



# The Ethics of Realism in Virtual and Augmented Reality

Mel Slater<sup>1,2,3\*</sup>, Cristina Gonzalez-Liencrees<sup>1</sup>, Patrick Haggard<sup>4</sup>, Charlotte Vinkers<sup>5</sup>, Rebecca Gregory-Clarke<sup>1†</sup>, Steve Jelley<sup>6</sup>, Zillah Watson<sup>7</sup>, Graham Breen<sup>8</sup>, Raz Schwarz<sup>9</sup>, William Steptoe<sup>9</sup>, Dalila Szostak<sup>10</sup>, Shivashankar Halan<sup>11</sup>, Deborah Fox<sup>12</sup> and Jeremy Silver<sup>1</sup>

<sup>1</sup> Digital Catapult, London, United Kingdom, <sup>2</sup> Event Lab, Department of Clinical Psychology and Psychobiology, University of Barcelona, Barcelona, Spain, <sup>3</sup> Institute of Neurosciences, University of Barcelona, Barcelona, Spain, <sup>4</sup> Institute of Cognitive Neuroscience, University College London, London, United Kingdom, <sup>5</sup> Magic Leap, Plantation, FL, United States, <sup>6</sup> Dimension – Hammerhead VR, Wimbledon, United Kingdom, <sup>7</sup> BBC, London, United Kingdom, <sup>8</sup> HTC Vive, Slough, United Kingdom, <sup>9</sup> Facebook AR/VR, London, United Kingdom, <sup>10</sup> Jigsaw, New York, NY, United States, <sup>11</sup> Facebook AR/VR, Menlo Park, CA, United States, <sup>12</sup> Nesta, London, United Kingdom

## OPEN ACCESS

### Edited by:

Xueni Pan,  
Goldsmiths University of London,  
United Kingdom

### Reviewed by:

Sylvia Terbeck,  
University of Plymouth,  
United Kingdom  
Eugene Ch'ng,  
The University of Nottingham Ningbo  
China, China

### \*Correspondence:

Mel Slater  
melslater@ub.edu

### † Present address:

Rebecca Gregory-Clarke,  
StoryFutures Academy: The National  
Centre for Immersive Storytelling,  
London, United Kingdom

### Specialty section:

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

**Received:** 19 November 2019

**Accepted:** 11 February 2020

**Published:** 03 March 2020

### Citation:

Slater M, Gonzalez-Liencrees C,  
Haggard P, Vinkers C,  
Gregory-Clarke R, Jelley S, Watson Z,  
Breen G, Schwarz R, Steptoe W,  
Szostak D, Halan S, Fox D and Silver J  
(2020) The Ethics of Realism in Virtual  
and Augmented Reality.  
*Front. Virtual Real.* 1:1.  
doi: 10.3389/frvir.2020.00001

**Keywords:** virtual reality, augmented reality, ethics, realism, VR, AR, XR

“That which is hateful to you, do not do to your fellow. That is the whole Law; the rest is the explanation; go and learn it.” (The Golden Rule of Reciprocity, Negative Form, Hillel).

## INTRODUCTION

The golden rule of reciprocity (“treat others as you would have them treat you”) is present in most philosophical traditions and religions, and can be thought of as a fundamental human moral imperative. The first and most positive aspect of virtual reality (VR) is that it is possible to give people the experience of the golden rule in operation. For example, VR can place people virtually in the body of another, such that an “ingroup” member can temporarily occupy the body and position of an “outgroup” member: a person with pale skin can temporarily have dark skin (Maister et al., 2013, 2015; Peck et al., 2013; Banakou et al., 2016) or vice versa, an adult can become a child (Banakou et al., 2013; Tajadura-Jiménez et al., 2017), and someone can experience a world where they are taller or shorter than their real height (Yee and Bailenson, 2007; Freeman et al., 2013).

Besides changing bodies, VR enables one to have a myriad of possible experiences from a first-person perspective. One can for instance be exposed to a virtual representation of a phobic agent (for example spiders), the participant *knowing* it is not real but *feeling* it as if it were. Thanks to this, VR has become increasingly used for therapeutic purposes including pain management (Matamala-Gomez et al., 2019) and treatment of phobias and anxiety disorders (Freeman et al., 2017). The therapeutic potential in other realms has already been experimentally tested, such as for physical rehabilitation, for example, (Levin et al., 2015), for the rehabilitation of violent offenders (Seinfeld et al., 2018), and for the assessment of symptoms and neurocognitive deficits in people experiencing or at risk of psychosis (Rus-Calafell et al., 2018). Its use for training purposes in several areas including military, medicine, surgery, and disaster response, among others, is also gaining popularity (Spiegel, 2018; Vehtari et al., 2019). All these advantages rely on the extent to which the experience is perceived as real. It is reasonable to imagine that more realism in these VR scenarios increases their effectiveness.

In augmented reality (AR), virtual features are added to the real environment through some sort of device (for example goggles or a smartphone) and the information presented often requires the actual location of the user. For instance, when visiting some ruins, one could see a depiction of what the site used to look like superimposed over the remains. This is useful not only for historical representations but also for educational purposes (for example architects and engineers). With AR, one can also visualize a product before purchasing it—even try it on virtually—or see

relevant information on the car windshield. AR also offers a huge value for companies that employ it for marketing aims. Similar to VR, augmenting the realism of AR technology is likely to boost its impact.

In addition, VR and AR (XR) systems can be employed for data visualization, for industrial design in architecture and urban planning and, naturally, for entertainment—the gaming industry has enormous potential in this field (Brey, 1999, 2008; Wassom, 2014). Moreover, it is commonplace today to be able to have a conversation in a virtual (VR) or real (AR) space with another person who is physically somewhere else but whose virtual representation is in that same space, which eventually might reduce the need to travel for meetings.

Despite all the benefits, however, XR technology also raises a host of interesting and important ethical questions of which readers should be aware. For instance, the fact that XR enables an individual to interact with virtual characters poses the question of whether the golden rule of reciprocity should apply to fictional virtual characters and, with the development of tools that allow for more realism, whether this should also extend to virtual representations of real people.

Thus, along these lines, is it wrong to do immoral acts in VR? This is explored in a play called “The Nether” (2013) by Jennifer Hayley<sup>1</sup>, where in a fully immersive virtual world a man engages in pedophilia. When confronted by the police in reality (in the play), he argues that this is a safe way to realize his unacceptable drives without harming anyone at all. As stated by Giles Fraser writing in *The Guardian* newspaper<sup>2</sup>, “Even by watching and applauding the production I felt somehow complicit in, or at least too much in the company of, what was being imagined. Some thoughts one shouldn’t think. Some ideas ought to be banished from one’s head.” But on the other hand, “Policing the imagination is the ultimate fascism. Take Orwell’s *Nineteen Eighty-Four*. But the point is surely this: imagination is not cut off from consequence. We all end up being shaped by what we imagine.”

The latter point was part of an argument by Brey (1999), who considered ethical issues associated with virtual reality. Following Kantian Duty Ethics (a version of the golden rule), he argued that it is a fundamental moral principle “that human beings have a duty to treat other persons with respect, that is, to treat them as ends and not as means, or to do to them as one would expect to be treated by others oneself.” But does this apply to virtual characters? He gave two arguments suggesting that it does. First, following Kant in relation to treatment of animals, we should treat virtual characters with respect because if not we may end up treating people badly too (note that this is a philosophical rather than an empirical argument). Second, if we treat virtual characters with disrespect or act violently toward them, this may actually cause psychological harm to people that those characters might represent. Of course, this happens in movies all the time (think of the “bad guys” in movies, they are often typified as members of particular ethnic groups or social

class). In XR this is different though—in movies it is other people who treat other people badly whereas in XR it could be ourselves doing so, or other (virtual or online) people may treat us badly. While this already takes place in video-games, particularly when the character in the video-game is seen and controlled from a first-person perspective, XR goes one step further in the sense that it can feel more real if the participant is fully embodied as that character. Therefore, Brey concludes that designers of VR applications—also applicable to AR—must take into account the possible immoral actions that they might depict or allow their participants to carry out.

It should be noted that causing harm in itself may not always be objectionable. For example, there has been considerable discussion in law about whether consensual harm, where a perpetrator claims that the victim agreed to the harm, can be exonerating (Bergelson, 2007). As another example, children may be required by law to be vaccinated against an illness, for the greater good, even if the parents consider this to be potentially harmful to the children. The utilitarian philosophy of choosing actions that maximize happiness and minimize pain for the greatest number can also justify the causing of harm, again for the greater good. There is, however, also research suggesting that moral judgements may depend not on the outcome but the action involved in achieving the outcome. For example, in the famous trolley problem (Thomson, 1985) a runaway trolley car on a track is about to kill 5 unaware people, but could be diverted onto another track where it would kill just 1 person, thus saving the 5. Utilitarianism would suggest that diverting the trolley is the right action, even though it involves harming the one person. However, people find personally pushing a “heavy man” off a bridge to block the trolley (Hauser et al., 2007) more objectionable than pulling a lever to throw the heavy man off the bridge, even though both actions would result in exactly the same outcome. Several experiments that demonstrate this result are discussed in Miller et al. (2014). What is interesting is that VR is proving to be an excellent method for finding out how people might behave in practice in these types of circumstance, rather than how they *think* that they might behave in answer to a questionnaire (Pan and Slater, 2011; Navarrete et al., 2012; Friedman et al., 2014; Skulmowski et al., 2014).

## Ethics of XR Use

Before we delve into the ethics of a specific aspect of XR—superrealism—we should consider ethical matters that have already been debated concerning XR use in general. As discussed later on, some of these issues are exacerbated with increasing realism of the virtual experience.

In a scientific context, the use of XR technology is controlled by ethics guidelines and laws that vary across countries but that tend to abide to some general principles. In the United Kingdom, for example, typical research ethics requirements in a scientific context include respect for autonomy and dignity of persons, scientific value, social responsibility, and maximizing benefit and minimizing harm<sup>3</sup>.

<sup>1</sup><http://www.thenetherplay.com>

<sup>2</sup><http://www.theguardian.com/commentisfree/belief/2015/mar/06/virtual-reality-paedophilia-not-harmless-victimless> (“Virtual-reality pedophilia is not victimless or harmless”, March 6th, 2015).

<sup>3</sup>[http://www.bps.org.uk/sites/default/files/documents/code\\_of\\_human\\_research\\_ethics.pdf](http://www.bps.org.uk/sites/default/files/documents/code_of_human_research_ethics.pdf), Code of Human Research Ethics, British Psychological Society, 2010.

On top of the risks in research in general (for example exposure of vulnerable people, exposure to sensitive topics, data-related issues, impact on the physical and psychological well-being, and on the social standing of the participants), XR research must also take into account risks specific to this technology. Behr et al. (2005) summarizes these risks in VR research as follows: (i) motion sickness; (ii) information overload; (iii) intensification of experience (any feeling may be intensified in a VR environment, potentially straining the participants' coping abilities thereby instigating adverse responses), and (iv) cognitive, emotional and behavioral disturbances after re-entry into the real world following the VR experience. Although these were described for VR, they are as well valid for AR (especially ii–iv).

The above though refers to what takes place in a scientific laboratory under strictly controlled conditions, subject to review and oversight by authorities. However, XR is on the verge of becoming a mass consumer product, and since we know that presence, first-person experience and agency are very powerful cues to the brain that “this is really happening,” careful attention needs to be paid to the presentation of violence or abusive behavior in these contexts.

There is already some literature on the ethics of VR and AR use. Some authors discuss this in detail and raise a number of issues of importance to XR industry and practitioners, and ultimately for regulatory authorities at various levels to consider (Wassom, 2014; Madary and Metzinger, 2016):

- Virtual embodiment can lead to emotional, cognitive, and behavioral changes. Although those investigated to date have been for what would generally be regarded as beneficial to the individual and society (for example against racial discrimination) there is the possibility that the same technique might be used for harmful applications.
- Exiting from VR may be problematic in some circumstances where individuals had been living in a virtual fantasy world with an enhanced virtual body. This is the downside of positive transfer effects known to occur from psychological therapy that employs VR.
- Long-term and frequent use of XR might lead to people prioritizing the virtual world over the real one.
- It should be clear what the legal and ethical responsibilities are for actions carried out at a distance if embodied in a virtual body or a remote robot controlled by some interface. Suppose the remote representation causes psychological or physical harm to others. Who is responsible—especially in a case where the participant might argue that her or his intentions were not properly realized through the interface, so that the harmful behavior was not intended? In the case of a physical robot, under which legal jurisdiction does the issue fall—that of the participant, the robot or the robot's manufacturer?
- It will be possible in XR to represent situations that might cause psychological harm such as the representation of deceased relatives with whom one will be able to interact. It is not clear whether this will affect, for example the process of acceptance after a loss or whether it could engender feelings like grief or anger.

- XR technology is highly persuasive—that is the whole point and that is how it exerts its benefits (for example training for disaster response in a virtual setting is a form of persuasion). Persuasion can nevertheless be used for ill-intended purposes, for example to incite someone to do something they would not naturally do or even to do something illegal or immoral.
- Personal data acquisition, use, and sharing with third parties is a vast topic that deserves careful attention. Because large amounts of personal data may be collected, so this data can be hacked and/or used for malicious reasons. Of particular relevance are data collection, including for example face recognition, data sharing policies (should the government or other third parties have access to what you do virtually?), scams that use someone's data or identity, and fake commercial transactions (for example you buy a product through a fake virtual store that steals your bank details).
- Virtual violence and pornography will be readily available—as they are currently in video games and on the internet—and it will feel more real. This might have significant social consequences.

The point of this list (and there are other issues) is to pose the challenges. Some of these are completely novel issues. While XR has been mostly confined to the lab, the clinic, and training/education institutions, these issues could be considered as worthy of academic and business discussion. Now that XR is about to become a tool widely used in society, they may become pressing problems. A particular problem set may be caused by what we refer to as “*superrealism*” where elements and even experiences in virtual or augmented reality may become indistinguishable from reality.

## SUPERREALISM

Very high quality visual and behavioral realism of virtual humans is becoming increasingly likely and available in the near future. For example, Facebook has been carrying out research and development in this area with impressive results<sup>4</sup>, and similarly Dimension Studios<sup>5</sup>. This will only improve over time as increasing resources are applied to this issue by researchers and companies. In this section we consider some of the implications.

### High-Quality Sensory Feedback

In a hypothetical superrealism we require first that sensory rendering becomes of such high quality that it becomes indistinguishable from reality. Advances in computer graphics such as real-time ray tracing, radiosity and, most powerful of all, light field rendering have reduced the gap between photographic realism and virtual realism enormously over the past three decades. However, the evidence (such as there is) suggests that in the context of how people respond to events and situations within XR, the **level of such visual realism is not so important as might be imagined**. People found VR compelling even in the late 1980s and 1990s when the quality was orders of magnitude worse

<sup>4</sup><https://www.wired.com/story/facebook-oculus-codec-avatars-vr/>

<sup>5</sup><https://www.dimensionstudio.co/work>

than now, and, for example, people became anxious talking to a poor-quality rendering of an audience (Pertaub et al., 2002), or standing in front of a virtual pit (Usuh et al., 1999). Zimmons and Panter (2003) found that **participants exhibited the same level of anxiety in front of a pit irrespective of which of five levels of rendering were used (ranging from wire frame through radiosity)**. In two experiments (Slater et al., 2009; Yu et al., 2012) it was again found that higher quality rendering (real-time ray tracing or a light-field based method), compared to lower quality rendering, did not influence the responses of participants. However, *dynamic* elements of the rendering, such as real-time shadows and reflections that moved with the movements of the participant, did enhance anxiety in response to an event within the virtual environment.

In today's XR systems, enhanced visual realism is increasingly facilitated by stereoscopic vision, head tracking and eye tracking to attain synchronization with the person's eye movements. Immersive sound rendering can also be highly realistic. However, there is still a massive way to go with haptic rendering; handshakes and light touches on the shoulder can be done, but not in a way that is going to be available to consumers in the near future. Advances in recent years have included air vortex generation to produce tactile feedback from a distance (Sodhi et al., 2013) and skin integrated wireless interfaces that present a potentially remarkable advance in tactile feedback (Yu et al., 2019). However, tactile feedback is only one half of the haptic interface. There is also a requirement for force feedback (for example, a virtual human character pushes you). Whilst there are advances in force feedback haptic devices in particular domains such as health care and surgery—e.g., (Vaughan et al., 2016; Rose et al., 2018), the problem with force-feedback haptics is the requirement for bulky and expensive robotic devices, and its lack of generality. With vision or sound, in principle, it is possible to render anything. Wherever participants look in VR they will see and hear something. However, an accidental collision of their knee with a moving virtual object requires a device that can generate contingent effects anywhere on the body. This is unlikely to be realized as a consumer product in the near future. Olfactory (odor) cues are also not available at the consumer level and are unlikely to be for some time, although there are advances toward this (Niedenthal et al., 2019; Yanagida et al., 2019). Therefore, primarily we are concerned with the visual and behavioral aspects of superrealism.

Sensory input and synchronization are far from being the only aspects of superrealism. For example, if humans are represented then not only must they look real (for example in terms of geometry, light reflection, light scattering, etc.) but their behavior must be realistic, ranging from subtle changes in facial expression, eye movements, body movements and gestures, to changes in folds of clothing as the characters move. Realism includes characters apparently seeing and looking at the participant, being able to engage in meaningful interactions even if not conversations. This is becoming possible to some extent with volumetric capture and rendering of people—certainly on the rendering side, if not yet with respect to interaction.

## The Device-Gap

Even if all this were achieved, there is still the further problem: the virtual representations in VR must be displayed through a device. Head-mounted displays (HMD) today, and into the foreseeable future, cannot display at a resolution anywhere near that of natural vision, together with the well-over 180-degree horizontal field-of-view and around 150-degree vertical field-of-view that humans have. Moreover, the fact of putting on the HMD itself demarcates reality from virtual reality—so that unless participants are induced to somehow forget that they are wearing the HMD, they will not *believe* that the virtual scenario is a real one. We refer to this as the *device-gap*, which provides a clear demarcation between reality and VR through the act of donning devices.

AR may be different with respect to the device-gap. We can imagine a future where AR devices become as ubiquitous as smartphones are today, with people typically wearing devices for long periods, for example in the street. Since “reality” would be experienced through the device, then virtual aspects may become indistinguishable from the real—assuming though that there are significant advances with respect to field-of-view and resolution. Accordingly, the device-gap is arguably diminished or may even be eliminated in an AR system where a “known ground truth” (the real world) is merged with virtual content that obeys the laws of physics and with which the participant can interact. AR systems thus create a paradox of visually validated truth comprised by the simultaneous appearance of real truth and possible (virtual) truth, possibly further challenging the separation of real and virtual worlds. This concept could introduce a variant to superrealism in which what the participant believes he/she knows about the real world can be altered in the virtual experience. On the other hand, the very reality of the ground truth may enhance the realism of virtual aspects apparently present in the physical reality.

## Physical vs. Psychological Realism

Despite the advances in realism in XR, it is extremely important to distinguish between *belief* and *illusion*. We do not envisage in the foreseeable future that people are actually going to *believe* that virtual situations and events are real. There are many studies over the past 25 years that show that people do nevertheless *respond* realistically in virtual environments, even when they know with certainty that nothing real is happening (Slater and Sanchez-Vives, 2016). Hence, many of the issues arising with respect to superrealism are likely to also apply even to today's XR systems. For example, someone may automatically, without thinking, try to sit on a virtual chair that has no counterpart in reality, possibly resulting in harm.

It is therefore worth noting that there is a difference between *physical* and *psychological* realism, the former referring to the physical appearance of the virtual features and the latter to the psychological sensation that what happens virtually in an XR world could be happening in reality. It is expected that (physical) superrealism in XR systems achieved through advances in computer graphics enabling more photographic realism, improvements in sensory feedback, and the possibility to interact with virtual elements, among others, also increases

the sensation that the virtual experience is real, i.e., the psychological realism.

## WORST CASE ETHICAL PROBLEMS OF SUPERREALISM

In this section we outline some possible ethical problems in XR that are exacerbated by the improvement in *realness* owing to superrealism. In other words, the issues described below might occur to a certain extent with the use of XR systems but are likely to be aggravated due to the sensation that what is happening virtually could be *really* happening. It is to be emphasized that these are *worst case scenarios, based not at all on evidence, but on speculation*. The intention here is to provoke debate and to highlight the need for further research as these represent concerns ahead of facts. The issues fall into a number of categories and we consider each in turn: the vulnerability of certain groups of people, the after-effects following XR use, the discrimination between real and virtual, data issues, XR as an interface to inflict physical harm, and the potential psychological and social implications. The ordering does not reflect levels of importance.

### Vulnerable Populations

An implicit assumption in the introduction was that participants in a virtual environment would typically be drawn from adult and non-patient groups, and generally non-vulnerable populations. However, as XR devices and applications become consumer products, there is no guarantee whatsoever of that being the case, unless subject to some regulatory controls (for example, like those applied to cigarette purchase, X-rated movies, and so on). For example, children or adolescents may not distinguish well between reality and virtual reality. This may also be the case for certain patient groups, such as those prone to psychosis. With such populations it is a reasonable assumption that even the device-gap would not necessarily operate, perhaps most especially for very young children. We have limited evidence regarding these possibilities although one study that concentrated specifically on postural stability and simulator sickness amongst children concluded that VR led to no changes from baseline (Tychsen and Foeller, 2018), supporting the idea that young children do not discriminate VR from reality as adults do.

### After-Effects

In the grand majority of use cases, the primary point of XR is to provide people with an experience that is apparently happening personally to them in the space in which they seemingly are right now. So although the experience is based on virtual sense data and virtual actions, it is nevertheless *real as an experience*. For example, when a virtual character smiles at a participant and the participant automatically smiles back—the “being smiled at” and the smiling—are real experiences (Chalmers, 2017). We change through our experiences: experiences produce changes in the body and the brain. In other words, **just as real-life experiences have after-effects, so virtual experiences may have physical, emotional, and cognitive after-effects** which may be beneficial or

harmful. For instance, motion sickness after XR use may result in an accident, or being insulted by a virtual character—be it fictional or an avatar controlled by a real person—may influence the person’s well-being in real life. Some of the consequences may be long-lasting.

Another key subject is how the perception of our body can be manipulated with XR—and the ensuing repercussions. In VR it is possible to give people the illusion that they have another body (Yee et al., 2009; Slater et al., 2010), and that their body has changed in some fundamental way. For example, adults can have the illusion of having a child body (Banakou et al., 2013), or white people a black body (Peck et al., 2013; Banakou et al., 2016), and these experiences change the participants—for example parents changing their behavior toward their children (Hamilton-Giachritsis et al., 2018), white people becoming more (Groom et al., 2009) or less implicitly biased against black (Maister et al., 2015), domestic violence offenders improving their recognition of fear in the faces of women after being embodied as a woman subject to abuse by a virtual man (Seinfeld et al., 2018), and so on. Scientific research has tended to explore positive benefits such as these. However, continued exposure to such embodied experiences may also cause confusion in people about their real body, leading to a type of body dysmorphia. The body may be changed in a dramatic way such as having a tail (Steptoe et al., 2013) or an additional limb (Laha et al., 2016), or a very long arm making the body asymmetric (Kilteni et al., 2012). It has been found that the disappearance of a virtual arm may elicit some cortical reorganization (for example changes in brain connections) after a short exposure (Kilteni et al., 2016). It is possible that repeated exposure to extra limbs or other dramatic body transformations may bring about unwanted changes, or even pain (inducing virtually caused phantom limb pain). However unlikely, such outcomes should be considered.

As people spend more and more time online in XR, their virtual bodies may tend to be evaluated as more beautiful or preferable in various ways in comparison to their real bodies. Just as present day social media such as Snapchat is apparently leading to higher rates of body dysmorphia (body dissatisfaction) leading to greater demands for cosmetic surgery<sup>6</sup> (Rajanala et al., 2018), so the same may occur with respect to future virtual bodies.

### Is It Real?

To the extent that a VR system supports natural sensorimotor contingencies (being able to use the body to perceive in a manner similar enough to perception in everyday reality) it will typically lead to participants experiencing “place illusion,” the illusion of being in the place depicted by the virtual reality. A VR system may support (i) credible responses to the actions of the participant, (ii) contingent events that are directed specifically and personally toward the participant (for example a virtual human character smiles at the participant), and (iii) scenarios that are faithful to expectations when they simulate events that could occur in reality in a domain in which the participant has expertise. To the extent that these three are supported, the

<sup>6</sup><https://www.theguardian.com/lifeandstyle/2019/jan/23/faking-it-how-selfie-dysmorphia-is-driving-people-to-seek-surgery>

VR experience may become a *plausible* one, where participants have the illusion that the depicted events are really happening (to them). These two illusions, place illusion and plausibility, provide the basis for people responding realistically in virtual environments (Slater, 2009). In AR, these illusions may be more easily attained because the virtual components are superimposed or inserted into the real world.

Imagine now repeated exposures to XR with strong place illusion and plausibility. The following are possible negative outcomes:

- **Uncertainty of past and current events:** Participants **remember virtual events as if they had been real**, and fail to distinguish over time events that really happened and those that happened in XR. This could also lead to mistrust of events that are actually occurring in reality. After spending some time in a scenario people forget the device-gap and become unsure about whether they are experiencing reality or virtual reality.
- **False attribution toward a specific group of people:** An event may have occurred in XR where a participant has a negative interaction with a representation of a particular type of person (for example another race or gender). Although this only happened in XR the participant generalizes beyond this, and attributes, for example harmful intents to real people of that type. This may occur even with representations of individual people known to the participant (see *Identity hacking* below).
- **Dangerous presuppositions leading to physical harm:** People carry out some physical action in XR that has no counterpart in the real world in which the XR is embedded. We have previously mentioned the chair problem where someone attempts to sit on a virtual chair that has no physical counterpart. Imagine that in VR or AR a participant sees others diving into a swimming pool, and decides to follow suit—and in reality they dive into a hard floor.
- **Difficult real-world transition:** After an intense and emotional experience in XR, you take the headset off, and you are suddenly in the very different real world. We are not good at rapid adjustment of behavior and emotion regulation. Re-entry to the real world (Behr et al., 2005; Lanier, 2017), especially after repeated XR exposure, might lead to disturbances of various types: cognitive (did something happen in XR or in real life?), emotional (cause of emotions is not *real*, for example your avatar was insulted by a fictional virtual character), and behavioral (for example actions accepted in XR may not be socially accepted in the real world).

## XR as an Interface to Physical Assault

A type of “VR” is typically used in drone strikes. The operator, thousands of kilometers away from an intended target, uses an interface to guide a drone which fires a weapon at designated hostile personnel. There is a debate in the military ethics literature about the ethical standing of such strikes (Braun and Brunstetter, 2013), with some arguing that they follow the

doctrine of proportionality (since typically there is less “collateral damage”) and others arguing that it nevertheless violates the principle of justice of force short of war (*jus ad vim*). Less dramatically than drone strikes, studies have been done where participants through VR become embodied in, and control in real-time, a remote physical robot. Such robots could also be used to inflict harm. One can also imagine in AR that a person is convinced by others, or by the situation, that a superrealistic avatar seen in physical space can be attacked, because it is only an avatar—yet it turns out to be a real person. It is not clear that these examples are ethical problems in the domain of XR. In the drone strikes, a type of VR is used solely as an interface. The case of a remote robot is just a modern version of a teleoperator system. The VR interfaces are used to deliver sensory information from the remote robot to the participant, and to track the participant to deliver movement instructions to the remote robot. Is this an ethical problem intrinsic to VR itself? The bigger problem may be the distancing and dehumanizing effects. In the AR example, it may happen by accident, or may be by design that a real person is attacked because the attacker had believed that the person was only virtual.

## PRIVACY AND DATA ISSUES

Superrealism can be enhanced by collecting personal data such as location, body movements, preferences, and actions in the virtual or semi-virtual environment. This has great implications for a number of applications, from storytelling to advertising to health, but it also raises important ethical issues related to privacy, data sharing, and the misuse of personal data for hacking and other criminal purposes.

### Personal Data

With the increase in realism may come an increase in personal data acquisition by the XR system, for instance to better articulate movements of a virtual representation of the participant, to personalize advertising or to enable features relevant to the geographical location in which you are. Traits including motor actions, patterns of eye movement, and reflexes (a person’s “kinematic fingerprint”) and information about preferences, habits and interests may be recorded (Spiegel, 2018). This type of personal data is not commonly collected by non-XR current products or experiences on the market today, and so new thinking and consideration will be required to address data collection specific to XR. It is also a critical issue for the uptake of XR. If by default XR devices collect and log such personal data, even if anonymously, then it would not be possible to use such systems in places such as hospitals without violating data protection rules, and in Europe especially the strict regulations of GDPR would have to be followed.

The right to privacy is the right to one’s identity in any form (including name, image, voice, preferences) remaining private, that is, not becoming publicly disclosed. Brey (2008) contrasts the right to privacy to the right to free speech, freedom of the press, and freedom of artistic expression. Whereas, the latter three deserve their own attention, it is crucial to maintain the right to privacy of individuals given that disclosure of private information

may be seriously harmful to the psychological well-being and social standing of the affected person. Legislation may have to be changed in order to accommodate the type of individual data that can be stored as a result of XR use. An example of misuse of disclosed private information is identity hacking, described below; another example could be misuse of deeply personal data, such as someone's phobias, for blackmail or other illegal purposes.

## Data Protection and Data Sharing

As happens with current technologies, so will data collected by XR systems be shared with third parties. The implications are similar to those already existing today in other forms of media except that the amount and type of information may put the individual whose data are being shared at a higher risk (as described in the next paragraphs). Additionally, because of the realism in XR worlds, if, for example, someone carries out an act in XR that would be illegal in reality and if that has been monitored and recorded, it might be later used in evidence about the character of that person in legal proceedings relating to acts in the real world.

## Identity Hacking

With superrealism it will be possible to make **virtual "copies" of people** that look, act, talk like a real person, even demonstrating aspects of personality (for example through the use of machine learning applied to behavior based on recordings of the real person). In this case, some potentially nefarious uses of this would include:

- **Fake news:** People could be portrayed as carrying out actions and saying things that they did not do. This is already powerful enough in photos and videos, but in XR could be even more dangerous because plausibility includes the automatic attribution of *realness* to virtual humans. Once having experienced a virtual rendition of someone carrying out an action, it may be difficult to remove this from memory, and may stimulate implicit changes of attitude toward that person. One step further is *defamation*, whereby a person is depicted in XR doing something immoral or ridiculous, consequently negatively affecting their social standing or reputation.
- **Deliberate mistaken identity:** In XR you are in a private conversation in your living room or in a virtual space with a significant other who is physically remote but apparently in the same space as yourself. You talk about private information or security issues that you would never mention to someone else. However, although the representation is of the significant other, in fact it is someone else who has hacked the avatar of that person.
- **Identity theft:** The same technique could be applied to virtual renditions of ourselves that are not "owned" or controlled by ourselves. We could be portrayed as carrying out virtual actions that we would never do in reality, with negative consequences in our relations with others generally, or with employers, or other authorities.

- **Body swapping:** The technique of body swapping in VR, where one person converses with themselves by successively occupying two different virtual bodies has thus far been used for positive means, such as solving personal problems, for example, people can alternately switch between describing a personal problem while embodying a virtual body closely resembling themselves, and offering themselves counseling while embodying a virtual representation of Dr. Sigmund Freud (Osimo et al., 2015; Slater et al., 2019). It is possible—if unlikely—that the same technology could be used to gain insight into another person's mind, insofar as the mind reflects in some sense the physical body, and thereby gain advantage. This could be very similar to role-play, and might not be considered an ethical problem intrinsic to VR.

## PSYCHOLOGICAL AND SOCIAL IMPLICATIONS

Using XR entails modifying our current perception of reality: entering VR necessarily involves paying little attention to physical reality (other than obvious aspects such as gravity and physical constraints such as walls), and using AR does the same albeit perhaps to a lesser extent since the virtual features are embedded in the real world. This is not particularly new—the same could be (and has been) said about TV viewing or playing of computer games. However, it could be argued that place illusion, plausibility, and transformed agency puts XR in a special category where the following should be considered:

- **Social isolation:** If the frequency of XR usage were to match or come close to current mobile use for example, it is possible that people's ability to interact in real life may be strongly hampered.
- **Preference for virtual social interactions:** Perhaps social interaction in XR could become more enjoyable and desirable than real-life interaction so that people withdraw from society (an extreme case being Hikikomori in Japan). Taking this to its extreme, we could eventually become an *abstract society*, as Karl Popper defines it, in which people never meet face-to-face (Popper, 2012, Chapter 10). As with any new technology that gains widespread use (for example, television, games, social media) questions will arise about the potential negative effects on mental health and social norms, and XR is expected to be no different.
- **Body neglect:** Extreme cases have been reported of people who have spent so much time playing video games that they end up neglecting their body and even their children—sometimes culminating in death<sup>7, 8</sup>. With more realism and more desirability for the virtual life, it is possible that body neglect also occurs in people who would overuse XR.
- **Imitative behavior:** The power of virtual experiences might encourage behavior that the person would not normally

<sup>7</sup><http://www.cnn.com/2015/01/19/world/taiwan-gamer-death>

<sup>8</sup><http://www.newsweek.com/2014/08/15/korean-couple-let-baby-die-while-they-played-videogame-261483.html>

carry out in reality. This could be through exposure—for example, it may be difficult for a person to carry out their first act of violence in XR, but eventually it becomes easy, and leads to a greater propensity for violence in reality—or it could also occur through copycat behavior—mimicking the harmful behaviors of other virtual characters, for example, peer group pressure seems to operate in VR (Neyret et al., 2020).

- **Persuasion:** VR and AR are necessarily persuasive in the sense that they provide the participant with an alternative experience that seems real and that can even change their perception, and even more so if the virtual world is superrealistic; however, persuasion directed at modifying someone's emotions or behavior for detrimental ends is highly unethical. Everyone may be at risk and particularly vulnerable populations.
- **Unexpected horror:** As part of, for example, an artistic virtual environment people may be exposed to horrors that they did not expect and of which they were not forewarned, resulting in a kind of post-traumatic stress response or, conversely, in desensitization for obscene sights.
- **Pornography and exposure to violence:** People will undoubtedly be exposed to realistic scenes with pornographic or violent content (this is already a fact in other forms of media). The consequences of such images being more realistic and being experienced from a first-person perspective (as already happens in video-games) is likely to have consequences for society. Nonetheless, these seem to be more attributable to pornography and violence themselves and not so much to XR technology.
- **Extreme violence and assault:** The realistic depiction of very obscene scenes portraying extreme acts of physical or sexual assault, including the representation of virtual characters with childlike features involved in any kind of sexual context, raises critical ethical concerns. Whether this would increase or decrease obscene behavior in real life is not clear and is very difficult to assess experimentally. On the one hand, engaging in or observing these acts carried out by virtual characters may trigger desensitization, which could normalize and thus increase these acts in real life; on the other hand, it may suppress the urges of aggressors to engage in such actions in the real world.
- **Lack of common environments:** Social science teaches us that our environment gives us norms for behavior and identity (defined, for example, by advertising in the media or fashion industry). The environments that we experience in XR may become the new normal, if we use XR enough. The particular ethical challenge here is that other people do not know or have access to an individual's XR environment, whereas everyone can see real-world environments and have public debates about them. Prolonged XR use on a large scale might challenge the normal public and societal mechanisms for monitoring, discussing, and improving the environments that we live in. The combination of immersion and personalisation could lead to a fracturing of what social and political thought calls “the public sphere.”

- **Lack of ground truth:** There are risks associated with the power of XR to provide convincing sensory evidence that people take as ground truth. For example, in legal settings, a witness may say “I saw the suspect leave the suitcase at the station entrance, look around, and then quickly walk away.” The visual experience of the witness is crucial for justice, and the law court trusts that the visual experiences of witnesses generally correspond to ground truth. XR potentially allows the people who control the system (i.e., the generated sensory data and possibilities for interaction) to control, reorganize, and manipulate the sensory experiences of others. Society is based on the premise that sensory experiences give ground truth. XR at societal scales has the capacity to decouple sensory experience from ground truth, potentially undermining some core elements of social fabric.
- **Persuasive advertising:** Potential negative manifestations of advertising content in XR should be considered. Up until recently advertising was public: everyone watching the same material on TV or reading the newspapers would see the same adverts. Later advertising on the web and social media became personal so that one person would see a set of personalized ads based on their own online profile and history. However, such advertising can be easily ignored. Now with AR it is possible that as we go about our daily lives (wearing AR headsets) we might be bombarded by advertising where virtual human characters continuously approach us acting out advertising scenarios, selling products, and directly trying to persuade us. It is also possible that we may not know that we are being actively persuaded in this manner. This cannot be ignored, and could be highly persuasive. Perhaps following certain types of web and games advertising, people will have to pay to stop such bombardment.

## PRINCIPLES FOR ACTION

Rather than try to deal with each of the above raised issues separately, here we outline some general principles that might be applied to each type of problem. Note that these principles are particularly relevant to superrealism in the context of XR rather than XR *per se*.

### Minimizing Potential Harm of Immoderate Use

First of all, it is essential to distinguish between the risks originating from immoderate use and those emanating from the content of XR applications. Spending 2 h a week in a virtual world is clearly not the same as devoting most of one's waking hours to creating and living in a virtual life. Indeed, a study of adolescents showed that moderate use of social media is not inherently harmful and may even be beneficial (Przybylski and Weinstein, 2017), so the same may be true about XR use. However, there is not yet a societal norm for what constitutes a reasonable frequency of use, or indeed an understanding of who is responsible for limiting or indeed enforcing the amount



of time the user spends in XR. For example, withdrawal from the public sphere of shared reality into a “private world” of individual experience that (although not real) is lived as if it is a private reality could be a real risk for the well-being of XR users. Yet, is living in this private world a *right*? Can we *require* people to be part of a shared public sphere? Can we justifiably *prevent* XR providers from providing the world into which they withdraw?

In fact, social norms are helpful here: we normally do not allow providers to supply a potentially harmful product and then devolve all of the ethical risk to the user. Instead, we regulate the supply of the product to ensure that the use is appropriate. For example, if a product is potentially addictive, we are cautious about providing it (think about tobacco or alcoholic beverages). It is therefore essential that developers are aware of the ethical implications that can arise as a consequence of how their products are constructed, that they recognize they have a major role in preventing dangers of immoderate use and that they must accept evidence-based regulation to minimize harm. Together with legal authorities, providers have a huge impact on how the *use* of their products is perceived by society. However, it is also recognized that, in order to do this, developers and authorities alike need access to more research on which to base their response and recommendations.

### Minimizing Content-Induced Risk

The other critical factor involves the risks posed by the content of XR applications. Again, one cannot compare racing cars with committing extremely violent crimes in XR. This is relevant particularly for how applications are designed, such as games, products for training or therapy, or applications for research. Brey states that designers should consider what kinds of actions are made possible within XR, how these actions are represented, and whether these actions are encouraged or dissuaded (Brey, 1999, 2008). In a game or another application in which killing is possible, for example, is this action encouraged or is it dissuaded? Is it rewarded or is it punished? Is the depiction of such action realistic or is it *toned down*? Is a specific social group (for example, a specific race) the target of such action? Whether one particular event taking place in XR is moral or immoral depends on multiple factors, some of which have been described here, and not on the event *per se*. Some authors have suggested that developers disclaim the potential effects of the content on the users. If developers are transparent, and openly and understandably transmit the possible effects on their users, they limit their legal liabilities on top of protecting individuals from potential harm (Brey, 2008; Spiegel, 2018).

In fact, certain principles that apply to other forms of media—such as broadcasting—are also appropriate for XR technology. For instance, many aspects relevant to XR systems are covered by existing BBC editorial policy guidelines<sup>9</sup>. In conventional media, there are clear warnings, for example, that what is being shown is a reconstruction, or that some material contains images that may be disturbing to some, so that people are not caught off-guard or misinterpret the veracity of what they observe. Trust is critical; if the BBC reconstruct something (for example, a crime scene), this

has to be very clearly labeled—to show visually that this is not the reality. However, XR has a different level of intensity—and often a different objective—that calls for the development of new conventions and sometimes the modification of existing ones (for example, clear depictions of violence may be necessary in military training with XR, and principles or guidelines intending to ameliorate the distress of the participants may not apply in this case). Clear warnings are always advisable and minimum age requirements may be adequate in some instances. Moreover, the short- and long-term effects are unknown; hence, guidelines of XR use will need to be modified as research unravels new findings.

### Selecting Levels of Deception

Along with the expectations from the XR industry, we should consider the nature of XR tools as well as which type of use society gives them. VR and AR are intrinsically “deceptive” in that they deliver virtual sense data that may be perceived by people as an alternate reality, and they provide the means to interact within that reality. Although this is “deceptive,” it is the point of XR. People are freely able to choose to enter into this deception, and there is an implicit contract with the designer of the virtual experience where the participant says “I want to experience your virtual world,” and the designer/implementer says “Suit-up in this way with these devices and you will experience it.” The question is then: how much should the contract go beyond this?

To reduce the level of realness, implementers (for example, researchers) and participants may be able to select a level of deception. For example, level 10 means that the XR should try its absolute best to completely convince participants that what they are experiencing is real. Level 1 might be “give me some experience, but do your best to keep reminding me that this is not happening, it is not real.” How this might be done is already problematic—for we have seen that, for example, rendering everything in wire frame (i.e., something that clearly appears unrealistic) in itself is almost certainly not sufficient to completely diminish place illusion and plausibility. An example might be that participants in AR may want to have a setting where virtual human characters are always with (for example) a halo, so that they always know that they are not real. What would the settings between 1 and 10 mean? We have no data that could shed light on this question. It would be important to uncover factors that influence the probability of people being able to distinguish real from virtual when they are wearing the device, and after they no longer wear the device again distinguish between real and virtual memories. Confusion is the heart of the problem given that the very idea of XR involves confusion. A simpler alternative to *selecting* the level of deception within an XR application would be to at least *be aware* of how realistic it is, perhaps based on some sort of standardized rating scale that allows users to select an XR application based on the level of deception. Broadcast and film have well-known and understood ratings but there is no known rating system for VR and AR.

### Educating Implementers and Participants

Education of implementers and participants about the power of the technology should be a fundamental principle and responsibility of producers of material. Realism is certainly vital

<sup>9</sup><https://www.bbc.co.uk/editorialguidelines/>

in extremely important applications such as for training (flight simulators are a good example of this). However, as well as emphasizing the positive aspects, potential negative effects should also be considered. Perhaps we may apply the same principles as for medicine: we take it for positive effect, but we are also warned of potential side-effects. Education should also take into account that although XR can result in surreptitious influences on behavior, this is nothing new. There are innumerable attempts in everyday life to influence our attitudes and behavior. The question though is whether people know that this is occurring. For example, in the 1950s there were attempts at subliminal advertising in cinemas (flashing an advert so fast it could not be consciously seen), which was eventually discovered and banned.

Education also includes training of end users. For example, when watching TV or playing a video game, if the content becomes uncomfortable or distressing to the viewers they can simply look away and immediately see the real world. In XR, the most obvious thing to do when in trouble would be to close your eyes and take off the device. However, it may not be so easy to disengage—precisely because place illusion and plausibility may lead some participants to simply forget that they can do this. Some form of training to remind users of their ability to “opt-out,” or else a stop button may therefore be an important concept, in order to always respect the participant’s right to stop. Additionally, some kind of post-experience “cleansing” may be needed.

As well as education, a related and fundamental issue is trust. Some people might be afraid of a hypothetical case in which reality and XR are not discernible, for example, they might be afraid about themselves or others becoming confused by using superreal XR before even trying it. This would be evidently assumed out of lack of experience but can nonetheless be resolved if there is trust. How can consumers of virtual experiences be assured that they can trust the content? A way forward on this is to develop industry standards or even a cross-industry code of conduct to which producers of virtual content must adhere. A technological solution may involve some concept such as a “watermarking” equivalent of XR. For any approach toward avoidance of negative influences, there have to be standards developed that are agreed upon across industry, with education amongst participants about what particular effects mean. As a simple example, if virtual characters always have a halo, then the meaning of this convention needs to be understood.

## Protecting Personal Information

Finally, data issues should be carefully addressed. Some authors have proposed that companies publicly disclose what kind of personal information they obtain and share with third parties, some encouraging “no share” data laws or options for the user to opt out (Pase, 2012; Spiegel, 2018). In Europe this is almost certainly already covered by the GDPR legislation, in particular Article 6<sup>10</sup>. The benefits of these legal restrictions, they defend, would outweigh the harm imposed on personal liberty, like the right to privacy. In the context of superrealism in which large amounts of personal data may be used, this option seems at the

very least cautious. If this were to be done, the disclosures for the lay public should be made simple and comprehensible.

In all cases, it seems that legal authorities in particular may benefit from considering the implementation of the *precautionary principle*, whereby discretionary measures are taken when the consequences of a new situation are not yet known; in this case, the effects of XR use are not yet fully understood, it is not clear how the content, and in particular superreal content, may influence individuals and society as a whole, and data issues remain a debatable topic. Research is thus warranted to bring insights into these matters.

## SCIENTIFIC QUESTIONS

As we have seen, there is essentially no data that can help in addressing these ethical issues. The problem is that while XR was confined to the lab and industry, it was under tight control through the standard ethical procedures of the institutions, which were guided by the rules and principles briefly outlined in the introduction. Now that XR technology is being released for mass consumption, there are no controls, and no relevant data.

Moreover, it is important to understand that ethical problems do not end with a particular experience—what happens in the longer term is critical. An after-effect might be prolonged. Even a single traumatic episode can have lasting consequences. After watching a movie, you move around in real space where other people are visible, you interact with the real world, and maybe that process diffuses the experience. But it might be the case that this does not work in XR—since as we argued earlier, an XR experience is a real and personal experience, even though the source of the experience is virtual. What was experienced was not about someone else (as it is in a movie) but personal.

Short-term after-effects should be experimentally tractable now. Suitable behavioral tests and measures could compare participant behavior in simple cognitive and social tasks immediately after a brief period in VR or AR, perhaps comparing two XR scenarios that elicit contrasting emotions.

Long-term acculturation effects are not easy to study experimentally, at least not at the moment. We do not know how much exposure is required, and we cannot control for the additional stimulation the participant gets while not in the XR.

Sensory grounding could be studied now. This could start by investigating whether VR or AR can be successfully used to manipulate memory for an event. In a pre-test, for example, I might experience that Bill gave me an apple, and Jane asked to borrow my phone. Can a subsequent session of XR overwrite, erase, or change those memories? This type of research has huge ethical implications for the field of “false memory” and historic child abuse, and would generate a lot of ethical discussion. It would be highly morally and politically sensitive. It would open up a debate on whether AR and immersive VR should or should not be used in situations of recovered memory and historic child abuse; if this use is not yet present, it seems likely to develop. It would be important to involve appropriate academic and clinical researchers in any experimental work, and to think carefully about stakeholders.

<sup>10</sup><https://gdpr.eu/article-6-how-to-process-personal-data-legally/>

We can consider additionally the following issues for experimentation, presented as a series of questions:

- Do people trust virtual characters more if they are more realistic?
- Does greater realism lead to greater confusion between the real and the virtual?
- Does greater realism lead to greater behavioral and emotional impact?
- Does greater realism lead to a greater chance of negative after-effects?
- Can people already today be confused between reality and virtual reality?
- Will there be greater plausibility (illusion that the events are really happening) in interactions with superrealistic characters?
- What, if any, are public perceptions of these issues today?
- How can there be longer term follow-ups of the effects of a virtual experience?
- What are the long-term cultural effects of superreal XR usage?

The other way to think about this might be to explore the concept of discernment in virtual environments. We are familiar with the uncanny effects of viewing avatars and even though animation is capable of producing more and more lifelike figures, we can still tell what is real and what is not real. So, is there a skill of discernment that allows people to learn to distinguish between the real and the virtual? Under some circumstances, some consumers might be more able to discern than others, some might be able to be taught to recognize—just as some people can be taught to tell fake news from real news online—but many would not. There may be longer term questions about the speed with which such education could be developed and extended into the community: what might the lag be between creation of virtual content and development of discernment skills?

## CONCLUSIONS

The development of increasingly realistic virtual worlds allows for advancements in XR technology to be used in training, education, psychotherapy, physical and mental rehabilitation, marketing, entertainment, and for further applications in research. The benefits of superrealism are clear: realistic virtual scenarios can make XR applications more efficacious. For example, aviators can be better trained because the virtual

## REFERENCES

- Banakou, D., Groten, R., and Slater, M. (2013). Illusory ownership of a virtual child body causes overestimation of object sizes and implicit attitude changes. *Proc. Natl. Acad. Sci. U.S.A.* 110, 12846–12851. doi: 10.1073/pnas.1306779110
- Banakou, D., Pd, H., 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
- Behr, K.-M., Nosper, A., Klimmt, C., and Hartmann, T. (2005). Some practical considerations of ethical issues in VR research. *Presence* 14, 668–676. doi: 10.1162/105474605775196535

simulation in which they operate is more accurate and closer to reality; exposure therapy in which a patient is presented with a realistic virtual version of the agent they are afraid of (for example, a spider) may be more efficient if the agent seems real, and so on. As occurs with most things in the world, with benefits come potential misuse, abuse or neglect, all of which bring about ethical concerns.

We started with a version of the golden rule: “That which is hateful to you, do not do to your fellow. That is the whole law; the rest is the explanation; go and learn it.” This is not at all about “empathy,” but very practical guidance. When we construct experiences for others, we need to think about whether we would want to have this experience—without prior warning, education, training, and assured compliance with a generally agreed and debated code of conduct. The challenge now is for researchers, content creators, and distributors of XR systems to determine what should be within this code of conduct.

## AUTHOR CONTRIBUTIONS

MS wrote the first draft of the paper. PH and CV provided further first-hand writing. CG-L systematically contributed to, edited, and organized the paper. All other authors contributed to and reviewed the paper. The ideas of the paper were formulated through a series of meetings to which all authors contributed.

## FUNDING

This work was initiated and funded by Digital Catapult, London, UK. Individual members of Digital Catapult took part in the research and writing. MS was Immersive Fellow at Digital Catapult, and CG-L was employed by Digital Catapult for this purpose. Digital Catapult has no financial interest in the publication of this paper.

## ACKNOWLEDGMENTS

This paper was produced as a result of meetings of the Digital Catapult Working Group on Ethics of Realism in XR. In addition we thank Andrew Fitzgibbon of Microsoft Cambridge UK for valuable input, and Naima Camara, Paul Childs, Mandy Mazliah, Cordelia O’Connell, and Philip Young of Digital Catapult for editing. MS led this work while Digital Catapult Immersive Fellow. MS is also supported by the ERC Advanced Grant MoTIVE (#742989).

- Bergelson, V. (2007). Consent to harm. *Pace L. Rev.* 28, 683–711. Available online at: <https://heinonline.org/HOL/LandingPage?handle=hein.journals/pace28&div=37&id=&page=>
- Braun, M., and Brunstetter, D. R. (2013). Rethinking the criterion for assessing CIA-targeted killings: drones, proportionality and jus ad vim. *J. Milit. Ethics* 12, 304–324. doi: 10.1080/15027570.2013.869390
- Brey, P. (1999). The ethics of representation and action in virtual reality. *Ethics Inf. Technol.* 1, 5–14. doi: 10.1023/A:1010069907461
- Brey, P. (2008). “Virtual reality and computer simulation,” in *The Handbook of Information and Computer Ethics*, eds K. Himma and H. Tavani (Hoboken, NJ: John Wiley & Sons, Inc.), 361–384.

- Chalmers, D. J. (2017). The virtual and the real. *Disputatio* 9, 309–352. doi: 10.1515/disp-2017-0009
- Freeman, D., Evans, N., Lister, R., Antley, A., Dunn, G., and Slater, M. (2013). Height, social comparison, and paranoia: an immersive virtual reality experimental study. *Psychiatry Res.* 213, 348–352. doi: 10.1016/j.psychres.2013.12.014
- 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
- Friedman, D., Pizarro, R., Or-Berkers, K., Neyret, S., Pan, X., and Slater, M. (2014). A method for generating an illusion of backwards time travel using immersive virtual reality—an exploratory study. *Front. Psychol.* 5:943. doi: 10.3389/fpsyg.2014.00943
- Groom, V., Bailenson, J. N., and Nass, C. (2009). The influence of racial embodiment on racial bias in immersive virtual environments. *Soc. Infl.* 4, 231–248. doi: 10.1080/15534510802643750
- Hamilton-Giachritsis, C., Banakou, D., Garcia Quiroga, M., Giachritsis, C., and Slater, M. (2018). Reducing risk and improving maternal perspective-taking and empathy using virtual embodiment. *Sci. Rep.* 8:2975. doi: 10.1038/s41598-018-21036-2
- Hauser, M., Cushman, F., Young, L., Kang-Xing Jin, R., and Mikhail, J. (2007). A dissociation between moral judgments and justifications. *Mind Lang.* 22, 1–21. doi: 10.1111/j.1468-0017.2006.00297.x
- Kilteni, K., Grau-Sánchez, J., Veciana De Las Heras, M., Rodríguez-Fornells, A., and Slater, M. (2016). Decreased corticospinal excitability after the illusion of missing part of the arm. *Front. Hum. Neurosci.* 10:145. doi: 10.3389/fnhum.2016.00145
- Kilteni, K., Normand, J.-M., Sanchez Vives, M. V., and Slater, M. (2012). Extending body space in immersive virtual reality: a very long arm illusion. *PLoS ONE* 7:e40867. doi: 10.1371/journal.pone.0040867
- Laha, B., Bailenson, J. N., Won, A. S., and Bailey, J. O. (2016). Evaluating control schemes for the third arm of an avatar. *Presence* 25, 129–147. doi: 10.1162/PRES\_a\_00251
- Lanier, J. (2017). *Dawn of the New Everything: Encounters With Reality and Virtual Reality*. London: Vintage Publishing.
- Levin, M. F., Weiss, P. L., and Keshner, E. A. (2015). Emergence of virtual reality as a tool for upper limb rehabilitation: incorporation of motor control and motor learning principles. *Phys. Ther.* 95, 415–425. doi: 10.2522/ptj.20130579
- 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. Robot. AI* 3:3. doi: 10.3389/frobt.2016.00003
- Maister, L., Sebanz, N., Knoblich, G., and Tsakiris, M. (2013). Experiencing ownership over a dark-skinned body reduces implicit racial bias. *Cognition* 128, 170–178. doi: 10.1016/j.cognition.2013.04.002
- 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
- Matamala-Gomez, M., Donegan, T., Bottiroli, S., Sandrini, G., Sanchez-Vives, M. V., and Tassorelli, C. (2019). Immersive virtual reality and virtual embodiment for pain relief. *Front. Hum. Neurosci.* 13: 279. doi: 10.3389/fnhum.2019.00279
- Miller, R. M., Hannikainen, I. A., and Cushman, F. A. (2014). Bad actions or bad outcomes? Differentiating affective contributions to the moral condemnation of harm. *Emotion* 14, 573–587. doi: 10.1037/a0035361
- Navarrete, C. D., McDonald, M. M., Mott, M. L., and Asher, B. (2012). Virtual morality: emotion and action in a simulated three-dimensional “trolley problem”. *Emotion* 12, 364–370. doi: 10.1037/a0025561
- Neyret, S., Oliva, R., Beacco, A., Navarro, X., Valenzuela, J., and Slater, M. (2020). An embodied perspective as a victim of sexual harassment in virtual reality reduces action conformity in a later milgram obedience scenario. submitted.
- Niedenthal, S., Lundén, P., Ehrndal, M., and Olofsson, J. K. (2019). “A handheld olfactory display for smell-enabled VR games,” in *2019 IEEE International Symposium on Olfaction and Electronic Nose (ISOEN)* (Fukuoka: IEEE), 1–4. doi: 10.1109/ISOEN.2019.8823162
- Osimo, S. A., Pizarro, R., Spanlang, B., and Slater, M. (2015). Conversations between self and self as sigmund freud – A virtual body ownership paradigm for self counselling. *Sci. Rep.* 5:13899. doi: 10.1038/srep13899
- Pan, X., and Slater, M. (2011). “Confronting a moral dilemma in virtual reality: a pilot study,” in *BCS-HCI '11 Proceedings of the 25th BCS Conference on Human-Computer Interaction* (Swindon), 46–51. doi: 10.14236/ewic/HCI2011.26
- Pase, S. (2012). “Ethical considerations in augmented reality applications,” in *Proceedings of the International Conference on e-Learning, e-Business, Enterprise Information Systems, and e-Government (EEE): The Steering Committee of The World Congress in Computer Science, Computer* (Las Vegas, NV), 1.
- 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
- Pertaub, D.-P., Slater, M., and Barker, C. (2002). An experiment on public speaking anxiety in response to three different types of virtual audience. *Presence* 11, 68–78. doi: 10.1162/105474602317343668
- Popper, K. (2012). *The Open Society and Its Enemies*. London; New York, NY: Routledge. doi: 10.4324/9780203439913
- Przybylski, A. K., and Weinstein, N. (2017). A large-scale test of the goldilocks hypothesis: quantifying the relations between digital-screen use and the mental well-being of adolescents. *Psychol. Sci.* 28, 204–215. doi: 10.1177/0956797616678438
- Rajanala, S., Maymone, M. B., and Vashi, N. A. (2018). Selfies—living in the era of filtered photographs. *JAMA Facial Plast. Surg.* 20, 443–444. doi: 10.1001/jamafacial.2018.0486
- Rose, T., Nam, C. S., and Chen, K. B. (2018). Immersion of virtual reality for rehabilitation-Review. *Appl. Ergon.* 69, 153–161. doi: 10.1016/j.apergo.2018.01.009
- Rus-Calafell, M., Garety, P., Sason, E., Craig, T. J., and Valmaggia, L. R. (2018). Virtual reality in the assessment and treatment of psychosis: a systematic review of its utility, acceptability and effectiveness. *Psychol. Med.* 48, 362–391. doi: 10.1017/S0033291717001945
- Seinfeld, S., Arroyo-Palacios, J., Iruretagoyena, G., Hortensius, R., Zapata, L. E., Borland, D., et al. (2018). Offenders become the victim in virtual reality: impact of changing perspective in domestic violence. *Sci. Rep.* 8:2692. doi: 10.1038/s41598-018-19987-7
- Skulmowski, A., Bunge, A., Kaspar, K., and Pipa, G. (2014). Forced-choice decision-making in modified trolley dilemma situations: a virtual reality and eye tracking study. *Front. Behav. Neurosci.* 8:426. doi: 10.3389/fnbeh.2014.00426
- Slater, M. (2009). Place Illusion and Plausibility can lead to realistic behaviour in immersive virtual environments. *Philos Trans R Soc Lond. B. Biol. Sci.* 364, 3549–3557. doi: 10.1098/rstb.2009.0138
- Slater, M., Khanna, P., Mortensen, J., and Yu, I. (2009). Visual realism enhances realistic response in an immersive virtual environment. *IEEE Comput. Graph. Appl.* 29, 76–84. doi: 10.1109/MCG.2009.55
- Slater, M., Neyret, S., Johnston, T., Iruretagoyena, G., Crespo, D. L. C., Alabèrnia-Segura, M., et al. (2019). An experimental study of a virtual reality counselling paradigm using embodied self-dialogue. *Sci. Rep.* 9:10903. doi: 10.1038/s41598-019-46877-3
- Slater, M., and Sanchez-Vives, M. V. (2016). Enhancing our lives with immersive virtual reality. *Front. Robot. AI* 3:74. doi: 10.3389/frobt.2016.00074
- Slater, M., Spanlang, B., Sanchez-Vives, M. V., and Blanke, O. (2010). First person experience of body transfer in virtual reality. *PLoS ONE* 5:e10564. doi: 10.1371/journal.pone.0010564
- Sodhi, R., Poupyrev, I., Glisson, M., and Israr, A. (2013). AIREAL: interactive tactile experiences in free air. *ACM Trans. Graph.* 32:134. doi: 10.1145/2461912.2462007
- Spiegel, J. S. (2018). The ethics of virtual reality technology: social hazards and public policy recommendations. *Sci. Eng. Ethics* 24, 1537–1550. doi: 10.1007/s11948-017-9979-y
- Stephoe, W., Steed, A., and Slater, M. (2013). Human tails: ownership and control of extended humanoid avatars. *IEEE Trans. Vis. Comput. Graph.* 19, 583–590. doi: 10.1109/TVCG.2013.32
- Tajadura-Jiménez, A., Banakou, D., Bianchi-Berthouze, N., and Slater, M. (2017). Embodiment in a child-like talking virtual body influences object size perception, self-identification, and subsequent real speaking. *Sci. Rep.* 7:9637. doi: 10.1038/s41598-017-09497-3
- Thomson, J. J. (1985). The trolley problem. *Yale Law J.* 94, 1395–1415. doi: 10.2307/796133

- Tychsen, L., and Foeller, P. (2018). Effects of immersive virtual reality viewing on young children: visuomotor function, postural stability and visually induced motion sickness. *JAAPOS* 22, e5. doi: 10.1016/j.jaapos.2018.07.011
- Usoh, M., Arthur, K., Whitton, M. C., Bastos, R., Steed, A., Slater, M., et al. (1999). "Walking > walking-in-place > flying, in virtual environments," in *Proceedings of the 26th Annual Conference On Computer Graphics And Interactive Techniques (SIGGRAPH)* (New York, NY), 359–364. doi: 10.1145/311535.311589
- Vaughan, N., Dubey, V. N., Wainwright, T. W., and Middleton, R. G. (2016). A review of virtual reality based training simulators for orthopaedic surgery. *Med. Eng. Phys.* 38, 59–71. doi: 10.1016/j.medengphy.2015.11.021
- Vehtari, A., Simpson, D. P., Yao, Y., and Gelman, A. (2019). Limitations of "limitations of bayesian leave-one-out cross-validation for model selection". *Comput. Brain and Behav.* 2, 22–27. doi: 10.1007/s42113-018-0020-6
- Wassom, B. (2014). *Augmented Reality Law, Privacy, and Ethics: Law, Society, and Emerging AR Technologies*. Waltham, MA: Syngress. doi: 10.1016/B978-0-12-800208-7.00003-X
- Yanagida, Y., Nakano, T., and Watanabe, K. (2019). "Towards precise spatio-temporal control of scents and air for olfactory augmented reality," in *2019 IEEE International Symposium on Olfaction and Electronic Nose (ISOEN)* (Fukuoka: IEEE), 1–4. doi: 10.1109/ISOEN.2019.8823180
- Yee, N., Bailenson, J. N., and Ducheneaut, N. (2009). The Proteus effect: implications of transformed digital self-representation on online and offline behavior. *Communic. Res.* 36, 285–312. doi: 10.1177/009365020830254
- Yee, N., and Bailenson, J. N. (2007). The Proteus effect: the effect of transformed self-representation on behavior. *Hum. Commun. Res.* 33, 271–290. doi: 10.1111/j.1468-2958.2007.00299.x
- Yu, I., Mortensen, J., Khanna, P., Spanlang, B., and Slater, M. (2012). Visual realism enhances realistic response in an immersive virtual environment - Part 2. *IEEE Comput. Graph. Appl.* 32, 36–45. doi: 10.1109/MCG.2012.121
- Yu, X., Xie, Z., Yu, Y., Lee, J., Vazquez-Guardado, A., Luan, H., et al. (2019). Skin-integrated wireless haptic interfaces for virtual and augmented reality. *Nature* 575, 473–479. doi: 10.1038/s41586-019-1687-0
- Zimmons, P., and Panter, A. (2003). "The Influence of Rendering Quality on Presence And Task Performance in a Virtual Environment," in *Proceedings of IEEE Virtual Reality* (Los Angeles, CA: IEEE), 293–294.

**Conflict of Interest:** MS was a consultant for the company Digital Catapult as Immersive Fellow in the carrying out of this work. CG-L was a consultant for Digital Catapult in the carrying out of this work. CV was employed by the company Magic Leap. RG-C and JS were employed by the company Digital Catapult. SJ was employed by the company Dimension – Hammerhead VR. ZW was employed by the BBC. GB was employed by the company HTC Vive. RS, WS, and SH were employed by the company Facebook. DS was employed by the company Jigsaw. DF was employed by the foundation Nesta.

The remaining author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Slater, Gonzalez-Liencre, Haggard, Vinkers, Gregory-Clarke, Jelley, Watson, Breen, Schwarz, Steptoe, Szostak, Halan, Fox and Silver. 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.