Ecological and ethical issues in virtual reality research: A call for increased scrutiny

Abstract: We argue that moral judgment studies currently conducted utilizing virtual reality (VR) devices must confront a dilemma due to how virtual environments are designed and how those environments are experienced. We begin by first describing the contexts present in paradigmatic cases of naturalistic moral judgments. We then compare these contexts to current traditional (vignette-based) and VR-based moral judgment research. We show that, contra to paradigmatic cases, vignette-based and VR-based moral judgment research often fails to accurately model the situational features of paradigmatic moral judgments. In particular, we compare and contrast six recent VR studies to support our view that only simulations high in context-realism and perspectival-fidelity can produce 'virtually real experiences.' After analyzing the constituents of a virtually-real experience, we go on to propose guidelines for the creation of VR studies. These guidelines serve two purposes. First, we aim to increase the ecological validity of such studies in order to advance our understanding of moral judgments. Second, we believe that such guidelines should inform how Institutional Review Boards assess VR research. We show that our guidelines are urgently needed given the current lax review standards in place.

Virtual reality (VR) technologies have rapidly entered the consumer market and their growing availability has provided researchers in many fields access to new tools and techniques for generating and gathering data. The promise of these technologies has not gone unnoticed, however, in this paper we show that the perils of VR research are equally deserving of attention. We survey recent literature on the nature of virtual experiences to show that VR simulations must be approached with caution. Not only are VR simulations likely to introduce issues relating to the ecological validity of data gathered from them, they threaten to cause real harm to their human subjects. We close by proposing a set of guidelines for the creation of VR experiments. Specifically, we will argue for two claims.

First, we argue that much of the current research utilizing VR simulations to study human judgments contain issues with their ecological validity. Many VR simulations fail as simulations of moral decision-making and thus cannot claim to accurately represent the roles of emotion and reason in generating moral judgments. We examine the nature of VR simulations in six recent studies in order to show that only those simulations that produce what we call virtually-real experiences manage to fully avoid ecological issues. However, we go on to argue that virtual environments capable of producing virtually real experiences, those that avoid ecological issues, deserve closer ethical scrutiny. Some virtual simulations are capable of generating at least two distinct forms of harm to their subjects. We end the paper by urging greater caution on the part of both researchers and Institutional Review Boards (IRBs) when it comes to the assessment and application of VR simulations in human subjects research. We introduce a principle we call The Equivalence Principle that we believe should guide the development and approval process of human subjects research utilizing VR simulations especially with respect to assessing the risk of harm that VR protocols may introduce.

Before saying much more about how VR simulations can generate different sorts of experiences, we need to more carefully lay out what it would mean for a study to generate concerns about ecological validity. Additionally, we need to also first map out how we understand the context of natural moral judgments in order to contrast them against not only vignette-based judgment studies but also more recent PC and VR based research.

Ecological validity: The theoretical concern

To say that an experimental protocol has issues with ecological validity is to argue that the experimental situation created for the study (scenario A) is sufficiently unlike the real-world scenario it aims to model (scenario B) to cast doubts about our ability to generalize from data collected in A to situations like B. In cases like these, although we might feel confident that researchers have operationalized their construct well and that they are measuring it reliability, we have doubts about the congruity between the subject's experiences, reactions, or decisions in the experimental context and their experiences, reactions, or decisions in a more naturalistic context.

For example, suppose that a researcher wanted to assess whether autonomic arousal influenced moral judgments during time-sensitive decisions. Assume that the researcher had taken adequate steps to record information relevant to ANS activity and that time-sensitivity was adequately defined and controlled in the study environment. However, if the researcher presented his or her subjects with long-form stories to prompt them with moral choices, we might rightfully worry about the ecological validity of such a study. Our moral decisions, especially time-sensitive moral decisions, are not typically prompted via long-form written narratives.[1] They are typically encountered as part of real-time interactions with other agents in the world.[2] As such, data generated from our hypothetical study may not accurately generalize to its intended context even if we do not have any concerns about the quality of the data gathered or of the statistical analyses applied to the data. We believe that a problem very much like this impacts much of the current VR research on moral decision-making. To see why, we need to say more about moral decision-making both theoretically and in the wild.

First, a minor word of caution. Social psychologists, among others, are currently wrestling with issues relating to experimental replication (Earp & Trafimow, 2015). As such, we should be careful to avoid giving undue credence to any specific experimental finding. However, even in the face of this controversy, it seems right to say that moral decisions can be analyzed in terms of doxastic (conscious) and subdoxastic (non-conscious) elements. Doxastic elements are perhaps most familiar. How does the subject conceptualize the world around her and how does she understand the decision she is being

[1] This is true even though moral judgments may occasionally be primed in this way (i.e. via e-mail or letter). For such a study to escape the ecological validity question, the study would have to narrow its scope such that it aims to model only those moral judgments prompted by such long-form narratives.
[2] We say more about the naturalistic contexts of moral judgments in just a moment.

asked to make? For example, does the subject construe the situation as dangerous? To herself or to others? Does she conceptualize the situation as one in which she can have real consequences or one in which her actions are unlikely to change the results? Doxastic elements are necessarily *conscious* in this sense and represent the phenomenological aspect of a moral judgment (what the situation looks and feels like to the agent making the decision).

Subdoxastic aspects of a decision are equally important, though often more subtle. Although subdoxastic features of an experience often (and importantly) *affect* our conscious experience of a situation, they are not themselves a part of that conscious awareness. For example, our affective states (positive, negative, and neutral) appear to trigger particular suites of heuristics and biases that affect how a subject consciously analyzes the situation around her (Alloy & Abramson, 1979; Kahneman, 2003; Ramirez, 2017). Positive affect biases appear to make one more likely to rely on stereotyped information about others (Lount, 2010) while broadening the scope of attention and thought-action repertoires (Fredrickson & Branigan, 2002).

To say that these effects are subdoxastic is to say that the effects are not present as a component of the conscious experience of subject while she makes her decision. Returning to the earlier example, positive affect can have its effects, insofar as it has them at all, without being part of a subject's conscious deliberative experience.[3] Indeed, in order to be subdoxastic, they *cannot* be a part of the subject's conscious experience. Additionally, other elements of a situation (like time-pressure) or the triggering of implicit biases (Avenanti, Sirigu, & Aglioti 2010) can alter a subject's rational calculus in ways that are not typically aspects of conscious decision-making.[4]

Consider the following prototypical moral scenario: on the way to give an important lecture, you run across a homeless person in need of help. You are running late for the lecture and you are not sure how much time it would take you to render assistance to the person (Darley & Batson, 1973). Such a situation is rich in doxastic and subdoxastic elements. For example, the subject must see the homeless person as someone in need of aid, must see herself as someone capable of helping, and a myriad of other conscious appraisals that go into recognizing a situation as calling for moral action must be present. The situation is also characterized in part by the effects of subdoxastic features. Does the sight of a homeless person trigger any emotional biases that impact the subject's judgment? How strongly does time-pressure decrease her desire to engage in helping behavior?[5] Other implicit biases (stemming from the subject's perception of the race, gender, or identity of the person in need of help) may also impact how the subject consciously assesses the situation without doing so as a part of their conscious

[3] In other words, the subject does not need to realize that she is feeling good and therefore that she is now more likely to rely on stereotypes when she makes judgments.

[4] As mentioned earlier, readers who are dubious about the effects or magnitude of implicit bias are free to discard this aspect of the example. What matters is that at least some of the features of a moral judgment are subdoxastic.

[5] Darley and Batson (1973) found that the greater time-pressure their subjects were under, the less likely they were to help the homeless confederate.

assessment (Avenanti, Sirigu, & Aglioti 2010). To accurately model a situation like this experimentally, researchers should aim to ensure that subjects bring the right doxastic elements of moral choice to the scenario; additionally, however, they must also aim to recreate the relevant subdoxastic elements that would be present in the naturalistic context (Ramirez, 2017).

One ecological problem may therefore arise immediately for traditional vignette-based studies of moral decisions: if the subdoxastic aspects of a subject's experience in a laboratory setting fail to match the subdoxastic elements of a naturalistic judgment, then the ecological validity of the study is compromised. Given the stark differences between the experience of participating in a traditional vignette-based judgment study and naturalistic moral judgements, such a concern seems warranted. We have good reason to think that such concerns are not merely theoretical, but actual.

For example, subjects under laboratory conditions tend to make very poor predictions of their own behavior outside of the laboratory context. None of Stanley Milgram's subjects predicted that they would be capable of delivering a 150-volt shock to their learners even though *all* of them did so under his experimental conditions (Zimbardo, 2004).[6] Additionally, although thought-experiment research appears capable of triggering both System 1 and System 2 responses to moral questions (Gendler, 2007), we do not have any way of knowing whether these responses generalize to real-world moral judgments of the same type in the same way. As a result, many researchers have argued that VR simulations can more faithfully recreate the environments and situational subdoxastic pressures under which moral decisions are made. Researchers hold out hope that "[a] virtual environment provides the researcher with an ecologically valid platform for presenting dynamic stimuli in a manner that allows for both the veridical control of laboratory measures and the verisimilitude of naturalistic observation of real life situations" (Parsons, 2015, p.7).

However, here we must pause to say more about why VR simulations appear to hold such promise and to say something about some of the elements that researchers must keep in mind as they think through and design their simulations. First, we should distinguish the features pertaining to the realism of virtual *environments* (which we expand upon in detail in the next two sections) from the features of a simulation that represent the consequences of an individual's virtual decisions.

These two features are important to keep distinct because, while the first feature, we argue, will always be important with respect to accurately representing (and hence generalizing to) naturalistic moral judgments, the significance of the second will depend on a study's design.[7] To further clarify, in a

[6] Subjects were, in other words, not capable of imaging the relevant subdoxastic influences present in the experimental context (Ramirez, 2017).

[7] For example, a within-subjects design where subjects make several moral judgments either of the same iterated scenario or of several different scenarios will be one in which an accurate representation of the consequences of an individual's decisions may affect the nature of future decisions. We thank an anonymous reviewer for pushing us here to distinguish between these features and to explain their relative significance.

one-off moral between-subjects judgment study, one where the experimenter is primarily interested in the specific *judgment* that a subject would make, an accurate representation of the consequences of a subject's judgment should have little to no impact on the data being collected (once the decision has been made, relevant data has been collected and the consequences of the decision need not be represented naturalistically-or at all-by the experimenter).

Third, for reasons we will elaborate on in our discussions of actual experiments, iterated within-subject judgment studies of the kind we discuss here are likely to have diminished context-realism deriving from the study design itself (though counterbalancing can help mitigate the degree to which this is a problem).

Virtual reality moral judgment studies thus generate a dilemma. If VR simulations accurately capture the essential elements of a moral decision and its environment then such simulations may (and, we argue later, in many cases *will*) be unethical to conduct. If, on the other hand, VR simulations fail to accurately reproduce the relevant situational features of a moral decision then the ecological validity of such a study is compromised and the data may fail to generalize well to real moral judgments. In order to say more about how such a problem can be generated, we need to say more about the components of VR simulations and how those components can alter a subject's virtual experiences.

Virtual experiences and virtually-real experiences

It seems clear that subjects who engage with VR simulations of different sorts have correspondingly different kinds of experiences. In this section, we hope to provide a scaffolding to structure these experiences and to pick out one sort of experience in particular that we will call a 'virtually-real' experience as deserving special attention. Virtually-real experiences play important roles in the ecological validity of a study and raise heretofore underappreciated ethical issues.

Virtually-real experiences are those VR experiences that subjects treat *as if* they were real. Although the degree to which an experience is treated as virtually-real is dimensional, we will argue that virtually-real experiences are more likely to occur in environments high in what we will call *context-realism* and *perspectival-fidelity* than their virtually un-real counterparts. To give an example of a virtually un-real experience: a subject playing a first-person shooter which involves gunning down scores of virtual agents is highly unlikely (for reasons we discuss in greater detail below) to be treating those experiences *as if* they were real.[8] Compare that case with the carefully designed therapeutic instances of VR

[8] As this example should help to make clear, virtually-real experience is not identical with the concept of immersion or of 'presence' as it is understood by psychologists (Cummings & Bailenson, 2016; Ahn et. al., 2016). The first-person shooting experience may well be immersive and may include the feeling of 'being present' in the environment. However, if the experience contains the sort of context-unreal design elements that we discuss below, such a simulation is *unlikely* to produce a *virtually-real* experience (indeed the lack of subject trauma is perhaps the most tell-tale sign that she did not have virtually-real experiences of killing her virtual enemies).

simulations intended to help a patient overcome post-traumatic-stress disorder (Rizzo et.al., 2010). Subjects in these therapeutic contexts are much more likely to treat these simulated experiences *as if* they were real (i.e. as virtually-real). That such therapeutic simulations allow users to have full control over their virtually-real experiences is in fact one reason why they appear to be useful in gaining a sense of control over their triggers.

Although empirical literature on the question of virtually-real experience is nascent, we believe that the research that exists licenses us to make specific conjectures about the aspects of a VR simulation that make it more likely to produce virtually-real experiences in their subjects. We say "more likely" to produce a virtually-real experience because of the dimensional nature of virtually-real experiences (some experiences can be more virtually-real than others). We also wish to make clear that we believe that the constituents of a virtually-real experience are themselves dimensional. Virtually-real experiences are constituted by two major elements of a VR simulation which we call context-realism and perspectival-fidelity. We turn our attention to these components next.

To say that a simulation is high in perspectival-fidelity is to say something about the *structural* design features of the simulation and not about the *content* of the simulation itself.[9] Perspectival-fidelity refers to the degree to which a simulation accurately represents the subjective point-of-view of human observers. For example, the relative height of the virtual subject should roughly match the height of a typical user as should the depth-of-field made accessible to the subject. A simulation is higher in perspectival fidelity if only diegetic sound effects are used in the simulation. Non-diegetic soundtracks and sound effects are not aspects of our everyday perspective and hence work against perspectival-fidelity. They help to remind the subject that they are in a simulation. Similarly other elements like the representation of neurotypical color-vision and perception enhance the perspectival-fidelity of a simulation. For example, a simulation that presents a first-person point-of-view to the subject is higher in perspectival-fidelity than one presenting the same content from a third-person perspective. It will increase the probability that a subject will have virtually-real experiences in the context of the simulation.

Another element of perspectival-fidelity is captured by the degree to which the VR hardware itself is invisible to the subject's experience (Sanchez-Vives & Slater, 2005). For example, heavy head-mounted displays (HMDs) can detract from the perspectival fidelity of a virtual experience by intruding into the experience itself. The particular ways in which developers solve for locomotion in a VR simulation can also affect the degree of perspectival fidelity (e.g., teleportation is a less perspectivally-faithful solution than simply moving around in one's actual environment). Lastly, simulations which overlay

[9] As we will show, some structural features of a simulation will also affect its degree of context-realism. For example, the presence of non-diegetic voice-over in a simulation will likely detract *both* from the perspectival-fidelity of the simulation and from the degree of context-realism present in it.

information for the subject (as many games do) will be less perspectivally-faithful than those which lack such overlays or 'meta-information.'[10]

One consequence of perspectival-fidelity is that only simulations of human perspectives can be high in perspectival-fidelity. Simulations that claim to give subjects the ability to experience what it is like to be a cow at a slaughterhouse (Ahn et.al., 2016), for example, generate less perspectival-fidelity (and context-realism) than one that attempted to simulate what it would be like for *us* to visit a slaughterhouse. Simulations that claim to provide a user with the experience of being radically re-embodied or disembodied (Won et.al., 2015) are, for the same reason, lower in perspectival-fidelity than those which aim to replicate a typical human perspective.[11] To see why, we need to now say more about context-realism and its role in producing virtually-real experiences.

Context-realism, like perspectival-fidelity, is dimensional. An experience can be more context-real than one experience and less context-real than another. Context-realism refers specifically to the actual *content* of the simulation. Context-realism is a way of signalling the degree to which a simulated environment is believable to the experiencing subject as a possible/real environment. The more that the rules of a virtual world cohere with those of the subject's world, the greater the context-realism of that simulation.[12] For example, although photorealism appears to play only a small role in terms of the degree of context-realism in a simulation (Sanchez-Vives & Slater, 2005; Slater et.al, 2006), the *behavior* of virtual agents appears to matter much more in terms of whether a subject experiences the world *as if* it were real (Slater et.al, 2006). We conjecture that simulations containing virtual agents (non-player characters or NPCs) who demonstrate what some philosophers have referred to as moderate receptivity-to-reason (Fischer & Ravizza, 1998) will be experienced as more context-real than those whose virtual agents lack receptivity to reason. Additionally, a world in which subjects can perform superhuman actions is less context-real, and is likely to be experienced as less real, as one where subjects are bound by the same rules as in our world.

---

[10] It seems plausible to suggest that perspectival-fidelity will vary as our perspectives change. For example, in a world where most of us make use of augmented reality (AR) devices, a simulation that includes AR meta-information may therefore be *more* perspectivally-faithful than one without. Additionally, perspectival-fidelity will be dependent on an individual's *mode of conceiving the world*. A 'color-blind' subject may experience a simulation as higher in perspectival-fidelity than a neurotypical subject. A subject who is deaf may find elements of an experience less faithful than a subject who is not, depending on how well the simulation recreates their own perspectives.

[11] Here again we press the distinction between immersion, presence, and virtually-real experience. Won et.al. (2015) claim, for example, that subjects can successfully feel in control of a virtual lobster body (that the experience is one in which they feel present), we argue that such experiences are unlikely to be experienced as virtually real. Although presence and virtually-real experience may overlap (a highly present experience is also likely a virtually-real experience) they are not identical.

[12] Context-realism is subjective in one important sense. It is dependent on how a particular subject understands the possibilities available to her in the world. For example, a subject who believes in angels and demons is likely to experience a simulation of a Christian supernatural nature as *more context-real* than a subject who does not believe in angels and demons, belongs to a different religious group, or who is an atheist.

For this reason, most first-person shooting games are highly context-unreal even though they may be somewhat high in perspectival-fidelity. Why is this? In a typical first-person shooter, the player's screen is not only typically overlaid with significant 'meta-information,' the world itself behaves little like our own world. The player can survive injuries that would kill a normal human subject and can perform athletic feats that would be impossible for the player to perform. Such simulations also typically are designed so that power-ups and other items emerge from the corpses of defeated foes and virtual enemies do not typically behave as normal human antagonists would. Non-diegetic soundtracks are also common. This lack of perspectival-fidelity and context-realism may explain why consumption of, and enactive engagement, with violent simulations is treated differently by subjects than real-world violent experiences. Most simulated experiences like these are not experienced as virtually-real: they are not experienced *as if* they are real.

In the following section, we will apply these concepts to four recent VR judgment studies to show that only VR simulations that are high in perspectival-fidelity and context-realism are free from ecological problems because they produce virtually-real experiences. However, in the final section, we argue that such studies should be subject to the same scrutiny as non-simulated studies utilizing the same methods.

VR Research: Recents studies and ecological validity

We here briefly survey the extant literature on moral judgments using VR to show the explanatory value of perspectival-fidelity and context-realism and to explain how the concept of a virtually-real experience can generate ecological validity problems for VR simulations. In particular, we will contrast Navarrete et.al.'s (2012) VR trolley problem experiment and Patil et.al's (2014) virtual moral dilemmas against Slater et.al's (2006) virtual replication of Stanley Milgram's (1963) original obedience study and Pertaub et.al.'s (2002) virtual public-speaking studies. We argue that the design of the VR simulations of the latter, but not the former, pair of studies produced virtually-real experiences in their subjects. The former, we argue, have issues relating to their ecological validity that should lead us to be careful about generalizing to real-world moral judgments. In the following section we narrow our aim even further and examine two state-of-art studies on the same dimensions.

The VR simulations in Navarrete's (2012) and Patil's (2014) studies share many features. Both attempt to create a virtual environment that functions as an analogue to classic thought experiments. Such thought experiments have formed the backbone of psychological research on moral judgments (De Nucci 2012; Liao et.al., 2012; Pastötter et.al., 2012). Both studies also arrived at roughly similar results about moral judgment regarding virtual trolley problem simulations. Both claimed that subjects were more likely to make utilitarian judgments in virtual scenarios than in traditional vignette-based studies. We believe that the structural features of the VR simulations used in both studies should lead us to ask questions their ecological validity.

Both Navarrete and Patil's results can be better understood by examining how the degrees of perspectival-fidelity and context-realism present in their simulated environments may have combined to diminish the degree to which subjects experienced their environments as virtually-real. For example, Navarette's study contains several elements that diminish its context-realism: subjects materialized into their decision-making simulation immediately. They were not provided with any context that explained their materialization into the simulation nor did the simulation itself provide this information to them in a naturalistic way. These features not only reduce the context-realism of the simulation, they make it more likely to trigger a subject to view the simulation as a game, as unreal, and to adjust their moral expectations accordingly. Additionally, Navarette's subjects were not given any logical explanation for their presence in the environment (why were they there? why are they responsible for switching tracks?). This lack of context-realism is also likely to adjust the sort of perspective that subjects bring to the decision-making task.

Additionally, context-realism was further diminished by the fact that the virtual agents in the path of the simulated trolley lacked indications of rational agency (they behaved robotically throughout the experience without any sign that they recognized or understood the danger they were in). Lastly, the world that subjects were presented with was barren and inexplicable. Subjects materialized into a featureless desert that contained a single track splitting into a valley. The environment lacked any other notable features that would make it a believable space in which a moral-decision would be naturalistically made. All of these elements combine to diminish the degree to which subjects would engage with the simulated environment as virtually-real. The same features of the study should lead us to doubt that the moral decisions that subjects made in the scenario generalize to moral decisions subjects would make under more naturalistic conditions.

By contrast, Patil et.al.'s (2014) environments were more plausibly context-real. Their recreation of the trolley problem, for example, placed subjects into the middle of a train yard, in an urban setting. In comparison to the virtual (featureless) desert, the train yard provides a sensible space in which subjects might encounter trolleys and also people who might find themselves in the path of those trolleys. Subjects in the simulation could immediately see six virtual agents facing (and walking) away from them in the traditional five-vs-one arrangement of the classic trolley problem.[13]

Though the setting of Patil et.al.'s (2014) study represents an improvement over Navarrete et.al.'s (2012) environment, other elements of their simulation diminish the context-realism of their scenario. Subjects were again not provided with any reason for their presence on the train yard (are the subjects employed at the train yard? did they stumble upon it?). Such features may crucially affect (doxastically

[13] Patil et.al.'s (2014) study examines many different versions of trolley problem-like judgments (involving flaming cars, sharks, and so on). Though we only examine their trolley problem environment, our concerns with ecological validity extend, in very similar ways, to all of their virtual environments.

and sub-doxastically) how subjects understand their duties in the simulation. Additionally, and oddly, Patil et.al (2014) chose to position their subjects such that they were floating in an impossible position within the virtual environment itself (hovering above the tracks with a god-like perspective over the train yard). Such a design feature diminishes not only the context-realism of the simulation but also its perspectival-fidelity.

Additionally, the structure of the study itself may have contributed to a decrease in context-realism. Because Patil et.al. utilized a within-subjects design exposing subjects to 8 different virtual dilemmas (and 8 control 'dilemmas'), the context-realism of the study is potentially decreased (2014, p. 97-98). From the point-of-view of the subject, such an experience is largely unnatural and will, we argue, contribute to their viewing the experiences as virtually unreal.

Patil. Et.al.'s decision to graphically depict the deaths of virtual subjects (by collapsing into bloody piles upon contact with trolleys, burning cars, etc.) is likely to further detract from the context-realism of the subject's experience for two reasons. First, although the virtual subjects perish with some degree of (bloody) accuracy, none of the remaining virtual subjects subjects react in a moderately reasons-responsive way to these deaths (they remain motionless, unaware, and unconcerned by the carnage around them). Second, the experimental subjects themselves are not invited to take these deaths as bearing moral weight (because they are ushered quickly into more iterated dilemmas regardless of their choices in the prior scenarios) which we believe is likely to detract from their seeing the experience as virtually real (transforming it instead into something more like the experience of killing a virtual agent in a pc-game).

Adding to these concerns is the fact that Patil's (2014) study, though taking place in a three-dimensional virtual environment, did not require its subjects to wear a VR HMD. Instead, the study relied on traditional (i.e., screen bound) PC equipment. The screen-bounded nature of the simulation therefore also diminishes the perspectival-fidelity of the study. These features resulted in the creation of a simulation that is less likely to generate virtually-real experiences in the subject and more likely to lead the subject to treat the decision-making scenario as a decidedly unreal game or low-stakes task (something that a naturalistic trolley situation would be unlikely to do). As with Navarette et.al.'s (2012) trolley simulation, the ecological validity of these scenarios falls short of the mark if the intention is to accurately map our real-world decision-making mechanisms.

David-Paul Pertaub and his colleagues conducted an early, but important, study using VR simulated environments. In it, they hoped to assess whether self-reports of public speaking confidence were affected by the behavior a virtual audience. Although conducted when the technology was much more cumbersome to use (and program for) than systems available today, Pertaub et.al.'s (2002) design environment, and results, suggest that designing a simulation with perspectival-fidelity and

context-realism in mind can impact the ecological validity of VR data. For example, Pertaub describes their simulation in the following way:

> A virtual seminar room was populated with an audience of eight male avatars seated in semi-circular fashion facing the speaker, all dressed in suits as if attending a formal meeting. These avatars were continuously animated, displaying random autonomous behaviors such as twitches, blinks, and nods that were consciously designed to foster the illusion of a real-life presence…The room in which they were seated was a virtual counterpart of the real seminar room in which subjects completed their questionnaires. (Pertaub et.al., 2002, p.71)

The congruence between the virtual seminar room and the room that subjects were physically present in enhances the perspectival-fidelity and the context-realism of the virtual scenario. As we will see in Slater et.al.'s (2006) simulation, the graphically primitive audience managed to elicit strong emotional responses from subjects. This is because virtual subjects in both studies behaved in a reasonably realistic way. As described above, virtual agents were animated with the intention of recreating plausible human responses demonstrating what we earlier called moderate responsiveness to reasons. In response to the negative audience, Pertaub et.al.'s (2002) findings are instructive:

> [S]ubjective reports confirmed that the negative audience was a strong anxiety-provoking experience and that it was capable of generating a range of very powerful emotions in the speaker. (Pertaub et.al., 2002, p.77)

The high degree of context-realism and perspectival-fidelity of Pertaub et.al.'s (2002) simulation thus suggests not only that such simulations are capable of producing virtually-real experiences in subjects but that they are capable of generating psychological and emotional discomfort. This was especially true for subjects confronted with a negative virtual audience. We see even stronger evidence for these claims emerge from Slater et.al.'s (2006) simulations.

Like Pertaub et.al.'s (2002) public-speaking simulations, Slater et.al's (2006) VR recreation of Milgram's obedience experiment demonstrate greater context-realism and perspectival-fidelity than either Navarette's (2012) or Patil's (2014) virtual environments. We argue that data generated by Pertaub and Slater speak most strongly in favor of our thesis. For example, consistent with our claims about the relatively unimportant contribution of photorealism to a simulation's ability to produce virtually-real experiences, Slater's virtual environment, like Pertaub's, was graphically primitive. The virtual learner was represented as a low-resolution female avatar seated in front of the subject. The virtual audience designed by Pertaub was similarly composed of low-resolution avatars.

Slater's environment was otherwise designed to be high in context-realism. Subjects, wearing a VR HMD, sat in front of a real console, designed to mimic Milgram's (1963) original console, with which

they administered virtual shocks to a virtual learner. Subjects were also given a clear and logical explanation for their presence in the virtual environment (they were there to be part of a virtual learning study).

Although the virtual environment that subjects found themselves in was sparse, it was appropriate to the experimental context (i.e., a bare office-space). Additionally, the virtual learner behaved as if she were moderately reasons-responsive (she demonstrated pain in ways that a subject would expect any normal human to).[14] Perspectival-fidelity of the simulation was also high. Subjects were presented with a view consistent with what a normal subject would expect given the experimental design without additional meta-information or non-diegetic sounds. As a result of these features, subjects in this experiment reported strong virtually-real experiences that replicated much of Milgram's original (1962) behavioral and physiological results. Consider one subject's response to delivering a high shock to the virtual learner:

> [W]hen the Learner failed to answer at the 28th and 29th questions, one participant repeatedly called out to her 'Hello? Hello? ...' in a concerned manner, then turned to the experimenter, and seemingly worried said: 'She's not answering ...' In the debriefing interviews many said that they were surprised by their own responses, and all said that it had produced negative feelings. (Slater et.al, 2006)

None of Navarette et.al's (2012) nor Patil et.al.'s (2014) subjects behaved in such a strong manner to their virtual experiences. We argue that this is because the context-realism and perspectival-fidelity of Slater et.al.'s (2006) and Pertaub et.al's (2002) simulations worked together to generate the virtually-real experiences their subjects felt. Furthermore, Slater's (2006) study largely replicated Milgram's (1963) original obedience study, supporting our claim that simulations high in context-realism and perspectival-fidelity are likely to avoid the issues with ecological validity we have suggested impact simulations like the ones designed by Navarette and also Patil.

Before moving on to consider the ethical issues that may be raised by studies high in perspectival fidelity and context-realism, we want to briefly consider what we believe to be two very recently conducted VR moral dilemma studies. We do this, in part, to demonstrate that researchers with an interest in investigating moral judgments using VR are better served by designing their studies not only with 'presence' in mind but also with context-realism and perspectival fidelity.

---

[14] Although Slater et.al's study is a one-shot within-subject design, the fact that subjects had to make many decision to shock (or not shock) the virtual learner may appear to make it like Patil et.al.'s (2014) study. However, although in the case of Patil et.al. we argued that a visual depiction of the bloody consequences of their decisions worked to *detract* from the context-realism of their environment, we believe the opposite is true in Navarette et.al's environment. It is part of the nature of the Milgram environment itself to test how far subjects are willing to go when delivering increasingly painful shocks to a learner, the same is not the case in trolley problem-like scenarios. We thank an anonymous reviewer for helping us to make this distinction.

VR Experiments - The current state of the art

In the previous section, we examined four landmark VR judgment studies carried out between 2001 and 2014. In this section, we want to briefly examine two very recent VR judgment studies, published within the last twelve months, to make clear not only that the perspectival fidelity and context-realism of a simulation continue to have effects on the ecological validity of experimental work but also to make clearer why experimentally minded philosophers and psychologists would benefit from designing their virtual environments with these dimensions in mind.

Kathryn Francis and her colleagues have recently (2016) conducted a study to explore whether moral judgments ("is it morally acceptable to?") differed from subject responses to moral actions ("would you do it?") prompted in a VR environment. In order to do this, Francis et.al. constructed a VR simulation of the Bridge variant of the trolley problem. Although we find the question regarding moral differences between actions and judgments interesting, we believe that several elements of Francis et.al.'s virtual environment may have contributed to reducing the perspectival fidelity and context-realism of their scenario and thus introducing potential issues with ecological validity. For example, Francis describes one aspect of her experiment in the following way:

> After 30 seconds, verbal instructions informed participants that a trolley car was approaching ("Look behind you, a train is coming."). After a further 25 seconds, a second verbal dialogue then followed ("Hey I am too far away but if you want to save the people you could push the large person on to the tracks and derail the train. If you're going to push him, do it now, but it is your choice."). (Francis, et.al. 2016, p.5).

There are several issues with the way Francis et.al. created their Bridge scenario. First, as with Navarette et.al.'s (2012) iteration of the trolley problem, Francis et.al.'s world is left barren and featureless except for elements specifically related to subject choice. Although this may make some sense (given a desire to avoid introducing potential confounding stimuli) it simultaneously works against the context realism of the scenario. Second, the use of non-diegetic voice over to both inform the subject of the train *and* of the possibility of pushing the man onto the path of the train affect not only the perspectival fidelity and context-realism of the scenario but also, we believe, alters the Bridge variant in an important way by moving the source of the idea to push away from the subject.[15]

[15] In our own work, we have found it very difficult (we believe next-to-impossible) to design a context-real and perspectivally faithful VR simulation of the Bridge variant of the trolley problem. Specifically, it is difficult to design a version of this variant that does not break context-realism in a way that undermines ecological validity in the ways that Francis et.al's (2016) own version does. One pressing question with respect to the Bridge variant is how to create a simulation such that subjects spontaneously generate the idea that pushing the man onto the tracks (as the original thought-experiment is written) is a viable option to save the five. By resorting to non-diegetic voice-over to prompt the choice in their subjects, the context-realism and perspectival fidelity of the scenario are both diminished (we conjecture it is

Contrast Francis et.al.'s experiment with an even more recent study conducted by Leon Sütfeld and his colleagues (Sütfeld et.al. 2017). Sütfeld et.al. interested in exploring how human subjects respond to trolley problem scenarios that emerge while driving a vehicle. Sütfeld et.al. constructed their scenario in the following way:

> The participants were sitting in the driver's seat of a virtual car heading down a suburban road. Eventually, two obstacles, one on either lane, blocked the car's path and the participants had to choose which of the two obstacles to hit. Using the arrow keys on the keyboard, the participants were able to switch between the two lanes at all times, up to a point approximately 15 m before impact. (Sütfeld et al.2017, pX)
>
> To indicate how much time was left to make a decision at each point in time, a low-to-high sweep sound was played as an acoustic cue. (Sütfeld et.al. 2017, p.X)

Several features of the virtual environment in Sütfeld et.al.'s (2017) experiments can be analyzed in terms of their effects on perspectival fidelity and context-realism: first, their subjects are not provided with any context as to why they are in the vehicle, where they are going, and so on, thus decreasing, to an extent, the context-realism of the scenario (driving is typically, thought not always, a goal-directed activity). Second, subjects were instructed to use a keyboard (instead of a steering wheel) in order to move the car from one lane to another (it is unclear if subjects also controlled the speed of the car). Because control inputs affect both the perspectival-fidelity of the task as well as the context-realism, we believe this was a significant design choice with respect to the production of virtually real experiences. Third, Sütfeld et.al. used an audio tone to alert their subjects to the presence of a moral decision and to indicate how much time subjects had in order to make their decisions also works against both the context-realism and perspectival-fidelity of the scenario.

However, Sütfeld et.al.'s (2017) environment also contains many features that appear designed to increase the degree of context-realism and perspectival fidelity present in the task. For example, the scenario itself represents a fairly naturalistic first-person perspective of being behind the wheel of a vehicle, increasing the perspectival fidelity of the simulation. Also, participants were sitting during the entire task, lending additional perspectival-fidelity to the scenario given the task involved (driving). The authors do not provide additional information on the behavior of the virtual human agents in the experiment and so we cannot assess the degree of reasons-responsiveness they demonstrated. Nor do the researchers provide information about other aspects of the simulation (diegetic sounds, the nature

---

diminished to such a degree that it affects the generalizability of the data but this is merely a conjecture). We believe that this speaks to problems embedded in the structure of the Bridge variant of the trolley problem itself thought explaining exactly why lies beyond the scope of this paper.

of the car's movement, the exact contexts leading up to the creation of the dilemma, etc). As such, our assessment of the simulation's degrees of perspectival fidelity and context-realism are truncated.

From the information provided, it does appear that a study like Sütfeld et.al's (2017) aims to create a virtual environment aiming to produce virtually-real experience in a way that Francis et.al's (2016) virtual environment does not. Not only is Francis et.al's simulated environments sparse in the same way as Navarette et.al's (2011) (affecting the context-realism of the environment), like Patil et.al.'s (2014) study, Francis et.al.'s subjects are not provided with any naturalistic explanation for their presence on the bridge or why pushing the man is a viable option, or who the non-diegetic voice is meant to represent.

As these two studies help demonstrate, researchers appear to only be partially attuned to the importance of context-realism and perspectival fidelity in their virtual simulations. In the next, and final, section, we return to the dilemma we noted that experimentally minded philosophers and moral psychologists face when it comes to designing virtual environments. They can construct environments without paying specific attention to context-realism and/or perspectival fidelity, raising concerns about the ecological validity of their results, or they can aim to generate virtually real experiences in their subjects, raising possibly ethical concerns about the research itself. We turn our attention to these ethical concerns now.

Ethical Concerns: The equivalence principle

We believe that VR simulations should be subjected to greater scrutiny by IRBs than they currently receive. Psychologists themselves do not currently appear to appreciate the potential harms that context-real and perspectivally-faithful VR simulations can impose on their subjects. Instead, many appear to believe that the virtual nature of a simulation renders it incapable of causing any ethically significant harms to their human subjects.

Slater himself, for example, argued that VR simulations make possible "all social and psychological research where, for ethical or safety reasons, it is not possible to immerse experimental participants into the actual phenomena to be studied" (Slater et.al., 2006, p.6). Other researchers strike a similar tone: "[t]he addition of virtual environments allows affective neuroscience researchers to move beyond the ethical concerns related to placing participants into real-world situations with hazardous contexts" (Parsons, 2015, p. 10). Patil himself has argued that "[m]oral dilemmas are especially difficult to create realistically in laboratory settings because of the ethical problems associated with violent and harmful experimental situations. Virtual reality (VR) helps to take a step forward in studying such situations in a more ecologically valid manner" (Patil et.al. 2014, p. 95).

For the reasons we have generated in this essay, we believe that claims like these all make the same mistake. Simulations that are capable of generating virtually-real experiences deserve much closer scrutiny than they are currently given. A virtually-real experience of killing a virtual learner in a Milgram-style experiment is, as we saw above, capable of traumatizing its subject despite its virtual nature (we say more about this below). Pace Slater et.al. (2006), Parsons (2015), and Patil et.al. (2014), we argue that the mere use of a VR simulation should not allow experimenters to conduct virtually-real experiments that would be otherwise unethical to conduct in actuality. Because current guidelines on the use of human subjects do not yet appear to recognize the nature of virtually-real experience, we believe that it is especially urgent to enact safeguards while the technology is relatively young.

In this section we propose and defend the following principle, *The Equivalence Principle*, as a guide for IRBs to use while they assess the relative risk of harm in VR simulations.

> **TEP**: If it would be wrong to subject a person to an experience, then it would be wrong to subject a person to a virtually-real analogue of that experience. As a simulation's likelihood of inducing virtually-real experiences in its subject increases, so too should the justification for the experimental protocol.

The justification for TEP is straightforward and relies on a more basic ethical principle requiring parity of reasoning that like cases be treated alike. TEP recognizes that, while cases will always differ in many ways with respect to one another, two cases which are alike in morally relevant respects should be approached similarly. Such a commitment grounds many of our moral and legal practices. If a real-world experimental protocol would raise concerns about possible harm to its subjects (physical or psychological) then a simulation with an identical protocol that is likely to generate a virtually-real experience should be subjected to increased scrutiny (increasing in proportion to its degree of context-realism and perspectival-fidelity).[16]

Conversely, our analysis of the components of virtually-real experience can also be used to avoid causing unintentional or scientifically unnecessary harm to subjects. For example, adjusting the

[16] Specifically, we argue that TEP should modify how IRBs interpret 45 CFR 46 §46.111 (2009). In particular, we believe that the way that IRBs currently assess risk, as evidenced by the approval of studies like Slater's (2006) and Sütfeld (2017), indicate that VR experiences which include perspectivally faithful and context-real designs involving death or harm to virtual agents are not seen as posing ethically significant harms on participating subjects. The Belmont Report cautions researchers and IRBs to ensure that "[t]he method of ascertaining risks should be explicit, especially where there is no alternative to the use of such vague categories as small or slight risk. It should also be determined whether an investigator's estimates of the probability of harm or benefits are reasonable, as judged by known facts or other available studies" (National Commission 1978). We argue that the methods for assessing the probability of harm in such studies is not reasonable and that TEP should be used as a tool to help revise these estimates. We thank an anonymous reviewer for asking us to clarify this aspect of our argument.

context-realism markers in a simulation or altering its degree of perspectival-fidelity can alter the simulation enough so as to make it unlikely to generate virtually-real experiences (as Navarette et.al. (2012) and Patil et.al.'s (2014) experiments unintentionally do). Depending on the aims of the study, such a reduction in context-realism and perspectival-fidelity may lower the risk of accidental subject harm enough to make the study ethically permissible to conduct without necessarily invalidating data gathered.

One might propose at this point that the harms incurred under a VR simulation might be of a different kind than those imposed in real-life settings and hence that TEP may be too strong. For example, one might argue that, in many, most (or perhaps even all), VR simulations, it will be impossible for a subject to *actually* believe that they are harming virtual subjects and hence that, unlike many now classical case studies of unethical experimental protocols, subjects in VR experiments will not suffer from the harm of thinking that they are causing real harm to others.

Here it is important to attempt to distinguish between the sorts of harms that moral judgment studies (of any kind) might impose on their subjects. First, we can distinguish between the harm one might experience as a result of believing that they caused actual harm to subjects from the harm of learning potentially ugly things about oneself or what one is capable of as a result of their participation in a study.

With that distinction in mind, we can see that both harms exist in the context of Stanley Milgram's (1963) classic obedience studies. Not only did subjects in that study believe that they might have actually seriously harmed someone, they also learned that they may be capable of killing someone in ways that might harm or otherwise damage their self concepts. One thing to point out, however, is that this second kind of harm is separable from the first.

In a similar way, we can see why subjects in Mel Slater's (2006) VR simulation of the Milgram experiments might also be disturbed or traumatized even if they thought that they were not really hurting real people but only virtual people. Notice that the moral problems in Milgram's original study design do not disappear once subjects are immediately debriefed and told that nobody was really harmed, that the 'learner' was in on it all along. They would still have learned the same distasteful, arguably traumatizing, lessons as Milgram's own subjects and incurred many of the same harms.[17]

[17] In realistic-seeming nightmares, for example, we might be relieved, upon waking, to realize that "it was all a dream" and that "none of that really happened." However, while that fact might diminish (or extinguish) the first sort of harm (the harm of believing that you hurt someone), it often does little to diminish or extinguish the other sort of harm. For example, realistic dreams in which we disrespect our parents, cheat on our partners, or intentionally harm or kill others might continue to bother us long after we awake. If they do this, it is likely because we believe, or are at least worried about the possibility, that we might have learned something ugly or distasteful about the kind of people we are given our behavior in the dream. We thank [omitted] for pressing us on this issue.

Lastly we argue that we should not be so quick to accept the claim that subjects in a VR simulation know that their experiences are not real and thus that they are not capable of being harmed in our first sense. The subjects in Pertaub and Slater's experiments produced behavioral and physiological responses that are consistent with treating their experiences *as if* they were real.

There's a real sense, we think, in which they treated their experiences as if they were real even if they, in some other sense, had access to beliefs about where they really were and what they were really doing. Even though they may have had access to such knowledge, they did not seem to make conscious use of that knowledge while having their virtual experiences (subdoxastic effects might have been having much more impact on their experiences here than we might normally acknowledge). In that sense, we want to push against the claim that subjects do not really believe that they are making decisions in a virtual context. We acknowledge that this claim is grounded on difficult phenomenological and empirical dimensions and hence is subject to revision as more data is collected on the nature of virtual experience.

Although VR hardware and virtual simulations are not yet capable of producing experiences that are literally indistinguishable from real-world experiences, the reactions of Slater et.al. (2006) and Pertaub et.al's (2002) subject's demonstrate that the VR simulations available *today* are capable of generating experiences that are virtually-real enough to cause ethically concerning responses in subjects and which deserve more scrutiny by both researchers and IRBs.

Because context-realism is less dependent on photorealism than it is on the design features of a virtual simulation, it is possible to build perspectivally-faithful and context-real simulations even with relatively primitive VR hardware. This problem will only grow in urgency as the technology (and especially as the tools for building perspectivally-faithful and context-real simulations) become more powerful and user-friendly.

Conclusion

VR hardware is just beginning to enter the commercial marketplace. In 2016, the VR marketplace generated some $6.2 billion dollars in sales; this figure is expected to rise to $162 billion by 2020 ("Virtual Market", 2016). The technology holds great promise in terms of providing researchers ecologically valid data regarding the mechanisms that ground moral judgment and moral reasoning. They thus hold the potential to dramatically expand our knowledge of moral capacities by moving beyond the thought-experiment vignette paradigm that has dominated moral judgment research. However, this great promise carries with it equally large risks.

Many researchers are currently operating under what we argue is the false belief that simulations are incapable of harming their subjects in ethically significant ways. Even in its current, relatively

primitive, state we believe that evidence is mounting that virtually-real experiences can be harmful (indeed as harmful as real experiences) and that ethical guidelines, like TEP, need to be put into place to guide the use of studies making use of VR simulations on human subjects.

ACKNOWLEDGEMENTS

[Omitted for review]

References

Ahn, S. J., Bostick, J., Ogle, E., Nowak, K., McGillicuddy, K., & Bailenson, J. N. (2016). Experiencing nature: Embodying animals in immersive virtual environments increases inclusion of nature in self and involvement with nature. *Journal of Computer-Mediated Communication*, doi:10.1111/jcc4.12173

Alloy, L. B., and L. Y. Abramson. (1979). Judgment of contingency in depressed and nondepressed students: Sadder but wiser? *Journal of Experimental Psychology*, 108, 441–85

Avenanti A., Sirigu A., and Aglioti S.M. (2010). Racial bias reduces empathic sensorimotor resonance with other-race pain. *Current Biology*, 20 (11), 1018-1022

The virtual and augmented reality market will reach $162 billion by 2020. (2016, August 22). Retrieved from http://www.businessinsider.com/virtual-and-augmented-reality-markets-will-reach-162-billion-by-2020-2016-8

Cummings, J., & Bailenson, J. (2016). How immersive Is enough? A meta-analysis of the effect of immersive technology on user presence. *Media Psychology*, 19 (2), 272-309

Darley, J.M. and Batson D.C. (1973). "From Jerusalem to Jericho": A study of situational and dispositional variables in helping behavior. *Journal of Personality and Social Psychology*, 27 (1), 100-108

Earp, B., and D. Trafimow. (2015). Replication, falsification, and the crisis of confidence in social psychology. *Frontiers in Psychology,* 19:1–11

Fischer, J.M., & Ravizza, M. (1998). *Responsibility and control: A theory of moral responsibility*. Cambridge: Cambridge University Press

Francis, K.B., Howard, C., Howard, I.S., Gummerum, M., Ganis, G., Anderson, G., Terbeck, S. (2016). Virtual morality: Transitioning from moral judgment to moral action? *PLoS ONE*, DOI:10.1371/journal.pone.0164374

Fredrickson, B.L., & Branigan, C. (2002). Positive emotions broaden the scope of attention and thought-action repertoires. *Cognition and Emotion*, 19 (3), 313-332

Gendler, T. S. (2007). Philosophical thought experiments, intuitions, and cognitive equilibrium. *Midwest Studies in Philosophy*, 31 (1), 68-89

Kahneman, D. (2003). Maps of bounded rationality: Psychology for behavioral economics. *The American Economic Review*, 93 (5), 1449-1475

Lount, R. B., Jr. (2010). The impact of positive mood on trust in interpersonal and intergroup interactions. *Journal of Personality and Social Psychology*, 98 (3), 420-433

Milgram, S. (1963). Behavioral study of obedience. *Journal of Abnormal and Social Psychology* 67, 371–378

National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research. (1978). *The Belmont report: Ethical principles and guidelines for the protection of human subjects of research*. [Bethesda, Md.]: The Commission.

Navarrete C. D., McDonald M. M., Mott M. L., & Asher, B. (2012). Virtual morality: Emotion and action in a simulated three-dimensional 'trolley problem'. *Emotion*, 12 (2), 364-370

Patil, I., Cogoni C., Zangrando N., Chittaro L., & Silani G. (2014). Affective basis of judgment-behavior discrepancy in virtual experiences of moral dilemmas. *Social Neuroscience,* 9 (1), 94–107

Parsons, T.D. (2015). Virtual reality for enhanced ecological validity and experimental control in the clinical, affective and social neurosciences. *Frontiers in Human Neuroscience,* 9 (650)

Parsons, T.D., Gaggioli, A., Riva, G. (2017). Virtual reality for research in social neuroscience. *Brain Sciences,* 7 (42), doi:10.3390/brainsci7040042

Pertaub, D.P., Slater, M. & Barker, C. (2002). An Experiment on public speaking anxiety in response to three different types of virtual audience. *Presence Teleoperators and Virtual Environments,* 11 (1), 68-78

Ramirez, E. (2017). Empathy and the limits of thought experiments. *Metaphilosophy*, 48 (4), 504-526

Rizzo, A., Difede J., Rothbaum B.O., Reger, G., Spitalnick, J., Cukor, J., & Mclay, R. (2010). Development and early evaluation of the Virtual Iraq/Afghanistan exposure therapy system for combat-related PTSD. *Annals of the New York Academy of Sciences*, 1208, 114-125

Sanchez-Vives, M.V., & Slater, M. (2005). From presence to consciousness through virtual reality. *Nature Reviews: Neuroscience*, 6, 332-339

Schlenker, B.R. & Forsyth, D.R. (1977). On the ethics of psychological research. *Journal of Experimental Social Psychology*, 13 (4), 369-396

Slater, M., Antley A., Davison D., Swapp D., Guger C., Barker C., Pistrang N., & Sanchez-Vives M.V. (2006). A Virtual Reprise of the Stanley Milgram Obedience Experiments. *PLoS ONE*, 1, e39

Sütfeld, L.R., Gast, R., König, P., & Pipa, G. (2017). Using virtual reality to assess ethical decisions in road traffic scenarios: Applicability of Value-of-Life-Based models and influences of time pressure. *Frontiers in Behavioral Neuroscience*, 11, 122. DOI: 10.3389/fnbeh.2017.00122

Protection of human subjects, 45 CFR 46 §46.111 (2009)

Won, A.S., Bailenson, J., & Lanier, J. (2015). Homuncular flexibility: The human ability to inhabit nonhuman avatars. *Emerging Trends in the Social and Behavioral Sciences: An Interdisciplinary, Searchable, and Linkable Resource*, DOI: 10.1002/9781118900772.etrds0165

Zimbardo, P. 2004. A Situationist Perspective on the Psychology of Evil: Understanding How Good People Are Turned into Predators. In *The Social Psychology of Good and Evil*, edited by A. G. Miller, 21–50. New York: Guilford Press