being inside an IVE, allows users to more deeply understand perspectives other than their own. Since then, companies and institutions like Facebook, HTC, and the United Nations have allocated millions of dollars toward programs such as Oculus's "VR for Good" and HTC's "VR for Impact" [8–9], which try to use VR to promote empathy and social welfare.

Interest in VR as an empathy tool has led to the increased production of IVEs designed with the sole purpose of increasing empathy. These experiences place users in novel environments, showing them what it would be like to experience a specific situation from someone else's perspective. Extensive research shows that taking the perspective of someone else (i.e., imagining what it would be like to be someone else) can be an effective method of promoting empathy and motivating prosocial behaviors [2–5, 10–12]. However, there is a need for more empirical evidence validating perspective-taking in VR as an effective way to promote empathy, especially now that multiple empathy-driven IVEs are publicly available online, and past research demonstrates that traditional perspective-taking does not always lead to positive outcomes [13]. Thus, the current investigation focuses on 1) comparing the short and long-term effects of a traditional perspective-taking task against a VR perspective-taking task, and 2) exploring the technological and psychological mechanisms that make perspective-taking an effective method of promoting empathy and prosocial behaviors.

Positive and negative effects of perspective-taking

Traditional perspective-taking tasks (i.e., where participants are asked to imagine what it would be like to be someone else under specific circumstances), have been used extensively in the literature, and can effectively increase the empathy a perceiver feels for a specific social target. These tasks can be particularly powerful since they often result in feelings of inclusion and self-other merging [10], reduce prejudice and attenuate negative stereotypes [14], can help create and maintain social bonds [15], facilitate social interactions [16], and motivate altruistic helping behaviors [12]. In addition, perspective-taking does not only increase empathy for specific people, but can increase empathy toward entire stigmatized groups. For example, Batson and colleagues [3] found that participants who were asked to take the perspective of a member of a stigmatized group (e.g., people diagnosed with AIDS or convicted criminals) reported more positive attitudes not only toward a specific member of that group, but toward the stigmatized group as a whole.

The literature suggests that there are differential effects depending on the type of perspective-taking employed. For example, past studies have demonstrated that *imagine-other* perspective-taking tasks, where participants are instructed to imagine how someone else feels about a specific situation, lead to empathy and an altruistic motivation to help. On the other hand, *imagine-self* perspective-taking tasks, where participants are instructed to imagine how they would feel if they were in someone else's situation, lead to empathy, personal distress, and an egoistic motivation to help [17–18]. Lamm, Batson, and Decety [19] have also found that there are differences in the neural mechanisms that are activated when imagine-self vs

imagine-other perspective-taking tasks are used. Batson and colleagues [17] suggest that if the goal of the perspective-taking task is to maximize motivation to help, employing an imagine-self perspective-taking task may be more effective than an imagine-other approach.

Despite the general consensus that perspective-taking leads to positive outcomes [3, 13], under certain circumstances, perspective-taking can backfire and not only negate the positive outcomes outlined above, but also increase the stereotyping of others [20–22]. Perspective-taking may even cause people to blame members of a stigmatized group for their own situation even when they may not be at fault [3], and generate adverse effects in competitive settings. For example, Pierce and colleagues [13] demonstrated that taking the perspective of a competitor led to more unethical behaviors, further confirming that the positive effects of perspective-taking are contextually bound.

Media and empathy

Apart from traditional perspective-taking, researchers have tried to promote empathy through a variety of mediated perspective-taking tasks (i.e., print media, interactive narratives, video games, and online interfaces that prompt users to take on a perspective other than their own). Mediated perspective-taking tasks provide additional information to the participant or user about a specific situation instead of relying solely on the user's imagination. This may be particularly helpful when participants have had limited contact with or have very little or erroneous information about a social target [23]. However, it is important to note that each type of media provides information through specific channels that rely on one or a combination of the human senses. A book, for example, provides visual content and therefore relies on the visual system. A video game, however, provides visual, auditory, and sometimes haptic stimuli to the user, and therefore relies on the visual, aural, and proprioceptive systems.

Different types of media also vary in their level of immersion and interactivity. Immersion is an objective and descriptive measure of the extent to which a particular medium is able to create and sustain a "vivid illusion of reality" [24]. Interactivity refers to the extent to which users can exert influence over the content of the mediated environment in real-time [25]. The fact that different types of media such as books, TVs, computers, and VR fall under different levels of immersion and interactivity, may help explain why extant research has mixed results regarding mediated perspective-taking tasks and their effects on empathy and prosocial behaviors.

When it comes to print media, specifically the sharing of information via text, past research shows that providing new information about specific events (e.g., natural disasters) or groups of people (e.g., Native Americans) reduces social distance, increases empathy, and often results in prosocial behaviors [26–27]. Oliver and colleagues [28] highlight that the way in which information is formatted affects reader's empathic responses, and found that information formatted as a narrative elicited more empathic responses from readers than fact-driven stories. However, providing new information does not always change people's stereotypes of others or mitigate implicit biases [29].

Interactive interfaces such as virtual narratives (e.g., online role playing games) and video games have also been considered effective platforms to promote empathy [30–31]. For example, Hailpern and colleagues [32] created an interactive system that emulated the effects of aphasia through the real-time distortion of written text. This system gave participants a first-hand experience of what it would be like to have aphasia. Participants who interacted with the system reported feeling more empathetic and being more understanding of people with speech impediments than their counterparts who did not interact with the system at all. These interfaces are more immersive than print media because they give the user the ability to interact

with digital representations of other users or algorithms (i.e., avatars or agents) and with their online environment in real-time. However, it is important to note that not all simulations of disabilities lead to positive results. For example, Brown [33] simulated schizophrenia by having participants listen to intrusive and distracting sounds for 15 minutes and found that participants' attitudes toward schizophrenics deteriorated after the simulation. In a different study, researchers simulated blindness by covering participants' eyes. Their results found that even though the simulation led to empathic concern, it also spread misinformation and reinforced stereotypes [34].

Higher interactivity has been linked to higher feelings of presence, or the feeling of "being there" [25], and can provide a more engaging and personalized experience [35]. Vorderer, Knobloch and Schramm [36] explored the effects of three different levels of interactivity (none, little, high) on empathy and entertainment while participants watched a movie. Results showed that, for participants with higher cognitive ability, increased interactivity led to greater empathy toward the movie's protagonist. This study suggests that higher levels of interactivity, as a feature of immersion, can increase empathy. Using a different media platform, Behm-Morawitz, Pennel, and Speno [37] created a digital gaming app designed to reduce prejudice toward African Americans. White participants, who interacted with the game using Black avatars reported having more favorable beliefs toward African Americans and were more willing to support "pro-minority" initiatives compared to participants who used White avatars. The collective results of these studies provide evidence in support of mediated perspective-taking as an effective way to promote empathy.

Perspective-taking tasks and VR

An IVE is a fully immersive and interactive computer-generated environment that gives the user the feeling of being somewhere other than where they are in the physical world. VR systems block out the perceptual input from the real world and replace it with perceptual input from a virtual environment that surrounds the user, is fully responsive to the user's actions, and elicits feelings of presence. Because of these affordances, VR allows users to vividly and viscerally experience any situation as if it were happening to them from any perspective [38]. Unlike traditional media, the high level of immersion, feeling of presence, and the ability to vividly experience any situation from any perspective, may uniquely position VR as an effective perspective-taking medium.

Additionally, imagining what it would be like to be someone else is cognitively taxing [39–40]. When performing a VR perspective-taking task, however, fewer mental resources are required to create a specific environment or situation because it is all rendered digitally, and the user can solely focus on acting and reacting within the experience. This positions VR perspective-taking at an advantage in terms of accurate content, in comparison to traditional perspective-taking, since participants do not have to rely on their preexisting schemas or biases [41]. It is also methodologically advantageous since it makes it possible for all participants to undergo the exact same experience, as opposed to the lack of experimental control that occurs when participants use their imagination during traditional perspective-taking tasks [42].

Embodied cognition theory postulates that cognition is an interaction of the body and mind [43] that takes place within the context of a specific environment [44]. Past studies show that physical movement can improve a participant's performance while completing cognitive tasks, and that the physical experience of a particular environment can have an effect on both perceptions and behaviors [45–46]. VR allows users to move and interact with their surroundings as if they were actually there through a combination of physical body movements (e.g., walking, extending their arms to reach an object, or turning their head around to examine

their surroundings) and button presses using controllers. These affordances, not available in less immersive media, allow users to gather spatial information about the virtual environment using the same perceptual systems they would use to gather spatial information about the real world. Furthermore, past research has shown that the gathering of more perceptual information results in a more accurate mental representation of the physical environment [47]. Thus, the user's ability to actively engage with and move around inside an IVE may result in improved cognition due to the additional information users are able to collect through physical movement [48], and may have a positive effect on empathy and prosocial behaviors.

In previous work, embodying the perspective of other groups (e.g., colorblind, schizophrenic, or elderly individuals) has increased helping behaviors and decreased prejudice against outgroup members [5, 49–50]. One study specifically compared the effectiveness of a traditional perspective-taking task and a VR perspective-taking task at mitigating ageism [40]. The results showed that the negative effect of threat on ageism was more effectively reduced by the VR perspective-taking task than by imagination alone. However, when the participants felt they were personally threatened by the elderly, neither VR perspective-taking nor imagination mitigated ageism.

Past studies have also shown that when light-skinned participants embody a dark-skinned avatar (i.e., a digital representation of the user) and interact in a virtual environment, implicit biases toward dark-skinned people are significantly reduced [51–52]. In a different line of studies, Ahn, Bailenson, and Park [53] compared the effects of cutting down a tree in a traditional perspective-taking task, a VR perspective-taking task, and a less immersive, mediated perspective-taking task where participants watched a video of someone cutting down a tree from the first-person perspective. Results showed that participants in the VR condition reported higher environmental behavior intentions and locus of control than their counterparts. After a week, these effects persisted for participants in the VR condition but deteriorated for the other conditions. A different study showed that embodying animals (e.g., a cow going to the slaughter house or a coral suffering the effects of ocean acidification) led to higher self-nature overlap when participants performed a VR perspective-taking task rather than just watching a video. After one week of treatment, however, the positive effect of VR mitigated [38]. To our knowledge, these are the only extant studies that have looked at the effect of VR perspective-taking past the day of treatment [38, 51–53].

Despite the encouraging results outlined above (e.g., VR perspective-taking leads to more positive outcomes than traditional perspective-taking or less immersive mediated perspective-taking tasks), these studies are limited by their small sample size (on average, less than 30 participants per condition). In fact, 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 [54].

In order to address the methodological issues of sample homogeneity and external validity, this investigation differs from past studies in that 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. For this investigation, in addition to recruiting participants in a medium-sized western university, the mobile lab unit was taken to museums, shopping centers, senior citizen centers, schools, and tech fairs with the goal of recruiting participants who represent a part of the population that has not yet been represented in psychological VR studies.

Additionally, previous studies have used self-report data to measure future intentions (e.g., 40 and 53), and while these measures often align with actual behavior, that is not always the case. For example, Rosenberg, Baughman, and Bailenson [55] measured intention to help and actual helping behaviors, and although the participants in the experimental condition exhibited more helping behaviors than those in the control condition, there was no significant

difference in self-reported intention to help between the two conditions. Similarly, Bailenson and colleagues [56] found no differences in their self-report measures but detected significant differences when examining nonverbal behavioral measures. In order to address these issues and gain a more thorough understanding of the effects of different types of empathy interventions, this investigation employs both self-report and behavioral measures.

Overview of studies

Two studies were conducted in order to compare the effects of different types of perspective-taking interventions on empathy and prosocial behaviors. Study 1 was a longitudinal investigation that compared the effects of a VR perspective-taking task against a traditional, narrative-based perspective-taking task. In Study 2, we sought a more nuanced approach to further explore the mechanisms that caused differences, or lack thereof, between the traditional and the VR perspective-taking tasks. In order to do so, we allowed for contrasts between a larger set of control conditions and developed four different empathy interventions: a fact-driven information intervention, a narrative-based perspective-taking task, and two mediated perspective-taking tasks (low vs. high immersion) in order to more accurately explore the effect of immersion and type of empathy intervention on elicited empathy and prosocial behaviors toward the homeless.

The homeless population was chosen as the social target in this work because the homeless are considered an extreme outgroup that people often struggle to empathize with [57–58]. In addition, the distinction between the in-group and outgroup is clearly defined and, unlike race or gender, it applies to a larger population because, under specific circumstances, anyone can become homeless.

Study 1

Study 1 compared the short and long-term effects of a traditional, narrative-based perspective-taking (NPT) task and a VR perspective-taking (VRPT) task at the time of the intervention and over the course of eight weeks. As outlined above, an extensive line of research shows that perspective-taking is a powerful exercise that often results in increased empathy and prosocial behaviors toward a specific social target (e.g., stigmatized groups). However, little is known about the duration of these effects. Past research examining the longitudinal effects of perspective-taking tasks (both traditional and mediated) has only looked 1 to 2 weeks after the time of the intervention [3, 38, 51–52].

Since VR affords vivid experiences in a cognitively effortless way, a perspective-taking task through the use an IVE may be more effective at promoting empathy and prosocial behaviors than a traditional, narrative-based perspective-taking task. Additionally, participants in IVEs are surrounded by virtual stimuli that they can interact with, which may make them feel greater presence inside the virtual environment. Thus, we predict VRPT will be more effective than NPT at eliciting empathy and prosocial behaviors at the time of the intervention and over the course of eight weeks.

Method

All procedures and materials were approved by the Ethical Committee of the Institutional Review Board at Stanford University, IRB-34677. Written consent was obtained from all participants. The individual in this manuscript has given written informed consent (as outlined in PLOS consent form) to publish these case details.

Participants

Individuals recruited from a medium-sized western university and from around the San Francisco Bay Area formed an initial sample of 130 participants. Participants received a $100 Amazon gift card for completing the four parts of the study. All participants who agreed to participate were asked to come into the laboratory to complete the first part of the study. Of the 130 participants, 13 participants failed to complete some part of the study and were excluded from the analysis. The final sample (N = 117) consisted of 40 men, 75 women, and 2 participants who identified as other. The ages ranged between 15 and 57 (M = 22.94, SD = .95). Of these 117 participants, 33 (28%) were White Caucasian, 7 (6%) were Hispanic, 8 (6.8%) were Indian, 49 (41.8%) were Asian, 7 (6.8%) were African American, and 13 (11.1%) were multiracial.

Design and procedure

In this mixed design, participants were randomly selected into one of two conditions, either the NPT (n = 56) or VRPT (n = 61) condition. All participants completed a pre-intervention questionnaire which included demographic questions, the Interpersonal Reactivity Index (IRI), and the Beliefs about Empathy scale. These scales were used in order to conduct sample checks and make sure that random assignment was successful. Participants then completed either the NPT or the VRPT task.

In both conditions, participants performed an *imagine-self* perspective-taking task. An imagine-self perspective taking task was chosen over an imagine-other task to ensure that the NPT and VRPT tasks were qualitatively the same apart from the medium being used to deliver the intervention. In the NPT condition, participants imagined what it would be like if they became homeless, and in the VRPT condition, participants experienced what it was like to become homeless inside an IVE from the first-person perspective. However, in order to control for the level of detail, accuracy of information, implicit biases against the homeless, and the scenarios that the participants would either imagine in a traditional perspective-taking task or virtually experience in the VRPT task, we standardized a narrative across the two conditions to guide participants through their respective perspective-taking tasks.

The narrative begins with the participants sitting in their apartment after losing their job and realizing rent is due. Despite selling most of their belongings, participants are not able to raise enough money to pay rent and are evicted from their apartment. Forced to live out of their car, they prepare themselves for the night by trying to find their toothbrush and other items needed to brush their teeth. Participants suddenly hear a police siren and are approached by a police officer who discovers the participants are living out of their car. Due to a city ordinance prohibiting the use of cars as homes in public spaces, the car is impounded. Participants are now traveling on a bus at night for shelter and warmth, when they are warned that there are two men onboard who may serve as a threat to them. One man may try to get unpleasantly close to the participant while the other may try to steal the participant's backpack. In the bus, participants also interact with other non-threatening homeless people and learn about their experiences. The events in this narrative were adapted from *Hotel 22*, a documentary depicting how the homeless use public transportation for shelter at night [59], and existing interviews with homeless people in the local area. The narrative represents the lived experiences of veterans, families crippled by medical bills, victims of domestic violence, and drug addicts.

In the NPT condition, participants were given a narrative version of the storyline described above. The narrative was a written account from a first person point of view of what it would be like to become homeless (see Text A in S1 Appendix for full version). Before participants read the narrative, they were instructed to imagine the content of the narrative as if it was happening to them. The instructions and method were adapted from Batson et al. [3].

In the VRPT condition, participants experienced what it was like to become homeless inside an IVE. The narrative that the participants followed was the same narrative from the NPT condition. Participants were seated throughout the experience, but navigated the IVE by moving their head and leaning their torso to look where they wanted, and selected objects by first looking at the object they wanted to interact with and then clicking on their mouse.

Participants in this condition wore a lightweight head-mounted display (HMD) with a resolution of 960 x 1080 pixels and a refresh rate of 75 frames per second for each eye with an average latency of 13ms. The HMD used was an Oculus Rift DK2 (Fig 1). Participants' head translation (x, y, z position) and orientation (yaw, pitch, and roll) was tracked throughout the experience using an infrared camera (Oculus DK2 IR camera).

Unlike the NPT condition, where participants imagined what it was like to become homeless, participants in the VRPT condition experienced the narrative events described above inside the IVE. The VR experience consisted of three scenes that matched the three locations described in the narrative. This perspective-taking task differed markedly from the NPT condition in that participants in this condition were able interact with certain objects in the virtual environment. Each scene contained an interactive task where participants had the opportunity to make choices and see the consequences of those choices in real-time. For example, participants in the NPT condition were told to imagine how difficult it would be to find their toothbrush in their car, whereas participants in the VRPT condition had to actually find the toothbrush inside the car by moving their head to look around the car. None of the choices made by participants affected the length or the course of the overall VR experience. Fig 2 shows the three scenes from the participant's point of view.

Both perspective-taking tasks, regardless of delivery method, lasted approximately 15 minutes. Upon completion of the post-intervention questionnaire, participants left the laboratory and received three follow up surveys over the next eight weeks. The follow up surveys were sent via email exactly two, four, and eight weeks after the intervention took place. Each follow-up survey was automatically sent by the Qualtrics platform as a link that would expire after 24 hours of being sent. When participants completed the surveys, researchers were able to download the data from that same platform without any identifying information.

Participants who did not meet the deadline were excluded from the analysis. At Time 0 (immediately after the intervention) 2 participants were excluded due to technical difficulties. At Time 1 (2 weeks after the intervention), 5 participants failed to complete part two within the given time frame. At Time 2 (4 weeks after the intervention) 2 participants did not complete part three. At Time 3 (eight weeks after the intervention) 4 participants did not complete the survey within the allotted time. All participants were debriefed once the study had been completed.

Measures

Population variables. Interpersonal Reactivity Index. The Interpersonal Reactivity Index (IRI) is a 28-item scale with four subscales answered on a 5-point Likert scale ranging from "does not describe me well" to "describes me well". The IRI measures individual differences in empathy by assessing the participant's tendency to adopt the point of view of others (M = 3.86, SD = 0.57, *Cronbach's alpha* = .77), empathic concern for others (M = 3.97, SD = 0.59, *Cronbach's alpha* = .80), and personal distress in tense situations with others (M = 2.7, SD = 0.83, *Cronbach's alpha* = .87) [60]. Sample items include "I sometimes find it difficult to see things from the 'other guy's' point of view" and "I am often quite touched by things that I see happen". There were two main reasons why the fantasy subscale was excluded. The first and main reason for excluding the fantasy subscale was that this was a mobile experiment. One of the

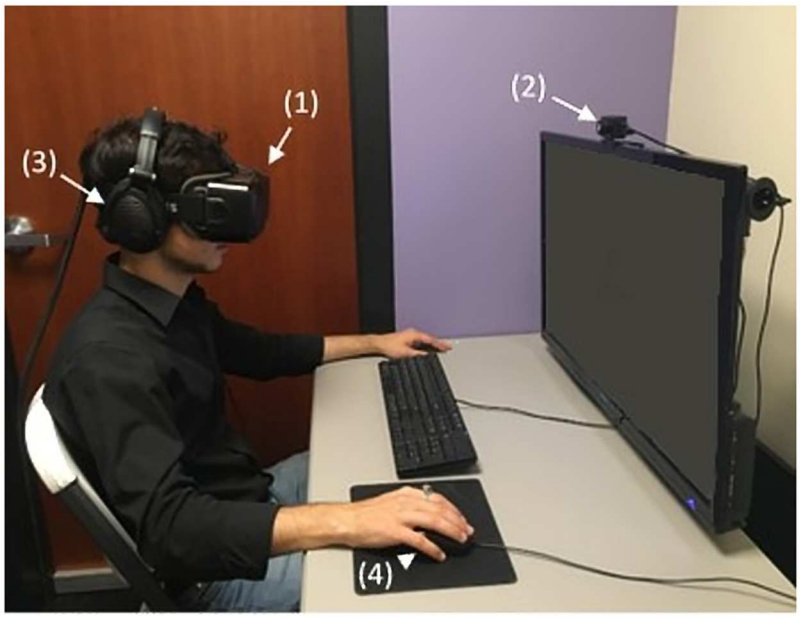


Fig 1. Equipment used for VRPT condition. (1) Oculus DK2 (2) Oculus DK2 infrared camera (3) headphones & (4) mouse. The individual in this manuscript has given written informed consent (as outlined in PLOS consent form) to publish these case details.

https://doi.org/10.1371/journal.pone.0204494.g001

limitations of having a mobile VR lab and running with volunteers (mostly composed of attendees at museums and events) is that time is an issue for most participants. Given we were already asking multiple questionnaires and the intervention lasted 15 minutes, we needed to be very selective about what questions to include. The second reason is that the fantasy subscale taps into a participant's ability to "transpose themselves imaginatively into the feelings and actions of fictitious characters from books, movies, and plays" [60]. The fantasy subscale specifically measures the ability to embody (through imagination) a fictional character and feel what they are feeling, something that was less relevant to our intervention. We did not expect differences in any of the IRI's subscales across conditions since they were collected before treatment.

Beliefs about Empathy Scale. The Beliefs about Empathy scale is a 12-item scale answered on a 7-point Likert scale (1 = strongly disagree, 7 = strongly agree) that measures the extent to which people believe that empathy is something that can be controlled. Sample items include "people can always change how much empathy they generally feel for others" and "when a person feels empathy for someone, they can't stop feeling empathy, even if they want to stop." The scale contains two different subscales, controllability (M = 4.31, SD = 1.06, *Cronbach's alpha* = .82) and implicit theories (M = 4.51, SD = 1.33, *Cronbach's alpha* = .92). The overall reliability of this scale was good, *Cronbach's alpha* = .87. This scale is used as a sample check to make sure that there are no significant differences in beliefs about empathy across conditions, as previous research has demonstrated that participants who believe they can control their empathic responses actually exert more empathic effort than those who believe they have no control [61].

Outcome variables. Manipulation Check. In order to assess if participants in the VRPT condition felt more spatially present than participants in the NPT condition, participant's open-ended responses to the question "What are your thoughts about your experience so far?" were coded. Two coders, blind to condition, coded the participants' responses independently.

Fig 2. Scene progression for VR perspective-taking task. Top: Participant's point of view in the apartment scene. Middle: Participant's point of view in the car scene. Bottom: Participant's point of view in the bus scene.

https://doi.org/10.1371/journal.pone.0204494.g002

The coding scheme employed was deductive and adapted from the presence questionnaires used in Nowak and Biocca [62] and Bailenson and Yee [63]. Participants received a 1, 0, or -1 depending on their responses.

Participants received a "1" if they explicitly mentioned that the experience or narrative felt real, was realistic, or immersive, if they felt immersed, "there", inside the environment, or if they were physically or spatially affected by what they experienced in VR or what they imagined. Participants received a "0" if their response was not related to presence or immersion, and a "-1" if they explicitly mentioned that the experience or narrative was not real or realistic, that they did not feel "there", immersed, or physically affected by what they experienced in VR or what they imagined (M = 0.15, SD = 0.47).

Inclusion of the Other in the Self. The Inclusion of the Other in the Self (IOS) scale is a single-item, pictorial measure of closeness and connectedness. In the IOS scale, participants select the picture that best represents their relationship with the average homeless person [64]. Pictures are Venn-like diagrams of two circles overlapping, with each circle representing a homeless person and the self, respectively. The pictures are coded from 1 to 7 with the larger numbers indicating a closer relationship with the homeless (i.e., a larger degree of overlap between the two circles) (M = 3.09, SD = 1.92).

Empathy. Using a 7-point Likert scale (1 = Not at All, 7 = Extremely), participants were asked to rate the extent to which they felt *softhearted*, *touched*, *sympathetic*, or *compassionate* throughout the intervention. The results of these four questions were used to create an index of empathic concern. This measure was adapted from Batson, Early and Salvarani [17]. The overall reliability of the index was good, *Cronbach's alpha* = .88 (M = 5.05, SD = 1.21).

Personal Distress. Using a 7-point Likert scale (1 = Not at All, 7 = Extremely), participants were asked to rate the extent to which they felt *uneasy*, *troubled*, *distressed*, or *disturbed* throughout the intervention. The results of these four questions were used to create an index of personal distress. This measure was also adapted from Batson, Early and Salvarani [17]. The overall reliability of the index was good, *Cronbach's alpha* = .87 (M = 4.58, SD = 1.33).

Dehumanization Scale. The Ascent of Man measure of Dehumanization is a single-item measure of blatant dehumanization [65]. This measure was chosen to examine whether or not traditional or VR perspective-taking tasks could prevent dehumanization toward the homeless since past research has demonstrated that people are willing to overtly describe outgroup members as less evolved than members of their own group [57–58, 65]. Participants are asked to rate how evolved the average member of the homeless population is by looking at the "Ascent of Man" evolution picture that depicts a primate evolving into a human (0 = not evolved/primate, 100 = fully evolved/human) (M = 89.06, SD = 16.09).

Social Presence Scale. The 6-item social presence scale assesses how present participants felt with virtual humans inside the IVE in the VRPT condition and was adapted from Nowak and Biocca [62]. Participants were asked to rate how strongly they felt the virtual humans were with them inside the IVE using a 5-point Likert scale (1 = not at all, 5 = very strongly). The reliability of the scale was good, *Cronbach's alpha* = .92 (M = 1.49, SD = 1.55). Sample items include "How strongly did you sense that the virtual humans were aware of your presence" and "How strongly did you sense that the virtual humans were present?"

Attitudes toward the Homeless. The 7-item attitudes scale was answered on a 9-point Likert scale (1 = strongly disagree, 7 = strongly agree). This scale adapted from Batson et al. [3] to gauge attitudes toward the homeless (M = 6.48, SD = 1.18, *Cronbach's alpha* = .83). Sample items include "Most homeless people could have avoided becoming homeless" and "Our society does not do enough to help homeless people." Two out of the seven items from the attitudes toward the homeless scale were reverse coded. Higher scores indicate more positive attitudes toward the homeless.

Behavioral measures. Agreement with Proposition A. Proposition A, a real proposition at the time the study was conducted, supported increasing affordable housing for vulnerable populations in the San Francisco Bay Area (Text B in S1 Appendix). After participants read Proposition A, they were asked the extent to which they agreed with the proposition using a 5-point Likert scale (M = 4.07, SD = .76) ranging from "not at all" to "completely". This measure was implemented at Time 0 (immediately after the intervention).

Signing Petition Supporting Proposition A. Participants were also asked whether or not they were willing to sign a petition supporting Proposition A. At this point, participants were reminded that the passing of the proposition would mean an increase in their taxes. Instead of simply selecting *yes* or *no*, participants either signed the petition or left the petition page blank (71.8% of all participants signed the petition). The researcher was not in the room while the participant read or responded to the petition, so as not to influence the participant's behaviors. This measure was also implemented at Time 0 and participants were debriefed that the petition was not a real petition upon completion of the study eight weeks later.

Donation Question. After either empathy intervention, participants were asked if they wanted to donate part of their compensation to a homeless shelter using an 11-point Likert scale ranging from "$0" to "$10" (M = 5.73, SD = 3.78). Participants were compensated in full regardless of whether they chose to donate any money or not. However, participants were not aware they were going to receive full compensation until they were debriefed after the study had been completed. Again, the researcher was not in the room with the participant while they chose whether or not to donate money. This measure was also implemented immediately after the intervention.

Letter Writing. Participants were asked to write a letter to their elected officials regarding the issue of homelessness at Time 1 (2 weeks after the intervention), and a letter to a friend describing how they felt and what they had learned about the issue of homelessness at Time 3 (eight weeks after the intervention). For instructions see Text C in S1 Appendix. These behavioral measures were added in order to 1) examine the extent of each participant's civic engagement regarding the issue of homelessness, and 2) conduct linguistic analysis to further understand participants' emotional states through their written language. Past research demonstrated that analyzing the way that people write (function words) rather than what they write about (content words), is a reliable measure of different psychological and emotional states [66]. We focused on analyzing 7 categories: word count, positive emotion, negative emotion, social, anxiety, I, and we. These categories were specifically chosen to quantify affect, and examine the extent to which participants included or excluded themselves as part of the solution when writing about the issue of homelessness [67]. All linguistic analyses were performed using Linguistic Inquiry and Word Count (LIWC) software.

Agreement with Measure B. Similar to Proposition A, Measure B also advocated for increasing affordable housing for vulnerable populations (Text D in S1 Appendix). This measure provided different, yet related, information about what can be done to help the homeless in order to assess support for helpful initiatives over time. This behavioral measure was implemented at Time 2 (4 weeks after the intervention). Participants were asked to read Measure B and report the extent to which they agreed with it using a 5-point Likert scale (1 = not at all, 5 = completely) (M = 4.19, SD = .92).

Results

Population variables

A one-way analysis of variance (ANOVA) showed that there was no significant difference in participants' trait-levels of empathy on any subscale of the IRI scale across the two conditions

(Perspective Taking: $t(112) = 0.30$, $p = .762$; 95% Confidence Interval (CI) [-.18, .24], Empathic Concern: $t(108) = -0.39$, $p = .697$; 95 CI [-.26, .18], and Personal Distress: $t(114) = 0.39$, $p = .695$; 95 CI [-.24, .36]). There were also no significant differences between the two conditions regarding beliefs about empathy (Controllability: $t(113) = 0.11$, $p = .912$; 95 CI [-.37, .41], Implicit Theories: $t(114) = -1.25$, $p = .214$; 95 CI [-.79, .18]). These results show that there was a balance across conditions in terms of the way that people think about empathy and in the way that they believe they are able to control their empathic responses, showing that random assignment was successful across conditions.

Outcome variables

Manipulation check. Cohen's Kappa was calculated in order to assess inter-coder reliability and agreement between the two coders [68]. There was substantial agreement between the two coders ($\kappa = .89$, $p < .001$, 95% CI [.81, .99]. Data from the two coders were averaged together. A Chi-Squared test was used in order to test whether or not there was a significant difference in the proportion of participants who reported feeling present (values equal to 1) and the participants who did not (values smaller than 1) across conditions. In the NPT condition 5.4% of participants received a score of 1 and in the VRPT condition 34.4% participants received a score of 1. Results showed that significantly more participants in the VRPT condition reported feeling spatially present than participants in the NPT condition ($X^2(1) = 15.06$, $p < .001$; 95 CI [14.98, 42.10]).

Continuous variables. The means and standard deviations for the outcome variables (i.e., IOS, Dehumanization, Empathy, Personal Distress, Social Presence, support for Proposition A, support for Measure B, and amount donated) by condition and time are summarized in Table 1.

All continuous outcome variables were analyzed using a linear growth curve model with fixed effect of condition (NPT vs VRPT) on the intercept (results right after the intervention) and linear terms (trend over the course of eight weeks). Random effect of individuals on the intercept and the slopes were included in the models as well. Growth curve modeling analysis was chosen since it accounts for inter-participant variability (between) and intra-participant (within) patterns of change over time [69–70]. All analyses were carried out in R version 3.0.2 using the nlme package.

IOS. There were no statistically significant differences in self-other overlap between the conditions at the time of the intervention (Time 0) ($X^2(1) = 2.22$, $p = .136$; 95 CI [-0.15, 1.08]). There was also no effect of condition on how connected participants felt toward the homeless in the eight weeks following the intervention ($X^2(1) = .1024$, $p = .749$; 95 CI [-0.23, .17]). Fig 3 shows the data and model fit. The general trend for both conditions is that participants felt moderately connected to the homeless immediately after the intervention, but experienced a slight decline of .03 points ($SE = 0.1$) in perceived connectedness at the time of each follow up.

Dehumanization. There was no significant effect of condition on blatant dehumanization at the time of the intervention ($X^2(1) = .32$, $p = .572$; 95 CI [-5.22, 5.68]). However, there was a significant effect of condition indicating that, over time, participants in the VRPT condition thought of the homeless as more evolved than the participants in the NPT condition who thought of the homeless as less evolved over time ($X^2(1) = 3.8$, $p = .051$; 95 CI [0, 2.66]). The difference was approximately 1.33 ($SE = .68$) points biweekly showing that although participants in both conditions thought of the homeless similarly at the time of the intervention, participants in the VRPT condition thought significantly higher of the homeless over time than those in the NPT condition. Fig 4 shows the data and model fit.

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

	NPT Condition							
Measures	Time 0		Time 1		Time 2		Time 3	
	Mean	*SD*	*Mean*	*SD*	*Mean*	*SD*	*Mean*	*SD*
IOS	2.84	1.99	2.42	1.44	2.65	1.63	2.68	1.73
Dehumanization	89.3	15.3	87.86	18.21	88.85	12.7	87.21	19
Empathy	4.8	1.2	4.57	1.05	4.53	1.14	4.43	1.12
Personal Distress	4.4	1.26	4.38	1.1	4.15	1.26	4.21	1.31
Attitudes	6.39	1.14	5.98	1.27	6	1.24	6.07	1.26
Support Proposition A	4.09	0.83	-	-	-	-	-	-
Amount Donated	5.44	3.72	-	-	-	-	-	-
Support Measure B	-	-	-	-	4.14	0.89	-	-
Petition Signing Proportion	35 out of 57							
	VRPT Condition							
Measures	Time 0		Time 1		Time 2		Time 3	
	Mean	*SD*	*Mean*	*SD*	*Mean*	*SD*	*Mean*	*SD*
IOS	3.33	1.87	2.97	1.68	3.08	1.83	3.1	1.8
Dehumanization	88.8	16.96	91.28	12.86	91.3	12.71	91.4	15.08
Empathy	5.29	1.18	4.9	1.07	4.87	1.2	4.99	1.22
Personal Distress	4.75	1.38	4.63	1.15	4.6	1.28	4.77	1.25
Presence	2.9	0.74	-	-	-	-	-	-
Attitudes	6.56	1.22	6.54	1.21	6.4	1.22	6.35	1.4
Support Proposition A	4.07	0.69	-	-	-	-	-	-
Amount Donated	6.02	3.86	-	-	-	-	-	-
Support Measure B	-	-	-	-	4.23	0.96	-	-
Petition Signing proportion	49 out of 60							

SD = Standard deviation.

https://doi.org/10.1371/journal.pone.0204494.t001

Attitudes toward the Homeless. Visual inspection of the residuals of the model testing the effect of condition and time on attitudes toward the homeless revealed a curvilinear relationship between these variables. The quadratic model fit the data significantly better than the linear model ($X^2(2) = 13.37$, $p = .001$; 95 CI [-5.02, -1.0]). Since two models were compared for this dependent variable, a Bonferroni correction ($p = \alpha/2$) was implemented and results reflect the adjusted p-values. There was no significant difference in attitudes toward the homeless between the two conditions at Time 0 ($X^2(1) = 2.37$, $p = 0.124$; 95 CI [-0.07, .75]). However, there was a significant effect of condition ($X^2(1) = 8.19$, $p = 0.033$; 95 CI [-5.02, -0.9]) over the course of the eight weeks. Even though attitudes toward the homeless deteriorated over time, the attitudes deteriorated at a significantly slower rate and were consistently more favorable for participants in the VRPT condition than the participants in the NPT condition. Fig 5 shows the data and model fit.

Empathy. Participants in the VRPT condition reported feeling significantly more empathetic than participants in the NPT condition immediately following the intervention ($X^2(1) = 4.53$, $p = .033$; 95 CI [.04, .77]). However, there was no significant difference between the two conditions in self-reported empathy two, four, or eight weeks after the intervention ($X^2(2) = 3.08$, $p = .214$; 95 CI [-.31, 4.16]). Fig 6 shows the data and model fit.

Personal Distress. Participants in the VRPT condition reported feeling significantly more distressed and troubled after the intervention than participants in the NPT condition ($X^2(1) = 4.48$, $p = .034$; 95 CI [.04, .76]). However, there were no significant differences between the two

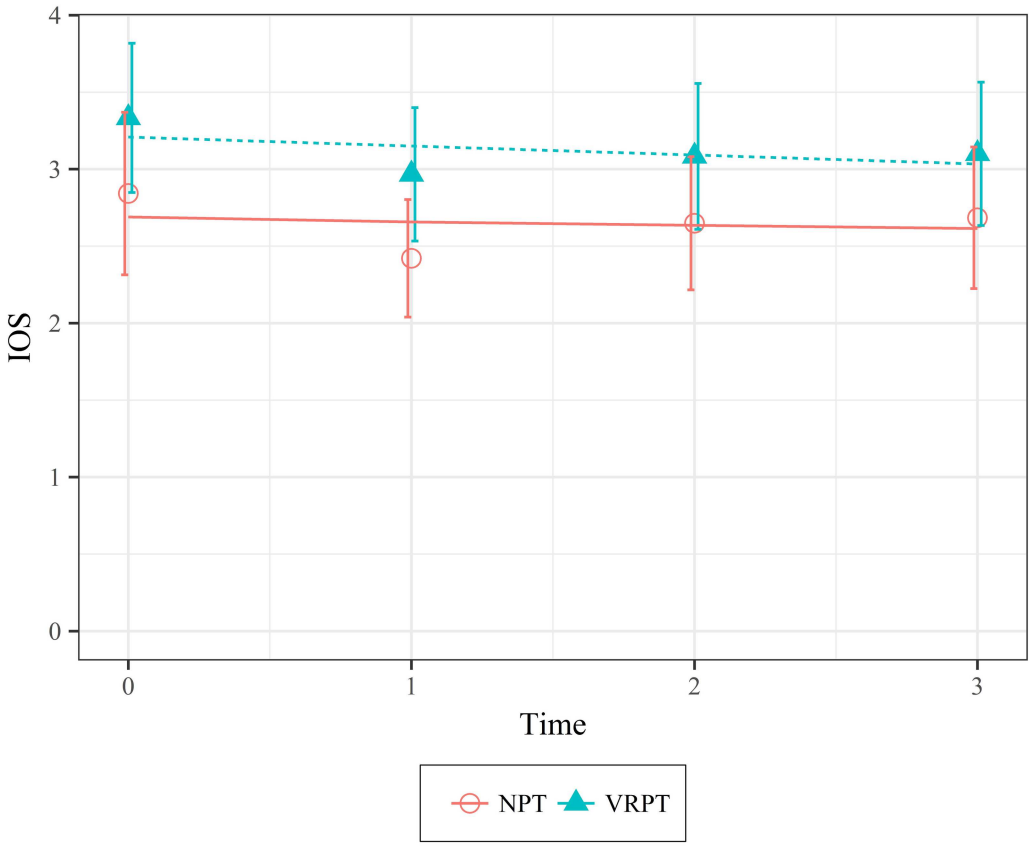


Fig 3. Mean values of self-other overlap with the homeless as a function of condition over time. Higher scores represent higher overlap between the self and the homeless. Error bars represent 95% Confidence Intervals.

https://doi.org/10.1371/journal.pone.0204494.g003

conditions over the course of eight weeks ($X^2(2) = 0.84$, $p = .359$; 95 CI [-1.78, 3.88]). These results indicate that significant differences in personal distress only occurred immediately after the intervention. Fig 7 shows the data and model fit.

Correlations among Dependent Variables. All simple Pearson correlations between dependent variables are included in Table 2. A Holm-Bonferroni correction was implemented to account for multiple comparisons. Results show the adjusted p-values.

The correlation analysis showed that IOS scores were not significantly correlated with attitudes toward the homeless immediately after the intervention. However, IOS scores and attitudes became significantly and positively correlated over the course of eight weeks (Time 0: r = 0.18, $p = 1.00$; Time 1: r = 0.25, $p = .039$; Time 2: r = 0.28, $p = .024$; Time 3: r = 0.30, $p = .008$) such that participants who reported feeling more connected to the homeless also reported more positive attitudes toward them. Dehumanization scores were also not correlated with attitudes toward the homeless immediately after the perspective-taking tasks but became significantly and positively correlated over the course of eight weeks (Time 0: r = 0.21, $p = .006$; Time 1: r = 0.32, $p = .003$; Time 2: r = 0.32, $p = .004$; Time 3: r = 0.31, $p = .006$). The analysis also revealed that attitudes toward the homeless scores were significantly and positively correlated with empathy (Time 0: r = 0.38, $p < .001$; Time 1: r = 0.61, $p < .001$; Time 2: r = 0.68, $p < .001$; Time 3: r = 0.55, $p < .001$) such that participants who reported more empathy also reported more positive attitudes toward the homeless.

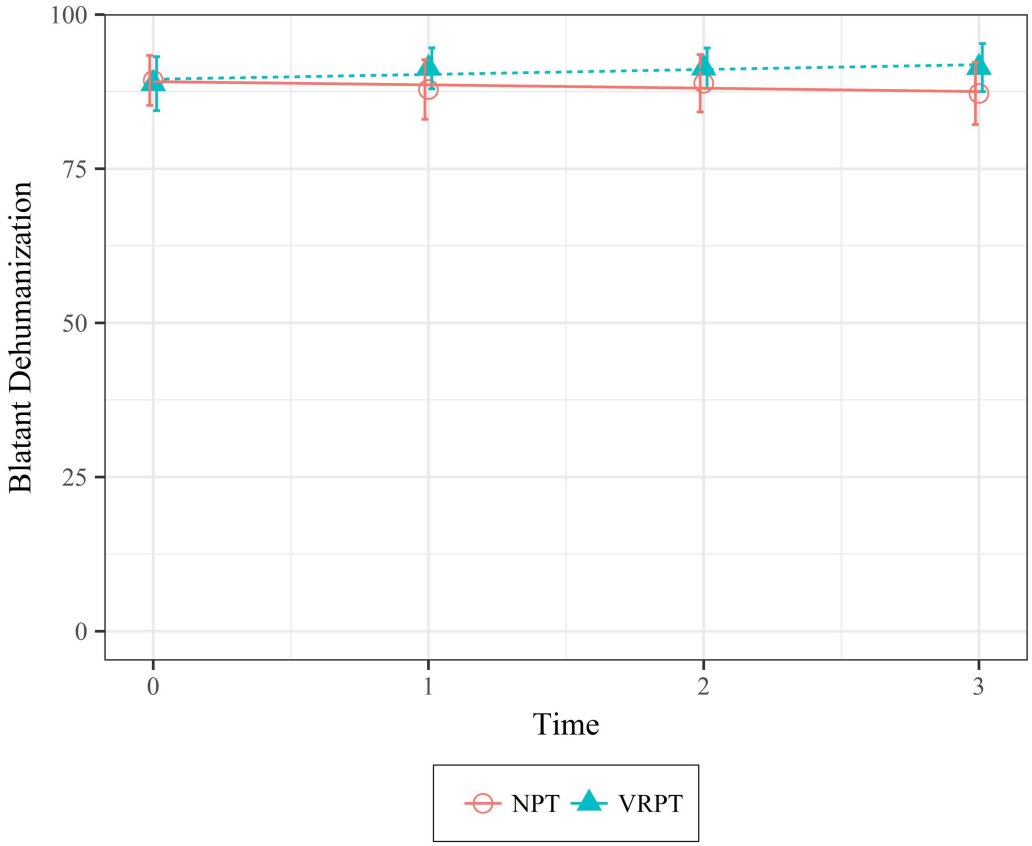


Fig 4. Mean values of blatant dehumanization of the homeless as a function of condition over time. Error bars represent 95% Confidence Intervals.

https://doi.org/10.1371/journal.pone.0204494.g004

Moreover, attitudes toward the homeless were not correlated with personal distress immediately after the intervention. However, the correlation between attitudes and personal distress became significant over the course of eight weeks (Time 0: r = 0.22, p = .489; Time 1: r = 0.27, p = .026; Time 2: r = 0.37, $p < .001$; Time 3: r = 0.28, p = .013). The analysis also revealed that self-reported empathy was significantly and positively correlated with personal distress (Time 0: r = 0.30, p = .032; Time 1: r = 0.41, $p < .001$; Time 2: r = 0.53, $p < .001$; Time 3: r = 0.46, $p <$ .001).

Furthermore, four weeks after the intervention, the correlation analysis revealed that attitudes toward the homeless were significantly and positively correlated with support for Measure B (r = 0.40, $p < .001$). In other words, participants who reported more positive attitudes toward the homeless also reported more support for affordable housing four weeks after the intervention. It is important to note that attitudes toward the homeless were not correlated with support for affordable housing at the time intervention (r = -0.04, $p > 1.00$). These results suggest that over time, the positive correlation between attitudes and support for helpful initiatives strengthens.

Behavioral measures. Behavioral Measures at time 0. There was no significant difference between conditions regarding self-reported support for Proposition A (t (108) = .15, p = .882; 95 CI [-.03, .3]). However, significantly more participants in the VRPT condition signed the petition in support of Proposition A (z (115) = 2.39, p = .023; 95 CI [1.12, 7.21]) in comparison to the NPT condition. Even though participants in both conditions claimed to support

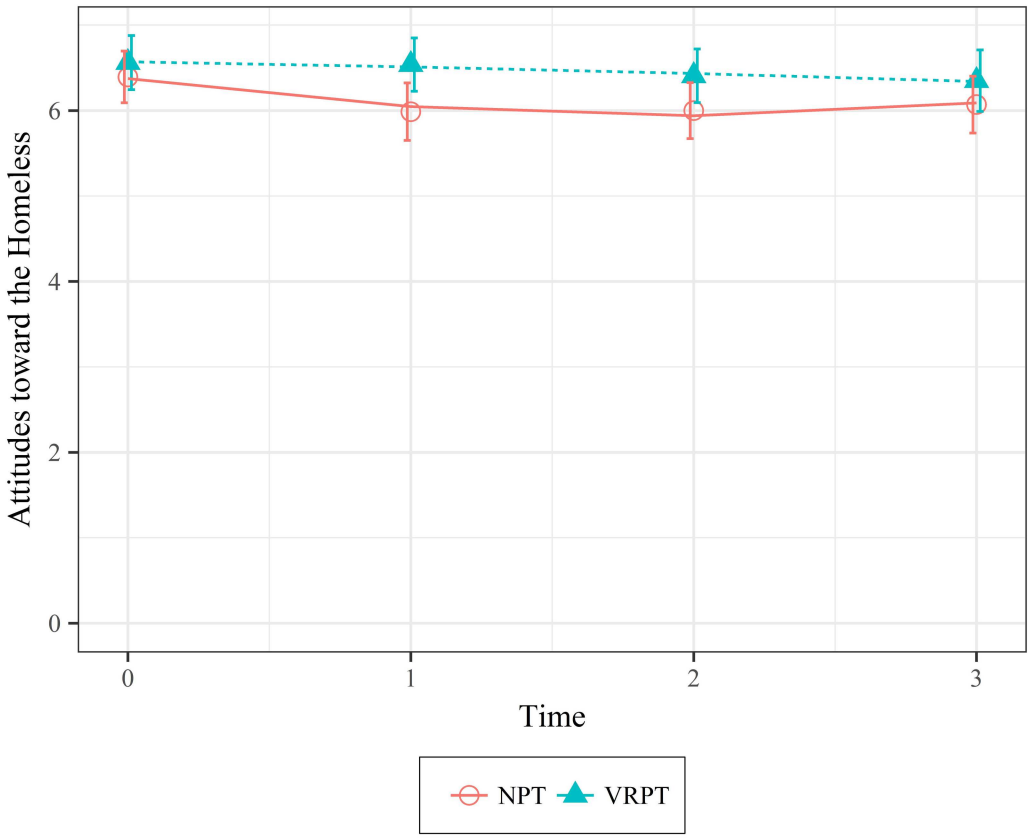


Fig 5. Mean values of attitudes toward the homeless as a function of condition over time. Error bars represent 95% Confidence Intervals.

https://doi.org/10.1371/journal.pone.0204494.g005

Proposition A to the same extent, a significantly higher proportion of participants who performed the VRPT task physically signed the petition in support of affordable housing. For the donation question, there was no significant effect of condition on amount donated to a homeless shelter (t (114) = -.82, p = .411; 95 CI [-1.97, .81]). On average, participants from both conditions donated about 5 dollars to a homeless shelter.

Behavioral Measures at Time 1. Two weeks after the intervention, participants were asked to write a letter to their elected official on the issue of homelessness. After analyzing the text of the letters, there were no significant differences in word count (t (108) = 0.64, p = 0.521; 95 CI [-15.58, 30.54]), positive or negative emotion (t(100) = 0.05, p = 0.960, 95 CI [-1.01, 1.06]; t (110) = -0.59, p = 0.557, 95 CI [-.87, .47]), or the 'I' dictionary (t(104) = 1.11, p = 0.269; 95 CI [-.38, 1.35]). However, there were marginally significant differences in the anxiety (t(80) = -1.78, p = 0.078; 95 CI [-.42, .02]) and social (t(107) = -1.85, p = 0.067; 95 CI [-3.94, .14]) dictionaries, showing that participants in the VRPT condition were more anxious while writing the letter, and used more social words than participants in the NPT condition. There was also a significant difference in the 'we' dictionary, indicating that participants in the VRPT condition wrote about the issue of homelessness and its possible solutions using more "we", "our", "us" pronouns than the participants in the NPT condition (t(102) = -2.01, p = 0.046; 95 CI [-1.53, -.01]). Means and standard deviations used for all the dictionaries by condition can be found in Table 3.

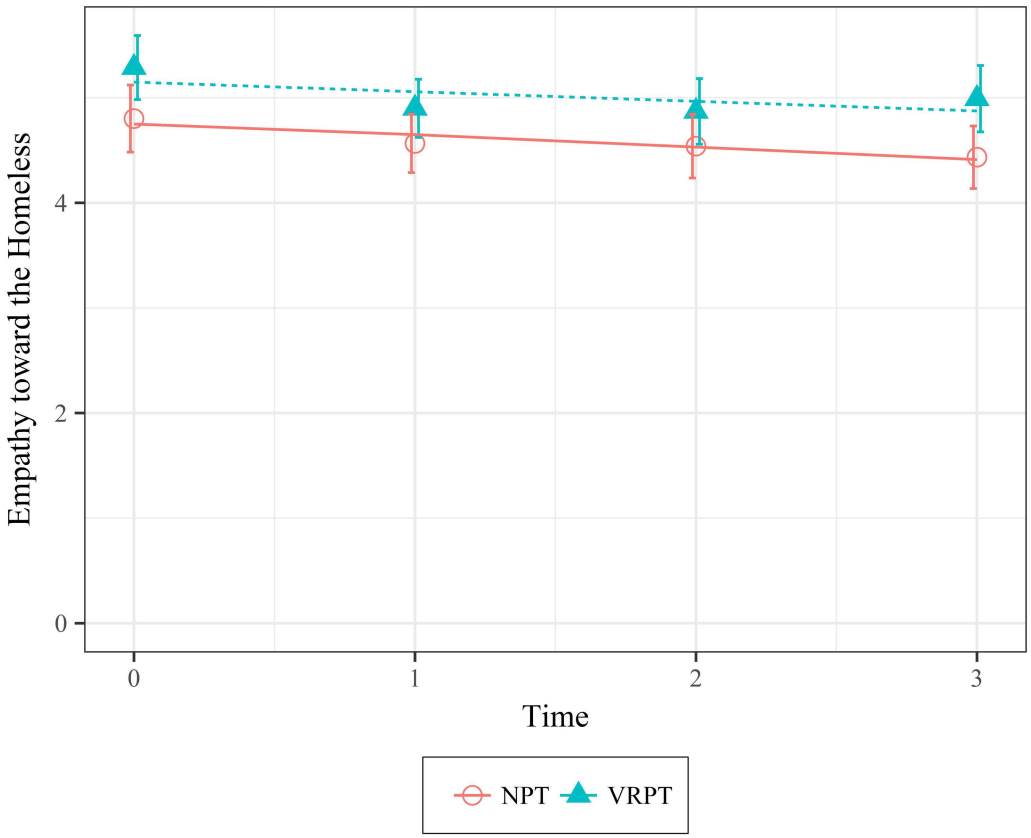


Fig 6. Mean values of empathy as a function of condition over time. Error bars represent 95% Confidence Intervals.

https://doi.org/10.1371/journal.pone.0204494.g006

Behavioral Measures at Time 2. Four weeks after the intervention, at Time 2, participants were asked to what extent they agreed with Measure B, a measure proposed in the San Francisco Bay Area supporting affordable housing. Participants in the VRPT condition reported more support for Measure B than participants in the NPT condition ($t(105) = -1.74$, $p = .08$; 95 CI [-.63, .04]). However, these results were only marginally significant.

Behavioral Measures at Time 3. Eight weeks after the intervention, participants were asked to write a letter to a friend telling them what they had learned and thought about the issue of homelessness. After analyzing the text of the letters, there were no significant differences in any of the 7 LIWC dictionaries (word count: $t(108) = -1.37$, $p = 0.176$, 95 CI [-59.78, 11.11]; positive emotion: $t(87) = 0.96$, $p = 0.340$, 95 CI [-0.53, 1.51]; negative emotion: $t(114) = -0.75$, $p = 0.454$, 95 CI [-1.11, 0.5]; we: $t(107) = -1.33$, $p = 0.189$, 95 CI [-1.2, 0.24]; I: $t(112) = -1.67$, $p = 0.097$, 95 CI [-1.75, 0.15]; anxiety: $t(115) = -0.67$, $p = 0.506$, 95 CI [-0.31, 0.15]; social: $t(92) = 0.34$, $p = 0.735$, 95 CI [-1.75, 2.47]).

Discussion

Both types of perspective-taking tasks led to similar results when it came to self-other overlap immediately after the intervention and over the course of eight weeks. Participants in the VRPT condition reported significantly more empathy and more personal distress immediately after the intervention. However, over time, participants in both the VRPT and NPT conditions reported similar rates of empathy and personal distress. The results for attitudes toward the

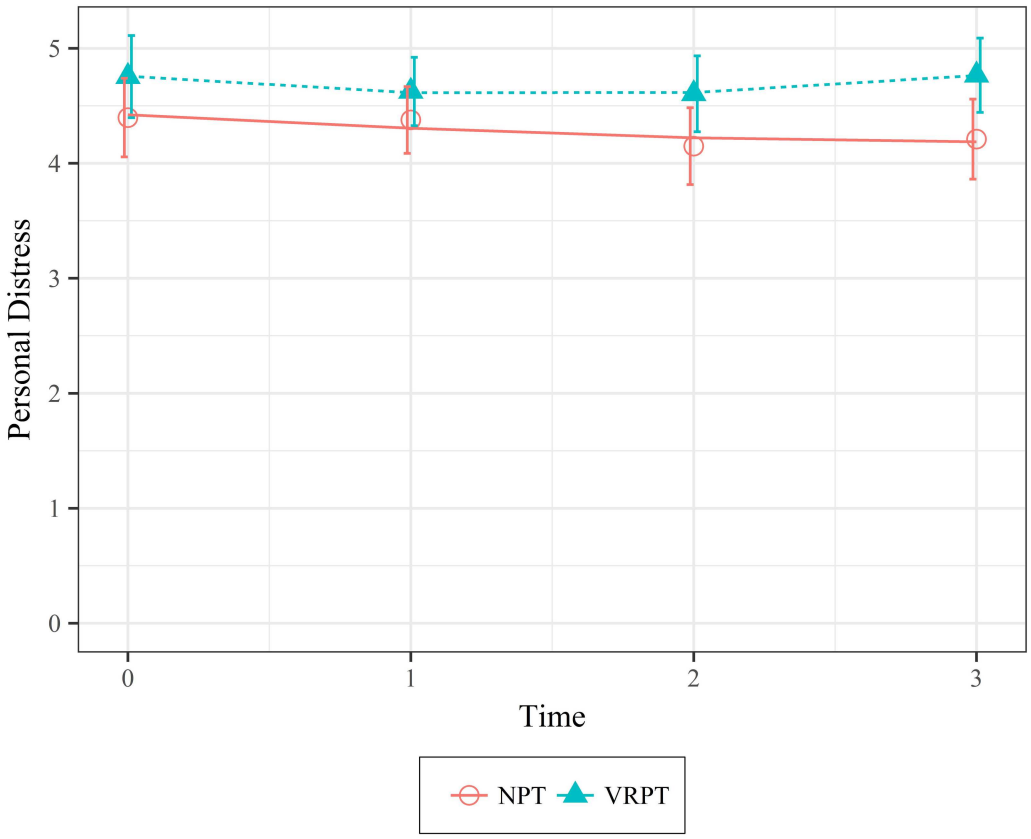


Fig 7. Mean values of personal distress as a function of condition over time. Error bars represent 95% Confidence Intervals.

https://doi.org/10.1371/journal.pone.0204494.g007

homeless and the dehumanization scale show the opposite pattern. Even though both conditions reported similar rates of dehumanization and attitudes toward the homeless at the time of the intervention, participants in the NPT condition thought of the homeless as less evolved over time, and the attitudes they had for the homeless deteriorated in the eight weeks that followed the intervention. In contrast, the VRPT condition, which allowed participants to interact with the virtual environment in real-time, led to more positive, longer-lasting attitudes toward the homeless up to two months after the intervention.

Past research has demonstrated that after a perspective-taking task, participants tend to feel empathetic toward a specific target immediately after the task, but over time, the empathic feelings wane while attitudes toward that target improve [71]. In a different study, Batson et al [3] found that participants who performed a perspective-taking task felt more empathy for convicted murderers than participants who were asked to remain objective. There were no differences in attitudes toward the convicts between the two conditions immediately after the task, but there were significant differences in attitudes 1–2 weeks after the intervention, indicating that feeling empathetic toward a member of a stigmatized group may lead to better attitudes over time rather than immediately after the perspective-taking task. Batson suggests that attitudes can "outlive the empathic emotion itself" ([3] p. 116). Our results replicate these findings since participants in the VRPT condition reported more empathy but similar attitudes when compared to the NPT participants immediately after the intervention, but significantly better

Table 2. Simple correlations among dependent variables.

Time 0—Immediately After the Intervention							
	Dehumanization	Attitudes	Empathy	Personal Distress	Social Presence	Support Prop A	Donation
IOS	0.10	0.18	0.08	-0.01	0.15	0.04	-0.01
Dehumanization		0.21	-0.07	-0.09	-0.03	-0.02	0.15
Attitudes			**0.38** [c]	0.22	0.14	-0.04	0.08
State Empathy				**0.30** [a]	0.25	-0.09	0.06
Personal Distress					0.16	-0.06	-0.05
Social Presence						-0.05	-0.12
Support Prop A							0.16
Donation							
Time 1—Two Weeks After the Intervention							
	Dehumanization	Attitudes	Empathy	Personal Distress			
IOS	0.13	**0.25** [a]	0.07	0.04			
Dehumanization		**0.32** [b]	0.11	0.13			
Attitudes			**0.61** [c]	**0.27** [a]			
State Empathy				**0.41** [c]			
Personal Distress							
Time 2—Four Weeks After the Intervention							
	Dehumanization	Attitudes	Empathy	Personal Distress	Measure B		
IOS	0.19	**0.28** [a]	0.24	0.18	0.16		
Dehumanization		**0.32** [b]	0.25	0.09	0.12		
Attitudes			**0.68** [c]	**0.37** [c]	**0.40** [c]		
State Empathy				**0.53** [c]	0.24		
Personal Distress					0.14		
Measure B							
Time 3—Eight Weeks After the Intervention							
	Dehumanization	Attitudes	Empathy	Personal Distress			
IOS	**0.26** [a]	**0.30** [b]	0.22	0.16			
Dehumanization		**0.31** [b]	0.12	-0.01			
Attitudes			**0.55** [c]	**0.28** [b]			
State Empathy				**0.46** [c]			
Personal Distress							

[a] positive at $p < .05$.

[b] positive at $p < .01$.

[c] positive at $p < .001$.

https://doi.org/10.1371/journal.pone.0204494.t002

attitudes toward the homeless over time. Additionally, improved attitudes lasted longer than the empathic feelings themselves.

It is also important to note that while there were no significant differences when it came to self-reported support for Proposition A between the two conditions, significantly more participants in the VRPT condition signed the petition in support of affordable housing even when it meant an increase in their taxes.

Overall, these results show that immediately after the intervention, VRPT led to more self-reported empathy and personal distress. Over time, VRPT did not lead to more self-other overlap, self-reported empathy, personal distress, or donations to a homeless shelter than more traditional perspective-taking tasks. However, VRPT did result in more positive, longer-

Table 3. Means and standard deviations for all LIWC dictionaries used.

	Time 1				Time 3			
	NPT		VRPT		NPT		VRPT	
Dictionaries	*Mean*	*SD*	*Mean*	*SD*	*Mean*	*SD*	*Mean*	*SD*
Word Count	111.98	65.1	104.5	58.16	98.32	81.4	122.65	110.49
Positive Emotion	3.87	3.12	3.84	2.37	3.46	3.39	2.96	1.91
Negative Emotion	1.95	1.67	2.15	1.92	2.11	2.18	2.41	2.21
Anxiety	0.15	0.35	0.35	0.77	0.25	0.61	0.32	0.64
Social	12.18	5.82	14.08	5.07	12.28	6.88	11.92	4.22
We	1.37	1.65	2.14	2.37	1.23	2.17	1.71	1.72
I	2.34	2.54	2.08	1.85	2.45	2.32	3.25	2.86

SD = Standard Deviation.

https://doi.org/10.1371/journal.pone.0204494.t003

lasting attitudes toward the homeless and significantly more signatures supporting helpful initiatives than the NPT condition.

Study 2

The results of Study 1 showed that, over time, participants in both the NPT and VRPT condition reported feeling empathetic and connected to the homeless at similar rates. However, VRPT led to more signatures supporting affordable housing and more positive, longer-lasting attitudes. These results show that perspective-taking aided by immersive VR media can effectively improve attitudes toward the homeless. However, it is possible that less immersive media, such as desktop computers and laptops, produce similar effects without the need to fully surround participants with stimuli. Study 2 expands on Study 1 by comparing the effect of four different types of empathy interventions: a fact-driven information intervention (Information), a traditional, narrative-based perspective-taking task (NPT), a VR perspective-taking task (VRPT), and a less immersive mediated perspective-taking task using a desktop computer (Desktop) in order to more accurately assess the effect of perspective-taking and examine the role immersion plays when attempting to promote empathy and prosocial behaviors.

Since past research demonstrates that perspective-taking leads to increased empathy and helping behaviors, we predict that any type of perspective-taking would be more effective at eliciting empathy and prosocial behaviors than receiving information (i.e., the three perspective-taking conditions vs. Information). Given the results obtained in Study 1, we also predict that mediated perspective-taking would be more effective than NPT (i.e. Desktop and VRPT vs NPT), and that the most immersive perspective-taking task would be more effective than the less immersive perspective-taking tasks (i.e. VRPT vs Desktop).

Method

Participants

A total of 452 participants were recruited to participate in this study. Thirteen participants were excluded from the analysis because they did not complete the study in its entirety. Of the remaining 439 participants (189 men, 250 women), 190 participants were students recruited from a medium-sized western university and 249 were recruited at mobile sites such as schools, museums, and senior citizen centers in the San Francisco Bay Area. Participants provided informed consent and were compensated with either course credit or a $10 Amazon gift card for participating. The mean age of the participants was 29.2 (*SD* = 14.8) and ranged

between 15 and 88 years old. Of these 439 participants, 22 (5%) were African American, 9 (2%) were Middle Eastern, 27 (6%) were Indian, 5 (1%) were Native American, 86 (20%) were Asian, 32 (7%) were Hispanic, 215 (48%) were White Caucasian, 40 (9%) were multiracial, and the rest (1%) declined to answer. Participants recruited at mobile sites participated in the study inside a 3.96m x 2.13m x 2m tent where they were given complete privacy in public areas. Apart from the physical setting, all of the participants used the same equipment and followed the same procedures. There was no overlap between participants from Study 1 and Study 2.

Design and procedure

In this between-subjects design, participants signed a consent form and were randomly selected into one of four conditions: 1) Information, 2) NPT 3) Desktop, or 4) VRPT. After random assignment, all participants completed a pre-intervention questionnaire which included demographic questions, the Interpersonal Reactivity Index (IRI), and the Beliefs about Empathy scale. These scales were used in order to conduct sample checks and make sure that random assignment was successful. Upon completing this questionnaire, participants in each condition received a different empathy intervention.

In the Information condition (n = 107), participants were asked to read a packet of information and statistics about the homeless population in the Bay Area in 2015 [72]. The packet was a combination of written information, graphs, tables, maps, and pie charts. The packet provided information on the main causes for homelessness, obstacles faced by the homeless when trying to get a job, as well as the percentage of the population that were children, sick, or had a history of foster care. To read all of the information provided in this packet please see Text E in S1 Appendix. After reading the packet, participants completed a post-intervention questionnaire consisting of self-report and behavioral measures. This fact-driven intervention was chosen in order to examine the effectiveness of perspective-taking (mediated or not) against a rigorous, real-world intervention that does not utilize perspective-taking at all and was specifically designed to increase awareness about the homeless.

The NPT (n = 104) and the VRPT (n = 115) conditions in Study 2 were the same as the conditions described Study 1.

In the 'Desktop' condition (n = 113), participants sat in front of a computer with a flat TV screen (39") and experienced what it was like to become homeless through a 2D interactive narrative (Fig 8). The narrative that the participants followed was the same interactive narrative from the VRPT condition, except participants were only able to view the environment on the screen, navigate the online environment using the arrow keys on the computer keyboard, and select objects with their mouse. The functionality of this interactive narrative resembled that of an online game. Participants in this condition were given the same interactive tasks as the VRPT condition, and were able to make their own choices, navigate the virtual environment, and receive immediate feedback. Fig 2 also shows the three scenes from the participant's point of view.

All of the perspective-taking tasks, regardless of type, lasted approximately 15 minutes. All participants completed the same post-intervention questionnaire immediately after the intervention. All participants were debriefed once the study had been completed.

Measures

The same manipulation check employed in Study 1 was used in Study 2 to assess how spatially present participants felt across the three perspective-taking conditions. The self-report measures used in Study 2 (described in Study 1) were IRI, Beliefs about Empathy, IOS, Dehumanization, Empathy, Personal Distress, and Social Presence. The behavioral measures used were

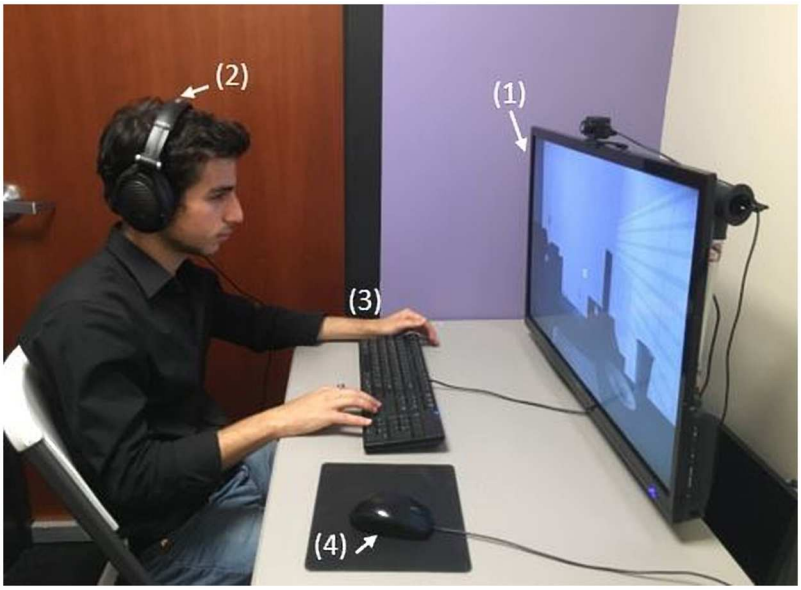


Fig 8. Equipment used for desktop condition. (1) TV screen (2) headphones (3) keyboard & (4) mouse. The individual in this manuscript has given written informed consent (as outlined in PLOS consent form) to publish these case details.

https://doi.org/10.1371/journal.pone.0204494.g008

self-reported support for Proposition A, whether or not participants signed a petition supporting Proposition A, and the homeless shelter donation question. The overall means and standard deviations, as well as the reliability of the scales (Cronbach's alpha) used in Study 2 can be found on Table 4.

Results

Population variables

Replicating Study 1, a one-way analysis of variance (ANOVA) showed that there was no significant difference in participants' trait-levels of empathy on any subscale across the four conditions (Perspective Taking: $F(3, 435) = .89$, $p = .446$, $\eta^2 = .006$, Empathic Concern: $F(3, 435) =$

Table 4. Overall means, standard deviations, and reliability of scales used in Study 2.

Measures	*Mean*	*SD*	*α*
IRI: EC	3.92	0.59	0.76
IRI: PT	3.70	0.59	0.74
IRI: PD	2.64	0.72	0.81
Beliefs about Empathy: C	4.15	1.09	0.81
Beliefs about Empathy: IT	4.28	1.20	0.89
IOS	2.86	1.75	-
Dehumanization	89.95	16.09	-
Empathy	5.09	1.22	0.88
Personal Distress	4.15	1.43	0.85
Presence	2.88	0.82	0.80

SD = Standard deviation.

α = Cronbach's alpha.

https://doi.org/10.1371/journal.pone.0204494.t004

1.09, $p = .353$, $\eta^2 = .007$, and Personal Distress: $F(3,435) = 1.275$, $p = .282$, $\eta^2 = .009$). There was also no significant difference across conditions when it came to beliefs about empathy (Controllability: $F(3, 435) = .339$, $p = .805$, $\eta^2 = .002$, Implicit Theories: $F(3,435) = .461$, $p = .709$, $\eta^2 = .003$). The lack of significant differences across conditions for our population variables suggests that all of the conditions were well balanced in terms of individual differences, and that random assignment to the four conditions was successful.

Outcome variables

Manipulation check. Cohen's Kappa was calculated in order to assess inter-coder reliability and agreement between the two coders [68]. There was substantial agreement between the two coders ($\kappa = .93$, $p < .001$, 95% CI [.87, .98]. Data from the two coders were averaged together. A Chi-Squared test was used in order to test whether or not there was a significant difference in the proportion of participants who reported feeling present and the participants who did not between conditions. In the NPT condition 2% of participants received a score of 1, in the Desktop condition 9.2%, and in the VRPT condition 30% of participants received a score of 1. Results showed that significantly more participants in the VRPT condition reported feeling spatially present than participants in the Desktop condition ($X^2(1) = 13.37$, $p < .001$; 95 CI [9.82, 31.30]) and the NPT condition ($X^2(1) = 28.72$, $p < .001$; 95 CI [18.43, 37.67]). Participants in the Desktop condition also reported feeling more spatially present than participants in the NPT condition ($X^2(1) = 4.78$, $p < .029$; 95 CI [.55, 14.73]). These results confirm that our experimental conditions are significantly different from each other in terms of presence with the VRPT condition being the most immersive, followed by the Desktop condition, and then the NPT condition. The Information condition responses were not coded because participants in this condition did not perform any kind of perspective-taking task.

Continuous variables. The means and standard deviations for the continuous variables (i.e., IOS, Dehumanization, Empathy, Personal Distress, Presence, support for Proposition A, and amount donated) are summarized in Table 5. A one-way, between-subjects analysis of variance (ANOVA) was carried out to compare the effect of condition on all of the continuous outcome variables. 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 any type of perspective-taking would be more effective at eliciting empathy and prosocial behaviors than receiving information (i.e., the three perspective-taking conditions vs. Information), 2) that mediated perspective-taking would be more effective than NPT (i.e., Desktop and VRPT vs NPT), and 3) that the most immersive perspective-taking task would be more effective than the less immersive perspective-taking task (i.e., VRPT vs Desktop). A Bonferroni adjusted alpha ($p = \alpha/k$) was utilized in order to account for multiple comparisons and avoid Type I errors.

IOS. There was a significant effect of condition on self-other overlap with the homeless: $F(3, 435) = 7.197$, $p = < .001$, $\eta^2 = .05$. There was a substantial difference in self-other overlap between the Information and the three perspective-taking conditions ($F(1, 435) = 18.97$, $p < .001$). Participants felt closer and more connected to the homeless after any type of perspective-taking than when they just received information about the homeless. This result replicates extant research highlighting the effectiveness of perspective-taking tasks on self-other overlap [73]. However, there was no significant difference between the NPT and the mediated perspective-taking conditions ($F(1, 435) = 1.66$, $p = .197$) or the mediated conditions themselves ($F(1, 435) = .96$, $p = .328$). These results replicate the findings from Study 1, highlighting that mediated perspective-taking tasks, regardless of how immersive they are, do not result in participants feeling more connected to the homeless than traditional perspective-taking tasks.

Table 5. Means and standard deviations for all outcome variables by condition in Study 2.

Measures	Information *M (SD)*	NPT *M (SD)*	Desktop *M (SD)*	VRPT *M (SD)*
IOS	2.45 (1.47)	2.89 (1.94)	3.27 (1.68)	3.04 (1.75)
Dehumanization	91.86 (15.17)	87.15 (20.46)	92.22 (12.74)	88.45 (14.82)
Empathy	4.79 (1.32)	5.31 (1.10)	5.21 (1.12)	5.16 (1.26)
Personal Distress	3.78 (1.47)	4.29 (1.41)	4.10 (1.34)	4.37 (1.45)
Presence	-	-	2.78 (.81)	2.98 (.82)
Support for Prop A	3.91 (.84)	3.75 (1.03)	3.89 (.97)	4.18 (.79)
Amount Donated	7.63 (3.69)	7.86 (3.61)	8.2 (3.28)	8 (3.42)
Petition Signing proportion	80 out 107	66 out of 104	75 out of 113	98 out of 115

M = Mean; *SD* = Standard Deviation.

https://doi.org/10.1371/journal.pone.0204494.t005

Taken together, these results showed that just receiving information about the current state of the homeless population did not result in participants feeling closer or more connected to the homeless population. However, perspective-taking tasks, regardless of delivery medium or level of immersion, had a positive effect in that participants felt more connected to the homeless after taking their perspective.

Dehumanization. The dehumanization scale measured how evolved participants felt the average member of the homeless population was at the time of the study. There was a significant effect of condition on dehumanization: F (3, 435) = 2.68, p = .046, η^2 = .02). However, there was no significant difference in how evolved participants reported the homeless to be between the Information and the three perspective-taking conditions (F (1, 435) = 2.10, p = .147), only a marginally significant difference between the NPT and mediated perspective-taking conditions (F (1, 435) = 2.85, p = .092) indicating that mediated perspective-taking led to higher (less dehumanizing) scores than NPT. There was also a marginally significant difference between the Desktop vs VRPT conditions (F (1, 435) = 3.09, p = .079) where participants in the less immersive perspective-taking task thought of the homeless as more evolved than the participants in the VRPT condition.

Empathy. There was a significant effect of condition on self-reported empathy: F (3, 435) = 3.14, p = .025, η^2 = .02. Participants in the three perspective-taking conditions reported feeling significantly more empathy toward the homeless than participants in the Information condition (F (1, 435) = 9.31, p = .002). However, there were no significant differences between the two types of mediated perspective-taking tasks (F (1, 435) = 0.032, p = .858) or between them and the NPT condition (F (1, 435) = 0.063, p = .802). Similar to the self-other overlap results, there was no difference between the participants in the NPT condition and the two mediated conditions (Desktop and VRPT).

Personal Distress. There was a significant effect of condition on how distressed participants reported feeling after the intervention: F (3, 435) = 3.743, p = .011, η^2 = .03. There was a significant difference in personal distress between the Information and the three perspective-taking conditions (F (1, 435) = 8.93, p = .003) with participants in the Information condition reporting less distress than their counterparts. As with the previous analyses of the self-other overlap and empathy scales, there was no significant difference in reported personal distress between the mediated conditions (F (1, 435) = 0.28, p = .595) or between them and the NPT condition (F (1, 435) = 2.01, p = .156).

These results show that participants in the three perspective-taking conditions reported feeling more empathetic and connected to the homeless than participants who only received

information about the current state of homeless population. Participants in the three perspective-taking conditions also reported feeling more personally distressed after taking the perspective of a homeless person than participants in the Information condition.

Social Presence. Social presence, or the feeling of being inside a virtual environment with others was only measured for the Desktop and VRPT conditions. There was only a marginally significant difference ($t(229) = 1.82$, $p = .070$; 95 CI [-0.4, .01]), with the Desktop condition ($M = 2.78$, $SD = .81$) being lower than the VRPT condition ($M = 2.98$, $SD = .82$).

Replication of Study 1. Study 1 compared NPT and VRPT immediately after the intervention. However, Study 2 does not directly compare NPT and VRPT, but rather compares NPT to both VRPT and Desktop conditions, making it unclear whether or not some of the findings in Study 1 were replicated in Study 2. In order to be able to compare the IOS, dehumanization, empathy, and personal distress findings from Study 2 with those of Study 1 the NPT and VRPT conditions were compared to each other.

In Study 2, there were no significant differences in IOS scores between the NPT and VRPT condition (t (211) = -1.59, $p = .554$; 95 CI [-0.64, 0.34]). There were also no significant differences immediately after the perspective-taking tasks in terms of dehumanization scores ($t(191) = -0.53$, $p = .594$; 95 CI [-6.10, 3.50]). Both of these results replicate the findings from Study 1. There were also no significant differences in empathy ($t(215) = -0.24$, $p = .812$; 95 CI [-6.10, 3.50]) or personal distress ($t(191) = -0.53$, $p = .594$; 95 CI [-6.10, 3.50]) between the NPT and VRPT conditions; this pattern deviates from Study 1 where there were significant differences between NPT and VRPT immediately after the intervention.

Correlations among Dependent Variables. All simple correlations are included in Table 6. A Holm-Bonferroni correction was implemented to account for multiple comparisons. Results show the adjusted p-values. The correlation analysis showed that IOS scores were significantly and positively correlated with self-reported empathy (r = .15, $p = .025$), and social presence (r = .20, $p = .049$). In other words, participants who reported feeling more connected to the homeless also reported more empathy, and felt more copresent with virtual humans in the mediated perspective-taking tasks (i.e., Desktop and VRPT). Self-reported empathy was significantly and positively correlated with personal distress (r = .39, $p < .001$) and social presence (r = .23, $p = .006$).

Behavioral measures. Support for Proposition A. There was a significant effect of condition on support for Proposition A: F (3, 435) = 4.29, $p = .005$, $\eta^2 = .03$. There was no significant difference between the Information condition and the three perspective-taking conditions (F (1, 435) = 0.107, $p = 0.744$). However, participants in the mediated perspective-taking conditions (i.e., Desktop and VRPT) supported Proposition A significantly more than participants in the NPT condition (F (1, 435) = 7.99, $p = .005$). There was also a significant difference between the VRPT and Desktop conditions (F (1, 435) = 4.77, $p = .029$) where participants in the VRPT condition reported supporting Proposition A significantly more than participants in the Desktop condition.

Signing Petition supporting Proposition A. There was a significant effect of condition on the proportion of petitions signed (Fisher's exact test: $p = < .001$). The first planned contrast (Information vs all perspective-taking conditions) showed that there was no significant effect of perspective-taking on petitions signed ($z = 0.36$, $p = .722$). However, the second planned contrast between mediated and traditional perspective-taking tasks showed that a significantly higher proportion of participants in the Desktop and VRPT conditions signed the petition in comparison to the participants in the NPT condition ($z = -2.53$, $p = .0113$). Finally, significantly more participants in the VRPT condition signed the petition than participants in the Desktop condition ($z = -3.25$, $p = .001$).

Table 6. Simple correlations among dependent variables for Study 2.

	Dehumanization	Empathy	Personal Distress	Social Presence	Support Prop A	Donation
IOS	0.13	**0.15** [a]	0.07	**0.20** [a]	0.08	0.01
Dehumanization		0.06	-0.03	-0.06	0.03	-0.04
State Empathy			**0.39** [c]	**0.23** [c]	0.03	0.04
Personal Distress				0.15	0.07	0.07
Social Presence					-0.02	-0.08
Support Prop A						0.11
Donation						

[a] positive at $p < .05$.

[b] positive at $p < .01$.

[c] positive at $p < .001$.

https://doi.org/10.1371/journal.pone.0204494.t006

These results show that mediated perspective-taking, regardless of immersion level, is more effective at encouraging political action in the form of signed petitions than NPT. However, increased immersion leads to significantly more signed petitions (significantly more participants in VRPT condition signed the petition than participants in the Desktop condition).

Donation Question. When it comes to the amount of money donated to a homeless shelter, there was no significant effect across conditions: F (3,435) = .528, p = .663, η^2 = .004. On average, participants in all conditions donated between $7 and $8 dollars to a homeless shelter, with 70% of all participants (n = 310) donating the maximum amount possible ($10). Since such a large percentage of participants chose the highest possible amount, this measure was at ceiling, and did not let us accurately conclude whether or not a specific type of empathy intervention was more effective at promoting prosocial behaviors in the form of donations than the others.

Replication of Study 1. In order to be able to compare the behavioral measures from Study 2 with those of Study 1, the NPT and VRPT conditions were compared to each other. Results showed that there was a significant difference in self-reported support for Proposition A (t (217) = 3.53, $p < .001$; 95 CI [.19, .68]). However, there was no significant difference in support for Proposition A in Study 1. When it comes to donations to a homeless shelter, there were no significant differences in amount donated between the NPT and VRPT participants (t(217) = 0.57, p = .567; 95 CI [-0.66, 1.21]), replicating the results from Study 1. Also replicating the results of Study 1, there was a significantly higher proportion of participants in the VRPT condition that signed the petition supporting affordable housing than participants in the NPT condition (t(218) = 3.61, $p < .001$; 95 CI [0.56, 1.87]).

Discussion

Overall, the self-report results show that the Information condition, which solely provided facts about the homeless population, was less effective at making participants feel empathetic and connected to the homeless than any of the perspective-taking conditions. Participants in the Information condition also reported feeling significantly less distressed and rated the homeless as less evolved than any of the perspective-taking participants. However, there was no significant difference between the Information condition and the VRPT condition in the proportion of people who signed the petition. These results show that information interventions can help promote prosocial behaviors, however, it is unclear whether empathy, social desirability, or increased awareness motivated these behaviors.

Showing the opposite pattern, participants in the NPT and Desktop condition reported feeling as empathetic and connected to the homeless as the participants in the VRPT condition. However, a smaller proportion of participants in the NPT and Desktop conditions signed the petition in support of Proposition A than participants in the VRPT condition. The immersive experience of becoming homeless in an IVE resulted in a significantly higher proportion of participants exhibiting helpful behaviors toward the homeless in the form of signing a petition when compared to traditional and less immersive perspective-taking tasks.

General discussion

Across two studies, we compared the effect of VR perspective-taking tasks against more traditional and less immersive perspective-taking tasks. We hypothesized that the more immersive the perspective-taking task, the more empathy and prosocial behaviors participants would exhibit toward the homeless.

Study 1 was a longitudinal investigation that compared the effects of a VR perspective-taking task against a more traditional, narrative-based perspective-taking task at the time of the intervention and over the course of eight weeks. Results showed that there was no significant difference in self-other overlap (i.e., the extent to which participants felt connected to the homeless) at the time of the intervention or over the course of eight weeks. At the time of the intervention, participants in the VRPT condition reported feeling more empathy toward the homeless and more personal distress than participants in the NPT condition. However, over the course of the eight weeks after the intervention, these differences dissipated and there was no significant difference between conditions in self-reported empathy or personal distress. Given we employed an imagine-self perspective taking task, the empathy and personal distress results were expected, and replicate Batson, Early, and Salvarani's [17] results. These findings add to the literature by providing empirical evidence showing that imagine-self perspective-taking tasks, regardless of delivery medium, result in a combination of other-oriented empathy and self-oriented distress.

Unlike reported empathy and personal distress, there were no significant differences between conditions in blatant dehumanization or attitudes toward the homeless immediately after the intervention. Over time, however, participants in the VRPT condition had more positive attitudes and thought of the homeless as more evolved than participants in the NPT condition. Even though there was no significant difference of blatant dehumanization at the time of the intervention, these results show that the positive effect of traditional perspective-taking on perceptions of the homeless (i.e., lack of dehumanization and better attitudes) is prolonged with the use of VR. These results are consistent with past research showing that attitudes toward a specific social target increase significantly over time even as empathic feelings dissipate [3,71].

Participants in the VRPT and NPT conditions reported similar rates of support for Proposition A immediately after their respective perspective-taking tasks. However, a significantly higher proportion of participants in the VRPT condition signed the petition supporting Proposition A. These results are consistent to those of Rosenberg, Baughman, and Bailenson [55], who saw no significant difference in intention to help but saw significant differences in actual helping behaviors.

Two weeks after the intervention, participants were asked to write a letter to their elected officials about the issue of homeless. Participants in the VRPT condition used significantly more first-person plural pronouns (e.g., we, our, us). Sample statements include "We must find ways to address the reasons why people become and stay homeless: job loss, mental health issues, high cost of living, lack of affordable housing" and "We must strive to be a community

[that] reaches out to support all of us, and we must think about ways to support those of us who currently do not have a place to stay." Another VRPT participant wrote "Homelessness is a pressing issue in our community, and we have clear moral and civic obligations to take action to help people in need." When writing to their elected officials about the issue of homelessness and its possible solutions, participants in the VRPT condition included themselves as part of the solution rather than telling the elected official what they or the government should do.

Four weeks after the intervention, participants were asked the extent to which they agreed with Measure B, a measure that advocated for affordable housing just like Proposition A. Participants in the VRPT condition reported more support for Measure B than NPT participants. These results showed that over time, participants in the VRPT condition continued to support political initiatives that could actually benefit the homeless population, whereas the level of support for these kinds of initiatives decreased significantly over time for the participants in the NPT condition.

Study 2 expanded on Study 1 by further exploring the mechanisms that caused differences between the narrative-based perspective-taking and the VR perspective-taking task. We compared the effect of three different types of perspective-taking tasks, each varying in levels of immersion, against each other and against a fact-driven information intervention in order to more accurately explore the effect of immersion and type of empathy intervention on elicited empathy and prosocial behaviors toward the homeless. Study 2 also differed from Study 1 in that we used a larger, more racially diverse sample with participants ranging from 15 to 88 years old. In line with our predictions, and replicating past studies [3], the results of Study 2 showed that participants in all three perspective-taking conditions reported feeling more empathetic toward the homeless compared to participants in the Information condition who did not perform a perspective-taking task at all.

When comparing the three different types of perspective-taking tasks, it was participants in the VRPT condition who reported feeling more connected and empathetic toward the homeless than the less immersive Desktop condition and the NPT condition. Replicating results of Study 1, a significantly higher proportion of participants in the VRPT condition signed the petition supporting efforts to increase affordable housing than participants in the Desktop or NPT conditions.

On most outcome variables (i.e., self-other overlap, reported empathy, and personal distress) participants in the information condition reported feeling less connected and less empathetic toward the homeless than participants in the three perspective-taking conditions. At first glance, these results were expected since past research has shown that giving people information does not always change their attitudes [29,74]. When it comes to petition signatures, more participants in the VRPT condition signed the petition than participants in the Information condition. However, this difference was not statistically significant. These results show that fact-driven interventions can also be successful at promoting prosocial behaviors. Additionally, the discrepancy in self-reported empathy results and signed petitions for participants in the Information condition suggests that empathy or self-other overlap were not the only mechanisms that led to prosocial behaviors.

In terms of dehumanization, there was no significant difference between the Information condition and the PT conditions. However, there was a significant difference between the immersive conditions and the NPT condition. This result was expected since we hypothesized that higher levels of immersion would lead to more empathy and more positive attitudes. There was also a marginally significant difference between the Desktop and the VRPT conditions, with Desktop participants evaluating the homeless as more evolved than the VRPT participants immediately after the intervention. This result was unexpected, but could potentially be explained by the fact that participants in the VRPT condition viscerally experienced

harassment in one of the bus scenes from a fellow rider. This experience may have influenced how participants thought of the homeless since they were personally accosted by someone inside the VR experience. While participants in the Desktop condition went through the exact same scene, they were not immersed in the environment, and their personal space was not violated. However, these results did not negatively affect prosocial behaviors. A significantly higher proportion of participants in the VRPT condition signed the petition in support of Proposition A than participants in the Desktop condition.

The dehumanization score results in Study 2 may be an example of how depicting a real scenario, such as being accosted by someone inside an IVE, can have undesirable or unintended effects when compared to depicting the same scenario through a less immersive medium (e.g., desktop computer or video). This concern becomes particularly salient as consumer adoption of VR systems continues to increase and empathy-driven VR experiences become more available to the public. Past research has demonstrated that short VR experiences can have visceral reactions that affect the way a person thinks, feels, and behaves [38, 53, 55]. There is also cogent evidence that virtual humans, such as the man in the bus scene, exert social influence over users [75]. In general, as VR scales up and more people have access to it, it is important for designers and researchers to pilot test their experiences to ensure the experiences they create are having the intended effect.

Study 2 replicated the IOS and dehumanization scale findings from Study 1. In both studies, there was no significant difference in IOS and dehumanization scores between the NPT and VRPT conditions. The results of both of these studies provide more evidence suggesting that when VR perspective-taking tasks are employed, valuations of the social target do not increase significantly at the time of the intervention, but tend to increase over time, and last up to two months after the intervention. However, when comparing the empathy and personal distress results, Study 2 did not replicate the findings from Study 1. Study 1 found that participants in the VRPT condition reported more empathy and personal distress immediately after the intervention, but Study 2 found no significant differences between these two conditions. One of the likely reasons the empathy and personal distress findings from Study 1 were not replicated is that the sample used in Study 2 was much larger and more demographically diverse than the sample used in Study 1. The Study 2 sample was also mostly composed of volunteers that had never used a VR headset before. It is possible that the novelty of the equipment acted as a distraction and prevented participants from focusing on the experience itself, resulting in lower empathy and personal distress scores.

In a longitudinal study, Bailenson and Yee [76] found that, over time, VR users change the way they behave inside virtual environments and suggest that the novelty of VR technology has an effect on virtual social interactions. They also propose that the way experienced VR users go through a VR experience and use the technology is different from the way that first-time users experience the technology [76]. As VR empathy-driven interventions begin to be used at scale, it will be important to control for the level of experience of the users. However, at this time, more research is needed in order to assess the full extent to which the benefits of VRPT are moderated by the participant's level of experience with VR technology.

Overall, these findings replicate past VR studies in which participants who embodied the perspectives of other groups (e.g., colorblind individuals and the elderly) performed more helping behaviors than participants who did not use VR [5, 40]. Results from the present investigation further confirm that VR, compared to other types of traditional or mediated perspective-taking, often leads to better attitudes and a higher proportion of users signing petitions in support of helpful initiatives for outgroup members.

Limitations and future directions

There are a number of limitations to these studies. First, the manipulation check assessing the level of spatial presence felt by participants between the different conditions in Study 1 and Study 2 was derived from participants' written responses to an open-ended question instead of a quantitative, self-reported measure of spatial presence. Additionally, this investigation only included a quantitative measure of social presence. In future studies, participants will be explicitly asked to rate the extent to which they felt spatially present, socially present, and self-present to better understand how these different dimensions of presence impact empathic and behavioral outcomes after VR perspective taking tasks.

Second, in Study 2, participants in all four conditions donated money to a homeless shelter at similar rates (between $7 and $8 dollars). Even though participants in the VRPT condition donated more money than the rest of the conditions, the donation question was at ceiling, meaning more than 70% of all participants across the four conditions chose to donate the maximum amount possible. The donation question was flawed in that it did not give a wide enough range of possible answers in order to address the possible differences between conditions or between different types of prosocial behaviors (i.e., signatures for Proposition A). Participants were also compensated with $10 dollars for their participation, therefore it is likely that the full burden of donating money was not felt by some of the participants. Even though on average most participants donated around the same amount of money, the standard deviations were high for each condition. Such high variability in amount donated across conditions prevented the researchers from being able to examine whether or not different types of perspective-taking tasks led to different levels of prosocial behaviors. Future studies should ask participants how much money out of their own pocket they would like to donate or allow for a wider range of possible answers (e.g., $0 to $100) in order to address these limitations.

It is important to note that the VR experience used immersed participants in an environment specifically designed to provide a visceral experience of what it would be like to be homeless from the first person perspective. However, the VR experience lasted approximately 15 minutes and was not able to simulate some of the psychological and physiological burdens that homeless people experience (e.g., desperation or hunger). These limitations prevent participants from actually experiencing what it would be like to become homeless. Another limitation of the technology is that despite the high level of interactivity within the experience, it still did not allow to participants to interact with the virtual world the way they naturally interact with the real world.

Another limitation of these studies is that attitudes toward the homeless were not measured before the intervention. Even though the participants were randomly selected into each condition, it is possible that participants already had set views regarding the homeless that the researchers were not aware of. Future studies should measure pre-existing biases and attitudes toward the homeless in order to more accurately assess the effect of the different types of empathy interventions. Additionally, the design of the present investigation lacked a pure control condition. In Study 2, the Information condition led to an unexpected number of donations and proportion of signatures supporting affordable housing. A pure control condition, without any kind of empathy intervention, would be necessary in order to address the mechanism that led to these prosocial behaviors.

Results of Study 2 found that a similar proportion of participants in the VRPT and the Information conditions signed the petition supporting affordable housing for vulnerable populations. Given participants in any of the perspective-conditions reported feeling more connected, more empathetic, and more personally distressed than participants in the Information condition, it is unlikely that empathy led to these behavioral outcomes. It is possible that social

desirability or increased awareness about the homeless led to these results. Future studies should compare a pure control condition, an information condition, and a perspective-taking task condition to try to understand the mechanism that makes information-driven interventions successful at promoting prosocial behaviors.

Despite these limitations, the results of this investigation provide encouraging evidence supporting the use of IVEs to promote empathy and prosocial behaviors toward extreme outgroup members. However, it is important to note the small effect sizes, and the fact that for most of the self-report measures in Study 2 there were no significant differences between the three perspective-taking conditions. This might be due to the fact that the content was as rigorous in all of these conditions, and that the level of immersion did not have as much of an effect eliciting more empathy. Future studies should try to replicate these results while targeting a different social group in order to assess the generalizability of VR as a perspective-taking tool, and test whether it is the content and context that allows perspective-taking to promote empathy or rather the modality of the medium by which the intervention is administered. Future studies should also consider using implicit measures of empathy in addition to explicit self-report measures in order to further understand the patterns exhibited between self-report and behavioral measures.

In the current version of the VR experience, participants are able to make choices about what they want to sell, what search strategy they implement in the car, how they react and respond to the men in the bus, and who and how they interact with the other homeless people in the bus. Participants are constantly seeing how their decisions affect the narrative and this, in turn, provides a highly individualized and interactive experience for the participant. Future studies should compare this type of VR experience to one in which participants do not have agency (i.e., a VR experience without interactivity) in order to assess the role that interactivity plays in self-reported empathy, attitudes toward a specific social target, and prosocial behaviors. We speculate that highly interactive and responsive IVEs, where participants are able to see how their own actions manifest in the IVE in real-time, may lead to higher levels of spatial and social presence, more self-reported empathy, better attitudes, and more prosocial behaviors. However, whether or not this difference is significant requires more empirical evidence.

Moreover, future research should compare imagine-self and imagine-other VR perspective-taking tasks. The present study employed imagine-self perspective-taking tasks and found some effects on behavioral measures, however, the motivation that led to these behaviors is not clear. Past research suggests that imagine-self tasks evoke both other-oriented empathy and self-oriented distress, and that this combination of emotions can lead to a stronger motivation to help when compared to imagine-other tasks [17]. Thus, it would be expected that more helping behaviors would be performed with an imagine-self task than an imagine-other task. A study comparing these two types of perspective-taking would be able to discern whether the motivation to help was altruistic or egoistic, and compare the amount of prosocial behaviors performed by participants.

Finally, future research should examine the effect of novelty on experimental outcomes. Past research has demonstrated that users change the way they behave inside virtual environments once they have been exposed to the technology a number of times [76]. These results suggest that the level of experience or familiarity with VR technology may have an effect on the way users interact with each other and in the way that they experience VR in general. However, more research is needed in order to assess the impact of novelty on empathy-driven VR experiences and interventions. Future studies should control for the number of times participants have used VR technology to see if novelty moderates elicited empathy and prosocial behaviors after a VR perspective-taking task.

Conclusion

The present investigation found that over the course of eight weeks, participants who completed a VR perspective-taking task had more positive attitudes and signed a petition supporting helpful initiatives toward the homeless at significantly higher rates than the participants who just imagined what it would be like to become homeless or performed a less immersive perspective-taking task. The investigation also found that narrative-based and mediated perspective-taking interventions, regardless of immersion level, are more effective at increasing self-reported empathy than interventions without any perspective-taking tasks. The results of this investigation provide evidence suggesting that VR perspective-taking tasks may be more effective at improving attitudes toward specific social targets and motivating prosocial behaviors in the form of signed petitions in support of helpful initiatives than traditional and less immersive perspective-taking tasks.

Supporting information

S1 Appendix. Materials. Contains all of the referenced texts.
(DOCX)

S2 Appendix. Questionnaires. Contains all of the questionnaires used in Study 1 and Study 2.
(DOCX)

S1 File. Additional analyses. Contains additional analysis examining social presence as a moderator.
(DOCX)

S1 Dataset. Complete dataset for Study 1. Contains all of the raw data and open-ended responses used for analysis in Study 1.
(XLSX)

S2 Dataset. Complete dataset for Study 2. Contains all of the raw data and open-ended responses used for analysis in Study 2.
(XLSX)

Acknowledgments

This research was supported by a grant from the Robert Wood Johnson Foundation.

Author Contributions

Conceptualization: Fernanda Herrera, Jeremy Bailenson, Erika Weisz, Elise Ogle, Jamil Zaki.

Data curation: Fernanda Herrera.

Formal analysis: Fernanda Herrera.

Funding acquisition: Jeremy Bailenson.

Investigation: Fernanda Herrera, Elise Ogle.

Methodology: Fernanda Herrera, Jeremy Bailenson, Erika Weisz, Jamil Zaki.

Project administration: Fernanda Herrera, Elise Ogle.

Resources: Jeremy Bailenson.

Supervision: Jeremy Bailenson.

Writing – original draft: Fernanda Herrera.

Writing – review & editing: Fernanda Herrera, Jeremy Bailenson, Erika Weisz, Elise Ogle, Jamil Zaki.

References

1. Hoffman ML. Empathy and moral development: Implications for caring and justice. Cambridge University Press; 2001 Nov 12.
2. Batson CD, Dyck JL, Brandt JR, Batson JG, Powell AL, McMaster MR, et al. Five studies testing two new egoistic alternatives to the empathy-altruism hypothesis. Journal of personality and social psychology. 1988 Jul; 55(1):52. PMID: 3418490
3. Batson CD, Polycarpou MP, Harmon-Jones E, Imhoff HJ, Mitchener EC, Bednar LL, et al. Empathy and attitudes: Can feeling for a member of a stigmatized group improve feelings toward the group?. Journal of personality and social psychology. 1997 Jan; 72(1):105. PMID: 9008376
4. Batson CD, Ahmad NY. Using empathy to improve intergroup attitudes and relations. Social Issues and Policy Review. 2009 Dec 1; 3(1):141–77.
5. Ahn SJ, Le AM, Bailenson J. The effect of embodied experiences on self-other merging, attitude, and helping behavior. Media Psychology. 2013 Jan 1; 16(1):7–38.
6. Milk C. How virtual reality can create the ultimate empathy machine [Internet]. TED: Ideas worth spreading. 2015 [cited 2018Feb6]. https://www.ted.com/talks/chris_milk_how_virtual_reality_can_create_the_ultimate_empathy_machine
7. De la Peña N, Weil P, Llobera J, Giannopoulos E, Pomés A, Spanlang B, et al. Immersive journalism: immersive virtual reality for the first-person experience of news. Presence: Teleoperators and virtual environments. 2010 Aug 1; 19(4):291–301.
8. Durbin J. Vive To Spend $10 Million on VR Projects That 'Create Positive Impact' [Internet]. UploadVR. UploadVR; 2017 [cited 2017 Sep 6]. https://uploadvr.com/vive-vr-for-impact-10-million/
9. HTC.VR for Impact. [cited 2017 Jun 28]. https://vrforimpact.com/
10. Cialdini RB, Brown SL, Lewis BP, Luce C, Neuberg SL. Reinterpreting the empathy—altruism relationship: When one into one equals oneness. Journal of personality and social psychology. 1997 Sep; 73 (3):481. PMID: 9294898
11. Batson CD, Duncan BD, Ackerman P, Buckley T, Birch K. Is empathic emotion a source of altruistic motivation?. Journal of personality and Social Psychology. 1981 Feb; 40(2):290.
12. Cialdini RB. Altruism or egoism? That is (still) the question. Psychological Inquiry. 1991 Apr 1; 2 (2):124–6.
13. Pierce JR, Kilduff GJ, Galinsky AD, Sivanathan N. From glue to gasoline: How competition turns perspective takers unethical. Psychological science. 2013 Oct; 24(10):1986–94. https://doi.org/10.1177/0956797613482144 PMID: 23955353
14. Todd AR, Galinsky AD. Perspective-taking as a strategy for improving intergroup relations: Evidence, mechanisms, and qualifications. Social and Personality Psychology Compass. 2014 Jul 1; 8(7):374–87.
15. Galinsky AD, Ku G, Wang CS. Perspective-taking and self-other overlap: Fostering social bonds and facilitating social coordination. Group Processes & Intergroup Relations. 2005 Apr; 8(2):109–24.
16. Krauss RM, Fussell SR. Perspective-taking in communication: Representations of others' knowledge in reference. Social cognition. 1991 Mar; 9(1):2–4.
17. Batson CD, Early S, Salvarani G. Perspective taking: Imagining how another feels versus imaging how you would feel. Personality and social psychology bulletin. 1997 Jul; 23(7):751–8.
18. Batson CD, Lishner DA, Carpenter A, Dulin L, Harjusola-Webb S, Stocks EL, et al. ". . . As You Would Have Them Do Unto You": Does Imagining Yourself in the Other's Place Stimulate Moral Action?. Personality and Social Psychology Bulletin. 2003 Sep; 29(9):1190–201. https://doi.org/10.1177/0146167203254600 PMID: 15189613
19. Lamm C, Batson CD, Decety J. The neural substrate of human empathy: effects of perspective-taking and cognitive appraisal. Journal of cognitive neuroscience. 2007 Jan; 19(1):42–58. https://doi.org/10.1162/jocn.2007.19.1.42 PMID: 17214562
20. Tarrant M, Calitri R, Weston D. Social identification structures the effects of perspective taking. Psychological science. 2012 Sep; 23(9):973–8. https://doi.org/10.1177/0956797612441221 PMID: 22851441
21. Skorinko JL, Sinclair SA. Perspective taking can increase stereotyping: The role of apparent stereotype confirmation. Journal of Experimental Social Psychology. 2013 Jan 1; 49(1):10–8.
22. Hodges S, Wegner DM, Ickes WJ. Automatic and controlled empathy. Empathic accuracy. 1997.

23. Gehlbach H, Marietta G, King AM, Karutz C, Bailenson JN, Dede C. Many ways to walk a mile in another's moccasins: Type of social perspective taking and its effect on negotiation outcomes. Computers in Human Behavior. 2015 Nov 30; 52:523–32.
24. Slater M, Wilbur S. A framework for immersive virtual environments (FIVE): Speculations on the role of presence in virtual environments. Presence: Teleoperators & Virtual Environments. 1997 Dec; 6 (6):603–16.
25. Steuer J. Defining virtual reality: Dimensions determining telepresence. Journal of communication. 1992 Dec 1; 42(4):73–93.
26. Eckel C, Grossman PJ, Milano A. Is more information always better? An experimental study of charitable giving and Hurricane Katrina. Southern Economic Journal. 2007 Oct 1:388–411.
27. Brophy J. The development of knowledge and empathy. Social Education. 1999 Jan; 63(1):39–45.
28. Oliver MB, Dillard JP, Bae K, Tamul DJ. The effect of narrative news format on empathy for stigmatized groups. Journalism & Mass Communication Quarterly. 2012 Jun; 89(2):205–24.
29. Teachman BA, Gapinski KD, Brownell KD, Rawlins M, Jeyaram S. Demonstrations of implicit anti-fat bias: the impact of providing causal information and evoking empathy. Health Psychology. 2003 Jan; 22 (1):68. PMID: 12558204
30. Greitemeyer T, Osswald S. Effects of prosocial video games on prosocial behavior. Journal of personality and social psychology. 2010 Feb; 98(2):211. https://doi.org/10.1037/a0016997 PMID: 20085396
31. Gentile DA, Anderson CA, Yukawa S, Ihori N, Saleem M, Ming LK, et al. The effects of prosocial video games on prosocial behaviors: International evidence from correlational, longitudinal, and experimental studies. Personality and Social Psychology Bulletin. 2009 Jun; 35(6):752–63. https://doi.org/10.1177/0146167209333045 PMID: 19321812
32. Hailpern J, Danilevsky M, Harris A, Karahalios K, Dell G, Hengst J. ACES: promoting empathy towards aphasia through language distortion emulation software. InProceedings of the SIGCHI Conference on Human Factors in Computing Systems 2011 May 7 (pp. 609–618). ACM.
33. Brown SA. Implementing a brief hallucination simulation as a mental illness stigma reduction strategy. Community mental health journal. 2010 Oct 1; 46(5):500–4. https://doi.org/10.1007/s10597-009-9229-0 PMID: 19669675
34. Silverman AM, Gwinn JD, Van Boven L. Stumbling in their shoes: Disability simulations reduce judged capabilities of disabled people. Social Psychological and Personality Science. 2015 May; 6(4):464–71.
35. Hand S, Varan D. Interactive narratives: Exploring the links between empathy, interactivity and structure. InEuropean Conference on Interactive Television 2008 Jul 3 (pp. 11–19). Springer, Berlin, Heidelberg.
36. Vorderer P, Knobloch S, Schramm H. Does entertainment suffer from interactivity? The impact of watching an interactive TV movie on viewers' experience of entertainment. Media Psychology. 2001 Nov 1; 3(4):343–63.
37. Behm-Morawitz E, Pennell H, Speno AG. The effects of virtual racial embodiment in a gaming app on reducing prejudice. Communication Monographs. 2016 Jul 2; 83(3):396–418.
38. Ahn SJ, Bostick J, Ogle E, Nowak KL, McGillicuddy KT, Bailenson JN. Experiencing nature: Embodying animals in immersive virtual environments increases inclusion of nature in self and involvement with nature. Journal of Computer-Mediated Communication. 2016 Sep 14; 21(6):399–419.
39. Zaki J. Empathy: a motivated account. Psychological bulletin. 2014 Nov; 140(6):1608. https://doi.org/10.1037/a0037679 PMID: 25347133
40. Oh SY, Bailenson J, Weisz E, Zaki J. Virtually old: Embodied perspective taking and the reduction of ageism under threat. Computers in Human Behavior. 2016 Jul 1; 60:398–410.
41. Macrae CN, Bodenhausen GV. Social cognition: Thinking categorically about others. Annual review of psychology. 2000 Feb; 51(1):93–120.
42. Blascovich J, Loomis J, Beall AC, Swinth KR, Hoyt CL, Bailenson JN. Immersive virtual environment technology as a methodological tool for social psychology. Psychological Inquiry. 2002 Apr 1; 13 (2):103–24.
43. Barsalou LW. Grounded cognition: Past, present, and future. Topics in cognitive science. 2010 Oct 1; 2 (4):716–24. https://doi.org/10.1111/j.1756-8765.2010.01115.x PMID: 25164052
44. Wilson M. Six views of embodied cognition. Psychonomic bulletin & review. 2002 Dec 1; 9(4):625–36.
45. Oppezzo M, Schwartz DL. Give your ideas some legs: The positive effect of walking on creative thinking. Journal of experimental psychology: learning, memory, and cognition. 2014 Jul; 40(4):1142. https://doi.org/10.1037/a0036577 PMID: 24749966
46. Williams LE, Bargh JA. Experiencing physical warmth promotes interpersonal warmth. Science. 2008 Oct 24; 322(5901):606–7. https://doi.org/10.1126/science.1162548 PMID: 18948544

47. Blanke O. Multisensory brain mechanisms of bodily self-consciousness. Nature Reviews Neuroscience. 2012 Aug; 13(8):556. https://doi.org/10.1038/nrn3292 PMID: 22805909

48. Brisswalter J, Collardeau M, René A. Effects of acute physical exercise characteristics on cognitive performance. Sports medicine. 2002 Aug 1; 32(9):555–66. PMID: 12096929

49. Kalyanaraman SS, Penn DL, Ivory JD, Judge A. The virtual doppelganger: effects of a virtual reality simulator on perceptions of schizophrenia. The Journal of nervous and mental disease. 2010 Jun; 198 (6):437–43. https://doi.org/10.1097/NMD.0b013e3181e07d66 PMID: 20531123

50. Yee N, Bailenson JN. Walk a mile in digital shoes: The impact of embodied perspective-taking on the reduction of negative stereotyping in immersive virtual environments. Proceedings of PRESENCE. 2006 Aug; 24:26.

51. Peck TC, Seinfeld S, Aglioti SM, Slater M. Putting yourself in the skin of a black avatar reduces implicit racial bias. Consciousness and cognition. 2013 Sep 1; 22(3):779–87. https://doi.org/10.1016/j.concog.2013.04.016 PMID: 23727712

52. Banakou D, Hanumanthu PD, Slater M. Virtual embodiment of white people in a black virtual body leads to a sustained reduction in their implicit racial bias. Frontiers in human neuroscience. 2016 Nov 29; 10:601. https://doi.org/10.3389/fnhum.2016.00601 PMID: 27965555

53. Ahn SJ, Bailenson JN, Park D. Short-and long-term effects of embodied experiences in immersive virtual environments on environmental locus of control and behavior. Computers in Human Behavior. 2014 Oct 1; 39:235–45.

54. Oh SY, Shriram K, Laha B, Baughman S, Ogle E, Bailenson J. Immersion at scale: Researcher's guide to ecologically valid mobile experiments. InVirtual Reality (VR), 2016 IEEE 2016 Mar 19 (pp. 249–250). IEEE.

55. Rosenberg RS, Baughman SL, Bailenson JN. Virtual superheroes: Using superpowers in virtual reality to encourage prosocial behavior. PloS one. 2013 Jan 30; 8(1):e55003. https://doi.org/10.1371/journal.pone.0055003 PMID: 23383029

56. Bailenson JN, Aharoni E, Beall AC, Guadagno RE, Dimov A, Blascovich J. Comparing behavioral and self-report measures of embodied agents' social presence in immersive virtual environments. In Proceedings of the 7th Annual International Workshop on PRESENCE 2004 Oct (pp. 1864–1105).

57. Fiske ST, Cuddy AJ, Glick P, Xu J. A model of (often mixed) stereotype content: competence and warmth respectively follow from perceived status and competition. Journal of personality and social psychology. 2002 Jun; 82(6):878. PMID: 12051578

58. Harris LT, Fiske ST. Dehumanizing the lowest of the low: Neuroimaging responses to extreme outgroups. Psychological science. 2006 Oct; 17(10):847–53. https://doi.org/10.1111/j.1467-9280.2006.01793.x PMID: 17100784

59. Hotel 22. (2014). Directed by E. Lo. USA.

60. Davis MH. Measuring individual differences in empathy: Evidence for a multidimensional approach. Journal of personality and social psychology. 1983 Jan; 44(1):113.

61. Schumann K, Zaki J, Dweck CS. Addressing the empathy deficit: Beliefs about the malleability of empathy predict effortful responses when empathy is challenging. Journal of Personality and Social Psychology. 2014 Sep; 107(3):475. https://doi.org/10.1037/a0036738 PMID: 25133727

62. Nowak KL, Biocca F. The effect of the agency and anthropomorphism on users' sense of telepresence, copresence, and social presence in virtual environments. Presence: Teleoperators & Virtual Environments. 2003 Oct; 12(5):481–94.

63. Bailenson JN, Yee N. Virtual interpersonal touch: Haptic interaction and copresence in collaborative virtual environments. Multimedia Tools and Applications. 2008 Mar 1; 37(1):5–14.

64. Aron A, Aron EN, Smollan D. Inclusion of other in the self scale and the structure of interpersonal closeness. Journal of personality and social psychology. 1992 Oct; 63(4):596.

65. Kteily N, Bruneau E, Waytz A, Cotterill S. The ascent of man: Theoretical and empirical evidence for blatant dehumanization. Journal of personality and social psychology. 2015 Nov; 109(5):901. https://doi.org/10.1037/pspp0000048 PMID: 26121523

66. Pennebaker JW. The secret life of pronouns. New Scientist. 2011 Sep 3; 211(2828):42–5.

67. Oh SY, Bailenson J, Krämer N, Li B. Let the avatar brighten your smile: Effects of enhancing facial expressions in virtual environments. PloS one. 2016 Sep 7; 11(9):e0161794. https://doi.org/10.1371/journal.pone.0161794 PMID: 27603784

68. Fleiss JL, Cohen J. The equivalence of weighted kappa and the intraclass correlation coefficient as measures of reliability. Educational and psychological measurement. 1973 Oct; 33(3):613–9.

69. Bollen KA, Curran PJ. Latent curve models: A structural equation perspective. John Wiley & Sons; 2006 Jan 3.

70. Mirman D. Growth curve analysis and visualization using R. Boca Raton, FL: CRC Press; 2014.
71. Batson CD, Turk CL, Shaw LL, Klein TR. Information function of empathic emotion: Learning that we value the other's welfare. Journal of personality and social psychology. 1995 Feb; 68(2):300.
72. Applied Survey Research. San Francisco Homeless Count. 2015 Watsonville, CA.
73. Davis MH, Conklin L, Smith A, Luce C. Effect of perspective taking on the cognitive representation of persons: a merging of self and other. Journal of personality and social psychology. 1996 Apr; 70(4):713. PMID: 8636894
74. Bandura A. Self-efficacy: The exercise of control. Macmillan; 1997 Feb 15.
75. Blascovich J. Social influence within immersive virtual environments. In The social life of avatars 2002 (pp. 127–145). Springer, London.
76. Bailenson JN, Yee N. A longitudinal study of task performance, head movements, subjective report, simulator sickness, and transformed social interaction in collaborative virtual environments. Presence: Teleoperators and Virtual Environments. 2006 Dec 1; 15(6):699–716.

419.7k

How Virtual Reality Can Be Used To Fight Racism And Prejudice In Society

by **KENNETH SEWARD JR.** • APRIL 29TH, 2017

Virtual reality has the power to break boundaries.

I always thought I'd be more prepared for "the talk" than my parents were. It would be a priority of mine to have historic references on hand, pull from life experiences, maybe even pick out an educational video from the library – you know, make things easier on myself. I made sure to learn as much as I could about this topic so that, by the time it was my turn to have that serious sit down with my kids, I would be ready for their questions. Being a writer, at the very least, I figured I'd be more able to articulate what to expect when it happens to them. What I've discovered though, after actually becoming a parent, was that no amount of preparation will make this talk any less stressful.

I mean, just how do you tell a young child that some people might not like them because of the color of their skin?

What I do find solace in is the fact that I won't have to first try to convince my daughters that racism exists; being our children, they'll listen to what my wife and I have to say. I'd also foresee the girls being naturally empathetic to our shared experiences. This in turn takes some of the edge off as it promotes uninhibited dialogue. As a family, we'll never yearn to discuss such things at the dinner table. But we'd at least be comfortable enough to openly deal with whatever comes our way. That truth is very important given

the nature of this topic. Not everyone is going to be willing to listen to my children's concerns like my wife and I will when life becomes more challenging for them.

As I continued to mull this over, I wondered how I could surround my children with people who'd be willing to listen to their concerns. As of right now, a decent amount of my white friends would rather shy away from the conversation. It's just too uncomfortable a topic. Although to a lesser extent, some feel as though our struggles are fabricated. That we either unnecessarily lump race into every dispute or that most of our difficulties in life, regardless of how we're treated, are solely attributed to our own actions. They're basically closed off to what a lot of people like me face on a daily basis. Still, I wondered what could be done outside of being careful of who we trust.

As the virtual reality industry continues to evolve though, it can be used to help people see things from the perspective of a minority in America. You know, wear someone else's shoes and all that. This sort of thing has helped people see past their own experiences in the past (it's a popular idiom for a reason).

I found the work of Dr. Jeremy Bailenson, Ph.D., a Thomas More Storke Professor in the Department of Communication at Stanford University. Among his numerous accolades, which included being the founding director of Stanford's Virtual Human Interaction Lab, his doctorate in cognitive psychology from Northwestern University piqued my interest.

Dr. Bailenson's main area of research deals with *digital human representation, especially in the context of immersive virtual reality*, as per his bio. For years, he's researched how VR can help to facilitate change in a person's perceptions of others. For instance, in an academic paper he co-published, "Walk a Mile in Digital Shoes", Dr. Bailenson describes an intervention method developed by Galinsky, A.D. and G. B. Moskowitz called Perspective Taking. It deals with how people see themselves vs. others during social interactions – when judging ourselves, we tend to rely on outside forces to justify our actions. That's not the case when judging others. That said, extensive research has shown that when we're asked to take on the perspective of the person we're judging, in a virtual setting, we tend to give them the same benefit of doubt we'd give ourselves.

Then there's also Courtney D. Cogburn, Ph.D., an Assistant Professor at Columbia University's School of Social Work. Dr. Cogburn has partnered with Dr. Bailenson on a project called "Examining Racism with VR". The idea was to use an immersive VR experience that allows participants to embody a person of color (seemingly, more Perspective Taking).

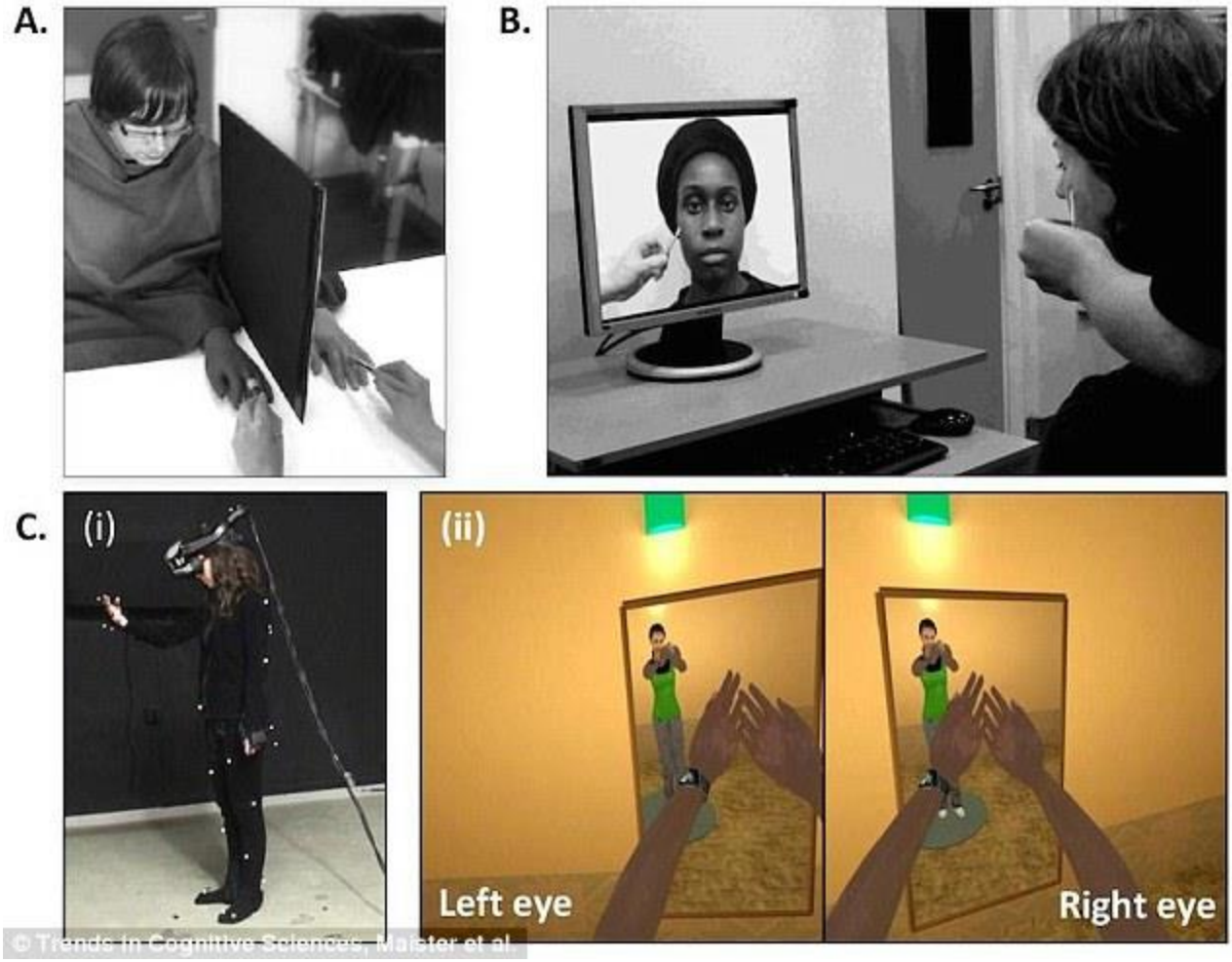
Dr. Cogburn explained how she was building on research that challenged the way people talked about or measured racism. "We tend to talk about racism as things we do to each other on the basis of race," Dr. Cogburn explained. "You know, being called a name, someone being followed in a store, or being directly harassed by a police officer.

[Dr. Bailenson and I] are trying to build on work that indicates that racism is actually much more complex and nuanced than those types of experiences…[Like] being praised for being articulate or someone being surprised that you didn't come from a single parent home or whatever stereotypes that might sneak their way into discourse in interactions."

Dr. Cogburn and Bailenson were also interested in how they could expose people to understanding the structural nature of racism. "[It's important to] understand the patterns of discrimination, policy and legislation that systematically disadvantages groups of people…if we hope to do anything about the effects that racism has on people's lives," Dr. Cogburn said. Unfortunately, something like that seems borderline impossible. It can be difficult to get people to listen when dealing with a single incident, let alone something more systemic.

She then detailed how they would be using qualitative data to help design content for the project. Everything is rooted and documented in reality, being careful not to take artistic license while constructing these scenarios. This is to help combat some of the resistance people will have towards the experiment; discounting its relevance as the scenarios "could never happen" in real life.

"Our team is comprised of about a dozen scholars, a majority of whom face issues of racial bias every day," Dr. Bailenson explained. "Part of the reason the design process is taking months is that we continually discuss the experiences of the team members, and spend time deliberating which types of experiences will actually be uniquely suited to VR."



Though their project is still in the design phase, using virtual reality as a means of evoking empathy is supported by sound research. Dr. Bailenson has worked multiple VR experiments dealing with a person's negative stereotypes. None of his studies found "the answer" to today's ills, given their complexity. Still, he was able to encourage a decrease in prejudiced thinking using different virtual programs. I particularly made note of the body transfer effect – where a person's brain helps them to believe they are the person seen in a mirror through their headset – created through diversity training software called the Minority Mirror. Used for businesses looking to improve social awareness, the software places people in the bodies of different minorities and in some cases, genders. All of a sudden, a white female is transformed into a black male. Once that happens, another avatar comes up to the person and starts saying negative things about the participant's race (that of which he or she sees in the mirror).

Interestingly, years of research revealed that the persons wearing the headset showed more empathy after coming in contact with the hateful avatar. When I asked if these feelings of empathy were short lived, Dr. Bailenson pointed out how his lab had "completed collecting data from an immersive empathy study that examines the effect of the treatment directly after the experience, two weeks out, four weeks out…we plan on analyzing the date in the next few months to examine how immersive VR compares to traditional role playing over time." In other words, the impact of the study doesn't end when the headset is removed.

The information I got from Dr. Bailenson and Dr. Cogburn was invaluable, in a sense that it gave me hope. Maybe people can have a change of heart and become more open to hearing what minorities have to say about racism. As uncomfortable as it may be, there's a real need for honest conversation among peers. I'm sure that their Examining Racism project, based on the progress made from similar software, will help to facilitate these difficult discussions given time.

While this is all true, I do fear that there are some unique challenges hindering the effectiveness of this sort of approach. Some of which, I noticed after talking with Nonny de la Peña.

Nonny de la Peña is an award winning journalist, having written for reputable news organizations like Newsweek and The New York Times, and has helped pioneer VR empathy projects. It could even be argued that she is more readily known for her work in immersive journalism, a genre she is credited in creating. As CEO of Emblematic Group – a digital reality media company that produces immersive VR content – she recreates events that people can experience in virtual reality. One of her major pieces, Project Syria, was the topic of our discussion. Not because of how powerful it was for those who experienced it in her studio, but because of the blow back she got after releasing it on Steam.

“We put Project Syria on Steam and the response was horrible,” de la Peña exclaimed. “Racist, vitriolic responses. I would expect that our game [wouldn’t be liked] as a game. We’ve [only] put up $35,000 to make it. So I would understand that kind of criticism. But no, it’s all racist.”

While she spoke, I headed over to Steam to check out the comments for myself. If you take away the comments about Project Syria not being an actual game, a valid point for

another discussion and perhaps more a result of Steam's user base consisting of almost entirely gamers, and people complaining about the visuals, a lot of what's left isn't appealing. With people claiming it to be "ISIS propaganda" and others asking "how to sink refugee boats", it's honestly disheartening.

For those that don't know, Project Syria (PS) is a piece that provides an immersive look at the effects of the Syrian civil war on the country's children using virtual reality. This is done by recreating a tragic event on the streets of Aleppo – a bomb goes off in refugee camp, wounding and killing many people living there – using found footage, photographs, and tons of documentation. Outside of improving the graphic fidelity of the piece (this was produced back in 2013), PS is probably the best chance anyone has of "experiencing" this sort of event other than actually visiting the country.

De la Peña didn't expect to receive so much push back from the Steam community. It's understandable really. Those of us who frequent the platform are more accustomed to the toxic behavior that the community can produce. I'd imagine it was even more shocking considering the opposite happened whenever she premiered her piece at conferences. For the most part, people seemed genuinely moved by what they saw. Some were brought to tears. Others came away asking how they could help in the relief efforts. I asked her if the experience was somehow more believable at these conferences. That maybe the gamers on Steam felt that she was exaggerating to push some sort of agenda.

"I could show you some of the art bible [from Project Syria] which shows how very carefully our reconstructed streets were created," De la Peña explained. "And I had people from Aleppo come in through it and just bawl because we even got the color of the bricks right – because they were a very special color from Aleppo. So it's always easy for me to counter those accusations with the unbelievable in-depth documentation that we do. And that documentation and the ability to find those pieces and what's true is based on my many years of being a careful, considerate, thoughtful, and to the best of my ability, a truthful journalist."

The truth of the matter is, there are always going to be those persons who are reluctant to hearing or even acknowledging a person of color's side of the story. Some people don't want to deal with anything that might disrupt their view on what life is like here in the United States (or any part of the world really).

"I think timing is always a factor," Dr. Cogburn remarks. "Socially and culturally we're in this unique space where a lot of people are alert and awake and paying attention, especially white people who might not have otherwise been paying attention…there are people who believe in the idea of justice and fairness but they don't understand racism. We're in a window of time where there's a lot of people [who] are very aware that

something is going on with race, even if they don't fully understand what it is and what is happening, specifically linked to police violence because of all of the media coverage that have people more aware and willing to talk and listen."

It's important to be able to share in the troubles of our fellow man. Shining a virtual light on certain issues can both educate people, potentially bringing about change, as well as more immediately provide vocal aid to those of us who may be suffering. During my interview with De la Peña, she informed me of a story she did about a Muslim family that was unfairly targeted by the FBI after 9-11. Because she needed a shot of their Mosque to go along with the piece, she contacted a representative to come and view the rough cut of the video.

“He came in, and he watched the piece and he started crying,” De la Peña said. “He kept saying, ‘Why would you do this? Why would [someone like me] be willing to help?’ I didn’t know how to answer that question.”

He wasn’t sad, but was instead overcome with emotion. Someone who didn’t look like him, who wasn’t Muslim, and who hadn’t been negatively affected by the treatment of his people, was willing to help tell this story. It was a powerful moment for both of them, but even more so for the representative of the Mosque. The world is changing. Despite the current political climate, things are slowly progressing. People are starting to listen.

I personally hope, for my children’s sake, that eventually they’ll do more than just listen.

Article can be found here: https://uploadvr.com/vr-fight-racism-prejudice/

Media Psychology, 16:7–38, 2013

ISSN: 1521-3269 print/1532-785X online
DOI: 10.1080/15213269.2012.755877

The Effect of Embodied Experiences on Self-Other Merging, Attitude, and Helping Behavior

SUN JOO (GRACE) AHN
Grady College of Journalism and Mass Communication, University of Georgia, Athens, Georgia, USA

AMANDA MINH TRAN LE
Metaio, Inc., San Francisco, California, USA

JEREMY BAILENSON
Department of Communication, Stanford University, Stanford, California, USA

Immersive virtual environment technology (IVET) provides users with vivid sensory information that allow them to embody another person's perceptual experiences. Three experiments explored whether embodied experiences via IVET would elicit greater self-other merging, favorable attitudes, and helping toward persons with disabilities compared to traditional perspective taking, which relies on imagination to put the self in another person's shoes. Trait dispositions to feel concern for others was tested as a moderating variable. Participants in the embodied experiences (EE) condition were exposed to a red-green colorblind simulation using IVET while participants in the perspective taking (PT) condition were exposed to a normal colored IVET world and instructed to imagine being colorblind. Experiment 1 compared EE against PT and found that EE was effective for participants with lower tendencies to feel concern for others 24 hours after treatment. Experiment 2 delved further into the underlying process of EE and confirmed that a heightened sense of realism during the EE led to greater self-other merging compared to PT. Finally, Experiment 3 demonstrated that the effect of EE transferred into the physical world, leading participants to voluntarily spend twice as much effort to help persons with colorblindness compared to participants who had only imagined being colorblind.

Address correspondence to Sun Joo (Grace) Ahn, University of Georgia, Grady College of Journalism & Mass Communication, 120 Hooper Street, Athens, GA 30602-3018. E-mail: sjahn@uga.edu

> "Tell me Craig, why do you like puppeteering?"
>
> "Well Maxine, I'm not sure exactly. Perhaps the idea of becoming someone else for a little while. Being inside another skin—thinking differently, moving differently, feeling differently." –Craig Schwartz, *Being John Malkovich*

This quote from a 1999 fantasy film, *Being John Malkovich*, refers to a mental process called perspective taking. In this film, people enter a portal that leads into the mind of John Malkovich and see, hear, and feel through his body, using Malkovich's body as a live puppet. The portal in the film is a fantastical depiction of the cognitive process of perspective taking, or imagining oneself in the shoes of another. Psychologists have discovered that perspective taking encourages a host of favorable outcomes, such as stereotype reduction (Batson et al., 1997) learning (Siegler, 1995), and improved interpersonal communication (Fussell & Krauss, 1989).

Despite the benefits of perspective taking, mentally putting oneself in the shoes of another requires extensive cognitive effort (Hoffman, 1982) and individuals may differ in their ability and motivation to engage in this activity (Davis and Kraus, 1997; Gehlbach, 2004). This article proposes using immersive virtual environment technology (IVET) to enable individuals to easily and effectively experience the world from another person's point of view. With novel affordances such as multisensory inputs and naturalistic control of point of view, IVET allows for a literal demonstration of climbing into another person's skin to embody his or her experiences first hand. Vivid, multilayer perceptual information simulated by digital devices enable individuals to see, hear, and feel as if they were undergoing the sensory experiences in the physical world—what we call "embodied experiences." Using IVET, embodied experiences allow the user to experience the closest realization of the portal to enter another person's mind and body.

Our main focus of investigation was on the potential of embodied experiences through IVET to foster greater self-other merging with persons with disabilities; increase favorable attitude toward them; and assess whether the influence of these experiences could transfer to the physical world, leading to actual helping behavior. These effects were compared against traditional perspective taking methods that rely on imagination to assess the strengths and weaknesses of embodied experiences through IVET.

PERSPECTIVE TAKING, A CATALYST TO HELPING BEHAVIOR

Humans are social animals who spend much of their lives interacting with others. Perspective taking facilitates social interaction by helping people establish common grounds and infer shared knowledge and beliefs between

interactants (Krauss & Fussell, 1991). Scholars argue that humans are hard-wired to help others in need, as are few other large-brained species such as orangutans or dolphins (de Waal, 2008). It has been demonstrated that sharing the same basis of feelings and thoughts of another person through perspective-taking can even lead to costly self-sacrifice during helping (Batson, 1991; de Waal, 2008).

People often engage in perspective taking through mentally merging the self with the other, a process in which the other is considered to be more self-like. For instance, upon being instructed to take the perspective of a stranger, individuals demonstrated a tendency to project positive self-relevant traits to the stranger and used the positive self-traits to describe him or her (Davis, Conklin, Smith, & Luce, 1996). Much in the same way, instructing individuals to take the perspective of a member of a certain social category (e.g., elderly) led them to describe that member with trait words that they used to describe themselves (Galinsky & Moskowitz, 2000). Cialdini, Brown, Lewis, Luce, and Neuberg (1997) suggested this merging is enhanced when individuals recognize cues related to genetic relatedness or close attachment, and Aron, Aron, and Smollan (1992) proposed that people in close relationships such as spouse or family have highly overlapping mental representations of one another.

This phenomenon can be mimicked even with unfamiliar persons by establishing close attachment cues via perspective taking. Cialdini and colleagues (1997) termed this sense of shared identity through self-other merging as *oneness*, and demonstrated that when individuals are led to feel oneness with another person via perspective taking, greater intentions for helping emerge. Goldstein and Cialdini (2007) also went on to show that when oneness is heightened through perspective taking, individuals begin to behave as the other person would, and that this change in behavior is fully mediated by self-other merging.

However, perspective taking is a controlled, effortful process that requires substantial cognitive resources (Davis et al., 1996) and can be challenging to achieve. When individuals fail to understand and share the other person's thoughts and feelings through perspective taking, they will not engage in helping behavior unless there are clear benefits from helping (Batson, 1991). The cognitive challenge of perspective taking may be even less appealing when individuals have low motivation to engage in an effortful mental task (Gehlbach, Brinkworth, & Wang, 2012; Hodges & Klein, 2001; Webster, Richter, & Kruglanski, 1996).

Even if a person attempts to take on the cognitive challenge of perspective taking to mentally put themselves in another person's shoes, he or she may fail to fully grasp the urgency or the reality of the other person's situation. For instance, it has been demonstrated that people underestimated how thirsty they would be after exercising when they were asked to merely use their imagination. But when they were asked to make the same estima-

tion after an actual exercise session, they expressed a significant increase in their estimation of how much they would crave water in similar future situations (Van Boven & Loewenstein, 2003). Regardless of best efforts to put themselves in the situation, people generally have a difficult time fully appreciating the true nuances of the situation unless they are living the situation in that moment. This article explores how IVET may aid people to overcome such shortcomings of perspective taking by allowing them to experience the vivid nuances of another person's situation, and how that may lead to more favorable outcomes than traditional perspective taking.

EMBODIED EXPERIENCES THROUGH IMMERSIVE VIRTUAL ENVIRONMENT TECHNOLOGY

IVET is a mediated environment simulated by digital computer technologies that blur the distinction between reality and its virtual representations (Blascovich et al., 2002). The system used for the current experiments was comprised of a head-mounted display (HMD), a headpiece with screens that provide a stereoscopic view of the computer-generated world, and devices that track physical movements of the head in three-dimensional space. Stereoscopic view enables users to see depth. Users also had head-controlled point of view, meaning that they were able to look around naturally in the virtual world as they would in the physical world.

These novel affordances offer an innovative way to facilitate perspective taking by allowing the user to experience vivid sensory information firsthand. With rich layers of sensory information, users feel a high sense of presence, or the perception that the mediated virtual environment is real (Loomis, 1992; Slater & Wilbur, 1997). Biocca (1997) notes that, by digitally recreating and extending the human sensory capabilities, virtual stimuli lead the mind to temporarily accept the illusion of sufficiently realistic experiences. Real experiences in the physical world become associated and stored with existing memories and these memories are later activated and recalled when the individual encounters or thinks about similar stimuli (Barsalou, 2009). In much the same way, the realism of virtual experiences is likely to produce mental schema about the simulated event as if he or she had firsthand experience of it, to be recalled later when necessary.

The concept of putting oneself in another's shoes to vicariously share experiences using media is not new. Even a print medium that presents no simulated sensory information can feel relatively realistic when an individual becomes deeply engaged (Green, 2004; Green, Brock, & Kaufman, 2004). However, no other medium to date has been able to replicate the degree of realism that IVETs offer. Furthermore, research exploring the potential of applying IVETs to perspective taking has been sparse and is even more so for motivating helping behavior.

Earlier IVET studies have investigated the effect of embodying only the physical traits of another person and found that it modifies behavior (Yee & Bailenson, 2007). Termed the "Proteus Effect," results demonstrated that spending several minutes in a virtual world embodying a tall virtual self-representation (i.e., an avatar) led participants to choose more aggressive strategies in a negotiation task compared to participants who were given short avatars. Similarly, participants given attractive avatars were more confident in interacting with a stranger compared to those given unattractive avatars (Yee & Bailenson, 2009; Yee, Bailenson, & Ducheneaut, 2009).

Kalyanaraman, Penn, Ivory, and Judge (2010) conducted a relevant study that motivated the current experiments, comparing the embodiment of schizophrenic experiences in a virtual schizophrenia simulator that simulated visual and auditory hallucinations against traditional manipulations of perspective taking. They compared four different conditions—perspective taking instructions only; virtual simulation only; both perspective taking instructions and virtual simulation; and a control condition that did not receive either treatment. Results indicated that having participants think about schizophrenia through perspective taking as well as having them experience the simulation was more effective in terms of constructing more favorable attitudes toward persons with schizophrenia compared to participants who engaged in just the virtual simulation or just perspective taking. This seems to be a different process of self-other merging than the process of self-perception as suggested by the Proteus Effect (Yee & Bailenson, 2009), which would anticipate that participants would act in a manner similar to a schizophrenic person after the virtual simulation. The current experiments aim to advance these findings by exploring the process of self-other merging elicited by embodied experiences through IVET in the context of disabilities, more specifically red-green colorblindness.

Moreover, self-report measures might not be sufficient evidence to conclude that helping behaviors will be exhibited as a result of perspective taking. We attempted to go beyond assessing the effect of embodied experiences based solely on surveys and gauged actual helping behavior. Earlier work exploring the power of embodied experiences in IVET demonstrated behavioral changes in terms of paper conservation (Ahn, 2011). However, helping a person is contextually different from helping the environment and may lead to unforeseen differences.

Finally, perspective taking is a cognitive activity, subject to individual differences in capacity and ability. Some people are inherently more likely to feel concern for another person in need (Matthews, Batson, Horn, & Rosenman, 1981; Rushton, Fulkner, Neale, Nias, & Eysenck, 1986). Not only does this individual difference in predisposition to care for others influence their tendency to engage in perspective taking (Davis, Luce, & Kraus, 1994), but it also impacts how individuals respond to IVET simulations; that is, individuals predisposed to feel greater concern for others tend to experience

higher levels of presence and feel that the virtual world is real (Sas, 2004; Sas & O'Hare, 2003; Wallach, Safir, & Samana, 2010). Thus, the current experiments will also take trait differences in feeling concern for others into account as a potential moderator.

OVERVIEW OF EXPERIMENTS

Three experiments compared traditional perspective taking against embodied experiences using IVET. Red-green colorblindness, a visual disability that causes the inability to perceive differences between red and green, was chosen as the context of perspective taking in all three experiments to simulate a novel experience to control for the confounding effects of prior experience. Thus, the question of interest was whether embodied experiences would better promote greater self-other merging compared to traditional perspective taking in an unfamiliar situation with an unfamiliar person. In addition, the experiments sought to measure attitude change in terms of reduced prejudice against persons with red-green colorblindness. Most importantly, we were ultimately interested in the possibility of encouraging actual helping behavior; that is, the current experiments aimed to investigate whether embodied experiences within the virtual environment would be powerful enough to transfer the effects into the physical world in the form of providing actual help toward people with red-green colorblindness. Individual predisposition to feel concern for others was included in all three experiments to study how it moderates the effects of embodied experiences or perspective taking.

Experiment 1 was an exploratory investigation that pitted embodied experiences against traditional perspective taking. The main aim of the first experiment was to test the effectiveness of embodied experiences using IVET and whether it would outperform traditional perspective taking in terms of self-other merging (i.e., oneness), attitude toward persons with colorblindness, and actual helping behavior. Building on these results, Experiment 2 delved into underlying mechanisms to study the process of self-other merging in greater detail. The main aim of the second experiment was to discover the element of embodied experiences that set it apart from traditional perspective taking. Finally, Experiment 3 leveraged some insights from the first two studies and improved the measure of helping people with red-green colorblindness.

In all three experiments, all participants first received basic information about red-green colorblindness. Participants assigned to the embodied experience condition wore a HMD with a colorblind filter applied over the objects on the screen, which allowed them to accurately experience being red-green colorblind. Participants assigned to the perspective tak-

ing condition also wore the HMD but viewed the screen in normal colors while being told to mentally put themselves in the shoes of the colorblind individual.

EXPERIMENT 1

Experiment 1 explored the effectiveness of embodied experiences by comparing traditional perspective taking (PT) against embodied experiences (EE). Based on the discussion above regarding cognitive challenges of traditional PT and the ability of IVET to deliver a vivid sensory-rich experience that mimics climbing into someone else's skin, we anticipated:

H1: EE will elicit greater perceived oneness with the target compared to traditional PT.

The result of better self-other merging, as demonstrated through greater oneness, was expected to be manifested in actual helping:

H2: EE will elicit greater helping behavior toward the target compared to traditional PT.

Furthermore, the vivid experience of the disability in IVET was anticipated to reduce prejudice toward colorblind individuals even after some time had passed:

H3: EE will elicit more favorable attitudes toward the target compared to traditional PT that last up to 24 hours following the experimental treatment.

Finally, individual differences in predispositions to care for others were expected to moderate the effect of EE and PT to some degree:

RQ1: How will individual predispositions to care for others impact the effect of EE and PT on oneness, attitude, and helping behavior?

METHOD

Participants

A sample was recruited from the student population of a medium-sized university. The sample ($N = 44$)[1] consisted of 20 males and 24 females aged 18 to 69 ($M = 22.57$, $SD = 8.75$).[2]

Apparatus

Three computer stations were involved in the experimental setup: the participant's computer, where the participant was exposed to the EE or PT treatments; the confederate's computer, where a research assistant was posing as a colorblind person; and the survey computer, where the participant filled out questionnaires.

At the participant's computer, participants donned a HMD through which they were able to view the stimulus. The HMD presented the virtual environment with 640 horizontal by 480 vertical pixel resolution panels for each eye. Participants' head movements were tracked by a three-axis orientation sensing system (Intersense IS250 with an update rate of 150 Hz) and used to continuously update the simulated viewpoint. The system latency, or delay between the participant's movement and the resulting update in the HMD, was no greater than 80 ms. Vizard 3.0 software was used to assimilate tracking and rendering.

The confederate's computer was equipped with the Sensable Phantom Omni haptic device with six degrees of freedom (x, y, z, pitch, yaw, and roll). The haptic device allowed confederates to 'touch' objects in the virtual world by providing mechanical resistance based on the position of the hand as it interacted with the virtual environment. The confederate and participant's computers were networked so that the participant at the computer could see any movement on the confederate's screen in real-time. Figure 1 depicts the experimental setup.

Design and procedure

Pretest. Approximately 48 hours prior to coming into the lab, all participants filled out a pretest (Interpersonal Reactivity Index [IRI]; Davis, 1980). This 28-item questionnaire is comprised of four 7-item subscales: perspective taking, fantasy, empathic concern, and personal distress. Combined, the four subscales yield a comprehensive picture of individual predispositions to put oneself in another's person's situation and react to that person's needs. Each item was measured by a fully labeled 5-point scale (1 = *does not describe me well*; 5 = *describes me very well*). The 28 items had a Cronbach's alpha level of .78 and were averaged to create a composite IRI score. The IRI scores ranged from 2.32 to 4.07 ($M = 3.16$, $SD = .35$).

Experiment. During the actual experiment, participants were randomly assigned to either the EE or PT condition. Participants in both conditions received identical explanations about colorblindness, such as basic statistics, a short description of how colorblind people are unable to differentiate between red and green, and how this could affect their lives.

The main experimental task was a color matching exercise in which participants matched red or green colored screws with red or green colored holes on a board. To clearly present the target of perspective taking, a

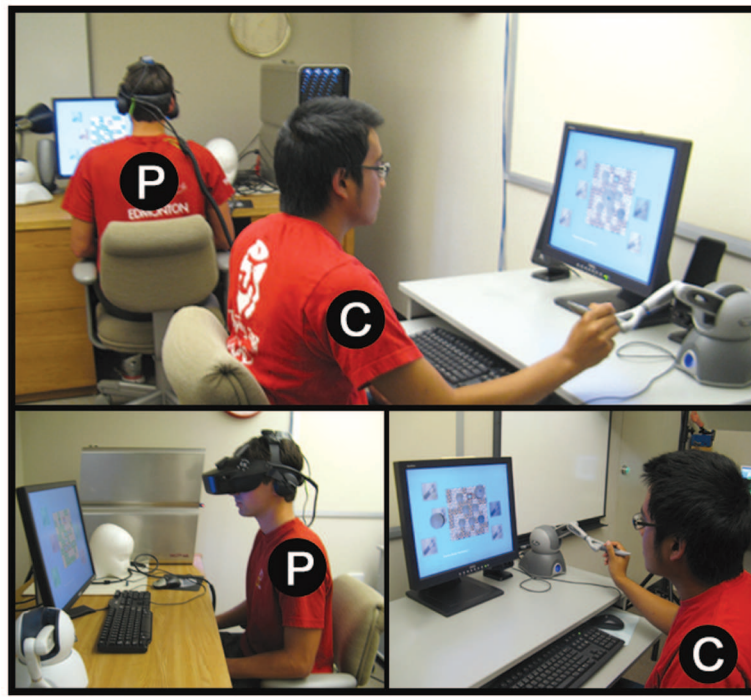


FIGURE 1 Top row: the experimental setup with participant (denoted P) and confederate (denoted C) computer stations. Bottom row: close-up of participant wearing the HMD (left), close-up of confederate using the haptic device. (Figure available in color online.)

same-sex confederate posed as a colorblind student. The cover story informed participants that the colorblind confederate was "training to differentiate red and green." The participants were asked to help this confederate in training since the colorblind student would be unable to differentiate the colors of the screws and the holes at first. Participants were instructed to verbally guide the confederate through the task (e.g., "move the screw up") while the confederate moved the screws with the haptic device.

Participants gave verbal instructions to the colorblind confederate for two sets of five boards, each board taking about one minute to complete. The first set, presented as the practice set, was the treatment. Wearing the HMD, participants in the EE condition completed the boards in the colorblind perspective, which displayed the objects treated with a colorblindness filter, rendering them as seen by a colorblind individual. In other words, in the EE condition, participants embodied the vivid sensory experience of being colorblind and were unable to discriminate red objects from green objects. Participants in the PT condition also wore the HMD and completed the same boards in the normal color perspective while imagining being colorblind.

Then, all participants gave verbal guidance to the confederate to complete the second set of boards which were presented to them as the actual training set. This second set was a filler task, administered to bolster the cover story of training the colorblind student after practice on the first set. Once the second set of boards was completed, participants were taken out of the HMD and led to the survey computer station where they filled out surveys. As thinking about the disability seemed to be an integral part of the manipulation that made the virtual schizophrenia simulation effective in the Kalyanaraman et al. (2010) study, all participants were asked to write down their thoughts on either embodying the experience of being colorblind or taking the mental perspective of being colorblind before moving on to answer survey questionnaires.

Upon completion of survey questionnaires, the researcher indicated that the experiment was over and that the participant was free to leave at any time. To confirm the termination of the experiment, all participants were paid at this point. The researcher then thanked the confederate for participating in the training, explicitly stating that they were welcome to stay as long as necessary to continue practicing with the color-matching program. Then, the researcher left while the participant was still in the room.

The confederates were instructed to wait until the participant was ready to leave, and then ask the participant for help: "I think I still need some practice on trying to match the colors. I know the experiment is finished but can you stay and help me some more?" It was evident that the experiment was over and the researcher was no longer overseeing the experiment. Therefore, the extra number of boards completed by the participant after this point is a measure of voluntary helping behavior. The confederates were blind to the experimental conditions and carefully followed a preconstructed script that was identical for all conditions. Further, confederates were instructed to allow the participant to leave whenever they desired.

Posttest. An online follow-up survey assessing the participant's prejudice toward colorblind people was administered 24 hours after termination of all experimental procedures.

Dependent measures

Manipulation check. Two items measured the success of the participant taking the perspective of the colorblind target through either EE or PT: "How colorblind did you feel during the practice session?" and "How difficult was it to complete the color matching task during the practice session?" Responses were measured with fully labeled 5-point Likert scales. The two items were highly correlated (Pearson's r and p-value = .88) and were averaged.

Oneness. The Inclusion of Other in the Self Scale measured how close the participants felt to the confederate. Developed by Aron and colleagues (1992), this scale depicts seven drawings of increasingly overlapping circles,

anchored by the first picture of two non-overlapping circles and the seventh picture of two almost completely overlapping circles. The participant was instructed to choose the picture that best depicted the extent to which he/she felt connected to and 'at one with' the confederate. This is a measurement of the participants' perceived similarity with the target (Maner et al., 2002).

Helping. This is the number of boards completed upon the confederate's request for help after the researcher had left. Unbeknownst to the participant, the program was still running and counting the number of completed boards. As the participant was well aware that this would be uncompensated work, the number of extra boards may be seen as voluntary helping behavior.

Attitude. This was a 10-item survey adapted from the Attitude Toward Disabled Persons Survey (Yuker, Block, & Younng, 1966), a widely used scale to determine the extent to which people perceive disabled people as inferior to people without disabilities. Ten items were taken from the original scale and the word "disabled" was replaced with "colorblind." For instance, if the original item was, "Disabled people are just as self-confident as other people," we adapted this to "Colorblind people are just as self-confident as other people." A fully labeled 6-point Likert-scale presented these statements manifesting either negative or positive attitudes toward colorblind people (−3 = *I disagree completely*; 3 = *I agree completely*). Thus, a higher score on this attitude scale implies greater reduction of prejudice, while a lower score on the scale implies lesser reduction of prejudice. The reliability of these 10 items had a Cronbach's alpha value of .77 and were averaged to create a comprehensive measure of general attitude toward colorblind people.

Results

Manipulation check

An independent samples t test compared means of the manipulation check scores between the EE and PT conditions. The Levene's test was significant and results were interpreted without assuming equal variances. Results of the t test was significant, $t(23.76) = 7.03$, $p < .01$, $d = 2.22$, confirming the success of the manipulation of EE in delivering a vivid sensory experience of being colorblind. Participants who embodied the experience of being colorblind felt significantly more colorblind ($M = 3.50$, $SD = 1.32$) and thus the color matching task became significantly more difficult for them than participants who imagined being colorblind ($M = 1.30$, $SD = .47$).

Descriptive statistics of all other dependent measures are given in Table 1, and detailed results of all regression tests to be discussed in the following pages are displayed in Table 2.

TABLE 1 Descriptive Statistics for Dependent Measures in Experiment 1 ($N = 40$)

		M	*SD*	Minimum	Maximum
Oneness	EE	4.35	1.79	1.00	7.00
	PT	4.00	1.81	1.00	7.00
	Total	4.18	1.78	1.00	7.00
Helping	EE	19.50	16.80	2.00	55.00
	PT	23.80	17.34	2.00	55.00
	Total	21.65	16.99	2.00	55.00
Attitude	EE	1.81	.51	1.00	3.00
	PT	1.51	.81	.10	2.60
	Total	1.66	.69	.10	3.00

Oneness

Next, a linear regression was conducted with oneness as the dependent variable and the experimental conditions (EE vs. PT), IRI, and their interaction term as the predictors. The experimental condition variable was dummy coded with 0 (PT) and 1 (EE), and IRI scores were centered. As shown in Table 2, both the main effects of experimental condition and IRI had positive but nonsignificant coefficients. The interaction between experimental condition and IRI was significant. The change in incremental R^2 value after adding the interaction term was significant with the interaction term adding 12% of R^2 change. Following the guidelines set forth by Aiken and West (1991), the effect of IRI's moderation on oneness is depicted in Figure 2. Figure 2 and the negative coefficient value of the interaction term imply that the difference in oneness from EE and PT was larger for participants with lower IRI. For these participants, EE was more effective than PT in eliciting higher oneness. The difference in oneness fostered by EE or PT was smaller for participants with higher IRIs and PT was more effective than EE for them. Therefore, Hypothesis 1 was only partially supported.

Helping

A regression was conducted with the number of extra boards completed as the dependent variable and the same predictors as above. As shown

TABLE 2 Regression Analyses for Dependent Measures in Experiment 1 ($N = 40$)[a]

	Experimental condition	IRI	Condition × IRI
Oneness	.11	.30	−.56*
Helping	−.12	−.64*	.69**
Attitude	.21	.60*	−.52*

[a]Cell entries refer to the standardized regression coefficient.
*$p < .05$, **$p < .01$.

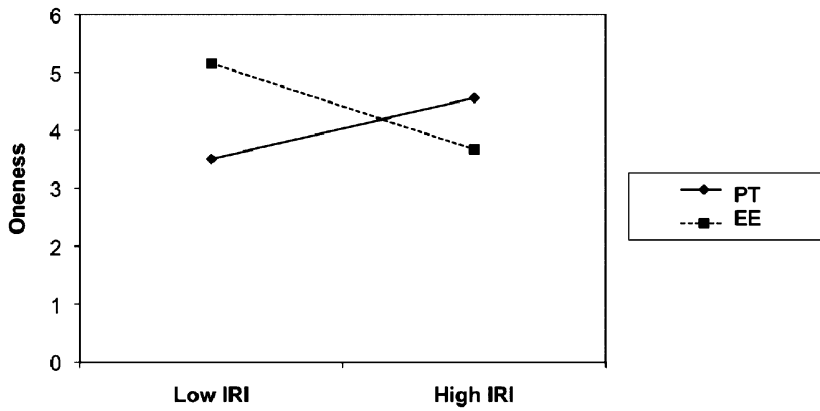


FIGURE 2 Interaction between experimental condition and IRI in predicting oneness.

in Table 2, the main effect of experimental condition had a negative but nonsignificant coefficient. The main effect of IRI was significant—participants with lower IRI scores offered more help to the colorblind confederate. The interaction between experimental conditions and IRI was also significant, as displayed in Figure 3. Adding the interaction term to the model significantly increased the R^2 value by 18%. The difference in helping behavior elicited by EE and PT was greater for participants with lower than with higher IRI. For participants with lower IRI, traditional PT elicited greater helping than EE. The difference between experimental conditions was smaller for participants with higher IRI, for whom PT seemed more effective than EE. Hypothesis 2 was not supported.

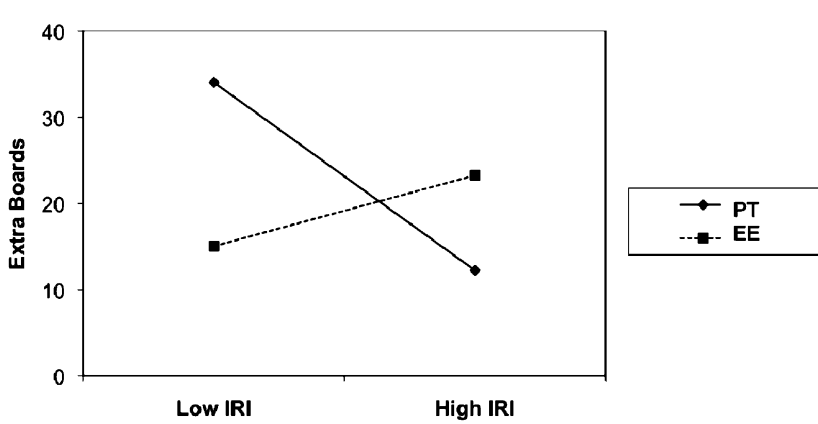


FIGURE 3 Interaction between experimental condition and IRI in predicting helping behavior.

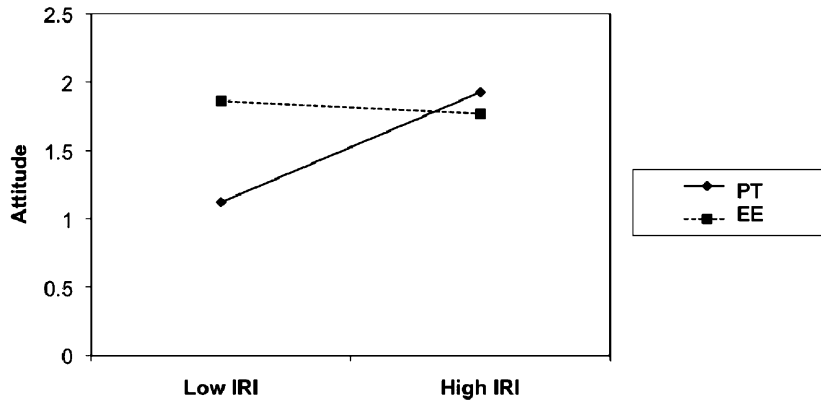


FIGURE 4 Interaction between experimental condition and IRI in predicting attitude toward people with colorblindness 24 hours after the experimental treatments.

Attitude

Finally, a regression was conducted with attitude as the dependent variable and the same predictors as above. The main effect of experimental condition had a positive but nonsignificant coefficient. On the other hand, the main effect of IRI was a significant predictor of attitude, implying that participants with higher IRI scores had more favorable attitude toward people with colorblindness (i.e., greater reduction of prejudice) 24 hours after the experimental treatments. The interaction was also significant, as shown in Figure 4. Adding the interaction term to the model significantly increased the R^2 value by 10%. The effect was largely driven by participants with lower IRI in the PT condition, who displayed the least favorable attitude toward people with colorblindness 24 hours after the treatment. This also implies that EE fostered relatively high levels of attitude regardless of individual differences in IRI. On the other hand, participants exposed to the PT condition had to have higher IRI to demonstrate an equally high level of attitude. Hypothesis 3 was partially supported.

Discussion

These results imply that embodying the experience of being colorblind has notable effects, particularly on participants with lower IRI. For those participants, EE enhanced the feeling of merging with the colorblind confederate compared to traditional PT. Furthermore, participants with lower IRI exposed to EE indicated more favorable attitudes—or reduced prejudice—toward people with colorblindness 24 hours after the treatment compared to participants with higher IRI. It should be noted that this longer-term effect resulted in positive attitudes toward not just the colorblind confederate that the participants interacted with in the laboratory, but toward the colorblind

population in general. Considering that the duration of the EE treatment was approximately five minutes, the fact that the effect transferred outside the laboratory 24 hours after the treatment is meaningful.

Individual differences in IRI also moderated the effect of experimental treatments on helping behavior, but exposing participants with lower IRI to EE actually led to less helping compared to traditional PT. One possible explanation is our failure to construct a believable task; that is, the IVET manipulation may have been so vivid that the helping task seemed implausible particularly after the EE treatment. When participants saw firsthand in the EE condition how completely indistinguishable the colors are to a colorblind person, they may have felt that the confederate's state was far more serious than they had imagined and that extra practice would be futile. Conversely, participants relying on imagination in the PT condition may have not realized the severity of the disability and may have had unrealistic optimism in their ability to help the colorblind student, thus spending more time helping. We revisit this anomaly by redesigning the helping task in the next two studies.

Because IRI's moderation of the helping behavior was at odds with how it affected oneness and attitude, IRI's role as a moderator is inconclusive. Similarly, IRI's main effect is also inconclusive as participants with higher IRI were less likely to help, but displayed more favorable attitude 24 hours after the experimental treatment. Despite inconsistencies, the fact that we were able to obtain strong interaction effects in all three measures with a small sample size is encouraging and merits further investigation.

EXPERIMENT 2

A follow-up experiment was conducted with a larger sample to delve deeper into the underlying mechanism of EE and how its effects on oneness, attitude, and helping may be moderated by IRI. Based on the results of Experiment 1 and the earlier discussion on prior work demonstrating correlations between IRI scores and the perception of presence, we expected that the moderated effect of EE on increased oneness would be related to the level of presence; that is, the vivid sensory information of EE is expected to heighten the sense of realism of the experience. The heightened sense of realism is, in turn, expected to lead participants to feel that they have truly climbed into the skin of, or are at one with, a colorblind person. Thus, Experiment 2 first aims to replicate Hypothesis 1 from Experiment 1, but with IRI moderating the effects as Experiment 1 implied that EE was more effective in eliciting oneness among participants with lower IRI:

H4: EE will elicit greater oneness with the target compared to traditional PT in participants with lower IRI.

The following is also hypothesized regarding the underlying mechanism of EE:

H5: Perceived level of presence will mediate the relationship between experimental conditions and oneness.

Furthermore, we chose an improved assessment of helping with greater feasibility than what was used in Experiment 1—helping to identify and enhance Web sites that would be problematic for colorblind individuals. This task was more believable in comparison to the task in Experiment 1 because it did not involve any claims to treat or improve the disability. The help provided by participants would offer realistic aid to people with colorblindness who use the Internet:

H6: EE will elicit greater helping behavior compared to traditional PT in participants with lower IRI.

Finally, we also expected that EE's effect on attitude toward people with colorblindness 24 hours after the treatment will be replicated and that IRI will moderate the effect, as demonstrated in Experiment 1.

H7: EE will elicit less prejudice toward colorblind people compared to traditional PT in participants with lower IRI.

Method

Participants

A sample was recruited from the student population of a medium-sized university. The sample ($N = 97$) consisted of 44 males and 53 females aged 18 to 36 ($M = 21.59$, $SD = 3.00$).

Apparatus

The same devices from Experiment 1 were used.

Design and procedure

Pretest. The same IRI scale from Experiment 1 was used as a pretest. The 28 items had a Cronbach's alpha value of .79 and were averaged to create a comprehensive IRI score. The IRI scores ranged from 1.89 to 4.25 ($M = 3.31$, $SD = .39$).

Experiment. The same design from Experiment 1 pitting EE against PT was used, but rather than using an in-person confederate, Experiment 2 introduced the same-sex colorblind confederate from Experiment 1 with a color photograph to provide a specific target of perspective taking for the participants. Numerous prior studies have used static forms of target representation such as photographs (Galinsky, Wang, & Ku, 2005). Because they did not have a live confederate in the same room, participants were told to complete the boards themselves for the color-matching task instead of giving verbal guidance to the confederate as in Experiment 1. Thus, all participants wore the HMD and had control of the haptic pen. All other instructions and procedures were identical with Experiment 1 with the exception of the new helping task.

The helping task in this experiment involved participants reading a short text message on the computer from a student group that was supposedly trying to build colorblind friendly Web sites. The task required them to view screenshots of Web sites and to write about why the Web site may be inaccessible to colorblind individuals and how it may be improved. It was made clear to the participants that this activity was not a part of the experiment and that it would be volunteer work. The researcher left the room at this point, instructing the participant to leave whenever they wished to go.

If a participant agreed to participate, they clicked on a next button. The computer walked the participant through instructions and presented a picture of a Web site with a textbox. There was a quit button on the screen that participants could click at any point to leave the experiment. Time stamps were gathered from the point the participant clicked on "next" to begin the activity to the point where the participant clicked on "quit." Also, anything typed into the textbox was recorded for further analyses. The time spent and the number of words written for this uncompensated activity was a measure of helping toward colorblind people in general.

Posttest. The same attitude scale from Experiment 1 was administered 24 hours following the experimental treatments.

Dependent measures

Manipulation check. The same two items from Experiment 1 measured the success of the participant taking the perspective of the colorblind target through either EE or PT. The two items were highly correlated (Pearson's r and p-value = .92) and were averaged.

Oneness. The same Inclusion of Other in the Self Scale from Experiment 1 was used.

Presence. Five items gauged how vivid and realistic the experiment treatment felt by asking participants the extent that they felt the colorblind was happening to them; they were in the colorblind person's body; they were

colorblind; they felt they could reach out and touch the red and green screws; they felt that the screws and the board were real. These items were culled from several sources (Bailenson & Yee, 2007; Nowak & Biocca, 2003; Witmer & Singer, 1998) and have been tested in other comparable experiments (Fox, Bailenson, & Binney, 2009). Participants used fully labeled 5-point Likert scales (1 = *not at all*; 5 = *completely*) to rate their perceptions of realism. The items had high reliability (α = .79) and were averaged to create a single composite measure.

Helping. The total number of seconds that the participants invested in the volunteer activity was calculated to determine the degree of helping behavior. The duration of time invested began the moment a participant finished reading the instructions and started viewing the screenshots of Web sites and terminated the moment the participant hit the 'quit' button. The total number of words written to help identify the problematic areas of the Web sites was also counted as a degree of helping behavior. Measuring the total number of words ensured that we were measuring actual effort to help rather than time spent passively viewing Web sites.

Attitude. The same 10-item attitude scale from Experiment 1 was used. The reliability of these 10 items had a Cronbach's alpha value of .77 and were averaged to create a comprehensive measure of attitude toward colorblind people.

Results

Descriptive statistics of all dependent measures are given in Table 3, and detailed results of all regression tests are displayed in Table 4.

TABLE 3 Descriptive Statistics for Dependent Measures in Experiment 2 (N = 97)

		M	*SD*	Minimum	Maximum
Oneness	EE	3.98	1.84	.00	7.00
	PT	2.39	1.48	1.00	6.00
	Total	3.26	1.86	.00	7.00
Presence	EE	2.72	.86	1.00	4.40
	PT	1.46	.52	1.00	3.40
	Total	2.15	.96	1.00	4.40
Helping (time)	EE	91.34	226.55	.00	1411.78
	PT	79.48	152.65	.00	655.91
	Total	86.04	196.10	.00	1411.78
Helping (words)	EE	151.89	447.42	.00	2775.00
	PT	105.91	217.97	.00	1147.00
	Total	131.29	362.09	.00	2775.00
Attitude	EE	1.62	.78	−0.11	3.00
	PT	1.43	.91	−1.56	3.00
	Total	1.53	.84	−1.56	3.00

TABLE 4 Regression Analyses for Dependent Measures in Experiment 2 ($N = 97$)[a]

	Experimental condition	IRI	Condition × IRI
Oneness	.43**	.07	−.03
Helping (time)	.04	.11	.01
Helping (words)	.07	.04	.09
Attitude	.11	−.19	.31*

[a]Cell entries refer to the standardized regression coefficient.
*$p < .05$, **$p < .01$.

Manipulation check

An independent samples t test was conducted to compare means of the manipulation check scores between the EE and PT conditions. The Levene's test was significant and results were interpreted without assuming equal variances. The t test was significant, $t(80.02) = 19.05$, $p < .01$, $d = 4.01$, confirming the success of the manipulation of EE in delivering a vivid sensory experience of being colorblind. Participants who embodied the experience of being colorblind felt significantly more colorblind ($M = 4.15$, $SD = .91$) than participants who imagined being colorblind ($M = 1.26$, $SD = .46$).

Oneness

Next, a linear regression was conducted with oneness as the dependent variable and experimental conditions, IRI, and their interaction term as the predictors. Again, the experimental condition variable was dummy coded with 0 (PT) and 1 (EE), and IRI scores were centered. As shown in Table 4, results revealed that the main effect of experimental condition is significant, indicating that participants in the EE condition felt more at one with the colorblind person compared to participants in the PT condition. R^2 values revealed that experimental condition accounted for 18% of the model. IRI and the interaction term were not significant variables in the model. Therefore, Hypothesis 4 was not supported but the main effect of experimental condition is notable.

Presence–mediation analysis

Following the guidelines set out by Baron and Kenny (1986), a linear regression was conducted with experimental condition as the independent variable and oneness as the dependent variable, $\beta = .43$, $p < .01$. Next, a regression was conducted with experimental condition as the independent variable and presence as the dependent variable, $\beta = .66$, $p < .01$. Another regression was run with presence as the independent variable and oneness as the dependent variable, $\beta = .55$, $p < .01$. Finally, a linear regression with experimental condition as the independent variable and oneness as

the dependent variable, controlling for presence was run. Results indicated that presence ($\beta = .48$, $p < .01$) completely mediates the relationship between experimental condition and oneness ($\beta = .11$, $p > .05$). Therefore, Hypothesis 5 was supported.

Helping

A linear regression was conducted with helping time as the dependent variable and experimental conditions, IRI, and their interaction term as the predictors. None of the variables in the model were significant. Another linear regression was conducted with helping words as the dependent variable. Again, none of the variables in the model were significant. Therefore, Hypothesis 6 was not supported.

Attitude

Finally, another regression was conducted with attitude as the dependent variable and the same predictors as above. As shown in Table 4, the main effect of experimental condition had a positive but nonsignificant coefficient. The main effect of IRI had a negative but nonsignificant coefficient. The interaction term was significant, with the effect largely driven by participants with higher IRI, as shown in Figure 5. Adding the interaction term to the model significantly increased the R^2 value by 4%. Among participants with higher IRI, EE had lasting effects 24 hours following the experimental treatment leading to more favorable attitudes toward people with colorblindness than those of participants with lower IRI. Participants with higher IRI in the traditional PT condition demonstrated attitudes more negative than participants with lower IRI in the PT condition. Therefore, Hypothesis 7 was not supported.

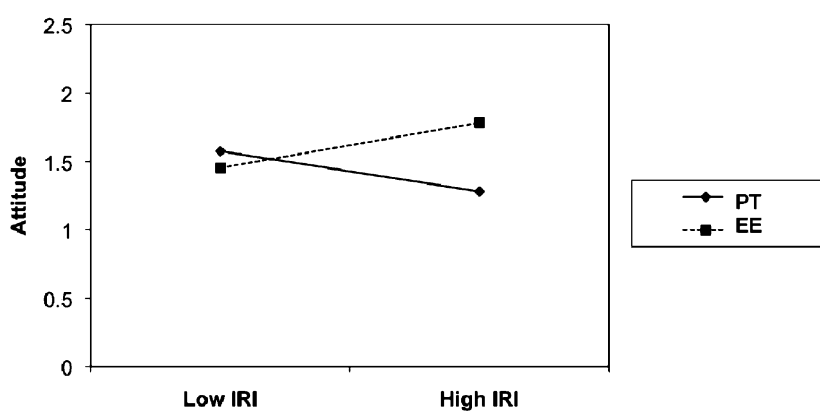


FIGURE 5 Interaction between experimental condition and IRI in predicting attitude toward people with colorblindness 24 hours after the experimental treatments.

Discussion

Experiment 2 built upon the results of the first experiment by shedding light on the mechanism of EE. First, it partially replicated the finding from the previous study that EE resulted in more oneness than PT. Moreover, the results demonstrated that presence mediated the relationship between experimental condition and oneness. In other words, the findings imply that participants in the EE condition perceived the experience of being colorblind to be more realistic than participants in the PT condition. As a result, these participants felt more merged with the colorblind person.

However, the moderating effect of IRI remains inconclusive in that everyone, regardless of their individual differences in IRI, felt higher self-other merging in the EE condition than in the PT condition. Also, the interaction between experimental condition and IRI was manifested in the direction opposite to our hypothesis in attitudes toward people with colorblindness—participants with higher, rather than lower, IRI drove the effect. Thus, although combined results of Experiments 1 and 2 imply that IRI moderates the effect of EE, the directions of the interaction are inconsistent, making it difficult to reach a conclusion.

The newly implemented helping measure failed to yield any significant results. However, when examining post experiment interviews, we were able to gain anecdotal evidence during the course of the experiment that the many experimental sessions were conducted later in the evening and at night when students often had prior engagements. Many participants expressed their regret in their inability to cancel these prior engagements despite their desire to help.

Despite some discrepancies, Experiment 2 fulfilled what we had set out to do by yielding insight into the underlying process of EE and finding out that EE leads to greater self-other merging than traditional PT through offering a more vivid and realistic experience of another person. However, because our ultimate interest in EE was whether it would encourage more actual helping behavior in the physical world than traditional PT, and because the attitudinal changes in the two earlier experiments were encouraging, a follow-up experiment was conducted specifically to test helping in the physical world.[3]

EXPERIMENT 3

To investigate whether EE could promote actual helping behavior outside of the virtual environment, Experiment 3 presented the improved helping task from Experiment 2 but also scheduled experimental time slots during the day when participants were less likely to have scheduling conflicts and would be able to volunteer their time. Situational variables such as scheduling

conflict are crucial barriers to manifested helping behavior as demonstrated in the seminal Good Samaritan study by Darley and Batson (1973), in which seminary students, who were ironically on their way to give a religious talk, failed to help an injured confederate. Therefore, the main purpose of Experiment 3 was to repeat the test of Hypothesis 6 from Experiment 2.

H6: EE will elicit greater helping behavior compared to traditional PT in participants with lower IRI.

As the moderating role of IRI was inconclusive from the earlier two experiments, Experiment 3 explored IRI as a moderator once again.

Method

Participants

A sample was recruited from two medium-sized schools. The sample (N = 57) consisted of 22 males and 35 females aged 18 to 56 (M = 22.54, SD = 7.29).

Apparatus

The same devices from Experiment 1 and 2 were used.

Design and procedure

Pretest. The IRI scale from Experiment 1 and 2 was used as a pretest. The 28 items had a Cronbach's alpha value of .74 and were averaged as a single IRI score. The IRI score ranged from 2.07 to 3.96 (M = 3.03, SD = .40).

Experiment. The same design from Experiment 1 and 2 pitting EE against PT was used, with the same-sex colorblind confederate photo used in Experiment 2.

The only procedural difference from Experiment 2 was that instead of completing the color-matching task on their own, they observed a video of the color-matching task being completed by someone else; that is, participants were told that in order to observe how a colorblind person would complete the task, the researchers had recorded a video of the colorblind person filling out a series of boards. The video stimulus was used in place of having participants move the haptic pen (as in Experiment 2) to control for the possible influence of interactivity on experimental treatments, as having participants interact with the virtual environment may lead to possible confounds. The participants were told that the video was a movie that captured the movements on the screen that the confederate made while filling out boards with the haptic device. The video was created so that it

would contain obvious mismatches, ostensibly committed by the colorblind confederate that they had been introduced to from a photograph. Participants donned a HMD and watched this video in stereoscopic view in either the colorblind perspective (EE) or in normal colors while imagining the confederate's perspective (PT).

Dependent measures

Manipulation check. The same two items (Pearson's r and p-value = .77) from Experiment 1 and 2 assessed how much the EE or PT treatment presented realistic experiences of being colorblind.

Helping. The same measure of helping time and words from Experiment 2 were used.

Results

Manipulation check

An independent samples t test was conducted to compare means of the manipulation check scores between the EE and PT conditions. The Levene's test was significant and results were interpreted without assuming equal variances. Results of the t test was significant, $t(38.61) = 6.54$, $p < .01$, $d = 1.71$, confirming the success of the manipulation of EE in delivering a vivid sensory experience of being colorblind. Participants who embodied the experience of being colorblind felt significantly more colorblind ($M = 3.25$, $SD = 1.34$) than participants who imagined being colorblind ($M = 1.52$, $SD = .53$).

Helping

Next, a linear regression was run with helping time as the dependent variable and experimental conditions, IRI, and their interaction term as the predictors. The same dummy coding (0 = PT; 1 = EE) was used and IRI scores were centered. Results revealed that the main effect of experimental condition is significant, $\beta = .27$, $p < .05$, indicating that participants in the EE condition ($M = 412.23$, $SD = 384.59$) spent significantly more time helping the colorblind person compared to participants in the PT condition ($M = 216.81$, $SD = 160.29$). R^2 values revealed that experimental condition accounted for 10% of the model. IRI and the interaction term were not significant variables in the model.

Another regression was run with helping words as the dependent variable. Again, results revealed that the main effect of experimental condition is significant, $\beta = .28$, $p < .05$, indicating that participants in the EE condition ($M = 149.13$, $SD = 141.25$) wrote significantly more words to help the colorblind person compared to participants in the PT condition ($M = 79.15$,

$SD = 62.33$). R^2 values revealed that experimental condition accounted for 9% of the model. IRI and the interaction term were not significant. Therefore, Hypothesis 6 was not supported. However, the main effects observed from both regressions are notable in that exposure to EE led participants to invest significant efforts in helping individuals inflicted with red-green colorblindness.

Discussion

Building upon the shortcomings of the first two experiments, Experiment 3 provided evidence of EE's ability to promote helping behavior in the physical world. Results indicated that participants in the EE condition invested approximately twice as much time and words to help colorblind people compared to the participants who relied solely on imagination. It is noteworthy that only a few minutes of exposure to EE triggered participants to spend twice as much time helping the general colorblind population (rather than the just the confederate) as participants who engaged in traditional PT. IRI did not moderate the effect of experimental treatments on helping behavior and its role as a moderator remains inconclusive.[4]

GENERAL DISCUSSION

Summary of Experiments

The current experiments explored the possibility of using IVET to allow individuals to embody vivid perceptual experiences and whether such experiences would encourage greater oneness, more favorable attitude toward people with colorblindness, and ultimately, helping behavior. Individual differences in feeling concern for others were measured and investigated as a moderating variable. Across the three experiments, there is evidence that IVET may be used as a tool to study perspective taking and promote helping.

Experiment 1 was an initial exploration of EE to test its effectiveness against PT and main findings demonstrated that the difference in the level of oneness and attitude elicited by experimental treatments (EE vs. PT) was greater for participants with lower IRI. EE was more effective for individuals with lower IRI: These people perceived greater oneness with the colorblind confederate and developed more positive attitude toward people with colorblindness that lasted up to 24 hours after the treatment. The difference between the effect of EE and PT on oneness and attitude was smaller for participants with higher IRI, for whom PT was more effective than EE. Our initial attempt at measuring helping by asking participants to help colorblind people train to overcome the disability backfired with people with lower

IRIs in the PT condition demonstrating more helping behavior than any other participants.

Experiment 2 was conducted to investigate the underlying mechanism of EE. Results indicated that participants in the EE condition found the treatment to be much more vivid and realistic compared to traditional PT. The increase in realism led to the increase in oneness. However, the measurement on attitude toward colorblind persons yielded results inconsistent with the first experiment in terms of the moderating role of IRI. This time, the difference in attitudes elicited by either EE or PT was greater for participants with higher IRI, for whom EE was more effective than PT. Also, a new helping measure was introduced with improved believability but no differences were found between conditions.

Experiment 3 was a final attempt at testing the ability of EE to promote helping behavior in the physical world. The helping task with improved believability from Experiment 2 was scheduled earlier on in the day so that participants would not have scheduling conflicts with their personal lives such as dinner appointments. Consequently, participants in the EE condition demonstrated twice as much helping behavior compared to participants in the PT condition. IRI did not moderate this effect.

Theoretical Implications

These results contribute to the field of communication and more specifically media psychology in largely three ways. First of all, the findings expand the Proteus Effect (Yee & Bailenson, 2007, 2009; Yee et al., 2009). In the original work, participants were asked to embody an avatar that differed in appearance with the participant in terms of height or attractiveness and found that participants behaved in ways that they believed people of such height or attractiveness would behave. The current experiments expanded on these results by allowing individuals to embody perceptual experiences. Another procedural difference from the Proteus Effect was that participants were explicitly instructed that they would be stepping into the shoes of another person, whereas participants in Yee and Bailenson's (2007, 2009) studies were told that the embodied avatar was the self. Consequently, the current experiments demonstrated that when participants embodied the perceptual experience of a colorblind person through the vivid sensory information provided via IVET, they felt similar to and spent time trying to help the colorblind person rather than simply acting like a colorblind person.

Second, Experiment 2 contributed to our understanding of EE by confirming that the perception of presence fully mediates the relationship between EE and oneness. The realism of the experience as measured by presence may well be a crucial factor of EE, particularly in the context of helping. Earlier studies (Cialdini et al., 1997; Maner et al., 2002) have evidenced the importance of perceived oneness with the target in eliciting

helping intentions; that is, these studies argued that people help based on selfish motivations and that people were only willing to help when they felt high oneness with the target (i.e., self is seen in the other when self-other merging occurs). Thus, the vivid sensory information provided in EE may have made the experience feel as if it were really happening to the self, leading to high perceptions of oneness. Based on the results of prior studies, higher perception of oneness may be what led to actual helping behavior in Experiment 3 and this relationship should be explored in future studies. These implications from Experiment 2 advance earlier work that explored the effects of embodying avatars and their experiences in immersive virtual environments (Ahn, 2011; Ahn & Bailenson, 2011; Fox, 2010) by yielding meaningful insights to the underlying mechanism of EE.

Finally, Experiment 3 expanded the findings of studies that deal with helping and altruism but stop at measuring self-reported intentions to help (e.g., Batson, 1997; Cialdini et al., 1997; Maner et al., 2002) and also studies that deal with virtual simulations (e.g. Jin, Ai, & Rasmussen, 2005; Kalyanaraman et al., 2010) by confirming that the effects of EE within the virtual environment can transfer into the physical world as actual helping behavior. The fact that participants in the EE condition, who were mostly undergraduate college students, spent an average of about 7 minutes alone (vs. 3.5 minutes in the PT condition), unsupervised, to help a complete stranger only after less than 5 minutes of exposure to EE is worth noting. This is even more so when these students clearly knew that they would not be compensated for their help and that they were free to leave whenever they wished.

There may be myriads of opportunities to practically apply EE to encourage favorable attitudes and helping behavior. The most representative application would be in diversity training in either educational or workplace contexts. For instance, a recent study demonstrated that diversity training with English speaking participants through traditional perspective taking and role-playing significantly increased favorable attitude toward non-English speaking individuals in a workplace scenario (Madera, Neal, & Dawson, 2011). Our findings imply that using EE in similar diversity training settings would amplify these positive attitudes. More importantly, the result of the diversity training through EE may be manifested as actual helping behavior toward members of different social groups.

Of course, the study is not without limitations. Helping, because it involves some extent of self-sacrifice, is difficult to demonstrate in laboratory settings when the participant has been introduced to a stranger in an unfamiliar location and situation. While it is meaningful that we were able to demonstrate actual helping behavior regardless of these restrictions, we can only conclude that EE encourages helping under qualified circumstances—when it is clear how the task will actually help the person in need and when the individual can spare the time. Although the helping tasks across all three experiments were designed to help the confederate and other

individuals sharing the same disability, the main difference in the helping tasks incorporated in the current experiments seems to be the degree of feasibility. The task in Experiment 1 claiming to help the confederate to learn how to differentiate colors may have seemed highly unfeasible whereas the task in Experiments 2 and 3 claiming to help the colorblind population by creating more accessible Web sites may have seemed more feasible. Indeed, earlier research showed that the confidence that the help-giver had with regard to successfully helping the individual in need moderated the amount of effort invested in helping (Oettingen, Stephens, Mayer, & Brinkmann, 2010). In Experiment 1, participants with lower IRI that were in the PT condition experienced less self-other merging and likely had greater unrealistic confidence of helping the confederate compared to participants who experienced greater self-other merging. This unrealistic confidence, in turn, may have elicited more helping and remains an interesting question for future research.

Also, the role of IRI as a moderator remains inconclusive. Although the effect of EE on people's attitude toward people with colorblindness seems to linger for up to 24 hours, the inconsistency in the moderating role of IRI makes it difficult to draw solid conclusions. Because the attitude measure was administered 24 hours following the experiment, there was little control over the participants during that time and situational variances that may have confound the results. Nevertheless, the fact that the effect of a few minutes of EE on attitudes toward individuals with disabilities was sustained for up to 24 hours is encouraging and merits further research.

Another limitation is the large standard deviations in the helping measure in Experiment 3. Because the measure of helping behavior is a ratio level of measurement with potentially large variability in the scale (i.e., participants could spend no time or as much time as they wanted to), the large standard deviation is almost inevitable. Despite the relatively small effect size in terms of R^2 change, no other study to the best of our knowledge has attempted to measure helping behavior in terms of actual time and effort spent. Considering the various factors that may have hindered helping behavior, such as the undergraduate population, a target who is a complete stranger, and no compensation for helping, our findings are meaningful in that we saw preliminary success in using EE to encourage actual helping behavior in the physical world, quantifiable in terms of the time and effort invested.

Conclusion and Future Directions

Based on these limitations, future studies should pay attention to both dispositional (e.g., IRI) and situational (e.g., EE or PT) variables and how they encourage interpersonal understanding and helping behavior. As the current experiments have provided a starting point for investigating the process of EE, future studies should probe boundary conditions from both the technical

aspect—amount of sensory information and its relationship to presence and oneness—as well as the situational aspect—whether some contexts induce more understanding and helping behavior than others.

With the advent of consumer technologies such as the Microsoft Kinect and Nintendo Wii, virtual environments of varying levels of immersiveness have become increasingly accessible. This means that people are able to enjoy the benefits of EE in their own livings rooms. Furthermore, technological advances in these consumer technologies are rapidly eliminating the need for relatively clunky devices such as the HMD and haptic device used for the current experiments, replacing them with motion sensing technology that allows users to control mediated contents with naturalistic gestures. This is meaningful as users are now able to enjoy virtual experiences without having to rely on costly equipment that restrict natural movement. The increased accessibility of hardware also implies an increase in accessibility of software, meaning that the content of EE may be tailored to the individual user.

In sum, the collection of three experiments compared EE against PT and investigated underlying mechanisms. More importantly, the current study makes a meaningful contribution as one of the few studies to demonstrate the increase of actual helping behavior through the use of IVET. With EE, the user is able to vividly, accurately, and realistically experience the sensations of another person and feel as if they have merged with that person. This sense of self-other merging in the virtual environment transfers to the physical world and translates into actual helping behavior, even when the other person is a complete stranger. In the words of a participant, "I did not realize how difficult it could be to do such an easy task [match colors]." These experiments demonstrate that too often, this realization seems to come only when you can see for yourself.

NOTES

1. Based on the participants' responses describing their thoughts on the EE or PT treatments, two coders were trained to code the thoughts for suspicion regarding the cover story (Cohen's $\kappa = .74$). Four participants demonstrated extremely high suspicion (e.g., "Is this person just another student looking at the 'colorblind' version of the board," "He's not really colorblind."). Data from these participants were removed from the final dataset.
2. Two of the male participants were much older than the student population—50 and 69 years old—but statistical analyses did not change after excluding them. Consequently. they were kept in the final dataset.
3. Chronologically, Experiment 3 was run before Experiment 2, but they are presented in reverse order in this article to enhance the comprehensibility and logic behind the experiments.
4. Sex of the participants was also tested as a moderator in an analysis of variance (ANOVA) to explore the effect of interactions between sex and experimental conditions on oneness, helping time, helping word, and attitude. None of the tests yielded significant results for the interaction and consequently, were not reported.

REFERENCES

Ahn, S. J. (2011). *Embodied experiences in immersive virtual environments: Effects on pro-environmental attitude and behavior* (Doctoral dissertation). Stanford University, Stanford, CA.

Ahn, S. J., & Bailenson, J. (2011). Self-endorsing versus other-endorsing in virtual environments: The effect on brand attitude and purchase intention. *Journal of Advertising, 40*, 93–106. doi:10.2753/JOA0091-3367400207

Aiken, L. S., & West, S. G. (1991). *Multiple regression: Testing and interpreting interactions*. Newbury Park, CA: Sage.

Aron, A., Aron, E., & Smollan, D. (1992). Inclusion of the other in the self scale and the structure of interpersonal closeness. *Journal of Personality and Social Psychology*, *63*, 596–612. doi:10.1037/0022-3514.63.4.596

Bailenson, J. N., & Yee, N. (2007). Virtual interpersonal touch: Haptic interaction and copresence in collaborative virtual environments. *International Journal of Multimedia Tools and Applications, 37*, 5–14. doi:10.1007/s11042-007-0171-2

Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51*, 1173–1182. doi:10.1037//0022-3514.51.6.1173

Barsalou, L. W. (2009). Simulation, situated conceptualization, and prediction. *Philosophical Transactions of the Royal Society B, 364*, 1281–1289. doi:10.1098/rstb.2008.0319.

Batson, C. D. (1991). *The altruism question: Toward a social-psychological answer*. Hillsdale, NJ: Erlbaum.

Batson, C. D., Polycarpou, M. P., Harmon-Jones, E., Imhoff, H. J., Mitchener, E. C., Bednar, L. L., et al. (1997). Empathy and attitudes: Can feeling for a member of a stigmatized group improve feelings toward that group? *Journal of Personality and Social Psychology, 72*, 105–118. doi:10.1037/0022-3514.72.1.105

Batson, C. D., Sympson, S., Hindman, J., Decruz, P., Todd, R., Weeks, J., et al. (1996). "I've been there, too": Effect on empathy of prior experience with a need. *Personality and Social Psychology Bulletin, 22*, 474–882. doi:10.1177/0146167296225005.

Biocca, F. (1997). The cyborg's dilemma: Progressive embodiment in virtual environments. *Journal of Computer Mediated-Communication, 3*(2). Retrieved from http://www.ascusc.org/jcmc/vol3/issue2/biocca2.html. doi:10.1111/j.1083-6101.1997.tb00070.x

Blascovich, J., Loomis, J., Beall, A. C., Swinth, K., Hoyt, C., & Bailenson, J. N. (2002). Immersive virtual environment technology as a methodological tool for social psychology. *Psychological Inquiry, 13*, 103–124. doi:10.1207/S15327965PLI1302_01

Cialdini, R. B., Brown, S. L., Lewis, B. P., Luce, C. L., & Neuberg, S. L. (1997). Reinterpreting the empathy-altruism relationship: When one into one equals oneness. *Journal of Personality and Social Psychology, 73*, 481–494. doi:10.1037//0022-3514.73.3.481

Darley, J. M., & Batson, C. D. (1973). "From Jerusalem to Jericho:" A study of situational and dispositional variables in helping behavior. *Journal of Personality and Social Psychology*, *27*, 100–108. doi:10.1037/h0034449

Davis, M. H. (1980). A multidimensional approach to individual differences in empathy. JSAS *Catalog of Selected Documents in Psychology, 10*, 85. doi:10.1037//0022-3514.44.1.113

Davis, M. H., Conklin, L., Smith, A., & Luce, C. (1996). Effect of perspective taking on the cognitive representation of persons: A merging of self and other. *Journal of Personality and Social Psychology*, *70*, 713–726. doi:10.1037//0022-3514.70.4.713

Davis, M. H., & Kraus, L. A. (1997). Personality and empathic accuracy. In W. J. Ickes (Ed.), *Empathic Accuracy* (pp. 144–168), New York, NY: Guilford Press.

Davis, M. H., Luce, C., & Kraus, S. J. (1994). The heritability of characteristics associated with dispositional empathy. *Journal of Personality, 62*, 369–391. doi:10.1111/j.1467-6494.1994.tb00302.x

de Waal, F. B. M. (2008). Putting the altruism back into altruism: The evolution of empathy. *Annual Review of Psychology, 59*, 279–300. doi:10.1146/annurev.psych.59.103006.093625.

Fox, J. (2010). *The use of virtual self models to promote self-efficacy and physical activity performance* (Doctoral dissertation). Stanford University, Stanford, CA.

Fox, J., Bailenson, J. N., & Binney, J. (2009). Virtual experiences, physical behaviors: The effect of presence on imitation of an eating avatar. *PRESENCE: Teleoperators & Virtual Environments, 18*, 294–303.

Fussell, S. R., & Krauss, R. M. (1989). The effects of intended audience on message production and comprehension: Reference in a common ground framework. *Journal of Experimental Social Psychology, 25*, 203–219. doi:10.1016/0022-1031(89)90019-X

Galinsky, A. D., & Moskowitz, G. B. (2000). Perspective taking: Decreasing stereotype expression, stereotype accessibility and in-group favoritism. *Journal of Personality and Social Psychology, 78*, 708–724. doi:10.1037//0022-3514.78.4.708

Galinsky, A. D., Wang, C. S., & Ku, G. (2005, June 12–15). *The defecting perspective-taker: The impact of stereotypes and perspective taking in a prisoner's dilemma*. Proceedings of the IACM 18th Annual Conference, Seville, Spain.

Gehlbach, H. (2004). A new perspective on perspective taking: A multidimensional approach to conceptualizing an aptitude. *Educational Psychology Review, 16*, 207–234. doi:10.1023/B:EDPR.0000034021.12899.11

Gehlbach, H., Brinkworth, M., & Wang, M. (2012). The social perspective taking process: What motivates individuals to take another's perspective? *Teachers College Record, 114*, 197–225.

Goldstein, N. J., & Cialdini, R. B. (2007). The spyglass self: A model of vicarious self-perception. *Journal of Personality and Social Psychology*, *92*, 402–417. doi:10.1037/0022-3514.92.3.402

Golin, S. (Producer), & Jonze, S. (Director). (1999). *Being John Malkovich* [Motion picture]. United States: Propaganda Films.

Green, M. C. (2004). Transportation into narrative worlds: The role of prior knowledge and perceived realism. *Discourse Processes, 38*, 247–266. doi:10.1207/s15326950dp3802_5

Green, M. C., Brock, T. C., & Kaufman, G. F. (2004). Understanding media enjoyment: The role of transportation in to narrative worlds. *Communication Theory, 14*, 311–327. doi:10.1093/ct/14.4.311

Hodges, S. D., & Klein, K. (2001). Regulating the costs of empathy: The price of being human. *Journal of Socio-Economics, 30*, 437–452. doi:10.1016/S1053-5357(01)00112-3

Hoffman, M. (1982). Development of prosocial motivation: Empathy and guilt. In N. Eisenberg (Ed.), *The development of prosocial behavior* (pp. 281–313). New York, NY: Academic Press.

Jin, B., Ai, Z., & Rasmussen, M. (2005, September 1–4). *Simulation of eye disease in virtual reality.* In Proceedings of the 2005 IEEE Engineering in Medicine and Biology 27th Annual Conference, Shanghai, China.

Kalyanaraman, S., Penn, D. L., Ivory, J. D., & Judge, A. (2010). The virtual doppelganger: Effects of a virtual reality simulator on perceptions of schizophrenia. *Journal of Nervous and Mental Disease, 198*, 437–443.

Krauss, R. M., & Fussell, S. R. (1991). Perspective-taking in communication: Representations of others' knowledge in reference. *Social Cognition, 9*, 2–24. doi:10.1521/soco.1991.9.1.2.

Loomis, J. M. (1992). Distal attribution and presence. *PRESENCE: Teleoperators and Virtual Environments, 1*, 113–119. doi:10.1117/12.136005

Madera, J. M., Neal, J. A., & Dawson, M. (2011). A strategy for diversity training: Focusing on empathy in the workplace. *Journal of Hospitality & Tourism Research, 35*, 469–487. doi:10.1177/1096348010382240

Maner, J. K., Luce, C. L., Neuberg, S. L., Cialdini, R. B., Brown, S., & Sagarin, B. J. (2002). The effects of perspective taking on motivations for helping: Still no evidence for altruism. *Personality and Social Psychology Bulletin, 28*, 1601–1610. doi:10.1177/014616702237586

Matthews, K. A., Batson, C. D., Horn, J., & Rosenman, R. H. (1981). "Principles in his nature which interest him in the fortune of others . . . " The heritability of empathic concern of others. *Journal of Personality, 49*, 237–247. doi:10.1111/j.1467-6494.1981.tb00933.x

Nowak, K. L., & Biocca, F. (2003). The effect of agency and anthropomorphism on users' sense of telepresence, copresence, and social presence in virtual environments. *Presence: Teleoperators and Virtual Environments, 12*, 481–494. doi:10.1162/105474603322761289

Oettingen, G., Stephens, E. J., Mayer, D., & Brinkmann, B. (2010). Mental contrasting and the self-regulation of helping relations. *Social Cognition, 28*, 490–508. doi:10.1521/soco.2010.28.4.490

Rushton, J. P., Fulkner, D. W., Neale, M. C., Nias, D. K. B., & Eysenck, H. J. (1986). Altruism and aggression: The heritability of individual differences. *Journal of Personality and Social Psychology, 50*, 1192–1198. doi:10.1037//0022-3514.50.6.1192

Sas, C. (2004). Individual differences in virtual environments. In M. Bubak, G. Dick van Albada, P. Sloot, & J. Dongarra (Eds.), *Computational science—ICCS 2004, Fourth International Conference, Proceedings, Part III. Lecture notes in computer science* (vol. 3038, pp. 1017–1024). Berlin, Germany: Springer-Verlag.

Sas, C., & O'Hare, G. M. P. (2003). Presence equation: An investigation into cognitive factors underlying presence. *Presence: Teleoperators and Virtual Environments, 12*, 523–537. doi:10.1162/105474603322761315

Siegler, R. S. (1995). How does change occur: A microgenetic study of number conversation. *Cognitive Psychology, 28*, 225–273. doi:10.1006/cogp.1995.1006

Slater, M., & Wilbur, S. (1997). A framework for immersive virtual environments (FIVE): Speculations on the role of presence in virtual environments. *Presence: Teleoperators and Virtual Environments, 6*, 603–616.

Van Boven, L., & Loewenstein, G. (2003). Social projection of transient drive states. *Personality and Social Psychology Bulletin, 29*, 1159–1168. doi:10.1177/0146167203254597

Wallach, H. S., Safir, M. P., & Samana, R. (2010). Personality variables and presence. *Virtual Reality, 14*, 3–13. doi:10.1007/s10055-009-0124-3

Webster D. M., Richter, L., & Kruglanski, A. W. (1996). On leaping to conclusions when feeling tired: Mental fatigue effects on impressional primacy. *Journal of Experimental Social Psychology, 32*, 181–195. doi:10.1006/jesp.1996.0009

Witmer, B. G., & Singer, M. J. (1998). Measuring presence in virtual environments: A presence questionnaire. *Presence: Teleoperators and Virtual Environments, 7*, 225–240. doi:10.1162/105474698565686

Yee, N., & Bailenson, J. N. (2007). The Proteus Effect: The effect of transformed self-representation on behavior. *Human Communication Research*, *33*, 271–290. doi:10.1111/j.1468-2958.2007.00299.x

Yee, N., & Bailenson, J. N. (2009). The difference between being and seeing: The relative contribution of self perception and priming to behavioral changes via digital self-representation. *Media Psychology*, *12*, 195–209. doi:10.1080/15213260902849943

Yee, N., Bailenson, J. N., & Duchenaut, N. (2009). The Proteus effect: Implications of transformed digital self-representation on online and offline behavior. *Communication Research, 36*, 285–312. doi:10.1177/0093650208330254

Yuker, H. E., Block, J. R., & Younng, J. H. (1966). *The measurement of attitudes toward disabled persons*. Albertson, NY: Human Resources Center.

THE PLUG

This VR Founder Wants to Gamify Empathy to Reduce Racial Bias

Clorama Dorvilias wants to remove shame and guilt from diversity and inclusion training.

By Tyler Young | Jul 20 2018, 9:00am

Clorama Dorvilias Image: Courtesy of Dorvilias

The tech industry, and companies like Starbucks, have spent millions of dollars to address implicit racial bias through diversity initiatives and racial sensitivity trainings. When Starbucks closed 8,000 of its stores in May for racial bias training, the company recognized it would take more than four hours to undo a lifetime of unconscious behaviors in just one of its baristas, let alone 175,000 employees. One barista said she felt "shaken" during the training while watching "Story of Access," a short film by award-winning documentary filmmaker Stanley Nelson highlighting the emotional tax for people of color in public spaces.

US companies invest an estimated $8 billion on diversity and inclusion training without clear metrics to determine the effectiveness and employee performance. But they don't always go as planned. Employees can walk away from these trainings with triggered emotions.

Clorama Dorvilias wants to remove shame and guilt from this experience. Her gaming lab, Debias VR, wants to remove bias in professional and educational settings. Dorvilias

"Companies can spend millions of dollars on extremely inaffective, and virtually useless trainings that can have an adverse effect which can hurt the company even more," she said. "Bias training shouldn't be there to shame. People should feel good about making others feel accepted. Debias isn't something that you can work out in a day. It's a behavior that you have to work through. We want to give people the capacity to work in a safe and comfortable space."

Before she developed the VR lab, Dorvilias says she enrolled in grad school to expand her front-end development expertise and planned to backpack across the country working with grassroots organizations to amplify their digital presence.

Dorvilias enrolled in the Interaction Design Master's program at the University of the Arts London in 2014. She continued working as an independent UX designer, becoming quickly immersed in emerging tech and motion sensors for everyday human interactions.

Dorvilias and her partner, Jessica Outlaw. Image: courtesy of Dorvilias

"That was really frustrating for me, and I was contemplating quitting because it began to affect my mental health and the self-image of the work I was doing."

were asked to work with technology to solve social problems. The served as the most fitting assignment for the unfair treatment she felt in the course. She used this as an opportunity to explore what causes implicit bias and how to solve it. “I stumbled upon empathy, which allows people to humanize each other. I knew the solution would have to be, how can I develop empathy for a professor like him for someone like me?”

Little did she know, this would forge her path to Teacher’s Lens.

She explored the language around implicit bias including microaggressions, a term first introduced psychiatrist Dr. Chester Pierce in the 1970s. The term has since evolved with Columbia University psychologist Dr. Derald Wing Sue defining racial microaggression as "everyday insults, indignities and demeaning messages sent to people of color by well-intentioned white people who are unaware of the hidden messages being sent to them."

During her research on empathy solutions, Dorvilias became intrigued watching a TED talk by Chris Milk titled, "How Virtual Reality Can Create the Ultimate Empathy Machine." Milk, founder and CEO of VR media company Within, explained how he used footage of day-to-day life in third world countries using 360 video to establish empathy. He produced a VR film, “Clouds Over Sidra,” placing users inside a Syrian refugee camp that follows a 12-year-old girl named Sidra.

When the Oculus Rift is in use, viewers experienced first-hand trauma with Sidra. Milk then took the machine to the World Economic Forum where he says the people who make decisions for third-world countries can insert themselves in the places where they're charged with creating change.

Dorvilias didn’t have the same resources but she knew she wanted a similar outcome. With the help of a defunct Oculus headset at school and video tutorials on building a virtual reality environment, she made her own in two days. “It bit me from there. For the next four months. I slept, ate, and dreamed about virtual reality and I was able to make my first project. I was able to gamify a diversity trainer using debiasing techniques.”

Along the way she, documented her findings.

She developed the game using minority avatars set in a tech environment and users had to rely on people of color or women to get closer to their end goal to win the game. Her game is backed by research from the University of Barcelona. In the experiment, 32 white women participated in a demo where half were portrayed in a white virtual body and the other half in a black virtual body. The participants unconsciously imitated gestures of avatars who shared their skin color. They responded to more interactions with avatars that resembled their virtual features. For instance, the white women who depicted dark-skinned avatars copied the movements of other black avatars as opposed to white ones. The result concluded positive social outcomes and proved implicit racial bias tends to diminish when users are assigned a virtual race, different than their own. Dorvilias implemented the study, successfully completed the project, and finished her graduate program in late 2015.

Dorvilias didn't expect to stay in VR because at the time the industry wasn't prominent in Europe and returned to in 2016. "When I came back, I saw there was a huge thriving VR community and I realized I was ahead of ahead of the game than people who were just getting started."

In 2017, Dorvilias earned a UX Researcher and Designer fellowship through Code for America, developing an app for job seekers in Anchorage.

Dorvilias recently developed Teacher's Lens, described by its developers as the first VR solution "that allows for measurable training towards diversity and inclusion goals." Teacher's Lens is a pilot program under Debias VR developed by Dorvilias and with research led by her app cofounder, Jessica Outlaw. Using the same guiding principles of Debias VR, Dorvilias built an accountability tool to track bias reduction in classrooms.

Outlaw, founder of The Extended Mind, a company that trains VR/AR companies through practical behavioral science insights to improve overall user engagement, met Dorvilias at Oculus Launch Pad in 2017. The event, held at Facebook HQ, had ended and the two women were waiting for their Lyfts to travel back to San Francisco. Outlaw describes their meetup like a classic scene from a romcom.

A few minutes into their chat, she remembers telling Dorvilias, “Just cancel your Lyft, let’s just go together.” They shared a ride and began a 45-minute discussion that would soon lead to the creation of an app correcting implicit racial bias using VR.

The two delved into classroom bias, where studies show educators have lower expectations for students of color and are less likely to encourage girls in STEM courses. For instance, researchers at the Proceedings of the National Academy of Sciences of the United States of America (PNAS) found recurring gender disparities exists the in sciences where educators tend to show more favor for male students that girls. What comes as the perfect prism for their cofounder relationship, Outlaw says six months prior to meeting Dorvilias she’d become interested in anti-discrimination tools and the research behind it. In six weeks, the concept became concrete and Outlaw says Dorvilias got to work completing the app in three weeks.


Outlaw and Dorvilias. Image: Courtesy of Dorvilias

Outlaw credits Dorvilias as the tech muscle behind the platform with her research contributions rounding out their team.

“Black and Latino students are disciplined at higher rates and tracked into AP courses at lower rates so it would be good if people would start making decisions based on data and examine some underlying structural issues,” Outlaw said. “This is an evidence-based approach, if you can change teacher expectations you can change student’s performance.”

disproportionately subject students of a different race to harsher treatment and are less likely to provide them support causing an achievement gap for students of color.

Outlaw says the “late bloomers” study also played a key role in their research. Formally known as The Pygmalion Effect, the experiment was conducted in the early 1960's by psychologist Robert Rosenthal and Lenore Jacobson. It found that, proving intellectual improvements take place when students were held to higher expectation by their teacher—which often doesn’t happen if a teacher has extreme biases.

“It affects society and our economy as a whole when you draw the line on what you think a person is capable of doing based on how they look,” Dorvilias said. “What limitations are we placing on society and our innovation when people in power put a cap on who gets to succeed and who doesn’t?”

Dorvilias and Outlaw won an Oculus Launch Pad scholarship in late 2017 and officially launched the app on May 25th, 2018.

While Teacher's Lens is primarily funded by Oculus, Dorvilias says the app is still being pitched to investors. “A lot of the investors that we meet will say they’re willing to mentor us. For some reason when it comes to money they’re not throwing it in our direction, ” she said.

There are similar VR tools on the market, but Dorvilias she feels that hers is more unique in a sense that it packages positive interactions by tracking progress over time with a sustainable outcome that doesn’t involve suffering. She’s continuing to pitch Debias VR and has plans to develop more virtual empathy tools to use in schools and businesses. “The nature of VR has limits to it. Until people try it, they won’t really understand why it’s so powerful. Once they try it, they’re transformed.”

This piece is part of a series of stories produced in partnership with The Plug .

Watch This Next

Walk A Mile in Digital Shoes: The Impact of Embodied Perspective-Taking on The Reduction of Negative Stereotyping in Immersive Virtual Environments

Nick Yee and Jeremy Bailenson
Stanford University
{nyee@stanford.edu, bailenson@stanford.edu}

Abstract

In social psychology, perspective-taking has been shown to be a reliable method in reducing negative social stereotyping. These exercises have until now only relied on asking a person to imagine themselves in the mindset of another person. We argue that immersive virtual environments provide the unique opportunity to allow individuals to directly take the perspective of another person and thus may lead to a greater reduction in negative stereotypes. In the current work, we report on an initial experimental investigation into the benefits of embodied perspective-taking in immersive virtual environments. It was found that negative stereotyping of the elderly was significantly reduced when participants were placed in avatars of old people compared with those participants placed in avatars of young people. We discuss the implications of these results on theories of social interaction and on copresence.

1. Introduction

A great deal of research in social psychology has focused on the nature of stereotypes and prejudice. For example, since Allport's classic work developing a theoretical framework of social categories [1], it has become apparent that strong beliefs about inter-group differences can be created with minimal and oftentimes arbitrary decisions [2, 3]. For example, individuals randomly assigned to one of two different groups will have more positive expectations of members in their own group and more negative expectations of members of the other group despite the fact that there is no rational reason to differentiate members of one group over another.

Researchers have also demonstrated that stereotype activation oftentimes occurs with an automaticity that is beyond conscious control [4] and that the presence of these stereotypes leads to prejudicial interactions unless conscious intervention is applied [5]. These stereotypes not only impact minority groups in social interactions due to the prejudicial treatment they receive from others, but also create cognitive burdens for these minorities themselves as well. For example, negative stereotypes can lead to systematic underperformance via a mechanism known as stereotype threat [6-8]. In a landmark study of stereotype threat, when black students were given a verbal test that they were told was a direct measure of their intellectual ability, they performed worse than another group of black students who were told the test was about understanding different problem-solving strategies [8]. The authors of that study argued that the extra pressure caused by the fear of reinforcing a negative stereotype causes systematic underperformance in ability tests.

These findings all point to the explicit and implicit hold that stereotypes and prejudice have on our society, and the difficulty in preventing tensions and conflicts due to the existence of stereotypes. In the current work, we explore potential interventions for decreasing the application of stereotypes, and in particular, we present empirical data from an experimental design that implemented an intervention method in immersive virtual reality (VR) with beneficial results.

1.1 Decreasing impact of stereotypes

The line of research into the nature of prejudice has led other researchers to explore ways to decrease the accessibility and application of stereotypes. The earliest method of intervention suggested was by Allport himself in his early work [1]. The *Contact Hypothesis* was the suggestion that social interaction between two groups of individuals would decrease the existing conflicts and tensions between them as mutual understanding occurred. The success of this form of intervention, however, depended on several factors, which Allport noted. For example, 1) the two groups must have equal status, 2) and share resources or power in such a way as to create a mutual interdependence, and 3) the context for interaction must be conducive to positive and friendly interactions. While these factors may be introduced in some artificially-created groups [2] to decrease inter-group tension, these factors may be hard to introduce in groups where power imbalances are deeply entrenched (such as with racial or gender stereotypes). Thus, the Contact Hypothesis offers one potential solution that is unfortunately highly situational.

Another intuitive intervention method is *thought suppression*. For example, individuals who try to avoid treating minorities prejudicially may attempt to suppress stereotypical references before the interaction. It has been shown, however, that deliberate thought suppression often backfires [9, 10]. Because a representation of the target stereotype must be articulated in order to suppress it, deliberate thought suppression has the effect of making stereotypical concepts and traits hyper-accessible. This search also functions as a form of

repetitive priming. In other words, suppressed stereotypes often become more, rather than less, accessible and salient after the intervention.

1.2 Perspective Taking

One intervention method that has yielded positive results derives from the concept of *perspective-taking*. In social interactions, the fundamental attribution error affects how we think about and evaluate ourselves and others [11]. When we judge ourselves, we tend to rely on situational factors (i.e., "I did poorly on the test because I didn't sleep well the night before."). On the other hand, when we judge others, we tend to rely on dispositional factors (i.e., "He did poorly on the test because he's not that bright."). Thus, when people are forced to observe their own actions (via a video tape), they tend to make more dispositional rather than situational attributions [12]. The reverse is also true. When participants are asked to take the perspective of the person they are observing, participants tend to make situational rather than dispositional attributions [13].

More importantly, it has been found that perspective-taking leads to an increased overlap between the self and other. In one study, participants rated themselves and another person more similarly on a set of trait words in the perspective-taking condition in which they asked to take the point of view of another person than in a control condition [14]. It was also found that participants felt the target was more similar to themselves than control participants and liked the other person more after the perspective-taking exercise. Thus, on an individual level, perspective-taking had been shown to generate positive interpersonal effects.

Galinsky and Moskowitz [15] extended this work on perspective-taking into the domain of stereotypes and prejudice. They hypothesized that the benefits of perspective-taking may extend to inter-group evaluations and interactions. In other words, by encouraging people to focus on situational rather than dispositional factors via perspective-taking, they may rely less on stereotypes in evaluating and interacting with members of minority groups. In their work, they worked with stereotypes about the elderly because it has been shown that college-age students automatically associate negative traits with the elderly [16]. Participants were shown a photograph of an elderly man and were asked to write a short narrative essay about a typical day in the life of this individual. Participants also performed an implicit association task involving recognition of words related to old age. In a lexical decision task [17], words are flashed briefly on a computer screen and participants are asked to categorize the flashes as words or non-words. It has been shown that concepts and associations that are more accessible in a person's mind will be recognized faster than concepts that are less accessible. Thus, if an individual has a strong negative association with the elderly, then they are more likely to recognize words with negative connotations quicker (i.e., frail, wrinkled, sick) than neutral words or worlds with a positive connotation. It was found that perspective-taking decreased the amount of implicit and explicit stereotyping while increasing the amount of self-other overlap. The significant contribution of their work was in showing that the positive effects of perspective-taking can extend to the group level (i.e., improving evaluations of all elderly people after taking the perspective of one elderly man) rather than simply on the individual level (i.e., improving the evaluation of Tom after taking the perspective of Tom).

1.3 Collaborative Virtual Environments

If it were possible to convincingly place an individual into the body of an elderly person, rather than simply asking them to imagine this, we may expect this perspective-taking exercise to have an even stronger effect. Collaborative Virtual Environments [CVEs, see 18, 19, 20] make this manipulation possible. CVEs are communication systems in which multiple interactants share the same three-dimensional digital space despite occupying remote physical locations. In a CVE, immersive virtual environment technology monitors the movements and behaviors of individual interactants and renders those behaviors within the CVE via avatars (digital representations of people). These digital representations are tracked naturalistically by optical sensors, mechanical devices, and cameras. Thus, CVEs offer unique opportunities for social science research [21, 22].

1.4. Transformed Social Interactions

Unlike telephone conversations and videoconferences, the physical appearance and behavioral actions of avatars can be systematically filtered in immersive CVEs idiosyncratically for other interactants, amplifying or suppressing features and nonverbal signals in real-time for strategic purposes. Theoretically, these transformations should impact interactants' persuasive and instructional abilities. Previously, we outlined a theoretical framework for such strategic filtering of communicative behaviors called Transformed Social Interaction [23]. In a CVE, every user perceives their own digital rendering of the world and each other and these renderings need not be congruent. In other words, the target may perceive his or her own avatar as being attractive while the perceiver sees the target as being unattractive.

Previous work on transformed social interaction has demonstrated quite resoundingly that changing one's representation has large implications on other's in terms of social influence. In other words, transforming Avatar A strategically causes Avatar B to behave consistently with the representation of Avatar A (as opposed to the actual representation of Avatar A). A recent review [24] summarizes a number of studies that show social influence resulting from transformations in facial similarity, mimicry, and eye gaze.

Other research has also shown that alterations in digital self-representation can have a large impact on how a person behaves in virtual environments - a phenomenon termed The Proteus Effect [25]. For example, participants given attractive avatars are willing to walk closer and share more personal information to a stranger in a virtual environment than participants given unattractive avatars. Also, participants in taller avatars are significantly more willing to make unfair offers to their own advantage than participants in shorter avatars in a negotiation task. In other words, previous work has shown that alterations in self-representation can lead to significant changes in behavior.

In the current study, we were interested in altering self-representation for a different goal. Specifically, we wanted to explore whether changing a person's self-representation may help in reducing the negative stereotypes against particular social groups. For example, we could place college-age users into an elderly avatar to test whether embodied perspective-taking increases the positive evaluations of the elderly.

Presence is the concept of "being there", a measure of how immersive an environment is [26-31]. We argue that the beneficial effects of perspective-taking would be evident in a high presence situation such as the one proposed. Instead of asking people to imagine the world from the perspective of another person, immersive virtual reality allows us to place a person directly into the body of another person. Thus, in our study, we immersed participants into a virtual reality environment and presented them with an avatar via a virtual mirror. Following the work by Levy [32] and Galinsky and Moskowitz [15], we worked with stereotypes of the elderly in the study. Thus, participants were either given an avatar of a young or an elderly person and were forced to interact with another person while wearing the old or young avatar. We predicted that participants given avatars of elderly people would come to have fewer negative stereotypes of the elderly.

2. Method

2.1. Design

In a between-subjects design, participants were randomly assigned to have an avatar of an elderly person or an avatar of a young adult of the same gender. Participants then interacted with a confederate of the same gender who was blind to condition. Finally, participants completed a survey measure of their attitudes towards the elderly.

2.2. Participants

Forty-eight undergraduate students (24 men and 24 women) participated in the study for course credit.

2.3. Materials

2.3.1 Face Pretest. We selected faces of old and young people based on the results of a pretest. First, we found 12 digital photographs of individuals in each of the gender and age conditions needed in the study (thus 48 images altogether). These digital photographs were frontal photographs of individuals in well-lit conditions consisting of at least 400 by 400 pixels. To reduce variation, we selected only photographs of Caucasians who had no facial hair and were not wearing glasses.

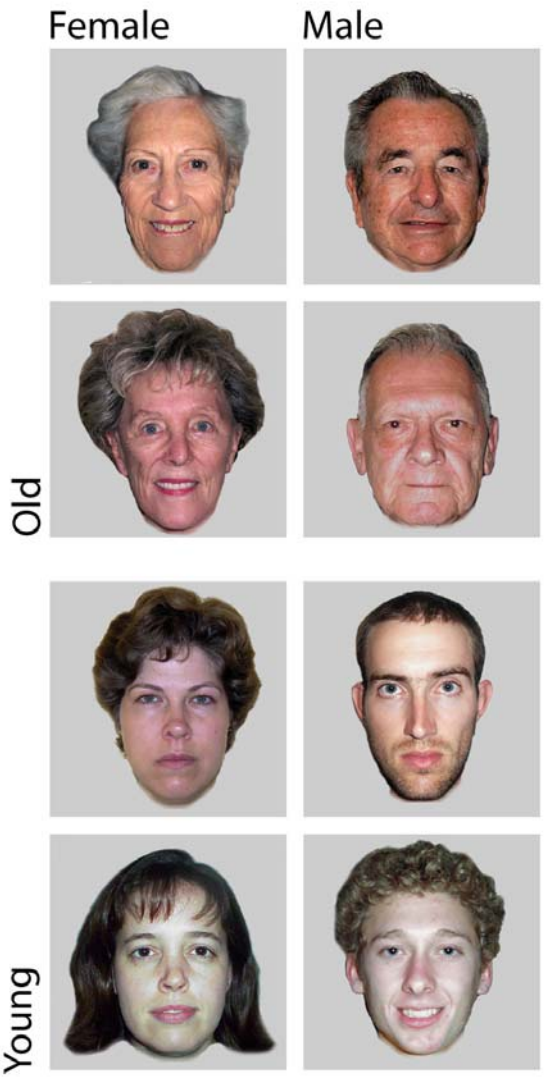


Figure 1 - Avatar faces selected from the pretest for the old and young conditions

Thirty-three undergraduate students who did not participate in the study were presented with each photograph sequentially in a randomized order. They were asked to estimate the age of the individual in the photo on a 6-point

Figure 2 - Equipment setup: A) head-mounted display, B) tracking camera, C) rendering machine.

scale (labeled as 16-25, 26-35, 36-45, 46-55, 56-65, 66-75) as well as rate the attractiveness of the individual on a 7-point fully-labeled scale (from "Extremely Attractive" to "Extremely Unattractive").

Our goal was to select photographs of individuals for both genders that had significantly different age ratings but non-significantly different attractiveness ratings near the mid-point of the attractiveness scale (labeled as "Average"). We selected the two faces for each gender and age condition (thus eight faces altogether) that were closest to the mid-point of the attractiveness scale. A repeated-measures analysis of these eight faces using face trial as the independent variable and attractiveness as the dependent variable was not significant ($F[7, 224] = 1.32$, $p = .24$). A repeated-measures analysis of these eight faces using age group and face trial as the independent variables and age estimate as the dependent variable showed that the effect of age group was significant ($F[1, 32] = 2693.60$, $p < .001$). In particular, the photographed individuals selected for the old condition were rated as significantly older ($M = 5.14$, $SE = .06$, thus between the labels of 55-65 and 66-75) than those in the young condition ($M = 1.67$, $SE = .04$, thus between the labels of 16-25 and 26-35). See Figure 1.

2.3.2. The Physical Lab Setting. The lab consisted of two rooms with an open doorway. In the room where the study took place, a black curtain divided the room. To ensure that confederates and participants were not biased by the attractiveness each other's real faces, confederates stayed behind this black curtain until the VR interaction began and thus never saw the participant's real face and vice versa.

2.3.3. The Virtual Setting. The virtual setting was a white room that had the same exact dimensions as the physical room participants were in. Two meters behind the participant was a virtual mirror that reflected the z-rotation (roll) of the head and body translation (translation on X,Y, and Z) of the participant with the designated face (See Figure 3). Thus, the mirror image tracked and reflected four degrees of freedom such that when the participant moved in physical space, his or her avatar moved in synchrony in the mirror. The confederate's avatar was located 5 meters in front of the participant, facing the participant, and remained invisible until the conversational portion of the experiment began. The confederate's avatar always had a young face of average attractiveness and also had an automated blink animation based on human blinking behavior.

2.4. Apparatus

Perspectively-correct stereoscopic images were rendered by a 1700 MHz Pentium IV computer with an NVIDIA 5950 graphics card, and were updated at an average frame rate of 60 Hz. The simulated viewpoint was continually updated as a

function of the participants' head movements, which were tracked by a three-axis orientation sensing system (Intersense IS250, update rate of 150 Hz). The position of the participant along the x, y, and z planes were tracked via an optical tracking system (WorldViz PPT, update rate of 60 Hz). The system latency, or delay between a participant's movement and the resulting concomitant update in the head-mounted display (HMD) was 45 ms maximum. The software used to assimilate the rendering and tracking was Vizard 2.17. Participants wore an nVisor SX HMD that featured dual 1280 horizontal by 1024 vertical pixel resolution panels that refreshed at 60 Hz. The display optics presented a visual field subtending approximately 50 degrees horizontally by 38 degrees vertically. See Figure 2.

2.5. Procedure

Two research assistants - one male and one female - were present for each trial. Because the confederate was always the same gender as the participant, one research assistant would greet the participant and guide the study while the other would be the confederate based on the gender of the participant.

After informed consent, participants were told that the goal of the experiment was to study social interaction in virtual environments and that they would be having a conversation with another person in a virtual environment. Participants were then led into the room with the black curtain and shown how to wear and adjust the HMD. Once the virtual world was loaded, participants saw themselves in a room that was exactly the same dimensions as the physical lab room, as depicted in Figure 3. The research assistant drew open the curtains at this point.

Figure 3 - Screenshot of participant's point of view in the virtual room with the mirror when he is embodied in an elderly avatar.

Participants were then asked by the lead research assistant to turn around 180 degrees and asked to verify that they saw a mirror in front of them. After verbal affirmation, participants were then told that this is how they appeared to others in the virtual room. Several procedures were used to make sure participants had enough time to observe their avatars' faces. First, they were asked to tilt or nod their head and verbally affirm whether the reflection was following them. Then they were asked to walk up closer to the mirror to get a good look at the face and verbally affirm that the mirror image responded correctly. They were then asked to tilt or nod their head again and verbally affirm that their mirror image followed them. And finally, they were asked to bend down at the knee and come back up and verbally affirm that the mirror image was following them. Every participant was thus exposed to the designated face for between 60 to 75 seconds.

Participants were then asked to turn back around to face the front (i.e., their original orientation). Slightly ahead of time, the research assistant had triggered the program to make the confederate's avatar visible to the participant in the virtual world. Participants were then reminded that others in the virtual room saw them as they had just seen themselves in the mirror. The lead research assistant then told the participants that the other person in the room would initiate the next portion of the experiment.

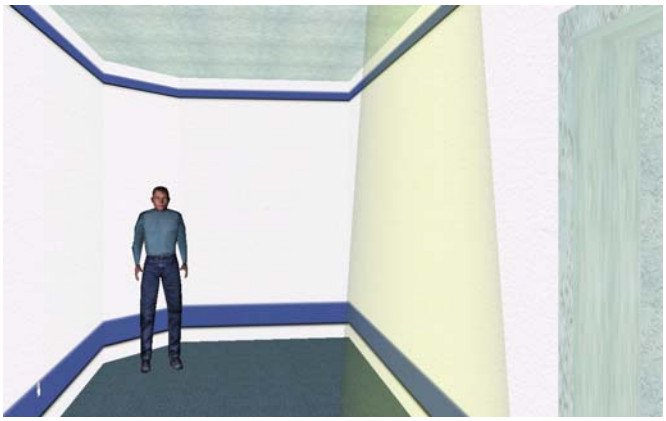


Figure 4 - Screenshot of participant's point of view with confederate across the room and beam of light for the movement task.

The confederate followed a strict script that was displayed in their HMD so they could follow the specific verbal and nonverbal procedures while interacting with the participant inside the CVE. In order to provide the subjects with the time and maximal opportunity to feel self-presence [26] from the designated avatar, subjects were asked to perform a variety of social interactions. First, the confederate asked the participant to walk to several locations in the room under the guise of getting used to moving in a virtual environment. These locations were shown to the participant as cones of light within the virtual environment. There were three such locations that

the participants walked to sequentially (see Figure 4). Then participants - now back at the starting location, facing the confederate again - were asked to answer several questions: 1) “Tell me a little about yourself.”, 2) “What makes you happy in life?”, and 3) “What do you think are the most important things to have in life?”. Participants also performed a brief memory exercise of trying to repeat a list of fifteen words. The purpose of the memory exercise was to reinforce the fact they were wearing an elderly avatar. Finally, the confederate asked participants to approach him or her to make the interaction more social. Participants were then taken out of the virtual environment and completed a questionnaire.

2.6. Measures

The questionnaire included three attitudinal measures towards the elderly. This collection of attitudinal measures were taken from Levy’s [32] study on the effect of self-stereotyping on attitudinal differences towards the elderly.

2.6.1. Word Association. In the word association task, participants were asked the open-ended question “When you think of somebody old, what are the first five words that come to mind?” Two individuals who were blind to condition rated the responses according to the scoring criteria first developed in an earlier study by Levy and Langer [6]. In brief, each of the five words was scored on three dimensions: 1) whether it described an internal or external trait (i.e., wise vs. wrinkled), 2) whether it had a positive or negative connotation (i.e., kind vs. agitated), and 3) whether the word was associated with activity or inactivity (i.e., gardening vs. slow). Finally, the scores for each of the three ratings (low was always negative; high was always positive) were averaged across the five words for each participant. The inter-coder reliability was .66.

2.6.2. Indirect Attitudes. The second measure of participant’s views on aging was derived from Palmore’s [33] Fact on Aging Quiz - a true or false quiz that could be used to measure underlying biases towards the elderly. Again, the scoring was based on the earlier Levy and Langer [6] study. We used both positive and negative bias items in this measure. In a positive bias question, a true negative statement about the elderly is made (i.e., “Old people usually take longer to learn something new.”). If participants choose “false” as the response, then it means they have a positive bias towards the elderly. Conversely, in a negative bias item, a false negative statement about the elderly is made (i.e., “The majority of people over the age of 65 reside in nursing homes.”). If participants choose “true” as the response, then it means they have a negative bias towards the elderly. In this measure, we had a total of 6 positive bias items and 6 negative bias items (see Appendix A). We calculated a bias score for each individual by subtracting their negative bias score from their positive bias score.

2.6.3. Ambiguous Story. Finally, we included an ambiguous story - the Margaret story - drawn from Levy’s [32] study (see Appendix B). The story described a 73-year-old woman named Margaret who moves in with her adult daughter and attends a college reunion. The story consisted of 21 sentences and presented information that could be judged in both positive, neutral, or negative ways. For example, Margaret’s inability to concentrate during the long reunion speech could be attributed to her age (negative dispositional attribution), her lack of sleep the night before (neutral situational attribution), or the tedium everyone experiences in listening to a long speech (a neutral dispositional attribution).

Participants were asked to summarize the story in their own words as well as to describe Margaret. Two coders blind to condition read through and scored each participant’s responses on four dimensions: 1) was Margaret dependent or independent in her relationship with her daughter?, 2) is Margaret described as being imaginative or losing touch with reality (becoming senile)?, 3) is Margaret forgetful because she is getting old or because of situational factors?, and 4) is Margaret described overall positively or negatively? The inter-coder reliability was .68. These four scores were then combined to form an aggregate rating.

3. Results

3.1. Word Association

We ran a t-test on the word association scores from the two Age Conditions (old and young). We found a significant difference ($t[44] = 2.60$, $p = .01$) where participants in the old condition ($M = 7.67$, $SD = 1.72$) associated traits that were significantly more positive to the elderly than participants in the young condition ($M = 6.42$, $SD = 1.54$).

3.2. Indirect Attitudes

We ran a t-test on the overall bias score from the two Age Conditions. The effect was not significant ($t[45] = .30$, $p = .77$).

3.3. Ambiguous Story

We ran a t-test on the aggregate rating from the two Age Conditions. The effect was not significant ($t[31] = .77$, $p = .45$). See Table 1 for the detailed results of all our significance tests.

Table 1. Means and standard deviations for all our dependent measures.

	Young *M* (*SD*)	Old *M* (*SD*)	*t*	*p*
Word Association	6.42 (1.54)	7.67 (1.72)	2.60	.01
Indirect Attitudes	-1.13 (1.61)	-1.17 (2.06)	.30	.77
Ambiguous Story	10.10 (2.52)	9.47 (2.12)	.77	.45

4. Discussion

While our results did not show a consistent effect across all three dependent measures, it was extremely encouraging to find that such a short virtual interaction can change a person's negative stereotypes at all. We found a significant effect in the word association task, but not in the indirect attitudes or ambiguous story tasks. On the other hand, similar inconsistencies were found in Levy's [32] original study, which only demonstrated significant differences on one of several dependent measures. Thus, perhaps these other dependent measures were not particularly sensitive to the manipulations we chose. Nevertheless, it is encouraging that a brief immersion into the avatar of an elderly person has a significant effect on attitudes towards the elderly in general.

There were several limitations to the current study. First of all, the dependence on explicit attitudinal measures made it impossible to assess whether the manipulation had an effect on implicit stereotype activation. For example, the use of a lexical decision task (with words associated with the elderly) may be used in future studies to explore the effect of embodied perspective-taking on implicit stereotype activation. And secondly, it is difficult to minimize demand characteristics in the current experimental paradigm. During debriefing, we demonstrated that many participants in the old condition guessed the goal of the study. On the other hand, intervention methods (such as the contact hypothesis or thought suppression) are typically explicit as well. So while there are demand characteristics in both the narrative and embodied perspective-taking tasks, it may be the case that the increased amount of presence from the latter type of simulation is a more effective tool in reducing stereotypes. Furthermore, it is not clear how one might conceal the nature of the intervention in practice. After all, diversity training in school or business settings are never covert interventions.

Future work can expand on the current findings in several ways. First of all, as mentioned before, implicit measures should be included. The inclusion of implicit measures may also yield more consistent results, as lexical decision tasks are less susceptible to social desirability bias among other demand characteristics. Thus, these implicit measures may improve upon the inconsistent results apparent in both Levy's study and the current study. Secondly, it would be interesting to compare the effects of embodied perspective-taking with the effects of narrative perspective-taking. In other words, an estimation of the added benefit of using this intervention in immersive virtual reality. And finally, future work might explore how embodied perspective-taking impacts other stereotypes, such as race or gender. Previous work has demonstrated that VR is an effective tool for treating many types of psychological disorders (such as phobias or post-traumatic stress syndrome) [34-37]. In future work, we hope to add prejudice to this list.

Social inequality caused by stereotypes and prejudice is a problem that has no simple solutions. Perspective-taking has been shown to be a practical intervention method that decreases both implicit and explicit stereotyping. While the current study found inconsistent results for the use of embodied perspective-taking, the findings do suggest that this intervention in immersive virtual reality can have a positive effect on reducing negative stereotypes.

5. Appendix A

The following items were drawn from Palmore's original work [33] on the stereotypes people have of the elderly.

5.1. Positive bias items

1. Physical strength tends to decline in old age.
2. The five senses (sight, hearing, taste, touch, and smell) all tend to weaken in old age.
3. Old people usually take longer to learn something new.
4. Older people tend to react slower than younger people.
5. Lung vital capacity tends to decline in old age.
6. A person's height tends to decline in old age.

5.2. Negative bias items

1. About half of the people over the age of 65 in the US have Alzheimer's Disease.
2. The majority of people over the age of 65 reside in nursing homes.
3. The majority of old people lose interest in and capacity for sex.
4. About half of old people live below the poverty line.
5. Drivers over the age of 65 are more likely to get into car accidents than drivers under the age of 65.
6. Old people are more likely to victims of theft, murder, and burglary than people under the age of 65.

6. Appendix B

Because the “Margaret story” was not reproduced in Levy’s earlier paper [32], we requested a copy from her and she kindly shared her original stimulus. Since the story has not been reproduced in full before, we include it here for reference.

6.1. The Margaret story

Margaret had just moved in with her daughter, son-in-law and grandchildren (aged 7, 5, and 1). Margaret's daughter, Anne had been worried about her mother. Margaret had been living alone since her husband died soon after her 70 th birthday, which was three years ago. Anne convinced her mother to move into her nearby home where she could ask her mother to baby-sit for her grandchildren and where she could better help her mother with meals and household chores.

The week Margaret moved in with her daughter happened to be the week of Margaret's 50 th college reunion. During Margaret's first night in her daughter's home, she woke up many times because her one year old grandson had an earache and cried throughout the night. The next day, Margaret was supposed to attend the dinner to honor the reunion of her college class. Margaret's college roommate Essie had convinced Margaret to attend the reunion which was to be held at the nearby Hyatt Hotel. Margaret tried to walk to the hotel but could not recognize the local streets and got lost. She finally wandered back to her daughter's house. When she walked in, her grandchildren started to giggle at their disheveled grandmother. Margaret's daughter decided to give her mother a ride to the reunion dinner and go as her escort.

When they arrived, Margaret introduced her daughter to some of her college acquaintances but could not recall many of their names. During the salmon dinner, Margaret found it difficult to concentrate on the speeches by the long-winded members of her class. She noticed that one of the speakers resembled a poodle. Margaret's attention shifted downward. She saw a napkin that had fallen under the table and thought how it looked very much like a squirrel. She then looked over at her daughter and started to think about how she used to take her to a playground with lots of swings and a big green sandbox. She smiled at her daughter and commented half out loud, “That sandbox - how you loved it.” Anne looked at her mother with a concerned look and whispered, “Shhh! Be quiet mom!” After dinner, Anne drove her mother home.

Acknowledgements

The authors would like to thank Alexia Nielsen and Keith Avila for their assistance in running the experimental study, Andy Orin and Anna Jayaram for their assistance in coding the open-ended responses, as well as Becca Levy for sharing the “Margaret story” stimulus from her earlier study. The current work was partially supported by NSF grant 0527377 and funding from Stanford’s Media X.

References

1. Allport, G., *The nature of prejudice*. 1954, Reading, MA: Addison-Wesley.
2. Sherif, M., et al., *Intergroup cooperation and competition: The Robbers Cave experiment*. 1961, Norman, OK: University Book Exchange.
3. Tajfel, H., et al., *Social categorization and intergroup behavior.* European Journal of Social Psychology, 1971. **1**: p. 149-178.
4. Bargh, J.A., *Automaticity in social psychology*, in *Social psychology: Handbook of basic principles*, E.T. Higgins and A.W. Kruglanski, Editors. 1996, Guilford: New York. p. 169-183.
5. Devine, P.G., *Stereotypes and prejudice: Their automatic and controlled components.* Journal of Personality and Social Psychology, 1989. **56**: p. 5-18.
6. Levy, B. and E. Langer, *Aging free from negative stereotypes: Successful memory in China and among the American deaf.* Journal of Personality and Social Psychology, 1994. **66**: p. 989-997.
7. Steele, C.M., *A threat in the air: How stereotypes shape the intellectual identities and performance of women and African Americans.* American Psychologist, 1997. **52**: p. 613-629.
8. Steele, C.M. and J. Aronson, *Stereotype threat and the intellectual test performance of African Americans.* Journal of Personality and Social Psychology, 1995. **69**: p. 797-811.
9. Macrae, C.N., et al., *Out of mind but back in sight: Stereotypes on the rebound.* Journal of Personality and Social Psychology, 1994. **67**: p. 808-817.
10. Wegner, D.M., et al., *Paradoxical effects of thought suppression.* Journal of Personality and Social Psychology, 1987. **58**: p. 409-418.
11. Jones, E.E. and R.E. Nisbett, *The actor and the observer: Divergent perceptions of the causes of behavior.* 1971, Morristown, NJ: General Learning Press.
12. Storms, M.D., *Videotape in the attribution process: Reversing actors' and observers' points of view.* Journal of Personality and Social Psychology, 1973. **27**: p. 165-175.
13. Regan, D.T. and J. Totten, *Empathy and attribution: Turning observers into actors.* Journal of Personality and Social Psychology, 1975. **32**: p. 850-856.
14. Davis, M.H., et al., *Effect of perspective taking on the cognitive representation of persons: A merging of self and other.* Journal of Personality and Social Psychology, 1996. **70**: p. 713-726.
15. Galinsky, A.D. and G.B. Moskowitz, *Perspective-taking: Decreasing stereotype expression, stereotype accessibility, and in-group favoritism.* Journal of Personality and Social Psychology, 2000. **78**: p. 708-724.
16. Perdue, C.W. and M.B. Gurtman, *Evidence for the automaticity of ageism. Journal of Experimental Social Psychology.* Journal of Personality and Social Psychology, 1990. **26**: p. 199-216.
17. Greenwald, A.G., D.E. McGhee, and J.K.L. Schwartz, *Measuring individual differences in implicit cognition: The*

implicit association test. Journal of Personality and Social Psychology, 1998. **74**: p. 1464-1480.

18. Normand, V., et al., *The COVEN Project: Exploring Applicative, Technical, and Usage Dimensions of Collaborative Virtual Environment.* Presence: Teleoperators and Virtual Environments, 1999. **8**: p. 1999.
19. Slater, M., et al., *Small Group Behavior in a Virtual and Real Environment: A Comparative Study.* Presence: Teleoperators and Virtual Environments, 2000. **9**: p. 37-51.
20. Schroeder, R., ed. *The social life of avatars*. 2002, Springer-Verlag: London.
21. Blascovich, J., et al., *Immersive virtual environment technology as a methodological tool for social psychology.* Psychological Inquiry, 2002. **13**(2): p. 103-124.
22. Loomis, J., J. Blascovich, and A. Beall, *Immersive virtual environments as a basic research tool in psychology.* Behavior Research Methods, Instruments, and Computers, 1999. **31**: p. 557-564.
23. Bailenson, J., et al., *Transformed Social Interaction: Decoupling Representation from Behavior and Form in Collaborative Virtual Environments.* Presence, 2004. **13**(4): p. 428-441.
24. Bailenson, J., *Transformed Social Interaction in Collaborative Virtual Environments.*, in *Digital Media: Transformations in Human Communication*, P. Messaris and L. Humphreys, Editors. in press, Peter Lang: New York.
25. Yee, N. and J. Bailenson, *The Proteus Effect: The effect of transformed self-representation on behavior.* under review.
26. Lee, K.M., *Presence, explicated.* Communication Theory, 2004. **14**: p. 27-50.
27. Lombard, M. and T. Ditton, *At the heart of it all: The concept of presence.* Journal of Computer-Mediated Communication, 1997. **3**.
28. Biocca, F., C. Harms, and J. Burgoon, *Towards a more robust theory and measure of social presence: Review and suggested criteria.* Presence: Teleoperators and virtual environments., 2003. **12**: p. 156-480.
29. Slater, M., Usoh, M., & Steed, A., *Depth of presence in virtual environments.* Presence: Teleoperators and Virtual Environments, 1994. **3**: p. 130-144.
30. Bente, G., et al. *Measuring behavioral correlates of social presence in virtual encounters*. in *International Communication Association Conference, May 27–31*. 2004.
31. Ijsselsteijn, W., J. Freeman, and H. de Ridder, *Presence: Where are we?* Cyberpsychology and Behavior, 2001. **4**: p. 179-182.
32. Levy, B., *Improving memory in old age through implicit self-stereotyping.* Journal of Personality and Social Psychology, 1996. **71**: p. 1092-1107.
33. Palmore, E., *Facts on Aging Quiz: A handbook of uses and results.* 1988, New York: Springer.
34. Rizzo, A., et al., *Design and Development of a VR Therapy Application for Iraq War Veterans with PTSD.*, in *Technology and Informatics*, J. Westwood, Editor. 2005, OIS Press: Amsterdam.
35. Riva, G., B. Wiederhold, and E. Molinari, *Virtual Environments in Clinical Psychology and Neuroscience, Methods and Techniques in Advanced Patient-Therapist Interaction*. 1998, Amsterdam: IOS Press.
36. North, M., S. North, and J. Coble, *Virtual Reality Therapy: An Innovative Paradigm.* 1997, Colorado Springs: IPI Press.
37. Wiederhold, B. and M. Wiederhold, *Clinical observations during virtual reality therapy for specific phobias.* CyberPsychology & Behavior, 1999. **2**: p. 161-168.

CHAPTER 3

WALKING IN THE SHOES OF ANOTHER

We walked for days crossing the desert into Jordan. The week we left my kite got caught in a tree in my yard. I wonder if it is still there. I want it back. My name is Sidra, I am 12 years old. I am in the 5th grade. I am from Syria in the Daara Province, Inkhil city. I have lived here in the Za'atari camp for the last year and a half.[1]

These words mark the heartrending beginning of the virtual reality documentary *Clouds over Sidra*, an eight-and-a-half-minute-long 360 immersive film that takes viewers inside the Za'atari refugee camp in northern Jordan, home to over 80,000 Syrians displaced by civil war. In a voice-over you hear the young girl Sidra describing the camp, while different scenes of daily life in the sprawling complex play out around you. You see her with her family in the small converted shipping container in which they sleep. Next, you are in a makeshift gymnasium where men are lifting weights to pass the time. Then you are among a group of young men, laughing and talking as they bake flat bread in a large open oven. Then you are standing in the middle of a soccer field, while young girls kick a ball around you. All the while, through a translator, Sidra

matter-of-factly narrates scenes of normalcy amidst this ongoing humanitarian tragedy.[2]

The first time I saw *Clouds over Sidra* was at the Tribeca Film Festival in April 2015. I was struck by the sense of intimacy the film created between viewers and the film's subjects. In one scene, for example, children line up for school in one of the muddy streets within the camp. Temporary dwellings stretch out beyond sight in all directions. As I looked around, I could see and feel a vastness that the limited field of view offered by a photograph or filmed images simply cannot convey. Eighty thousand is just another heart-numbing statistic—an abstraction—until you are standing in the middle of this enormous makeshift city in the middle of the desert. Some of the children, clearly amused during the filming by the large, strange-looking camera apparatus that has been plunked down in the middle of the street, approach it, laughing and making funny faces, or just scrutinizing it. It felt like they were interacting with me. Many of these children had walked to Jordan through the desert, fleeing death and destruction, just as Sidra had, and the reality of that fact, combined with their playful normalcy, is intensely emotional.

When the experience was over, I noticed others around me wiping tears from their eyes as they removed the Samsung Gear headsets from their faces. The film was clearly unlocking powerful emotions. But these reactions hadn't been produced by a dramatic score, or clever editing, or close-up shots that lingered on a particularly poignant face or detail. These aspects of traditional filmmaking, designed to intensify our emotional engagement, are virtually absent in this immersive film. Viewers of *Clouds over Sidra* are simply confronted with a series of ordinary moments—people baking bread, families laughing together, children at play and in school. The major difference is that the immersive video made us briefly feel as if we were there with them.

Cocreated by artist and filmmaker Chris Milk, working with the support of Samsung and the United Nations, *Clouds over Sidra* is not just an early experiment in VR documentary, it is also an example of explicit VR advocacy. "We're taking these films," Milk said in 2015, "and we're showing them at the United Nations to people that work there and people that are visiting there. And we're showing them to the people that can actually change the lives of the people inside of the films."[3] This was from a talk titled "The Ultimate Empathy Machine," and in it Milk expressed his conviction that the immersive properties of VR make it particularly suited to sharing the experiences of others, to deepening our understanding of lives outside our own. VR, Milk has said, "connects humans to other humans in a profound way that I've never seen before in any other form of media. And it can change people's perception of each other. And that's how I think virtual reality has the potential to actually change the world."[4] Indeed Milk's predictions came to fruition; according to the United Nations, the VR experience has literally doubled the number of people who donate.[5]

Milk's work in VR has reflected his commitment to this belief. One of the premiere directors in this new medium, Milk has gone on to create several VR short documentaries, including *The Displaced* for the *New York Times Magazine*, which describes the experience of three children from Syria, Ukraine, and Sudan as they attempt to rebuild their lives in the wake of war. The *New York Times* distributed over a million Google Cardboards to subscribers to accompany this featured piece of VR journalism. With these short films, Milk and some other early VR content creators join a venerable tradition in the history of art, using their medium to encourage empathic understanding. A social novel like *Uncle Tom's Cabin,* which employed narrative storytelling in text and illustrations to impress upon northern audiences the suffering

of slaves in the years before the American Civil War, is just one of countless examples of political art. Others include Francisco Goya's *Disasters of War* series, which captured shocking images of violence and conflict in the early nineteenth century (although they weren't widely shared until decades later), or Jacob Riis's photographs of New York City slums published in his book *How the Other Half Lives*. For as long as humans have created art there has been a preoccupation with understanding and conveying the experience of others, particularly suffering—we see it in painting, sculpture, photography, film, and more recently even in certain videogames.

Whether all this empathic art has actually changed human behavior is a subject of ongoing debate. Some have pointed to statistics showing large declines in war and violence throughout the world and suggested that this is in part due to improvements in communication, which have broadened our innately tribal outlooks and contributed to an "expanding circle" of moral concern.[6] As a consequence, our instinct to favor our family or tribe expands to include previous outgroups. There is an impressive amount of data to broadly make this case, but also significant exceptions to this view that the power of our rational minds, informed by empathic experiences, is leading to undeniable progress for humanity. Some, for example, point to recent examples of societal fragmentation and terrible violence, as seen in Nazi Germany or in Rwanda in 1994. In both these cases, modern communications media were actually used for propaganda to demonize others.

Empathy has been the subject of much psychological research and debate in recent years, but many psychologists agree that two different systems, psychologically distinct, appear to be at work when we experience empathy. One is an emotional system which reacts reflexively when we encounter the suffering of others. It is

activated when you flinch upon viewing footage of an athlete getting injured, for instance, or when you turn away from a particularly gruesome scene in a horror film. The second major ingredient of empathy is cognitive, the ability of your brain to form theories about what other people are feeling and what might be causing those feelings. For some psychologists, the basic model of empathy ends there, but I don't think that goes quite far enough. For instance, psychopaths, con men, torturers—people in these categories might be very good at the emotional and cognitive aspects of empathy, but still lack the prosocial quality we usually associate with empathic people. For this reason I favor the conception of empathy put forward by my Stanford colleague Jamil Zaki, who adds a third component to what he calls "full-fledged empathy": a motivational element. Does a person, after experiencing the emotional and cognitive reactions to another's suffering, have a desire to alleviate the suffering of that person? Does a person expose herself to a trying empathic emotional experience in the first place?[7]

Zaki's empathy model is also illuminating because it highlights how empathy is something we can *choose* to experience, rather than something that just happens unconsciously. In Zaki's view, we turn empathy on and off because feeling it can be emotionally taxing—a drain on our mental resources. As an example, Zaki asks us to imagine a scenario in which we are watching television and we learn a leukemia telethon is about to come on, featuring children suffering from the disease telling their story. Few would doubt that this experience would create empathy in viewers. But there would be competing motives over whether to watch it. In some situations, you may be curious to learn more and continue watching. In others, you might want to avoid feelings of guilt, or just the sadness that hearing the stories might cause you.

Another example: most of us don't spend the mental energy

required to imagine what it is like to sleep out on the street every time we pass by a homeless person. If we did so, it would be difficult to function, given how traumatizing that understanding would be for most of us. Instead, we usually ignore the homeless person, or push the depressing reality of his existence further out of mind by constructing a story about the person that makes it easier to dismiss their suffering, perhaps by attributing their plight to choices that individual might have made. It is sad that this person is homeless, we might tell ourselves, but he shouldn't have used drugs or gotten fired from his job. We put aside the many reasons beyond that person's control that might have put him on the street—a financially crippling medical condition or mental illness, for example.

Choosing to experience empathy is not without costs. In intense situations it can cause real lasting pain and psychological damage. We see this in the damaging psychological effects suffered by certain professionals, such as doctors, nurses, psychotherapists, and emergency workers—people who confront extreme human suffering on a regular basis. Psychologists call this "compassion fatigue," the strain caused by pushing our empathy to its limits. It can lead to symptoms like anxiety, nightmares, dissociation, anger issues, and burnout.[8]

Still other occupations may require the suspension of empathy in certain aspects of the job, especially ones that are extremely competitive. Politicians, soldiers, and professional athletes all have good reasons to switch empathy on and off. And none of us are immune—we all live in a world with stunning inequality and many demands on our emotions. It is the rare person who can make it through the day without ever consciously putting on blinders to the suffering and inequality around us.

Empathy isn't a fixed quality. Our capacity for empathy can be

changed by our culture and the media technologies that transmit the culture's values—for good and bad. Certain types of training can desensitize people to other people's feelings, even as other types can enhance our empathy. Research has shown that one of the best ways to foster empathy is the psychological process of "perspective-taking," or imagining the world from another person's point of view. Perhaps the scholar most associated with this work is Mark Davis, a psychology professor who has done most of his teaching and research at Eckhard College in St. Petersburg, Florida. Davis created the most popular tool for measuring perspective-taking, a questionnaire called the "Interpersonal Reactivity Index," which asks people to rate their agreement with statements such as "Before criticizing somebody, I try to imagine how I would feel if I were in their place," or "When I watch a good movie, I can very easily put myself in the place of a leading character."[9] In 1996, Davis also published what is likely the canonical paper showing how perspective-taking actually works. The title of the paper tells it all: "Effect of perspective taking on the cognitive representation of persons: A merging of self and other."[10] His work over the past few decades has shown that if one imagines the world from another's perspective, the gap between oneself and the other decreases. Thinking similarly to another person literally causes changes in cognitive structures, such that one's thoughts concerning the other become more "selflike."

Consider a famous study from the year 2000 by Adam Galinsky, now a professor at Columbia Business School, showing that perspective-taking causes empathy. Experimental participants—college students—looked at a photograph of an elderly man and were asked to write an essay describing a day in his life. One group of the participants were given no explicit instructions; they were simply told to write about the man's day. The second group were

given specific "perspective-taking" instructions and were specifically urged to take the first-person perspective of the individual in the photograph when writing their narrative essay. Compared to the control condition, perspective-takers showed more empathy. They took longer to recognize prejudice phrases associated with negative stereotypes in a psychological reaction-time task, and also expressed more positive attitudes toward the elderly when writing their essays. But how many people will actually be able to use this imagination technique in their daily lives?[11]

Taking into account VR's ability to create seemingly real experiences and to allow users to look at a virtual world from many different points of view, one would imagine that VR would be particularly suited for perspective-taking. First, it relieves users of the cognitive effort required to make a mental model of another person's perspective from scratch. This might help a user overcome a motivational hurdle that causes her to avoid taking perspectives of others.

This leads to a second advantage of virtual reality in enhancing empathy. Because the mental model of the perspective of the empathic subject can be created in great detail in VR, it can be designed to help avoid stereotypes and false or comforting narratives. For example, if a teenager has a negative stereotype of the elderly, then merely asking the teenager to imagine an elderly perspective might simply serve to reinforce those stereotypes. The teen might create images of elderly who are slow, frugal, and who tell boring stories. But formalizing the role-playing into a set simulation that counteracts stereotypical representations can avoid these negative stereotypes by creating scenes that show the strengths of the elderly. People might not have the right information to actually take a proper perspective, and VR can guide them through the process more accurately.

No medium, of course, can fully capture the subjective experience of another person, but by richly evoking a real-seeming, first-person perspective, virtual reality does seem to promise to offer new, empathy-enhancing qualities. We can read news accounts or watch a documentary about refugees, but these media require a lot of imaginative work on our part. Narrative can give us a lot of information about life in the camp, but it is poor at conveying what living in a camp feels like. We don't have a mental library of the appropriate sights, sounds, and story to imagine what it is like to be a refugee. VR can convey the feeling of the camp's environment, the smallness of the living quarters, the size of the camp. It can bring Sidra, and the other members of the camp, alive in a way that documentary footage can't.

It seems that the world has taken as a given that VR can foster empathy in unprecedented ways. But does it actually work? What does the research tell us?

THE VIRTUAL MIRROR

My lab at Stanford has been running and publishing experiments on VR and empathy since 2003. We have looked at ageism, racism, helping the handicapped, and other domains. And what we have discovered is that the answer to the question of whether VR is, indeed, "the ultimate empathy machine" is nuanced. It's true that VR, across studies, tends to outperform control conditions. But it is not a magic bullet. It doesn't work every time, and the "effect size," the statistical term for how strong an outcome is, varies. The first study we ever ran, published in 2006 but designed and conducted starting in 2003, looked at ageism, the bias that young people tend to have toward the elderly. We were inspired by the Galinsky study in which people looked at a photograph of an

elderly man, and imagined what it was like to see the world from his eyes. But we decided to take imagination out of the equation.

First, we designed a "virtual mirror" that allowed a subject to see her avatar reflected on the wall of the virtual version of the lab room. We asked each subject to walk up to this virtual mirror, spend about 90 seconds gesturing in front of the mirror, and closely observe her "reflection" as it moved with her. She would roll her head left and right, touching her ears to her shoulders. In the virtual mirror, her reflection would do the same. Then she would take a step forward physically and watch her mirror image get bigger. After some time, we would ask the subject to crouch outside of the mirror's frame so she could no longer see herself, and then to pop back up and watch her reflection rise with her.

VR mirrors had been around since long before I came on the scene. In fact, a virtual mirror is one of the demos that sold me on VR in the late 1990s. I was visiting UCSB for the first time and experienced the scenario in Jack Loomis's lab. Jack is one of the pioneers of VR, and he had long been using VR to study distance judgment, vision, and other aspects of perceptual psychology. But Jack had an intellectual curiosity that extended far beyond the human visual system, and his VR demo suite reflected that. In 1999, the mirror was fairly low-tech—a user's reflection was a clunky, robotic avatar, and only rotated its head and moved back and forth without having any hand or leg movements. But it still created a powerful effect. With enough time in VR, that clunky body began to feel like your body. It was a glimpse into the future, and one of the first experiences I built when I arrived at Stanford was the mirror.

The principle at work was based on the "rubber hand illusion." This now classic experiment was first performed in the 1990s by two scientists at Princeton University who had their subjects sit

at a table, positioned so that one of the subject's hands was hidden out of sight, underneath the table. Instead of seeing his actual hand on top of the table the subject would instead see a lifelike rubber one where his actual hand should be. Neuroscientists then gently stroked both the hidden hand and the visible rubber hand simultaneously with a paintbrush. If the two hands were stroked synchronously and in the same direction, then the subject began to experience the rubber hand as his own.[12]

How was this determined? One example was that when asked to point to his hand, most of the time a subject would point to the rubber hand instead of the one hidden under the table. Interestingly, if the strokes are not timed well, then the illusion fails. But if the touch is synchronous, then the hand gets incorporated into one's body schema. For example, if the neuroscientist subsequently threatens the rubber hand by making a stabbing movement toward it with a needle, fMRI scans shows activity in the part of the subject's brain that is normally activated when a person anticipates pain, along with the area that is normally activated when a person feels the urge to move his or her arm. This landmark study using rubber hands has informed us how to induce body transfer into avatars. For example, if someone sees his avatar get lightly poked with a stick, and also physically feels his chest getting poked synchronously, the avatar is treated as the self. People "transfer" their consciousness into it, according to dozens of studies. This mirror technique is now a fairly commonplace one—neuroscientists today call it "body transfer."[13]

When one sees her avatar—whether it's from the first person looking down at her digital body, or by looking in the virtual mirror—and events occur synchronously, her mind takes ownership of the virtual body. The brain has very little experience—from an evolutionary standpoint—at seeing a perfect mirror image that

is not in fact a reflection of reality. Because we can digitally create any mirror image we want in VR, VR offers a unique, surreal way to allow people to take ownership of bodies that are not their own. Because the brain incorporates the mirror image into the self, the self is malleable. The mirror is the ultimate tool to allow someone to become someone else, to walk a mile in his shoes. One's mirror image can look just like her, or be bigger, become male, grow a third arm, or even change species.

In our study on ageism, we made the mirror image older.[14] First, we induced body transfer in a group of college students by having each of them perform a series of motions as he looked at his avatar in a virtual mirror. The avatar matched his movements, as you would expect a mirror image to do. However, half the college students were looking at an avatar that resembled them in age and gender. The other half saw an avatar of the same gender that was clearly in his 60s or 70s. Once we had induced ownership of their new body, the students then turned away from the mirror and met a confederate—a person networked into VR with them, who was part of the experiment. From the subject's point of view, there was simply another avatar—this one of college age—in the room with him, and the two avatars could talk and move about the room together. We very clearly reminded subjects that the image they saw in the virtual mirror was exactly how they would appear to others in the virtual environment. In other words, from the subject's point of view, the confederates would believe the subject *was* older and had no idea the avatars didn't reflect their actual, young age.

The confederate then asked the subject to step forward and answer some questions to spur some social interaction, saying things like, "Tell me a little about yourself" or "What makes you happy in life?" In order to highlight the social roles in the simula-

tion, it is critical to have a personal interaction. Finally, the subject was asked to perform a brief memory exercise of trying to repeat a list of fifteen words. The purpose of the memory exercise was to reinforce the fact that he was wearing an elderly avatar, as one of the more prevalent stereotypes of the elderly is bad memory. We found it particularly powerful when a college student has his memory questioned by another young avatar while wearing an elderly body. The idea is to induce mild reminders of the negative stereotypes so that the subject gets to "walk a mile" in the shoes of another.

In the study, the subjects who wore the elderly avatar used more positive words to describe the elderly in general, compared to those who wore a younger avatar. For example, they were more likely to list the word *wise* than *wrinkled* when pressed to list the first thing that came to their minds. We looked at three common measures of bias, but only found significant expressions of ageism in one of the three measures, the Word Association Task. This is a previously developed psychological technique during which subjects were asked the open-ended question, "When you think of somebody old, what are the first five words that come to mind?" Then, two coders who were not a part of the study rated the response as having a positive or negative connotation. The difference was about 20%. This was a modest but noticeable improvement. While our results did not show a consistent effect across all three dependent measures, it was still encouraging to find that such a short virtual interaction—less than 20 minutes—can change a person's negative stereotypes at all.[15]

A few years later, in 2009, working with a doctoral student named Victoria Groom, we used the virtual mirror to examine race.[16] Victoria thought that racial empathy would be induced when white participants wore black avatars. She reasoned that if a white person took the perspective of a black person, her negative

racial stereotypes would break down. Groom ran about one hundred participants, half of whom approached a virtual mirror in a black avatar and half in a white avatar. This time, the confederate performed a job interview, where the subject had to answer questions about aptitude and previous work experience.

As in the ageism study, we looked at three separate, previously validated measures of bias. And again, just as in the ageism study, only one of them—in this case the Implicit Association Task, which gauges reaction time to positive and negative concepts—showed a significant difference. However, we noticed that inhabiting a black avatar had the opposite effect from what we observed in the ageism study—it actually caused people to score *higher* on standard measures of implicit racial bias than those who wore a white avatar. In other words, wearing a black avatar primed more racial stereotypes instead of creating empathy. Stunningly, this pattern was true not only for white participants but for black participants as well.[17] Regarding virtual racism, it seems that the story is complicated, and that wearing a black avatar actually reinforced stereotypes and made them more salient.

In the fall of 2016, I had a conversation with renowned Harvard psychologist Mahzarin Banaji, who helped invent the Implicit Association Task and is likely the world's expert on implicit racial bias. She had read this study and wanted to discuss possible future interactions. Her feedback on the study was clear—she would have predicted the same outcome, given our procedure.

Social psychological research has demonstrated that indicators of a stereotype's social group such as physical features indicating gender or race can activate concepts relating to those social groups. Often these concepts are widely held stereotypes—in many cases negative—that can directly affect cognition, attitudes, and behavior. For example, the well-documented stereotype held

by some Americans that blacks are prone to violence has been demonstrated in a number of studies. This implicit bias happens automatically, without intention or awareness. One explanation for our findings is that we didn't succeed in encouraging perspective-taking. Given the limitations of the early technology, we were using a very simple hardware system that did not track arm movement and only allowed rudimentary body transfer. It is hard to be certain if body transfer had actually occurred. Perhaps we didn't induce transfer but instead simply primed the negative stereotypes that subjects already had about race, which is why we produced the counterproductive results.

But it's not all bad news.

Mel Slater and his colleagues in Barcelona ran a similar study a few years later. Slater has some of the best VR technology in the world, and is literally the world's expert at inducing body transfer into avatars. In his system, he tracked the body movements much more accurately and fully than we did in our early study, and consequently demonstrated with a high probability that body transfer was induced. In his study, being embodied in a black avatar *reduced* bias among white participants as measured by the Implicit Association Task, compared to control conditions. Slater and colleagues concluded that "embodiment may change negative interpersonal attitudes and thus represent a powerful tool for exploring such fundamental psychological and societal phenomena."[18]

Perhaps the strongest case of this powerful effect can be seen in the work of Sun Joo Ahn, a former graduate student of mine who now is a professor at the University of Georgia. In a paper published in 2013, Sun Joo conducted three experiments to explore whether becoming colorblind in VR would improve empathy toward those who are colorblind.[19] Subjects first received basic

information about red-green colorblindness. They then entered VR to do a sorting task designed to be particularly difficult if one could not differentiate perceptually between red and green. Half of the participants wore a headset with a colorblind filter applied over the objects on the screen, which allowed them to accurately experience being red-green colorblind. The other half also wore the HMD and did the sorting task but only imagined that they were colorblind.

In one of the experiments, after leaving VR, subjects who became colorblind spent almost twice as much time helping persons with colorblindness compared to participants who had only imagined being colorblind. The helping task involved participants working with a student group that was supposedly trying to build colorblind-friendly Web sites. The task required them to view screenshots of Web sites and to write about why the Web site might be inaccessible to colorblind individuals and how it might be improved. It was made clear to the participants that this activity was not a part of the experiment and that it would be volunteer work. We measured the time spent volunteering and showed that VR increased helping.[20]

Informal reactions from participants were telling. "I felt like a colorblind person. I felt like I was in a whole different world. It made me realize how tough it is for them to do certain things in life, such as drive." This points to a unique strength VR has in sharing with users some of the challenges people with physical disabilities face. It is a powerful experience to inhabit the avatar body of a minority and experience a scenario in which you are discriminated against, but a VR scenario can't hope to capture all the subtle aspects of discrimination a person experiences in her life. But it is easier to illustrate the difficulties a person with perceptual or physical disabilities might have, although again, it's important

not to overstate findings that have not yet been extensively replicated. Research has shown that we still must take care with how these experiences of disability are implemented. Experiments by Arielle Michal Silverman into empathy for the blind, for instance, have shown that the initial disorientation caused by making a seeing subject blind might actually promote discrimination instead of empathy, because the newly blind subject experiences only the trauma of sudden blindness, rather than the realities of being blind for a long time, and all of the acquired skills that go along with it. Instead of perceiving the blind as empowered, able people, a person in this study might only focus on the difficulties of coping with the sudden loss of sight.[21]

We also sent the colorblind participants a questionnaire designed to gauge their opinions toward colorblind people 24 hours after the study. We demonstrated that, for people who in general have a hard time feeling concern for others—that is, people who scored low on the Interpersonal Reactivity Index described above—becoming colorblind caused more favorable attitudes toward colorblind people than imagining it. This is preliminary evidence that VR is a great tool for those who tend to have a hard time being empathetic. For people who were "naturals," that is, those who in general tend to engage in empathy, there was no difference in attitudes 24 hours later. But for those who struggle with empathy, the virtual tool allowed them to improve their perspective-taking abilities.[22]

EMPATHY AT SCALE

In 2014 a program officer from the Robert Woods Johnson Foundation was on a lab tour and experienced some of our empathy demonstrations. I explained to her the general methodology of

the research—inducing body transfer followed by an experience. She was intrigued, but skeptical. We began a conversation about the robustness of the effects. As should be clear from reading this chapter, most of the time VR empathy seems to work inside the lab. But how strong are the effects? How long do they last before the treatment wears off?

The conversation evolved into a three-year project called "Empathy at Scale," where we sought to test the results of our studies and attempted to understand how well this tool works out in the real world. In the first series of experiments, published in 2016, we sought to discover the boundary conditions of VR empathy. Finding boundary conditions is a popular strategy in validating experimental findings; basically psychologists change the conditions as they reproduce experiments to discover when their effect stops replicating, looking for the "ceiling." Very few findings generalize to every single situation in life, and understanding when a tool no longer works is important information. In particular, we wanted to look at threat. Perspective-taking and empathy are most important, arguably, in situations where there is tension. I get a few calls per month from folks suggesting to use VR to bridge the gap between warring nations. In these situations, there is a preexisting and deep tension; psychologists call this "threat."

We returned to ageism and once again used the virtual mirror to embody our subjects as elderly. In the first study, subjects either imagined being older, similar to the Galinksy study described earlier, or became older in the mirror. In addition, we also varied threat, high or low. In the high threat condition, before entering VR, subjects read a passage called "Elderly Pose Immediate Threat to Young Americans." In the low threat condition, they read an article titled "America Prepared for Changing Demographics." Both were about the implications of the demographic trends of

people living longer, but one framed them as a threat, while the other framed them as a situation America was prepared to deal with. We then made the manipulations more impactful by asking subjects to write an essay to reflect upon the demographic trends and how they would affect their own lives. To sum up, there were four conditions. Half the subjects imagined being older, while half became older. Within each of those conditions, half were threatened by the elderly while half were not.

In this study VR succeeded in buffering threat. For those who imagined being elderly, being threatened caused them to have less empathy for the elderly. This is what we would predict; it is difficult to demonstrate empathy for those who might harm us. But for those who became elderly via body transfer in VR, the opposite occurred. They actually had more empathy when they were threatened compared to those who were not threatened. One possible explanation for these results is that it may have been easier for participants to take the perspective of an elderly person when they were in VR compared to when they were relying entirely on their imagination. Similar to the colorblind study, VR may be particularly beneficial when the intergroup context makes it difficult for people to engage in perspective-taking.

But now, enter the boundary effect. We ran a second study where we made the threat personal. Instead of having the subjects read a passage about the elderly, we created two elderly avatars and had the subjects interact with them in VR. We used a classic ostracism task developed by psychologist Kip Williams called Cyberball. In this exercise, three people are to play catch with a ball, and as the play progresses, one of them is left out—nobody throws him the ball.[23] It may sound trivial, but there have been dozens of experiments using this task and the feeling of being ignored is significant. Ostracism hurts—some studies have shown in an fMRI that

Cyberball ostracism causes the areas associated with pain to light up. Others simply show it makes people feel saddened.

In our study, we again had the same four conditions—imagination versus VR, and high and low threat. Regardless of condition, after the perspective-taking exercise, subjects then put on the headset and played a ball toss game with two partners, both of whom were clearly elderly. In the high threat condition, subjects received a total of 3 tosses out of 30 tosses. In the nonthreat condition, they received a total of 10 tosses, which is what one would expect in a game with equal sharing.[24]

Results demonstrated that the threat manipulation was about twice as strong in this study as in the first study, when subjects simply read about the elderly. Those who only received three throws felt angry and offended according to self-report measures. Also, in almost all of the measures of empathy, subjects who were threatened were later less empathetic to the elderly in general compared to those who were included in the ball-tossing game. But the critical data concern the effect of VR. In contrast to the promising results in Study 1, in the second study that had more intense, experiential, and intentional threat, there was no difference between VR and imagination. In other words, increasing immersion was not enough to overcome empathic avoidance. Rather, when under threat, participants consistently showed negative attitudes toward the elderly. Comparing the two studies, this preliminary evidence suggests that perspective-taking via VR can be effective at fostering positive behaviors toward the outgroup when the intergroup threat is indirect, but not as effective when the threat becomes more concrete and experiential.

All of the studies I've described in this section, like the vast majority of all psychological studies, have a major limitation. The elephant in the room is that we psychologists present findings

taken from a limited and narrow sample group and make conclusions as if they apply to all people. However, the statistics we perform dictate that we can only draw inferences specifically about the sample used in our observations. As a result, a majority of studies in the field of psychology only truly inform us about upper-class, college-educated people below the age of 22 who are taking an "Introduction to Psychology" course.

In truth, research findings that appear to be robust can crumble when taken to the general population, as the foundational work of Stanford psychologist Hazel Markus—who has made a spectacular career out of showing cultural differences in psychological processing—has demonstrated. For example, during graduate school, I worked with both psychologists and anthropologists to take a "bedrock" psychological finding out into the real world. We used a reasoning task that is taught in cognitive psychology textbooks across the world, specifically, how people rely on similarity-based notions of typicality and central tendency in order to perform natural categorization. Consider the following two arguments:

Robins have sesamoid bones, therefore all birds do.
Penguins have sesamoid bones, therefore all birds do.

Which argument is stronger? Undergraduates will tell you the first one, and most readers will likely agree. Robins are the more typical bird, so a trait from them is more likely to generalize to all birds. But in our research, the preference for typical arguments failed to replicate when looking at those with high domain knowledge—for example, birders from the Chicago area and tribal members of the Itza Maya in Mexico, who interact with birds more often than college undergraduates.

There is reason to suspect that empathy in particular will vary

according to demographic. Jamil Zaki recently developed a framework to understand motivation for empathy. He focuses in particular on individual differences. At least three motives—trying to avoid emotional suffering, incurring material costs such as having to donate money to help, and worrying about competition by performing weakly at work or in social situations—drive people to avoid empathy. Similarly, at least three motives—doing a good deed, which results in positive affect; strengthening affiliation with in-group members such as friends and family; and wanting others to see one as a good person, called "social desirability"—drive them to approach empathy. To truly understand how a VR empathy treatment works, it is important to have a large enough sample size to seek out populations that would likely vary in these traits.[25]

Consequently we have developed the "Empathy at Scale" project, with the goal of running 1,000 subjects through our study to see how well VR works at promoting empathy, compared to typical media techniques like using narrative and statistics. The 1,000 subjects are unique not just for the magnitude of the sample, but also in the diversity of the sample. We have installed VR systems on the road—our Mobile VR Unit—in museums, near libraries, and at festivals and fairs to try to get people who are not just the typical college students.

As of September 2017, across two studies, we have data from over 2,000 subjects. It's been an adventure, and we actually published an entire paper summarizing the lessons we have learned from VR field studies, ranging from how to accommodate those who can't stand up for long periods of time, to which direction the line to the tent should form, to how to alert people in a public place that we are running a study. One of the more interesting differences between college students and nonstudents is the amount of time it takes to

complete a questionnaire. Students race through them, while non-students spend more time reading and answering. After pilot testing the study in the lab with college students, we thought it would only take 25 minutes. But in the field with "real people," it takes about 45 minutes.

Our simulation is called "Becoming Homeless," and premiered as a VR film at the Tribeca Film Festival in April 2017. Users experience the transition from having a job and a home to losing them and becoming homeless. In the experience, you endure the cascading series of events that often lead people to become homeless. First you lose your job and are forced to sell your belongings. Unable to pay rent, you are evicted from your apartment and are forced to sleep in your car while you continue to look for work. Then, one night, sleeping in your car, you are roused by police and fined for violating a local ordinance. The car is sold, and you next find yourself harassed by a stranger while trying to get sleep on a bus. We designed the experience not only to be moving and compelling, but also interactive. For instance, you are forced to choose between which items you want to sell in your apartment—the sofa, the television, your phone? We ask the user to navigate the confining space of the small car he is living in to brush his teeth. On the bus, you must watch your bag and fend off strangers who approach you as you try to sleep.

With 900 subjects run, we have begun to look at the results. In general, VR is working better than the control conditions, which involve reading a narrative about becoming homeless or reading statistics about the homeless. When compared to these controls, people who have had the VR experience have shown more empathy on questionnaires, and are more likely to sign a petition to support affordable housing. But as in the other studies, the effect is not uniform across dependent variables. The effects are consis-

tent across about half of our measures but are by no means definitive. It seems that VR is the best medium of the four we chose, but the effect size is modest. Of course, we are continuing to analyze the data.

VR AND EMPATHY APPLICATIONS

In 2003, I arrived at Stanford as an assistant professor. One of the biggest priorities was to get my new lab funded, and being in Silicon Valley, I turned to industry in addition to the traditional routes of the National Science Foundation and other government sources. The first grant I got from Stanford was a small gift fund from Cisco to study social interaction in VR.

The executive from Cisco was a big thinker, a woman named Marcia Sitosky. Her tour of the lab caused a light bulb to go off after she saw herself in the virtual mirror and viewed our empathy research. She urged me to think about using VR for training in diversity. So much effort is spent on diversity training in corporations and other organizations, but for all the resources, the methods fall short. Could a VR experience be the answer?

In 2003, I was head down, tunnel-vision focused on publishing academic work, which meant I had to forgo many great opportunities to build applications. Between 2003 and 2010 (when I received tenure), I published dozens of experiments that used the virtual mirror. Most of them were centered around the Proteus Effect, namely, that when one wears an avatar, he implicitly *becomes* like that avatar. People in taller avatars negotiate more aggressively, people in attractive avatars speak more socially, and people in older avatars care more about the distant future. This theoretical work sought to understand the psychological mechanisms of how avatars changed the people that wore them.

But in the back of my mind, Marcia's suggestion resonated. While in the lab, she reflected on how corporate diversity training was far from perfect. While the general idea was promising—having people learn about workplace harassment with the goal of preventing it—as implemented it was not very effective.

Indeed, the research backs up Marcia's intuition. A 2013 paper published by Harvard sociologist Frank Dobbins examined the existing data on diversity training. Dobbins and his coauthor concluded that "initiatives designed to quash managerial bias, through diversity training, diversity performance evaluations, and bureaucratic rules, have been broadly ineffective."[26] My own experience resonates with this assessment. At Stanford, every eighteen months I am asked to participate in one of two training methods. I either watch a presentation by an acting troupe in a large group, or I do an online assessment—driving-school style—where I read about case studies and then get tested on the legality of various behaviors. Each is better than nothing, but in my opinion, neither is changing the way I think, other than perhaps giving me some good advice on how to pragmatically deal with harassment in the office.

In 2003, the last thing I wanted to do was build corporate training software. I needed to work 80-hour weeks to publish papers, and distractions like this were not allowed. But my life post-tenure was really about scaling our lab's work. Stanford is a university that encourages us to face outward and deliver our research out into the world.

Adam Silver, the National Basketball Association commissioner, visited the lab in 2015. He came, with the rest of the executives from the NBA, during a technology tour of Silicon Valley. His initial motivation was to investigate using VR to allow fans to watch games from their living room, as if they were courtside, or even better, as one of the players. I politely let him know why I

didn't think this was a great idea—I think two hours in an HMD is a bad idea. And we'll further discuss the trials and tribulations of how to show narrative and information via VR later. But what really grabbed the group—especially Eric Hutcherson, the head of human resources for the NBA—was diversity training. We spent a solid hour in a conference room discussing possibilities, and the group was extremely excited to move forward.

A few months later, Roger Goodell, the National Football League commissioner, came to the lab expressly to talk about diversity training. The NFL, like the NBA, brought most of its executive committee. The massive corporation known as the NFL was surveying Silicon Valley technology to find ways to improve the fan experience with their product and improve the league. One of the best reactions from that visit came from their chief information officer, Michelle McKenna-Doyle, who got to experience becoming an NFL quarterback by doing a dozen plays in our QB training simulator. When she came out, she had an interesting reaction. She argued that she had long felt that, conceptually, she knew as much if not more about the game than any of her male colleagues, but they could dismiss her insights because "she has never been on the field before." With VR, she felt that she'd had that experience, and she left the lab truly feeling she had a new understanding of the game.

Former NFL quarterback Trent Edwards, who is also a cofounder of STRIVR, has actually made an interesting point related to McKenna-Doyle's. He notes that allowing fans to see the QB's perspective might actually improve their understanding of what is happening on the field, and perhaps give them a better appreciation of how difficult the job is. Athletes are people too, and even though they draw huge salaries, the volume of hate mail and threats they receive could be reduced by perspective-taking.

Over the past few years we have built and tested a diversity training system to implement in the NFL. The biggest supporter within the organization has been Troy Vincent, a former NFL cornerback and current executive vice president of Football Operations. When we visited Troy in 2016 at the NFL headquarters to plan this project, one of the things I was most impressed with was his vision for how to tackle the problem. Instead of starting with the players, Troy suggested we start at the top of the organization, with league executives, owners, and head coaches. The way to change a corporate culture—and the NFL is a huge organization—is at the top, so we decided to begin there. We have built and tested with NFL personnel an interview simulation that is designed to train skills to avoid racial and sexual bias. Trainees get to relive the simulation multiple times to practice how to manage the inherent bias we all have. The idea is to leverage "repetitions" to give executives the tools to manage if not overcome their biases. The system debuted in February 2017, where scouts practiced interviewing draft prospects just before the NFL combine, and soon will have a larger role in the NFL. In an interview with *USA Today*, Vincent described his vision: "We'll start using this as another teaching tool later this year. We want to be known as the best place to work."

THE (FURTHER) WIDENING CIRCLE

The capacity of human empathy extends beyond other humans, of course. Many individuals prize animals as much as—if not more than—other people, and will lavish their pets with expensive food and health care. As we've seen the circle of empathy expand over the centuries, we've seen a growing number of species that were once casually abused or mistreated become deserving of moral concern, although for those who argue for animal rights there is

still a very long way to go. Perhaps the most significant amount of animal abuse occurs in our system of factory farming, an industry driven by our strong demand for cheap meat. It was concern over this system that led to one of the most intriguing experiments in VR empathy we performed in our lab, led by one of my students in 2013.

Joshua Bostick was one of those rare undergraduates who insisted on calling me "Professor Bailenson" instead of Jeremy. A Master's student, he had worked in the lab on and off for a few years as an undergraduate, and after learning the ropes, began to help out on research projects and in ushering public tours of the lab. During this time, Joshua transformed from a research assistant into a scholar. As I got to know him I noticed he was inspired, perhaps even fixated, on a single idea: putting humans into the avatar of cattle in order to reduce meat consumption and the environmental consequences caused by our high dietary demand for beef.

When he first made his pitch, I was taken aback by the boldness of the idea, and politely suggested he go back to the drawing board. Though I make an effort to cut down on my meat consumption, I still enjoy a burger from time to time, and I didn't want to look like a hypocrite. "So do I," he replied, and he told me he loved steak. Eventually he convinced me that a study that embodied humans in cow avatars would not be about trying to convert people to vegans, but instead would seek to reduce the amount of meat people eat—resulting in less energy use, deforestation, and carbon dioxide production associated with raising cattle. He reminded me of the fascinating work of Temple Grandin, who introduced humane reforms into cattle slaughterhouses that reduced the stress and panic of animals before they were killed, changes that she credited to her ability, perhaps enhanced by her autistic mind, to imagine the experience from the animals' perspective. Joshua's passion

and persistence paid off, and after a few meetings, he convinced me to green-light the experiment.

For the next few months we pushed the envelope on the lab's technology—one of the challenges of doing VR research is that we have to do a lot of programming and engineering just to set up the studies. We created a system that could track, in real time, the arm, leg, back, and head movements of people as they crawled in our lab on all fours. Participants would wear a head-mounted display that allowed them to see themselves in the third person (or third cow?) as they walked about. In the meantime, every millimeter that their arms and legs moved was tracked and transformed into their cow's gait. In order to pull off this feat we designed a special vest a person wore that had LED tracking lights attached to it, and also bought knee pads to prevent subjects from chafing on the lab carpet.

To "become" the cow, people saw the cow they were controlling through the head-mounted display and heard farm sounds from the speaker array that surrounded them. We also made it so they felt virtual touch. When the subject saw his cow avatar getting poked by a cattle prod, he felt that poke in his actual side accompanied by a shocking sound from the direction of the prod. The illusion was completed by a vibration from the ground, created by "buttkickers"—low frequency speakers hidden in the floors—that rounded out the simulation of getting shocked. The cattle prod was not just for effect—as we saw with the rubber hand illusion, the sensation of touch enhances body transfer. If a person *feels* his chest being poked, and *sees* a stick poking his avatar, this facilitates the mental transfer from self to avatar.

The experience we ended up with was persuasive and stunning. We put 50 college students into the cow simulator. They experienced a day in the life of a cow, drank from a virtual trough, ate virtual hay, and finally were prodded to a truck headed for the vir-

tual slaughterhouse. Compared to others in a control condition, who just watched the cows walk around and get shocked, those who became a cow gained more empathy for the plight of cattle using the most common questionnaire to measure self-reported empathy. But more telling than the formal analyses were the spontaneous quotes provided by the subjects. Here are a few of them: "Horrifying to be a cow and be poked by the prong," one participant wrote, "I was constantly on edge with what I would be forced to do next and what would happen when I ran out of things to do." Another remarked that it "made the horrible and sad lives of livestock animals seem more real and less theoretical than when I read about it." The body transfer technique clearly accomplished its goal, as pointed out by one student, who said, "I really did feel like I was the cow and I really didn't like being prodded." Even with the older technology—the study was run long before the much-improved consumer-grade headsets had arrived—the subjects experienced high presence: "It was surprisingly realistic. I felt as if I was in the place of the cow." It was clearly an intense experience.[27]

Our goal with this research was to discover ways for people to be better connected with where our meat comes from. Farming and slaughtering animals now happens far away from us, the meat we eat tidily packaged in ways designed not to remind us that what we are consuming was once a living, breathing animal. With our empathic imaginations stunted by this distance, it is no surprise that we overconsume and waste so much of the meat we produce. Joshua's experiment helped subjects imagine the animal's suffering and sacrifice.

While we were, in fact, advocating that people should consider eating meat a bit less often, we weren't trying to turn the United States into a Vegan Army, as some members of the conservative press suggested. Love the idea or hate it, becoming a virtual cow

struck a chord with the public. Journalists—for example, the BBC, the Associated Press, Fox News, Yahoo, *The Daily Mail*, and newspapers from around the world—continue to call the lab to cover the study, and for every cattle rancher who has protested the idea, five others have applauded it, including restaurant owners, government officials, educators, and parents.

We can see how VR can foster empathy by enhancing our experiences of traditional perspective-taking techniques, but Joshua Bostick's cow study also shows us how VR's ability to introduce surreal experiences, such as allowing us to inhabit the embodied perspective of another species, can influence attitudes. As the use of VR grows, and we understand better how to leverage its unique affordances, we can expect to see more unusual and novel applications that open our minds and hearts—not only to other people but to ourselves. This was shown in an unusual study done by Caroline Falconer at University College of London and Mel Slater at Barcelona's Institution for Research and Advanced Studies.

People who suffer from depression are often intensely self-critical, unable to give themselves the patience and understanding that they might grant to other people. Falconer and Slater set up a scenario to see if VR could increase "self-compassion," and their research has shown some promising results.[28]

The scenario they studied worked like this. A depressive patient was placed in VR, where they interacted compassionately with a virtual child. The words the patient spoke to the virtual child were recorded. Some of the patients then reentered the scenario, but this time as the child, hearing their own comforting words spoken back to them as if from another person. Others heard their words from

a nonembodied third-person perspective. While self-criticism decreased in both conditions based on scales designed to measure self-compassion, self-criticism, and fear of compassion, the study showed a "significantly greater increase in self-compassion" in the embodied condition.[29]

Typically, to foster self-compassion, therapists might encourage individuals to perform imaginative exercises, for instance, to think of how they might treat a friend who is going through a challenging experience, to write letters from the perspective of a friend to themselves, or to role-play as a criticizer and a person being criticized. In Falconer and Slater's elegant experiment, these more abstract exercises are turned into an immediate encounter that appears to amplify the traditional therapeutic technique. We will see this principle in effect in later chapters, which discuss the use of VR in a therapeutic setting. For now, I want to stick with the theme of perspective, and how changing one's point of view can have profound effects on her mental relationship to the world around her.