



# Beyond Speculation About the Ethics of Virtual Reality: The Need for Empirical Results

Mel Slater<sup>1,2\*</sup>

<sup>1</sup>Event Lab, Department of Clinical Psychology and Psychobiology, University of Barcelona, Barcelona, Spain, <sup>2</sup>Institute of Neurosciences of the University of Barcelona, University of Barcelona, Barcelona, Spain

**Keywords:** virtual reality, augmented reality, alternate realities, ethics, dystopian

## INTRODUCTION

Stanislaw Lem's 1974 novel *The Futurological Congress*<sup>1</sup> describes an apparently wonderful utopian world. The problem is that in reality the world is in a state of ruin but government addition of hallucinogens to the water supply and air generates the illusion in people that they are living in a paradise. As virtual and augmented reality (VR, AR) become mass products it is likely that an alternate "cyberspace" world will be created, a parallel world much as the Internet, except that we can choose to embody ourselves in it as a full and shared virtual reality that offers alternative life experiences. This is also the vision introduced in classic novels as William Gibson's 1984 *Neuromancer*<sup>2</sup> and Neal Stephenson's 1992 *Snow Crash*<sup>3</sup>, and many other so-called "cyberpunk" novels. Will these worlds be dystopian, reflecting the vision of Lem?

Even before it entered the mass market there were warnings that the future created by VR will be a dystopian one. For example the short movie "UNCANNY VALLEY" by 3DAR, written and directed by Federico Heller<sup>4</sup> depicts a situation much like Lem's novel. People living in ruins and hopelessness in reality spend much of their time in virtual reality to escape the disaster of their real lives (*do not read further until after watching the movie if you prefer to find out what happens yourself; some violence depicted*). They spend their time living only for VR, in which they play continual war games. Unknown to these players they are actually controlling remote robots, that are carrying out their actions in reality. A player killing a virtual character corresponds, in reality, with the robot that they embody killing a real person. At some point the movie steps out of virtual to real reality to show that the virtual foes are real people in a catastrophic war situation.

Could this happen? Certainly yes, it is almost possible now. For example, in (Kishore et al., 2016; Kishore et al., 2018; Aymerich-Franch et al., 2019) the actions of people in VR controlled remote robots in which they were embodied<sup>5</sup>, and executed *via* thought in (Cohen et al., 2014).

The fundamental lesson here is that this is an extreme example where people may carry out actions in alternate realities that they consider to be virtual, but actually there can be significant consequences in reality for situations far less extreme.

## OPEN ACCESS

### Edited by:

Michael Madary,  
University of the Pacific, United States

### Reviewed by:

Marco Fyfe Pietro Gillies,  
Goldsmiths University of London,  
United Kingdom  
Rosa M. Baños,  
University of Valencia, Spain

### \*Correspondence:

Mel Slater  
melslater@ub.edu

### Specialty section:

This article was submitted to  
Virtual Reality and Human Behaviour,  
a section of the journal  
Frontiers in Virtual Reality

**Received:** 29 March 2021

**Accepted:** 02 June 2021

**Published:** 12 August 2021

### Citation:

Slater M (2021) Beyond Speculation About the Ethics of Virtual Reality: The Need for Empirical Results. *Front. Virtual Real.* 2:687609. doi: 10.3389/frvir.2021.687609

<sup>1</sup>[https://en.wikipedia.org/wiki/The\\_Futurological\\_Congress](https://en.wikipedia.org/wiki/The_Futurological_Congress)

<sup>2</sup><https://williamgibsonbooks.com/books/neuromancer/>

<sup>3</sup><http://www.nealstephenson.com/snow-crash.html>

<sup>4</sup><https://www.youtube.com/watch?v=UXX0TRtg5Vk>

<sup>5</sup><https://www.bbc.com/news/av/technology-18017745>

## THE POWER OF IMMERSIVE SYSTEMS

The Declaration of Helsinki<sup>6</sup> is the cornerstone of modern ethical research practices. It deals with issues such as the rights and safety of research subjects, their “dignity, integrity, right to self-determination, privacy, and confidentiality of personal information,” and care for vulnerable groups. There must be careful assessment of risks, and attempts to minimise these. Subjects have the right to as much information as possible about the research study, and the right to be able to give informed consent. Of course there is a lot more and it is addressed in particular to medical research, but institutional ethics committees for research on human subjects take this declaration as fundamental. While this regulates research with human subjects, *it has no consequences for commercial activity*. Consumers of VR products are obviously not covered by the principles of the Declaration, since they are not subjects of research. Companies operating in the realm of immersive media sell their products for commercial reasons, and while they must adhere to the law, they are not bound by any particular ethical codes.

Why is there a particular ethical issue about VR beyond that which might be applied to media such as TV, film or even typical computer games that are played on a console or desktop computer? VR can generate at least four unique illusions that are not possible with other media. The first two illusions are under the umbrella of “presence” (Held and Durlach, 1992; Sheridan, 1992; Sanchez-Vives and Slater, 2005). Presence has been decomposed into two dimensions—“Place Illusion” (PI) the illusion of being in the virtual world and “Plausibility” (Psi), the illusion that events are actually occurring (Slater, 2009). When both PI and Psi operate then people tend to respond realistically to situations and events in the VR, even though they know for sure that these are illusions and not reality.

The third illusion is “body ownership”. When participants wear a wide field-of-view head-tracked stereo head-mounted display and they look down towards themselves, they will see a life-sized virtual body substituting their own, from their first person perspective (1PP) (if this has been programmed). Utilising real-time body tracking, as the person moves their virtual body can be programmed to move synchronously and in correspondence with their own movements. If something is seen to touch their virtual body, then it can be arranged that a corresponding tactile stimulation is applied synchronously to their own real body. We refer to this as *embodiment*, which involves integration of vision and touch or motor activity of the 1PP view of the body which will typically lead to the illusion that the virtual body is their own—even though they know for sure that it is not.

The fourth illusion is a corollary of the previous three, referred to here as co-presence. This requires PI and Psi as necessary conditions, but the domain is interaction between remotely located people. When several participants are in the same virtual environment simultaneously, co-presence is the illusion of being in the same space as and directly interacting with the other participants.

Assuming that each participant is embodied in a virtual body, they can interact with another. PI is needed for the illusion of being in the same space as the others. Psi is needed for the representations of the actions of the others to be taken as valid human actions. Embodiment is needed for each person to be represented in the VR—see (Barberia et al., 2018) for a complex example.

Overall, it is clear that VR can have very powerful effects on participants. Evidence of this is that one of the most extensive research areas over the past 3 decades has been its application in clinical psychology - Freeman et al. (2017) provide a review and meta-study.

Ethical issues therefore centre around the issues of presence (PI and Psi), body ownership where embodiment in a virtual body can lead to various changes, and co-presence the illusion of being present and embodied alongside others. Several authors have considered ethical issues concerned with the use of VR or Augmented Reality (AR). A common aspect amongst all of the articles is that they are wholly speculative, i.e., based on opinion rather than evidence, recent ones being (Madary and Metzinger, 2016; Slater et al., 2020; Snijders et al., 2021).

## ADDRESSING ETHICAL ISSUES WITH EMPIRICAL STUDIES

In (Slater et al., 2020) we extensively considered many of the ethical issues that could arise from the widespread adoption of immersive systems. The following outlines a subset that could be addressed by experimental studies, to investigate whether they might constitute real problems.

### Attitudinal and Behavioural Change

Embodiment can lead to physiological, behavioural, attitudinal and cognitive changes in the participant as a function of the type of body. For example, multiple replications have shown that embodying “White” people in a “Black” virtual body will lead to sustained reduction in their implicit racial bias towards Black people (Peck et al., 2013; Maister et al., 2015; Banakou et al., 2016; Hasler et al., 2017; Banakou et al., 2020). Although embodied applications investigated to date have been for what would generally be regarded as beneficial to the individual and society (e.g., against racial bias) there is the possibility that the same technique might be used for harmful applications. The empirical question is whether it is possible to change neutral or implicit positive attitudes and behaviours of people towards an arbitrarily chosen group to become negative using embodiment. However, a challenge is that any study to test whether a bias can be deliberately introduced or reversed must be carried out in a way that it does not produce a new prejudice against real groups of people. There have been unexpected findings suggesting that if the social events in the embodied situation produce negative affect then implicit bias may increase (Groom et al., 2009; Banakou et al., 2020). Moreover, previous work has shown that high levels of body ownership can be reached over virtual bodies that are humanoid, but not human. For example, Peck et al. (2013) included a condition of embodiment in a purple virtual body, and Barberia et al. (2018) in humanoid avatars of red, green or blue colours with alien characteristics. An

<sup>6</sup><https://www.wma.net/policies-post/wma-declaration-of-helsinki-ethical-principles-for-medical-research-involving-human-subjects/>

interesting empirical study would be to first give participants an article to read that depicts the likely features of aliens who may inhabit the planet Proxima b (an inhabitable planet orbiting Proxima Centauri). Then a specially designed implicit association test (IAT) (Greenwald et al., 2003) would be administered to test fear of potential alien civilisations. At least 1 week later participants would be invited back to be embodied in an alien virtual body with an appearance based on the article they had read. One embodied group would be given a positive welcome reception by a virtual human crowd, and another group a hostile reception. Another week after that the IAT would be administered again. The prediction is that those who had the hostile reception would show an increased fear of extra-terrestrial aliens, and those with the positive reception would show either no change or a decrease. If this happens it would be evidence that indeed embodiment can produce prejudices where none existed, even with respect to an invented group (For ethical reasons, after data were collected, it would be prudent to expose those in the hostile group to another session where they received the positive reception). This would have important practical implications for embodiment employed for example in VR-based games, where embodiment in a body epitomising a recognisable social or racial group in stressful or negative experience should be avoided.

## Exiting VR Could Result in Negative Feelings

Leaving a VR session may be problematic in some circumstances where individuals had spent considerable time in a virtual fantasy world with an enhanced virtual body, and then exit to the less captivating real world. Counter to this is the anecdotal report by Lanier (2017) that in the early days of VR people, after spending some time immersed, would take off the head-mounted display and be pleasantly impressed by the vividness of physical reality - although perhaps this is no longer pertinent given the enhanced level of realism today. Following the movie *Avatar* (2009, directed by James Cameron) there were reports of viewers becoming distressed at their loss of the beauty and harmony of Pandora (Falquina, 2014). Note that the movie was shown in stereo 3D and thus had a level of immersion greater than typical 2D presentations. As another example, our own study involved groups of three people unknown to each other, embodied in alien humanoid bodies, interacting together in a beautiful island each day for 6 days, only to see their collaborators “die” one by one and eventually themselves have an out-of-body and near-death-experience (Barberia et al., 2018). When later all the participants were brought together for a physical meeting, many expressed regret that the experience was over. So it is certainly possible that participants might suffer on “return” to reality after some time in VR. This also connects with another possible negative effect where long term and frequent use might lead to people prioritizing the virtual world over the real one. Notice that all the evidence here is anecdotal. An empirical study would need to show a change in perception of, and attitudes towards reality after exposure in VR, although in a way that does not itself induce negative feelings in reality. A safe way to do this would be to immerse people in a rather poor virtual world, unattractive, for example, with dull colours,

uninteresting, empty of others, and so on. Here the test would be whether on exiting after considerable time in such an environment they would be elated by the richness of physical reality. This would be a doable and ethically safe experiment, and show whether or not immersion in VR can, under certain circumstances, lead to changed perceptions of reality. How great an ethical problem this is for VR remains a question for discussion: reading a book can also change perceptions of reality.

## Separate Realities

There is another critical problem that may be exacerbated through virtual and augmented reality. People are increasingly living in their own realities tailored to them individually through social media. A society functions through some level of shared reality. Of course there can be conflict and disputes, but today the very notion of “facts” is in question—with some politicians and commentators considering their “facts” to have the same status as anyone else’s “facts” even if they knowingly invented them. This is occurring through manipulation of speech, text, images and video. Imagine that it were also propagated through *experience*. VR and AR deliver actual first-hand experiences to people (Chalmers, 2017) so possibly social divisions may become solidified and enhanced not only through text, images and videos, but through actual lived experiences. With deep fake technology it is possible to give people experiences that would, for example, apparently confirm conspiracy theories as factual through direct first person experience. An experiment to study this would not be difficult but would be on ethically precarious grounds. Immersing someone in a VR where they are confronted in a personal conversation with, for example, a famous politician is technically feasible. The politician could be observed to engage in nefarious activities or argue convincingly for a conspiracy theory. How much would this influence the views of the person? A critical question in order to be able to run such an experiment in an ethical way would be whether a second exposure could undo any damage that might have been caused by the first, if this time the virtual politician undermines in a humorous way everything said and done in the first session.

## CONCLUSION

VR and AR are amazing tools to enrich experience, provide fantastic means to enhance education, training, therapy, surgery, learning, and entertainment (Slater and Sanchez-Vives, 2016). But just as any technology, they can be used for evil or good. The problem has to be tackled in a multidisciplinary way—ranging from engineering (including AI and Machine Learning) through to psychology and philosophy (ethics). Since immersive systems produce lived experiences, even though they are simulations, the range of disciplines is very broad.

However, as mentioned above, a critical problem is the *lack of data*; prognoses are based on speculation rather than empirical findings. So it is important to emphasise the need for experimental studies that tackle issues such as the extent to which anti-social changes in their attitudes and behaviours can actually be deliberately caused through unscrupulous use of the technology, whether people are disturbed in reality after a prolonged session in

VR, and whether they can be induced into separate realities. A major goal of research in the area of ethical considerations of immersive media should be *to provide experimental evidence that would contribute to our understanding of some of these ethical issues*, so helping to make immersive media safer, while *demonstrating that the experimental studies themselves stay within the bounds of what is ethically acceptable*.

## AUTHOR CONTRIBUTIONS

The author confirms being the sole contributor of this work and has approved it for publication.

## REFERENCES

- Aymerich-Franch, L., Kishore, S., and Slater, M. (2019). When Your Robot Avatar Misbehaves You Are Likely to Apologize: an Exploration of Guilt during Robot Embodiment. *Int. J. Soc. Robotics* 12, 217–226. doi:10.1007/s12369-019-00556-5
- Banakou, D., Beacco, A., Neyret, S., Blasco-Oliver, M., Seinfeld, S., and Slater, M. (2020). Virtual Body Ownership and its Consequences for Implicit Racial Bias Are Dependent on Social Context. *R. Soc. Open Sci.* 7, 201848. doi:10.1098/rsos.201848
- Banakou, D., Hanumanthu, P. D., and Slater, M. (2016). Virtual Embodiment of White People in a Black Virtual Body Leads to a Sustained Reduction in Their Implicit Racial Bias. *Front. Hum. Neurosci.* 10, 601. doi:10.3389/fnhum.2016.00601
- Barberia, I., Oliva, R., Bourdin, P., and Slater, M. (2018). Virtual Mortality and Near-Death Experience after a Prolonged Exposure in a Shared Virtual Reality May lead to Positive Life-Attitude Changes. *PLoS one* 13, e0203358. doi:10.1371/journal.pone.0203358
- Chalmers, D. J. (2017). The Virtual and the Real. *Disputatio* 9, 309–352. doi:10.1515/disp-2017-0009
- Cohen, O., Koppel, M., Malach, R., and Friedman, D. (2014). Controlling an Avatar by Thought Using Real-Time fMRI. *J. Neural Eng.* 11, 035006. doi:10.1088/1741-2560/11/3/035006
- Falquina, S. (2014). "The Pandora Effect:" James Cameron's "Avatar" and a Trauma Studies Perspective. *Atlantis* 36, 115–131.
- Freeman, D., Reeve, S., Robinson, A., Ehlers, A., Clark, D., Spanlang, B., et al. (2017). Virtual Reality in the Assessment, Understanding, and Treatment of Mental Health Disorders. *Psychol. Med.* 47, 2393–2400. doi:10.1017/S003329171700040X
- Greenwald, A. G., Nosek, B. A., and Banaji, M. R. (2003). Understanding and Using the Implicit Association Test: I. An Improved Scoring Algorithm. *J. Personal. Soc. Psychol.* 85, 197–216. doi:10.1037/0022-3514.85.2.197
- Groom, V., Bailenson, J. N., and Nass, C. (2009). The Influence of Racial Embodiment on Racial Bias in Immersive Virtual Environments. *Social Influence* 4, 231–248. doi:10.1080/15534510802643750
- Hasler, B., Spanlang, B., and Slater, M. (2017). Virtual Race Transformation Reverses Racial In-Group Bias. *PLOS ONE* 12, e0174965. doi:10.1371/journal.pone.0174965
- Held, R. M., and Durlach, N. I. (1992). Telepresence. *Presence: Teleoperators & Virtual Environments* 1, 109–112. doi:10.1162/pres.1992.1.1.109
- Kishore, S., Muncunill, X. N., Bourdin, P., Or-Berkers, K., Friedman, D., and Slater, M. (2016). Multi-destination Beaming: Apparently Being in Three Places at once through Robotic and Virtual Embodiment. *Front. Robotics AI* 3, 65. doi:10.3389/frobt.2016.00065
- Kishore, S., Navarro, X., Dominguez, E., De La Peña, N., and Slater, M. (2018). Beaming into the News: A System for and Case Study of Tele-Immersive Journalism. *IEEE Comput. Grap. Appl.* 38, 89–101. doi:10.1109/mcg.2017.2801407

## FUNDING

MS is funded by the European Research Council Advanced Grant 'Moments in Time in Immersive Virtual Environments' (MoTIVE) #742989.

## ACKNOWLEDGMENTS

This work is based on a contribution to the research scope of the European Commission's forthcoming work on Meaningful and Ethical Communication, Directorate H - Knowledge Management, Knowledge for Policy: Concepts and Methods.

- Lanier, J. (2017). *Dawn of the New Everything: A Journey through Virtual Reality*. London: The Bodley Head.
- Madary, M., and Metzinger, T. (2016). Real Virtuality: A Code of Ethical Conduct Recommendations for Good Scientific Practice and the Consumers of VR-Technology. *Front. Robotics AI* 3, 3. doi:10.3389/frobt.2016.00003
- Maister, L., Slater, M., Sanchez-Vives, M. V., and Tsakiris, M. (2015). Changing Bodies Changes Minds: Owning Another Body Affects Social Cognition. *Trends Cogn. Sci.* 19, 6–12. doi:10.1016/j.tics.2014.11.001
- Peck, T. C., Seinfeld, S., Aglioti, S. M., and Slater, M. (2013). Putting Yourself in the Skin of a Black Avatar Reduces Implicit Racial Bias. *Conscious. Cogn.* 22, 779–787. doi:10.1016/j.concog.2013.04.016
- Sanchez-Vives, M. V., and Slater, M. (2005). From Presence to Consciousness through Virtual Reality. *Nat. Rev. Neurosci.* 6, 332–339. doi:10.1038/nrn1651
- Sheridan, T. B. (1992). Musings on Telepresence and Virtual Presence. *Presence: Teleoperators & Virtual Environments* 1, 120–126. doi:10.1162/pres.1992.1.1.120
- Slater, M., Gonzalez-Liencreas, C., Haggard, P., Vinkers, C., Gregory-Clarke, R., Jelley, S., et al. (2020). The Ethics of Realism in Virtual and Augmented Reality. *Front. Virtual Reality* 1. doi:10.3389/frvir.2020.00001
- Slater, M. (2009). Place Illusion and Plausibility Can lead to Realistic Behaviour in Immersive Virtual Environments. *Phil. Trans. R. Soc. B* 364, 3549–3557. doi:10.1098/rstb.2009.0138
- Slater, M., and Sanchez-Vives, M. V. (2016). Enhancing Our Lives with Immersive Virtual Reality. *Front. Robotics AI* 3, 74. doi:10.3389/frobt.2016.00074
- Snijders, D., Masson, E., Doesborgh, S., Groothuizen, R., and Van Est, R. 2021. Fake for Real. Ethical and Societal Implications of Augmented Reality. Available at: <https://www.rathenau.nl/en/digital-society/fake-real>. doi:10.17487/rfc9003

**Conflict of Interest:** The author participated in the EU committee on Meaningful and Ethical Communication, Directorate H - Knowledge Management, Knowledge for Policy: Concepts and Methods. This paper is based on an earlier version written as part of the work for this group, for which MS received compensation. The committee was not involved in the formulation or writing of this paper.

**Publisher's Note:** All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Slater. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.