UNIVERSITY OF READING

A Philosophical Discussion of the Implications and Limitations of Using Virtual Reality Technology (VR) as an "Empathy Machine"

By

Sarra Bouabdeli

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy

To

The Department of Philosophy, School of Humanities

University of Reading

September 2023

Declaration

I confirm that this is my own work and that the use of all material from other sources has been properly and fully acknowledged.

Sarra Bouabdeli (11 September 2023)

Acknowledgement

Abstract

This thesis engages in a philosophical discussion on "empathy", "virtuality", and the use of virtual reality (VR) technology as an "empathy machine". Here, I define empathy as the intentional activity (or skill) of recreating aspects of another subject's emotional experience in one's imagination to reflectively and "experientially" *understand* what another is feeling. As opposed to isomorphically appropriating another's feelings to oneself, I identify empathy as third-personally "feeling *with*" others. After exploring the narrow and pluralistic approaches to understanding empathy, I argue that there are compelling pragmatic reasons for adopting the pluralistic approach, the proponents of which prefer to highlight *varieties* of empathy instead of a sole conceptualisation of "empathy proper". As for virtuality, I subscribe to a third view that can be located between "virtual realism" and "virtual irrealism", in that I understand virtuality as a sui generis mode of technological actualisation, where psychophysiological illusions, of virtual presence and embodiment, coexist with veridical elements, such as virtual social objects, without causing a defect in users' rational judgment. My main contention in this research is that VR's multisensory affordances can be instrumentally utilised as a complementary *extension* (but never as a replacement) for offsetting some of the limitations in attaining interpersonal empathy through imaginative perspective-taking alone. After discussing this contention in more depth, I then attempt to address some of the recurrent challenges and criticism raised against VR's use as an empathy machine. Finally, I highlight some of the limitations in VR technology's capability to capture and transmit a *full* representation of others' lived experiences.

Table of Contents

General Introduction

Ever since it was invented in the 1950s and popularised in the 1980s by Jaron Lanier, virtual reality technology (VR) has been met with mixed attitudes and reactions. It was first described as a "consensual hallucination", a phrase borrowed from William Gibson's dystopian novel *Neuromancer* (1984). Before that, this tone was endorsed in the writings of Jean Baudrillard, especially in *Simulacra and Simulation* (1981) where Baudrillard warned that technologies that perceptually mimic our everyday reality will eventually lead to a widespread "disappearance" of the real, which would then be replaced by "the hyper-real, causing us to be sheltered from any distinction between the real and the imaginary" (pp. 2–3). The writings of Paul Virilio also conveyed a similar warning tone, accusing new technologies, such as VR, of collapsing spatiotemporal characters of reality into temporal ones, which, in Virilio's view, would cage humans in a "prison house" of illusory transcendence. Similarly, in one of his articles for Forbes magazine (2000), philosopher Robert Nozick drew hasty affinities between VR technology and his famous "experience machine" thought experiment (1974). He worried that (illusory) pleasures experienced in virtual environments (VEs), can prompt people to spend large portions of their time "inside" these environments, gradually detaching them from reality. In Nozick's words: "even if everybody were plugged into the same virtual reality, that wouldn't be enough to make its contents *truly real* [emphasis added]. So call me old-fashioned".

On the other side of the debate, philosopher David Chalmers, in his theory of virtual digitalism (2016, 2017, 2019, 2022), famously suggests that "virtual reality is a sort of *genuine reality*, virtual objects are real objects, and what goes on in virtual reality is truly real" (2017, p. 1). Chalmers then concludes that virtual experiences are as valuable, and sometimes even more optimal, than "real experiences". He problematically predicts that "VR may allow many experiences that are difficult or impossible in physical reality, such as flying, inhabiting entirely different bodies, and new forms of perception" (2022, p. 312). It may be worth noting that Chalmers writes about hypothetical VR technologies that do not yet exist, but which Chalmers professes will exist in the remote future.

Therefore, amidst philosophical debates on the notion of virtuality, we can at least distinguish two main positions; namely virtual *realism* and virtual *irrealism*. Proponents of these two positions disagree on metaphysical questions pondering whether virtual content is "real" or "illusory", and then draw upon their responses to address questions on *the value* of virtual experiences. In this thesis, I subscribe to a third view that can be placed somewhere

between these two positions; in that, I identify virtuality as a sui generis mode of actualisation where both psychophysiological illusions and veridical components coexist. I construct my conceptualisation of virtuality and VR technology drawing upon the writings of philosophers Pierre Lévy (1998) and Michael Heim (1998) who focus on providing an accurate description of VR-specific multisensory affordances and technical features. The main positive upshot of subscribing to this conceptualisation is that it adequately accounts for state-of-the-art, teleological implementations of VR technology, which would not otherwise be possible had VR been purely illusory, as virtual irrealists speculate.

Similar to Chalmers, I argue that it is coherent to ascribe multifaceted types of value to virtual experiences. Nevertheless, I contend that this should be done while avoiding the claim that virtual reality is a sort of metaphysically "genuine reality". Instead, I emphasise that virtual media; i.e. the multisensory content projected through VR hardware, is purely *representational* and only existent *qua* digital objects; i.e. computational processes and data constructs, generated through hardware and software dynamics. That is to say, experiences in VR are not ontologically indistinguishable from their physical counterparts. Making this clarification is a crucial precondition for examining the value of VR paradigms, *in themselves,* moving past hyperbolic claims.

Among the various, constantly evolving implementations of VR technology, this research specifically examines VR's usage as an *"empathy machine"*, i.e. as a medium for vividly, and often interactively, portraying aspects of others' lived experiences in first or third-person point of view (PoV). The phrase "empathy machine" was first used by the late American film critic Roger Ebert in a speech he gave in July 2005, where he described film as "the most powerful empathy machine in all the arts". To Ebert, an artistically valuable film is one that allows viewers to experientially understand "*what it's like* to be a different gender, a different race, a different age, a different economic class, a different nationality, a different profession, a different person, with different hopes, aspirations, dreams and fears…". A decade later, presenting his movie *Clouds over Sidra* (2015), film producer Chris Milk borrowed the phrase to describe VR technology as "the ultimate empathy machine". Despite the recurrence of this phrase, the term "empathy" remains semantically obscure.

In academia, there is a prevailing tradition for research papers on empathy to begin by highlighting the overwhelming lack of conceptual consensus on the meaning of the term. This has bred a great deal of confusion and even crosstalk amidst discussions on empathy across

disciplines and in interdisciplinary research. Nevertheless, we can distinguish, at least, two approaches (or more accurately attitudes) to defining empathy; the narrow approach and the pluralistic approach. After discussing the upshots of both approaches, I argue that there are compelling reasons for adopting the pluralistic approach, which comes with an even greater need for articulately explaining what one means by empathy in a given context. The main commitment of the pluralistic approach is to not claim that there is a single definition of empathy, considering its polysemic nature and the several etymological shifts that the term has undergone. With that said, we can identify *varieties* of empathy, among which I discuss *aesthetic empathy* and *interpersonal empathy*. The latter will be the centre of my investigation for the remainder of this thesis. With that said, I define interpersonal empathy as *an intentional activity where one employs imaginative perspective-taking to acquire a degree of experiential understanding of what another is feeling, from the other's perspective.*

Initially, this understanding of empathy deviates from its common connotation in ordinary parlance. For, empathy is often associated with benevolence and compassion. As I explain in this research, these concepts are more akin to "sympathy" than empathy. Furthermore, empathy is also usually characterised as a reflexive or instinctual "feeling" rather than an intentional activity. A feeling that is not (consciously) instrumentally steered towards accomplishing a cognitive achievement such as interpersonal understanding. However, recent studies in psychology and neuroscience have shown that even the underlying reflexive and automated mechanisms often associated with empathy can be either restrained or stimulated under the influence of context-related criteria; such as pain avoidance, material costs, and competition (Jamil, 2014, p. 5). Also, as explained in the writings of Franz de Waal and Mark Fagiano, empathy can be used as a tool for acquiring interpersonal understanding, which can function as grounds for moral (prosocial) behaviour, immoral behaviour, or instead abstention from behaviour or avoidance of possible sources of distress. That is to say, empathy should not be associated with a particular pattern of behavioural (moral or immoral) outcomes, leading some to characterise it as a morally neutral, epistemically valuable skill, as discussed by Diana Sofronieva (2018). In clearer terms, empathy *can* but does not essentially motivate behavioural consequences.

Next, in addition to conceptual issues on the notions of empathy and virtuality, there is a plethora of objections raised against using VR's use as an empathy machine, not to mention the strong sense of scepticism often expressed against this technological implementation. In this thesis, I argue that while it is true that it may be impossible to know exactly what it is like

to be "in another person's shoes", being exposed to some VR paradigms can still reinforce perspective-taking, widely recognised as a precondition for empathy. Perspective-taking is said to be reliant on – restrictive – mechanisms of human imagination. Using imagination alone to speculate what others are feeling can sometimes lead to misinterpretations or egocentric shifts, i.e. imagining *oneself* in another's situation without sufficient consideration of key distinctive features of the other person's experience. Hence, the main argument in favour of using VR as an empathy machine is that it provides verifiable, evidence-based input for acquiring a more epistemically reliable degree of interpersonal understanding of others' experiences, to some extent, overcoming challenges such as imaginative resistance, spatiotemporal distance, linguistic barriers, and socio-cultural differences. VR experiences also place crucial aspects of a target experience at the forefront of the VR user's attention, which can mitigate egoistic shifts.

As I will discuss in the course of this thesis, provided that a few conditions and requirements are satisfied, using VR as an empathy machine can bridge the communicative gap between individuals by offsetting (in varying degrees) the *deficit* in understanding the underpinnings of others' emotions, thoughts, and experiences. This implementation is particularly promising and worth exploring because it can pragmatically serve as a tool for enhancing clearer communication and conflict resolution, as well as initiating collective conversations on socio-cultural and political issues. Nevertheless, it is important to note that VR should be used alongside other epistemically reliable, secondary and primary sources of information and treated as an "add-on" (i.e. as an extension) but never a replacement for "natural" perspective-taking mechanisms or face-to-face interaction,

In more detail, the structure of this thesis will be as follows. The first chapter will seek to put forth a clear conceptualisation of the notion of interpersonal empathy. As mentioned above, empathy is commonly regarded as a cryptic term due to the overwhelming lack of conceptual consensus on what it signifies. Some philosophers, perhaps most notably Amy Coplan, worry that the multiplicity of distinct definitions of empathy, guided by different research objectives and methodologies, runs the risk of rendering research findings on empathy "incommensurable" (2012, p. 4). This is suspected to breed confusion and crosstalk about empathy. As a remedy, Coplan invites researchers to adopt a narrow conceptualisation of "empathy proper".

Alternatively, proponents of the pluralistic approach, such as philosopher and XR (extended reality) content creator Mark Fagiano (2016, 2019) prefer to set broader and more

flexible frameworks for categorising and studying varieties (and sometimes stages) of empathy, instead of a strict, unified conceptualisation. The main reason is that a narrow conceptualisation of empathy can always be contested by other ones, and it can eventually collapse into yet another conceptual variety of empathy. To disambiguate a few preliminary mechanisms enabling empathy, Frans de Waal offers a helpful metaphor which describes empathy as a multi-layered "Russian doll", "with at its core, the ancient tendency to match others' emotional states. Around this core, evolution has built ever more sophisticated capacities, such as feeling concerned for others and *adopting their viewpoint*" [emphasis added] (2009, p. 209). De Waal suggests that cognitively demanding abilities such as empathy have developed from more "primitive", neurological processes such as emotional contagion, initially defined by Elaine Hatfield, John Cacioppo, and Richard Rapson (1994) as: "a tendency to automatically mimic and synchronize expressions, vocalizations, postures, and movements with those of another person, and, consequently, to converge emotionally" (1994, 153–54). More will be said on emotional contagion in the first chapter of this thesis.

After exploring both the narrow and pluralistic approaches, I suggest that there are convincing pragmatic reasons for adopting a pluralistic understanding of empathy, which does not neglect or marginalise contextually differing – coherent – connotations of empathy. The first chapter also discusses two common mechanisms that are often said to enable empathy; namely emotional contagion and self and other-oriented perspective-taking. Following the recommendations of many proponents of the pluralistic approach, I avoid discussing empathy's inner workings in isolation, as sometimes done across the literature. Instead, I argue that human empathy is almost always manifested through complex cognitive and affective systems that work *in unison*. I suggest that separating these two quintessential dimensions of the phenomenon would only result in breeding more confusion and unnecessary equivocations.

Next, with a focus on the *object-directedness* of empathy, I discuss two varieties of empathy: aesthetic empathy (or *Einfühlung*) and interpersonal empathy. I, then, go on to focus on interpersonal empathy for the remainder of this thesis. Drawing upon insights from both the narrow and pluralistic approaches, I define interpersonal empathy as an intentional (in the sense of conscious or purposeful) ability that is reliant on imaginative perspective-taking for experientially understanding other persons' experiences (often involving emotions) from their perspective. At the end of the first chapter, I examine some of the limitations to (non-mediated) "perfect empathy", as discussed in the phenomenological tradition.

Pursuing similar explanatory objectives, the second chapter will critically discuss virtuality and VR technology. The main aim of this chapter is to show that virtual environments consist of both illusory and veridical components that coexist side by side, without causing a defect in users' rational judgment of the ontological structure of perceived virtual objects and events. That is to say, VR users are not "tricked" into *believing* that elements of virtuality physically exist. With that clarified, I define virtuality as *a sui generis mode of technological actualisation*, involving the transformation of finite, selected properties of a given (real or purely imagined) entity into a virtual form. Virtual objects and events are mainly generated thanks to human operators and specific types of hardware.

The heart of the matter is that the resulting virtual objects of perception that VR users virtually interact with can never be ontologically identical to their physical counterparts. I argue that physical properties, such as mass, gravity, and friction cannot fully exist in virtuality. Hence, virtual entities cannot be bearers of causal powers beyond the confines of virtual displays. At best, following Neil McDonnell and Nathan Wildman's terminology (2019), we can ascribe a type of pseudo-causality to virtual media, thanks to the way virtual objects *seem* to interact with one another, and with the user's virtual body representations (such as cursors and avatars).

After examining some ontological issues about virtuality, the second chapter underlines features that are specific to VR technology, following Michael Heim's notions of immersion and interactivity. Next, I critically discuss two psychophysiological *illusions* often associated with VR; namely the illusion of virtual presence and virtual embodiment. Then, following John Searle's theory of social ontology and Philip Brey's account of virtual social ontology, I elaborate on the generation of *veridical* "social objects" in virtual environments. By the end of the second chapter, which tackles several philosophical problems revolving around the notion of virtuality and VR technology, my account of virtuality consisting of both illusory and veridical elements will be complete.

Next, in the third chapter, I begin by highlighting some important disanalogies between Nozick's "experience machine" and the current status quo of VR technology, especially in its usage as an "empathy machine". The main contention supported in this chapter is that, unlike Nozick's hypothetical experience machine, VR produces real harm and benefits. To support these claims, I respectively discuss paradigms of "know-how" knowledge transfer and the intriguing case of virtual theft. Then, I focus on VR's implementation as an empathy machine

for visual and interactive storytelling (also referred to as "story living"). Importantly, VR narratives (and more generally virtual stimuli) have been observed *to heighten viewers' emotional engagement and behavioural responses* in real-time (Kisker, Gruber, et al., 2021; Kisker, Lange, et al., 2021a; Slater, 2009). The emotional responses that virtual experiences trigger often include "basic emotions", in Paul Ekman's terms (1992), such as fear, anger, joy, sadness, disgust, and surprise, as well as more complex psychological responses such as empathy. In the third chapter, I also outline a few exemplary fictional and non-fictional VR paradigms, specifically designed for enhancing users' empathic responses and perspective-taking. Such paradigms are reported to trigger users' heightened emotional responses, enabling ample communicative "transfer" of different subjects' emotional experiences among VR users.

Continuing to focus on VR-stimulated emotions, the fourth chapter attempts to rationalise VR users heightened responses triggered by virtual content they do not *believe* to exist in their immediate, spatiotemporal environment. I draw upon in Neil McDonnell and Nathan Wildman's view called "virtual fictionalism", where they argue that proper engagement with virtual experiences requires "make-believe" that virtual objects and events truly exist in a virtual (fictional or non-fictional) world. McDonnell and Wildman's view is inspired by Kendall Walton's theory of fiction, designed to account for audience responses to representational (art)works. The two proponents of virtual fictionalism use Walton's insights to tackle ontological questions about virtual objects and events (and to argue against virtual realism). However, they do not utilise Walton' valuable contribution to account for the intricacies of audience responses to representational works, to which VR experiences belong. The chapter, hence, attempts to explain VR-stimulated responses from the lens of Walton's theory of make-belief. Subsequently, I examine possible measures of fittingness that can be applied to evaluate the fittingness of VR-stimulated responses. And finally, I discuss some of the ethical worries as well as benefits that can arise due to VR technology's ability to stimulate heighted users' responses.

In the fifth chapter, I explore more focused objections against VR's ability to function as an empathy machine, raised by Grant Bollmer (2017), Joshua A. Fisher (2017) and Paul Bloom (2016, 2017). Bollmer believes that, at best, VR can only trigger what he calls "empathy circuits", i.e. "nonconscious reflex response that mirrors the experience of another in one's brain" (2017, p. 64). This conceptualisation seems to align with the notion of emotional contagion; the reflexive, contentless, psychophysiological tendency to mimic other people's expressed emotions or "catch" the prevailing mood or ambience of a given situation or place,

discussed in the first chapter. The main reason why Bollmer argues that VR only triggers quasi-empathy (i.e. emotional contagion) is that he believes developing genuine empathy requires "universal transmissibility" and full access to all aspects of others' emotional experiences. Problematically, Bollmer also suggests that genuine empathy necessitates assimilation with other individuals' personal, political, and ethical stances. Since satisfying these conditions is often not possible (nor always desirable), Bollmer denies VR's ability to stimulate genuine empathy and calls on VR users to adopt radical compassion instead, which in his words: "refers to an ethical stance that refuses any attempt to experience, or to completely understand, the experience of another, but instead embraces an openness to understanding and refuses assimilation into one's own self" (p.71). As this quote shows, Bollmer's intuitions seem contradictory, switching between denying the very possibility of interpersonal (experiential) understanding and contending that radical compassion necessitates it.

In response to Bollmer, I concede that it is true that our mirror systems may be activated while undergoing a virtual experience, as they would be in ordinary circumstances. However, this is not all VR narratives offer. The complementary propositional and non-propositional (i.e. representational) constituents of VR narratives also foster an *intellectual* understanding of others' experiences, instead of merely triggering certain neurological circuits. Furthermore, empathy is often characterised as an ethically neutral, epistemically valuable tool that involves several consciously employed mechanisms (other-oriented perspective-taking) and thinking processes. It is therefore farfetched to presuppose that VR-mediated empathy necessitates full assimilation. That is to say, one can preserve their identifying features and belief systems while also being able to resort to perspective-taking and acquire an experiential understanding of what characterises given emotional experiences, without radically endorsing the political or ethical views of another person.

For instance, empathically understanding a psychopath or serial killer's perspective or worldview does not require also subscribing to or sharing their moral and political stances, let alone full isomorphic assimilation. Bollmer's conceptual suppositions about empathy, when extended to the context of VR, would characterise the perfect VR empathy machines as automated brainwashing machines, which are highly (luckily) implausible. It would be more accurate to describe empathic understanding as occurring in degrees in a way that may or may not influence one's values and decision-making mechanisms, depending on an interplay among several criteria discussed in psychology and many other disciplines. Therefore, I suggest that

14

Bollmer is at fault for considering empathy an all-or-nothing phenomenon that necessitates full assimilation.

The next objection I tackle in the fifth chapter is presented by Joshua A. Fisher (2017). Fisher worries that through VR narratives, we can only empathise with the VR content creators' egocentrically-shaped understanding and representation of other subjects' experiences. He also claims that VR's "aesthetics" often disrupt (especially non-fictional) narratives, where the original features of a subject's experiences will always be doomed to be lost in transit when portrayed in VR, either due to the audience or content creators' misinterpretations or suspected misrepresentative reshaping of the original subject's narrative to fit the experiential paradigms of the viewer. Most intriguingly, Fisher reminds us that VR narratives are representational, i.e. only experienced second-hand, either through animated virtual media or photographic content. He adds that attempting to capture and recreate subjects' emotional experiences will always be determined, and limited, by the content creator's degree of experiential understanding of the experience in question. The worry is that it is never a primary subject's emotional experience being represented in VR narrative, but only the content creator's interpretation that is supposedly liable to misinterpretations and mischaracterisations. Fisher's critique speaks directly to the epistemic reliability of virtual representations in transmitting aspects of other persons' emotional experiences.

In an attempt to respond to Fisher's worries, it can be argued that even in the case of face-to-face communication, emotional expressions are not always fully or even correctly interpreted and understood among individuals. As I mention in the first chapter, there are limits to "perfect empathy"; complete understanding of others' experiences from their viewpoint. Nevertheless, I concede that the issue Fisher raises on the central role that VR storytellers and content creators play in the interpretation and representation of target emotional experiences is a legitimate one. Evaluative questions on the "process reliabilism" of how virtual representations are created are beyond the scope of this research. But at least, it can be suggested that Fisher's worry further emphasises the significance of empathic understanding in the stages preceding the production of a given VR narrative. This is because the degree to which a content creator empathises with a subject's emotional experience shapes the final product and determines, to a large extent, the degree to which viewers can experience empathy. Therefore, the complex topic of empathically understanding others' emotional experiences is not only relevant to philosophers and scholars, but also to VR content creators who aspire to create narratives pursuing these objectives.

That is to say, the significance of this research can be made visible in terms of highlighting the promising potential as well as the limitations of VR narratives directed towards stimulating empathy. This endeavour is intended for providing a theoretical blueprint, so to speak, for identifying and subsequently tackling some of the most persistent issues in this genre, building upon insights on empathy and virtuality in philosophy of art and philosophy of technology. It is hence evident that developing a better understanding of empathy and virtuality is a crucial preliminary step for creating more interpersonally valuable VR narratives that successfully stimulate empathy.

Next, it is important to note that absolute empathy cannot be realistically met by, in Paul Bloom's words, "strapping a VR head-mounted device around someone's head" (2017). Nevertheless, the concept of absolute empathy is itself too idealistic to be pursued. A more attainable goal is not to fully appropriate others' emotional experiences to oneself, nor to experience them as one's own. Instead, it is to acquire an epistemically reliable, multisensory, and memorable input through VR narratives, to more effectively understand *aspects* of others' experiences that may not have been known or considered before. Profound interpersonal understanding can sometimes enable prosocial behaviour or cause alterations in VR users' judgment and belief systems. But, as this research explains, these outcomes remain contingent on subject-specific criteria; neither VR nor empathy can produce morally sound and prosocial outcomes by default.

Finally, as mentioned above, it is advisable to consider empathy as an ongoing affective and cognitive activity of pursuing an interpersonal understanding of other persons (and even animals') multisensory experiences. Therefore, interpersonal empathy does not end, but only begins, with the use of VR as an "empathy machine". Such implementation is merely complementary to primary and secondary epistemic sources devoted to attaining this type of cognitive achievement. Importantly, considering the concrete psychological effects produced by VR technology, especially the generation of heightened emotional responses, VR should not be dismissed as an equivalent to Nozick's hedonistic and escapist experience machine, but instead, its ramifications should be taken seriously, specifically when placed in the ethical context.

Chapter One: Conceptual Problems Surrounding the notion of "Empathy"

1. Chapter Overview

When seeking a clear, unequivocal conceptualisation of empathy, one is instantly faced with an overwhelming lack of conceptual consensus on the term. This issue has made it customary for academic writings on empathy to engage in terminological disputes, consequently producing a myriad of distinct formulas of definitional necessary and sufficient conditions for empathy. In ordinary parlance, empathy is sometimes confused with sympathy and associated with concepts such as benevolence and compassion. As mentioned in the introduction, I instead identify empathy as an intentional ability (or activity) to employ imaginative perspective-taking for understanding another subject's – emotional – experiences. Metaphors such as "walking a mile in someone else's shoes" are hence frequently applied when speaking of empathy. In *Mimesis as Make-Believe* (1993), Kendall L. Walton describes empathy (towards fictional characters) as: "an ability to look at things more purely from [a character's] points of view, from a perspective relatively uncontaminated by his [i.e. the reader's] own personal concerns" (1993, p. 237). This provides a useful starting point for understanding empathy as an ability that is greatly dependent on imagination and representation, through which we can experientially share certain aspects of others' experiences and worldviews. Several researchers, especially psychologists and neuroscientists, have found evidence of *mirroring* systems, shared by both humans and some non-human animals.[1] As this chapter discusses, these mirroring systems are said to developmentally underpin our ability to be empathic towards various objects of empathy.

Despite the prolific empirical and theoretical studies on empathy, striking conceptual disagreements on its connotation persist, pushing us to wonder which aspects are most essential for effectively developing empathy. In other words, which elements of others' emotional experiences should a subject simulate or understand (in an objectual[2] manner) to be legitimately

---

[1] See Rizzolatti and Craighero (2004) for an initial discussion of research findings on the hypothetically automatic, neurophysiological mechanisms, known as "mirror neurons" or "mirror processes", first discovered in the premotor cortex of macaque monkeys. As the term "premotor" implies, these neurons are said to fire prior to performing an action and also while observing an action being performed by another agent.

[2] To say that empathic understanding is "objectual", I mean that it involves object-directedness, in that sense that it is always directed towards a given object of empathy. It also means that it involves the understanding of "information chunks" as opposed to isolated propositions about another individual's emotional experience. See Kvanvig (2009) for a more in depth discussion of objectual understanding, where he shows that this type of understanding occurs in degrees and does not essentially entail a truth condition.

accredited with empathy? Is it sufficient for empathy that one simply experiences reflexive emotional contagion? Or does a subject need to develop a more intellectual understanding of others' emotional states, which would involve, for instance, understanding the object-directedness of a given emotional state and the *reasons* triggering it? Conversely, is a detached intellectual understanding of another's emotional state sufficient for empathy? Or is it necessary to emotionally engage in a sort of affective *sharing* (matching, simulation, or mimicry) of another's emotions for one's empathic response to be complete?

Questions on empathy's moral significance have also been raised across the literature. For instance, philosopher Michael Slote (2010) considers empathy the "cement of the moral universe", which helps us constitute moral approval or disapproval and make sense of moral claims, utterances, and judgments. Although these are some of the *functions* that empathy can be used for, it is crucial to point out that the notion in question should not essentially be associated with good moral judgement per se, nor with prosocial behaviour. As discussed by Fagiano (2016) and de Waal (2009), we can unproblematically identify a "dark side" of empathy, where it can be applied as a tool for the dissociative understanding of others' emotional experiences for the sole purpose of afflicting pain or suffering. We can find examples of this kind of "empathy" in serial killers' and psychopaths' emotionally detached, theoretical understanding of others' emotions, which is devoid of any signs of remorse or emotional engagement. These kinds of ethically problematic paradigms also lead us to question what empathy towards such subjects would look like, especially concerning sharing aspects of these emotional experiences as a potential prerequisite for developing empathy. Fagiano asserts that *context matters* when discussing the complex, multidimensional notion of empathy. These questions, and perhaps many others, are most relevant to content creators seeking to use VR technology as a tool for simulating empathy through fictitious as well as veridical (for example, journalistic) storytelling.[3]

With that said, this chapter will, therefore, be divided into four sections. The first section will discuss Coplan's conceptualisation of empathy proper in details and argue that it is more candid to collapse this conceptualisation into "cognitive empathy". To Coplan, what distinguishes emotional contagion, from so-called higher-level cognitive processes, is that the former is an automatic, involuntary *reflex,* that can only be triggered by direct, sensory engagement with visual and/or auditory stimuli, unlike empathy proper, which essentially

---

[3] For some examples, see de la Peña et al (2010).

involves *cognitive evaluation* and active imaginative reflection. Hence, Coplan argues that emotional contagion cannot count as empathy because it is devoid of any traces of cognitive evaluation.

After closely looking into Coplan's account, I suspect that there is no obvious rationale for incorporating what she terms "affective matching" in the conceptualisation of empathy proper, while also excluding emotional contagion, other than "dovetailing with recent psychological and neuroscientific research" on empathy (Coplan, 2012, p. 4). It might also be noteworthy that most studies and experiments in psychology and neuroscience that Coplan references (For example, Simone G. Shamay-Tsoory, Judith Aharon-Peretz, and Daniella Perry (2009) and Tania Singer (2006)), interchangeably use the term empathy with *perspective-taking*, which is "cognitive empathy" in a nutshell (Spaulding, 2017). So far, it can be argued that it is not problematic to interpret Coplan's conceptualisation of "empathy proper" as cognitive empathy unless we insist that the latter is the *only* legitimate variety of empathy there is.

Next, the second section of this chapter seeks to provide a rationale for subscribing to a pluralistic understanding of varieties of empathy instead of a single definition. I make sure to show that this comes with an even greater need for clarifying what one means by empathy to avoid crosstalk and facilitate a legible exchange of insights on the rich polysemic concept in question. Drawing upon Mark H. Davis's categorisation of the mechanisms and processes that are said to instantiate empathy (2006), as well as de Waal's "Russian dolls' model" of empathy, I discuss emotional contagion and self and other-oriented perspective-taking as key mechanisms for broadly defining empathy. That is to say, without a reference to the interconnected inner workings of these two processes, we might not be able to attain a (developmentally) thorough understanding of human, interpersonal empathy.

The third section will then discuss the object-directedness of two varieties of empathy; namely, aesthetic empathy (also referred to as *Einfühlung* or "feeling *into*"), understood as a tool for kinaesthetic and embodied projection of imagined bodily movements and sensations into perceived inanimate artefacts**,** such as artworks, landscapes, and fictional characters (Pinotti, 2017). Theorist Theodor Lipps also considered this early connotation of empathy as a *sine qua non* of aesthetic experience, understanding it as the projection of inner bodily feelings and movement into art forms and even nature (Jahoda, 2005; Lanzoni, 2012, 2018). In this

context, I discuss empathy towards fictional characters with reference to emotional contagion and perspective-taking discussed in the previous section.

In the fourth section, I focus on interpersonal empathy; practised in the social context of human interpersonal relations and communication and make sure to highlight some of its epistemic and phenomenological limitations. As shown above, interpersonal empathy heavily relies on *imaginative*, third-personal perspective-taking, which is bound to be limited by the empathiser's preconceptions and experiential paradigms. As Coplan argues, individuals often tend to assume more similarities between themselves and others than typically exists, causing them to reduce empathy into a projective process, solely characterised by self-oriented perspective-taking. At the other end of the spectrum, as shown in (Stueber, 2006), "non-projectionism" (p. 205) characterised by overemphasising individual differences such as cultural, historical, or biological differences can also obstruct empathy by considering the other as completely "foreign", marginalising the common features and cues that can otherwise aid the empathiser in forming a more accurate experiential understanding of others' states. Overall, this highlights the limitations of imaginative perspective-taking as an epistemically reliable tool for developing a *full* understanding of others' emotional states and the reasons behind them. In Karsten R. Stueber's words, empathy is not an epistemically self-verifying process. Mechanisms of inner imitation and imaginative perspective-taking are indeed shown to be frequently utilised for making sense of others' emotional states. However, the limitations of these mechanisms show that there is a need to use them with caution and preferably alongside additional epistemic sources for reinforcing accurate empathic understanding.

Furthermore, as Dan Zahavi suggests (2014), others' emotions can only be "accessible" to us third-personally through finite channels of communicative expression, such as facial and linguistic expressions. Nevertheless, even though we do not encounter others' emotions first-personally as we do our own – through methods such as proprioception and introspection – we still share "basic emotions" with others, at least making some aspects of others' emotions easily identifiable to us. These basic emotions are explored in Paul Ekman's line of research on "affect programs" (1992, 2005). Additionally, despite the asymmetry between our introspective, first-personal access to our own emotions and the third-personal restrictedness of others' emotions, the latter is still presented to us with a great deal of "givenness". For instance, when we perceive someone's laugh, cry, or frown, we do not see these emotional expressions as mere facial distortions and grimaces, instead, we see the joy, sadness, or anger *in* these expressions. Zahavi further explains that this asymmetry should be considered a feature and not a bug in our

interpersonal, empathic interactions. He states: "It is precisely because of this difference, precisely because of this asymmetry, that we can claim that the minds we experience are *other*…" (2014, p. 130). Nevertheless, it is important to point out that some features of others' experiences remain beyond the reach of empathic understanding. These features, for instance, include physiological sensations, such as physical pain, which cannot be fully made available to us by employing emotional contagion or imaginative perspective-taking.

In the concluding section of this chapter, I foreshadow that due to the epistemic and phenomenological limitations outlined above, technologically-mediated empathy can support a more tangible, evidence-based empathic understanding of others' emotional experiences and their underlying reasons. Through VR state-of-the-art storytelling, viewers are offered an expressively vivid explanation of other (fictional and non-fictional) subjects' perspectives, overcoming some of the restrictions of imaginative perspective-taking. However, this claim is widely contested across disciplines as I will discuss throughout this thesis.

2. The Narrow Conceptualisation of "Empathy Proper"

As Frédérique de Vignemont and Tania Singer famously write: "there are probably nearly as many definitions of empathy as people working on the topic" (2006, p. 435). More concretely, after amassing a critically appraised review of literature, examining some of the most influential research papers on empathy, Benjamin Cuff and his colleagues report that there are at least *43 distinct conceptualisations* of the term across psychology, neuroscience, marketing, politics, philosophy, as well as other disciplines (2016, pp. 146–147). Nevertheless, this apparent terminological disagreement on empathy – especially within interdisciplinary research – did not make research on the topic any less dynamic. For instance, in their paper "Towards a Consensus on the Nature of Empathy" (2021) Jakob Hakansson Eklund and Martina Summer Meranius report that in the last hundred years, approximately ten thousand scientific articles on empathy have been published (most of them in the 21st century).

Amy Coplan finds some aspects of the status quo of research on empathy to be quite problematic. Coplan worries that philosophical discussions, as well as some empirical findings, clustered under the header of empathy, have become "incommensurable" (2012, p. 4) and very difficult to keep track of. Coplan worries that this threatens to obstruct the flow of productive exchange of insights on the topic across disciplines. Put simply, distinct theoretical frameworks and research methodologies have produced numerous conceptual varieties of empathy, and continue to generate research findings that may not identically fit with one another. This, in

turn, created an ominous shadow of ambiguity, vagueness, and even crosstalk, often present in discussions on the phenomenon.

Sharing some of Coplan's complaints, some content creators and filmmakers, for instance in the field of VR content design, raise similar concerns about empathy being rendered a vague "buzzword" (Siegel & Dray, 2019) due to the many connotations ascribed to the term. For instance, a great number of filmmakers attending the 2017 Tribeca Film Festival – an annual event supporting the innovative integration of technology in filmmaking – told Adi Robertson that empathy (as a term) has been overused in the industry to the point of confusion and that sometimes it is merely utilised as a publicity stunt, without any profound or clear meaning (Robertson, 2017).

Proposing a solution to these issues, Coplan invites researchers to collectively adopt her narrow conceptualisation of what she recognises as "empathy proper" which she contrasts with "pseudo-empathies" (2011, p. 44). Coplan, hence, defines "genuine" empathy as "a complex imaginative process in which an observer simulates another person's situated psychological states, both cognitive and affective while maintaining clear self-other differentiation" (ibid., p. 1). Although Coplan assertively contends that emotional contagion as well as any sort of accidental emotional congruence are neither necessary nor sufficient for empathy (ibid., pp. 45–53), she still insists that there exist some kinds of lower-level affective processes that are essential for empathy; namely *affective matching* that is particularly reached through other-oriented perspective-taking, where an observer comes to imaginatively (and intentionally) simulate qualitatively identical emotions (or affects) of another person, through imagining herself in the position or situation of another person, *as other*.

Next, Coplan dismisses self-oriented perspective-taking as a type of pseudo-empathy and not empathy proper, because it often leads to personal distress and egocentric misattributions and errors in interpreting others' emotional states, caused by "our natural tendency to assume greater similarity between self and other than typically exists" (2012, p. 10). Therefore, Coplan rejects the characterisation of self-oriented perspective-taking as a reliable mechanism for empathy. This implies that she expects that there *ought* to be specific kinds of behavioural outcomes resulting from empathy, such a true and accurate understanding of others' emotional experiences and sharing others' emotions through precisely determined routes and not through others. In response, it can be suggested that since empathy relies on imperfect mechanisms, chiefly reliant on an imaginatively acquired understanding of

others' emotional experiences through perspective-taking, it might be too demanding and idealistic to suggest that it should always constitutively entail truth, accuracy, and identical emotional matching. Our "reading" of others' emotional expressions can be erroneous, even in the case of employing other-oriented perspective-taking. This description is also coherent with Coplan's suggestion that empathy occurs in *degrees*, instead of being an all-or-nothing phenomenon. Furthermore, self-oriented perspective-taking, in some cases, can be a tool for attaining a more accurate understanding of others' emotions, considering that we share "basic emotions" with others, as argued in Paul Ekman's theory of "affect programs", which Coplan acknowledges in her writings on empathy

To iterate, Coplan proposes that genuine empathy; a "unique kind of experiential understanding", requires three essential features: a) affective matching, b) other-oriented perspective-taking, and c) self–other differentiation. To Coplan, all these features are necessary but none of them is sufficient on its own for ascribing empathy to a given subject. To Coplan, true empathy occurs through

> …Other-oriented perspective-taking, where an observer imagines a target's situation, experiences, and characteristics as though he were the target. And an observer maintains self–other differentiation only if he continuously represents himself as distinct from the target, thereby avoiding confusion about their respective situations, experiences, and characteristics. (2011, p. 6)

Iff all these requirements are satisfied, Coplan asserts that, through genuine empathy, "we can experience what it is like to be another person" (ibid).

With that said, I believe it is particularly confusing that Coplan excludes emotional contagion, a commonly recognised, basic mechanism for empathy, *while also* arguing that affective matching (through other-oriented perspective-taking) is necessary for empathy proper. What is even more problematic is that Coplan's account requires that the type of affective matching necessary for empathy should involve qualitatively identical affective matching. Following Dan Zahavi, we can argue that the fact that we don't always *share* (or mimic) a target's emotional expressions doesn't make our response any less a case of empathy. As in one of Zahavi's examples: "to empathically understand that your friend loves his wife is quite different from loving his wife yourself; it doesn't require you to share his love for his wife" (2014, p. 150). Additionally, demonstrating identically matching emotional expressions to those of a target is not sufficient proof that one's response is empathic (or not empathic), regardless of the methods through which this matching has been reached, for this can still be a

mere indication of projection (purely achieved through self-oriented perspective-taking) or even hollow pretence. As mentioned before, reactive responses that personally belong to a given observer (or empathiser) towards a target are beside the point when discussing conceptual questions about empathy, unless we consider specific behavioural consequences as conceptually *constitutive* of empathy, which is a move that Coplan does not explicitly make.

Another issue in Coplan's account of empathy proper revolves around her vague, even procrustean, identification of affect. In her words:

> Affect is a broad category encompassing multiple mental states, all typically thought to involve feelings and some degree of physiological arousal. Emotion and mood are paradigm cases of affect. *Affective states are not necessarily directed at specific objects nor do they necessarily involve cognitive evaluations or appraisals* [emphasis added]. (2012, p. 5)

This characterisation of emotions is certainly dismissive of the ongoing debate about the desiderata of emotional experiences (Scarantino & de Sousa, 2011). In this vein, several questions are posed about the essential, constitutive elements of emotions. What does an emotional experience include at a minimum? Is it the physiological, autonomic, or bodily sensations? The motivation or cause behind experiencing a given emotion? Or the object-directedness of the emotion itself, in being *about* something or someone for instance? Or rather, is it the phenomenological experience, shaped by an amalgam of all of these features? Follow-up definitional questions about emotional experiences are also raised across the literature on the dynamically intertwined neurological systems, making it possible for some to experience a given emotion and impossible for others to do so in cases of neurological impairment. Of course, responses to these questions are objective-specific, i.e. steered by various methodologies, experimental frameworks, and research objectives.[4] However, it is undeniable that seriously acknowledging the complexity and multidimensional nature of emotional experiences is pivotal to addressing conceptual questions about empathy.

With that said, Coplan seems to conflate emotions and moods by claiming that affective states, overall, are "not necessarily directed at specific objects nor do they necessarily involve cognitive evaluations or appraisals" (ibid). This is an accurate characterisation of moods, which are indeed paradigms of objectless affects, but not of the broader category of emotions. For

---

[4] For sample responses to some of these questions, see James (1884), Lazarus (1991), Prinz (2004), and Solomon (1973).

instance, Robert Solomon alternatively argued that emotions involve cognitive evaluations and appraisals. To illustratively support his contention, he writes:

> If I do not find my situation awkward, I cannot be ashamed or embarrassed. If I do not judge that I have suffered a loss, I cannot be sad or jealous. I am not sure whether all emotions entail such judgments; moods (depression and euphoria) [and also phobias] surely present special problems. But emotions, in general, do appear to require this feature: to have an emotion is to hold a normative judgment about one's situation" (1973, p. 27)

Solomon's description of emotional experiences provides useful insights on a few elements, among potential others, which we need to consider when attempting to gain an experiential understanding of others' *emotional* experiences. More relevantly, it establishes that Coplan's condition that individuals *must* identically match each other's emotional states when striving to be empathic may be too demanding, and perhaps not even necessary, as shown in Zahavi's example above. Considering these points, we can suspect that there might not be a clear rationale for including affective matching in the conceptualisation of empathy proper, while also contending that emotional contagion is not necessary for empathy, perhaps other than "dovetailing with recent psychological and neuroscientific research" on empathy (Coplan, 2012, p. 4).

In this vein, Coplan claims that philosophers, in particular, are too often guilty of neglecting lower-level affective processes, which she acknowledges as crucial for accounting for human, mental states. She also recommends that whenever possible, philosophical theories should be constrained by empirical research and that "while we as philosophers should never accept the conclusions of empirical scientists uncritically, to ignore them is to render our work less relevant, less credible, and, ultimately, less meaningful" (2012, p. 4). Coplan then proposes to bracket processes such as emotional contagion and self-oriented perspective-taking from the overarching context of experiencing empathy. In some passages, however, she leans toward acknowledging that emotional contagion can be *relevant* to empathy, although she does not explain in what ways this relationship can be established.

Coplan then proposes that affective matching – reached through other-oriented perspective-taking – is causally (neurologically), and phenomenologically *distinct* from emotional contagion, drawing upon experiments in psychology and neuroscience that, contrary to Coplan, use the term "empathy" interchangeably with "perspective-taking" or other reductive processes. In other words, most empirical studies on empathy prefer using a

pluralistic understanding of the term, investigating the notion in terms of its varieties and levels, rather than identifying it as an all-or-nothing phenomenon that only occurs in clearly demarcated regions in the brain. Coplan emphasises that the empirical studies she references are consistent with her intuition that emotional contagion engages different areas in the brain than those activated when performing tasks that require intentional perspective-taking. For example, Tania Singer reports that "the abilities to understand other people's *thoughts* and to share their affects [through emotional contagion] display different ontogenetic trajectories reflecting the different developmental paths of their underlying neural structures" (ibid., 855). However, this does not at all provide any insight into what Coplan terms "affective sharing". That is to say, this study does not show that there exists a sui generis process that occurs in special brain regions where an empathiser identically matches another target's emotional state specifically through *other*-oriented perspective-taking. Also, it does not deny that emotional contagion can be involved (or not involved) in experiencing empathy. Another issue is that psychologists and neuroscientists often *presuppose* different theoretical frameworks for defining empathy to precisely pinpoint what they are looking to "measure" before conducting a given experiment. This can be taken as evidence that empirical findings on which regions are most active when experiencing empathy, in whatever way a scientist chooses to define it, might not weigh in conceptual debates on defining what empathy is and is not.

Therefore, there exists no conclusive empirical evidence for the special type of affective matching that Coplan claims to be exclusively reached through other-oriented perspective-taking. In hopes of mitigating the equivocation and conceptual ambiguity surrounding empathy, I hence postulate that "affective matching reached through other-oriented perspective-taking", can simply be compressed and described as "perspective-taking", at the risk of collapsing Coplan's account into yet another variety of empathy, known across the literature as *cognitive empathy*. This is unproblematic provided that we acknowledge that cognitive empathy is not the only level or variety of empathy there is. To Coplan, strictly focusing on higher-level, cognitive empathy is at fault for not taking into consideration lower-level affective processes, which Coplan already deemed essential for providing a full – philosophical – account of empathy. Ironically, however, Coplan does not thoroughly attest to the role of these lower-level processes, especially when discussing emotional contagion, *in the context of empathy.* To fill in this explanatory gap, I suggest, in the next section of this chapter, that we resort to adopting a pluralistic understanding of varieties of empathy instead of empathy

proper. One of the advantages of this approach is that the role of emotional contagion in experiencing empathy is not marginalised and is clearly accounted for.

3. The "Russian Dolls' Model" of Empathy *Mechanisms*

In this section, I argue that the polysemic nature of empathy, anchored by the etymological shifts the term has undergone throughout history, gives us good reasons to adopt a pluralistic understanding of the phenomenon. Proponents of this approach, such as Stephanie Preston, Frans de Waal (de Waal, 2009, 2021; Preston & de Waal, 2002) and Mark Fagiano (Fagiano, 2016, 2019) believe that it is not practical to study empathy within a "closed system of understanding" such as that offered by Coplan's conceptualisation of the term (Fagiano, 2019, p. 30). Fagiano explains that right from the onset when empathy was introduced to the English language as a translation of the German term *Einfühlung* (or "feeling into"), it referred to what Coplan terms a lower-level process. Moreover, in the first decade of the twentieth century, psychologists James Ward at the University of Cambridge and Edward B. Titchener at Cornell studied *Einfühlung* as a tool for kinaesthetic, embodied appreciation of inanimate artefacts and objects of empathy**,** especially artworks, through projecting certain properties of the observer (or the subject) into said objects (Lanzoni, 2012). I will say more about *Einfühlung* (aesthetic empathy) in the next section of this chapter.

One of the main points of disagreement between proponents of the narrow and pluralistic approach concerns whether or not it is correct to count basic processes, mainly emotional contagion, as "real" empathy. To answer affirmatively to this question means to ascribe empathy to newborns and several non-human animals shown to be neurologically and behaviourally capable of emotional contagion. In his paper, "Empathy, the Umbrella Term" (2021) and his book *The Age of Empathy: Nature's Lessons for a Kinder Society* (2009), de Waal suggests that we can ascribe a basic level of empathy to human infants and non-human primates. However, it is false to accredit them with higher-level interpersonal empathy, which necessitates cognitively, intellectually and linguistically demanding constituents. For more clarity, more will be said about emotional contagion in this section. But for now, I will be focusing on unpacking de Waal's helpful recommendation that we understand empathy in levels in terms of his "Russian Dolls' model", as a phenomenon that is developmentally constitutive of interconnected basic and higher-level processes. Along with Stephanie Preston, de Waal supports the "Perception-Action Model" (PAM) of empathy, maintaining that there exist integrated, ultimate and proximate mechanisms, such as emotional contagion, which

neurologically instantiate more cognitively-demanding empathic responses over time, sometimes – *but not essentially* – enabling prosocial responses and empathic concern, characterised by a desire for alleviating others' suffering.

De Waal makes sure to account for the developmental and anthropological interconnectedness between emotional contagion and higher-level cognitive empathy. He argues that it is more pragmatically productive to "lump" rather than "split" basic and more developmentally advanced mental processes underpinning empathic responses. Then, he importantly concedes that linking these processes does not keep them "from sometimes leading independent lives" (2021, p. 2). For example, de Waal explains that psychopaths can engage in cold-blooded, intellectual perspective-taking, which is necessary for interpersonal (higher-level) cognitive empathy. However, they do so without any signs of emotional engagement or manifestations of affect sharing or empathic concern. This shows that there are instances when some of the mechanisms enabling interconnected cognitive and affective empathy can exist in separation. But as de Waal stresses, empathy in its totality as he defines it (and to some extent as Coplan defines it) has evolved from basic mirroring mechanisms into more robust cognitive ones, that enable intellectual processes such as understanding others' emotions through intentionally engaging in imaginative perspective-taking.

Next, De Waal shows that non-human animals' mental mirroring abilities sometimes precede prosocial behaviours that are important for survival. Examples of these behaviours include "alarm to danger, social facilitation, vicariousness of emotions, mother-infant responsiveness, and the modelling of competitors and predators that are crucial for the reproductive success of animals living in groups" (Preston & de Waal, 2002, p. 1). In several parts of his writings, de Waal seems to use empathic concern interchangeably with empathy. However, across the literature, this is recognised as a sign of terminological conflation between empathy and sympathy (i.e. empathic concern). Alternatively, it is more conceptually accurate to agree with Lauren Wispé, who prolifically writes about both empathy and sympathy, in making a straightforward distinction between the two terms, suggesting that: "whereas sympathy is a way of relating, empathy is a way of knowing [or understanding]" (1986, p. 1). Wispé explains that empathy refers to the attempt of one self-aware self to understand the subjective experiences of another self, whereas sympathy differently involves what is referred to as empathic concern towards another self, without the essential need to acquire any sort of understanding. For instance, if I see that my friend is sad through simply "catching" her expressions of sadness, I would be concerned, i.e. sympathetic, without any need to deeply

understand other facets of her emotional experience, or engage in other-oriented perspective-taking that is alternatively central to empathy. In her paper entitled "A Theory of Narrative Empathy", Susan Keen (2006) provides a similar distinction between the two concepts through the following example:

Empathy:

    (a) I feel what you feel.

       I feel your pain.

Sympathy:

    (b) I feel supportive feelings about your feelings.

       I feel pity for your pain   (p. 209).

With that being said, the key distinction between sympathy and empathy is an epistemological one, pertaining to the extent of knowledge or understanding required to empathise or sympathise with another. If I merely know that S is sad, this would be sufficient to sympathise with her. Therefore, sympathy is often described as a less epistemically demanding feeling of concern, often followed by a desire to ease another agent's suffering. But when it comes to empathising, or understanding another's emotional state, *as other*,[5] a subject would be required to assemble a greater deal of understanding; for example, understanding the reasons why S is sad, to be more effectively able to imaginatively "inhabit" the perspective of S, as S.

Despite this conflation, De Waal usefully shows that human (cognitive) empathy, which essentially involves perspective-taking, developmentally flourished with reliance on automated *sensitivities* to faces, bodies, and voices, which we (sometimes unconsciously) experience in our everyday interactions.[6] Contradicting some of Coplan's assumptions, de Waal suggests that the term "empathy" is not exclusively preserved for referring to complex cognitive or intellectual functions, such as understanding the reasons behind someone's grin or working out the object-directedness of someone's sadness. He, therefore, deems it coherent to consider emotional contagion as a valid "layer" or level of basic empathy, which has been "modified" and augmented by evolution, similar to language for instance, to help us not only feel what others feel but also understand the contents of others emotional experiences. De Waal

---

[5] For more on empathy as a process of perspective-taking of another agent as other see (cf. Husserl 1950, p. 139) as mentioned in (Zahavi, 2018).
[6] See Dimberg et al. (2000) for an intriguing discussion on a experiment where participants demonstrate involuntary mimicry of facial expressions, even when displayed on a computer screen.

asserts that "no one denies the importance of these higher strata of [human] empathy, which develop with age, but to focus on them is like staring at a splendid cathedral while forgetting that it's made of bricks and mortar" (2009, p. 205).

Therefore, as mentioned above, de Waal metaphorically describes empathy as a multi-layered Russian doll, "with at its core, the ancient tendency to match others' emotional states. Around this core, evolution has built ever more sophisticated capacities, such as feeling concerned for others and adopting their viewpoint" (2009, p. 209). This characterisation brings about a great deal of clarity, as long as we make sure to conceptually separate empathy from sympathy, as explained above. To conflates the two concepts in the name of conceptual pluralism can result in errors, especially when discussing empathy in the context of ethics, as briefly discussed at the beginning of this chapter with reference to Michael Slote's supposition that empathy is "the cement of the moral universe".

Some may still worry that adopting a loose, pluralistic, conceptual framework to determine what counts as empathy, can cause confusion and lead to crosstalk, pushing some to favour a more narrow, and unified, conceptualisation. In response, Fagiano contends that it is neither practical, feasible, nor necessary to enforce a definition of "true empathy" over the myriad of already existing empirical and philosophical conceptualisations because he believes this does not necessarily dissolve the problem of lack of conceptual consensus, in that new conceptualisations can always still be introduced, contesting the definitional constituents narrowly fixed for empathy. Also, as shown in the previous section, narrow conceptualisations can collapse into one of the conceptual varieties of empathy. More importantly, Fagiano argues that adopting a pluralistic approach to defining empathy does not essentially lead to ambiguity and vagueness, as it comes with greater responsibility for clearly categorising and unequivocally explaining which variety of empathy one is referring to in a given *context*. As long as this is skilfully done, unfortunate consequences such as crosstalk can be avoided to a very large extent.

In this vein, and following Mark H. Davis (2006), I propose organising varieties and levels of empathy into a clearly defined conceptual framework that highlights key *mechanisms* as well as the *object-directedness* of empathic responses. In what follows, and with close reference to de Waal's Russian Dolls' model of empathy, I discuss emotional contagion and other-oriented perspective-taking as key mechanisms underpinning empathy. Next, with a focus on the object-directedness of empathic experiences and with consideration of the

etymological shifts empathy has witnessed and continues to witness, I further examine two main varieties of empathy; aesthetic and interpersonal (or inter-subject) empathy.

3.1. Emotional Contagion

Following psychologists Elaine Hatfield, John Cacioppo, and Richard Rapson (1994), Coplan defines emotional contagion as the hardwired "tendency to automatically mimic and synchronize expressions, vocalizations, postures, and movements with those of another person, and, consequently, to converge emotionally" (1994, 153–54). In addition to being automatic and heavily reliant on neurological systems of what is sometimes referred to as "motor mimicry" (Blair & Blair, 2013), Coplan adds that emotional contagion is often "involuntary", in the sense that a subject may not be aware that they are mimicking, or "catching" aspects of another person's emotional expressions. Examples of emotional contagion are provided in Stephan Davies' (2011) account of the phenomenon, not only in reaction to other people's emotional expressions but also as a visceral response to being affected by music and ambience.

Davies illustrates that "…people find laughing and yawning contagious. The same effect also occurs with emotions. We can become depressed when surrounded by sad people, or find our feelings chiming with a happy ambience" (2014, p. 1). Davies makes sure to argue that experiencing emotional contagion, developed as a result of listening to "sad" music, for instance, does not entail ascribing the quality of sadness to the music itself. In other words, he emphasises that music's *expressive quality* is distinct from our emotional responses to it. Arguing against cognitivist theories of emotions, which sometimes take it for granted that all emotional paradigms must involve object-directedness, in being *about* something or someone, Davies contends that experiencing emotional contagion is in some way equivalent to "catching a *mood*", therefore it does not necessitate object-directedness. For instance, using background elevator music as an example, Davies explains that "the music generates an *objectless mood* that reflects its own, calm expressive character" (2011, p. 137).

If one feels melancholy as a result of listening to Antonio Vivaldi's "Four Seasons (Winter)", on a gloomy day, this does not entail that one is sad *about* the music, because this would imply the irrational belief that music is suffering or is in distress. Instead, Davies explains that in these cases, a listener experiences a sad mood as a result of commonly associating a slow tempo with sadness. However, this is not always straightforwardly the case. For instance, some people experience calmness and clarity when listening to hard metal music, which generally involves a fast tempo and the use of "loud" musical instruments such as drums

and electric guitars. This shows that the effects of emotional contagion are not as homogeneous as some paint them to be.

Davies further argues that emotional contagion can either be *attentional* or *non-attentional.* That is to say, a subject can intentionally focus their attention on the source stimulating their experience of emotional contagion, consequently amplifying or restraining their responses. Examples of attentional contagion can be seen in art critics who attentively react to a musical piece to detect false notes of discrepancies. Non-attentional emotional contagion can be, for instance, detected in the spontaneous responses of moviegoers, who flinch upon seeing a "jump scare" in a horror movie. In this context, Davies recites Darwin's example where he reports that "he flinched in fear of a striking snake even as he knew it was separated from him by an unbreakable glass barrier" (ibid,.p.135). Davies posits that the more attentive a subject is towards a given source of emotional contagion, the more amplified the emotional effect of the contentless contagion is. However, even in cases of attentional emotional contagion, where individuals can be aware of the source of their emotion, or even more, in the example of moviegoers and art critics who purposefully attend to cues of emotional expressions, it is still accurate to describe emotional contagion as involuntary and devoid of object-directedness. Emotional contagion is described as involuntary across the literature because it is shown to occur due to tacit mechanisms that the subject experiencing emotional contagion cannot account for or fully control while experiencing this phenomenon.

De Waal reports that emotional contagion is manifested in newborns and some non-human primates. For instance, a child's crying can elicit tears in another, without any awareness or understanding of the reasons causing the other child to cry. In agreement with de Waal, Coplan points out that emotional contagion occurs due to "distinctive causes, a distinctive phenomenology, and distinctive effects, and it relies on a distinctive neural system…" (2011, p. 53). Nevertheless, in other parts of her paper, Coplan vaguely *merges* emotional contagion with self-oriented perspective-taking, by identifying the former as a type of the latter. This highlights some contradictions in Coplan's account, for she reports that human infants exhibit the capacity for emotional contagion moments after they are born, whereas their ability to engage in perspective-taking does not develop until the age of four or five. Therefore, how can we consider infants' tendency for emotional contagion as reliant on self-oriented perspective-taking?

In the context of interpersonal empathy, it might be helpful to think of both emotional contagion and self and other-oriented perspective-taking as *preconditional inner workings* for experiencing empathy. As De Waal elucidates, perspective-taking can indeed be done without resorting to any forms of emotional engagement or affect sharing (as done by psychopaths) (2021, p. 181). However, this amputates an important source of input and emotional cues that can be used to enhance our contextual and experiential understanding of others' experiences and emotions.

With that said, I argue that the ability to experience emotional contagion *while* attempting to empathically understand other targets' emotional experiences can enhance our experiential understanding by replicating, to some extent, some aspects of what others are emoting. For instance, upon seeing that another person is in distress, our emotional contagion mechanisms can assist us in gathering more experiential input on their sadness, to closely understand their emotional experiences in clearer terms. Nevertheless, as argued by Remy Debes (2009), the inner workings of emotional contagion can be limited when it comes to attaining an intellectual *understanding* of the content of others' emotions. In other words, merely catching and mimicking a given emotion is not necessarily guaranteed to tell us what the emotion consists of. Hence, it is more reasonable to speculate that emotional contagion functions in parallel with other cognitive mechanisms. In a similar vein, I will now briefly discuss self and other-oriented perspective-taking, which I identify as an essential mechanism for interpersonal empathy.

3.2. Self and Other-Oriented Perspective-Taking

Davis defines perspective-taking as: "the attempt by one individual to understand another by explicitly imagining the other's perspective". He adds that "it is typically considered an effortful process, involving both the suppression of one's own egocentric perspective on events and the active entertaining of someone else's" (2006, p. 6). There is a wide consensus across the literature on the contention that perspective-taking is the backbone of cognitive (interpersonal) empathy demonstrated by human adults. Coplan eloquently suggests that it is important to differentiate between other-oriented perspective-taking and self-oriented perspective-taking. The latter simply refers to imagining oneself in a different situation, *as oneself,* without making any mental efforts to take into consideration aspects of others' experiences, such as contextual input or personality traits (pertaining, for instance, to historical,

cultural or biological differences) when attempting to understand others' experiences from their perspective.

In a crucial part of Coplan's discussion of empathy proper, she contends that self-oriented perspective-taking often leads to errors in prediction, misattributions, and most intriguingly, to personal distress, which Coplan regards antithetical to empathy. This contention is guided by the empirical assumption that individuals have a "natural tendency to assume greater similarity between [themselves] and others than typically exists" leading to "egocentric biases" (Coplan, 2012, p. 10). Coplan demonstrates that there are instances when subjects *assume* they have an experiential understanding of others' emotions, while in fact, their assumptions could not be more wrong. For instance, if Camellia, who is diagnosed with borderline personality disorder, exclaims that she is feeling depressed, and her colleague Susan, who is only having a really bad day, responds that she "knows exactly how she feels", while she does not (at least physiologically), Coplan believes that this would not be sufficient to accredit Susan with empathy towards Camellia, specifically because Susan fails to suppress her own self-perspective and emotions.

Although Coplan's claims may seem plausible at first, I doubt that suppressing one's own beliefs, emotions, and other features of the self, can, on its own, lead to bridging epistemic and phenomenological gaps in understanding others. That is to say, such suppression should also be supported by the intentional effort of understanding the contents of another subject's experience. Moreover, I argue that detecting *similarities* between oneself and others can sometimes *aid* in developing an even more thorough – empathic – experientially supported understanding. Coplan does not necessarily disagree with this position, as she considers a fusion between self and other-oriented perspective-taking helpful for empathy. What she rightly rejects is what aesthetician Theodor Lipps terms "projective empathy". As quoted in Zahavi (2014), Lipps wrote:

> The other psychological individual is consequently made by myself out of myself. His inner being is taken from mine. The other individual or ego is the product of a projection, a reflection, a radiation of myself—or of what I experience in myself, through the sense perception of an outside physical phenomenon—into this very sensory phenomenon, a peculiar kind of reduplication of myself. (1905, p.17)

Lipps claims that presupposing similarities between others and oneself, rather than differences, is sufficient for experiencing empathy, through inference from an analogy, which to Lipps is a source of the "best explanation" we have of others' emotional experiences, especially when attempting to make sense of basic emotions, as shown in Paul Ekman's account of "affect programs". There, Ekman proposes a universal association between certain facial expressions and emotions; namely fear, anger, joy, sadness, disgust, and surprise. He drew these conclusions as a result of an experiment he carried out on participants from 22 different countries (including isolated tribesmen from Papa New Guinea), who were shown images of different facial expressions and were asked to predict the emotional state that each of these facial expressions entails. The participants drew similar associations between those facial expressions and corresponding emotions. Therefore, since we seem to universally share *some* similar features of emotional expressions, it may be sufficient to infer someone's sadness, from seeing them cry, drawing upon a past instance of crying as an expression of sadness. Here, egoistic, self-oriented perspective-taking seems to be sufficient for comprehending the emotional experience of others.

However, emotions are rarely this simple, for as previously mentioned, they are part of a multidimensional net of widely debated features, including reasons causing a given emotion, psychophysiological factors, object-directedness, and phenomenal experiences that are compounded by an amalgamation of all these features. Therefore, it is useful to consider observable similarities and to sensibly watch out for possible differences that can make another's emotional experience different from our own. Even then, as argued by Max Scheler (1954) interpretative errors can still occur due to the fact that even first-hand, introspective perceptions of our own emotional experiences can be subject to errors and misinterpretations, as dynamically discussed in fields such as psychotherapy, let alone emotional experiences that are only available to use third-personally.

Before closing this section, it is important to restate that drawing upon observable, especially linguistically or physically expressed similarities does not by definition lead to errors or misattributions. Equally, suppressing features of one's perspective does not always lead to clarity on others' perspectives. For instance, Scheler denies that we can empathise with another's experience *only if* we had had the same experience ourselves. To assume otherwise would not leave room for accounting for the way we come to understand *new* emotional experiences that are completely "foreign" to us, such as "understanding from the wagging tail

of a dog that she is happy to see us" (Scheler, 1954, p. 11) despite never having experienced this ourselves.

This seems plausible, as we can understand what it is like to be in a given emotional state with sole reliance on "the givenness", in Edmund Husserl's terms, provided to us by third-personally witnessing another target go through the said emotional state. Therefore, while first-hand experience is not an indispensable requirement for empathy, it can be helpful to know what it is like to be in a given emotional state first-hand, as long as the empathiser keeps the other person at the centre of their attention. For example, it would be easier for an individual who has already experienced the grief of losing a child to empathise with another parent going through the same tragic event. This is not to say that another person who never had or lost children would have a complete deficit in understanding what the tragic experience is like, as they can establish such understanding third-personally through interacting with the grieving parents, then subsequently resorting to imaginative other-oriented perspective-taking to develop a more empathic understanding of the experience. All in all, saying that empathy makes use of projection and detection of similarities, is not to say that empathy cannot do without it, but this can be helpful in some cases.

4.   A Brief Overview of *Einfühlung*: "Feeling into" Aesthetic Empathy

So far, in an attempt to mitigate the conceptual confusion surrounding empathy across disciplines, I have argued that it is useful to thematically discuss empathy by identifying some of its main underlying mechanisms. In this context, I have discussed emotional contagion and self and other-oriented perspective-taking. Following proponents of the pluralistic approach, it is deemed useful to discuss the affective and cognitive inner workings of empathy in unison, instead of reductive nit-picking at separate micro (especially neurological) processes in isolation. With that said, I have argued that it is more coherent to not discuss lower-level and higher-level processes in isolation, and then contemplate which subset of processes is true empathy. Instead, I consider both categories of processes developmentally and pragmatically necessary for empathy, with close reference to de Waal's Russian dolls' model of empathy.

In this section, I discuss a variety of empathy namely *aesthetic empathy* (also referred to as *Einfühlung* or "feeling into"), which can be defined as a tool for kinaesthetic and embodied projection of imagined bodily movements and sensations into perceived inanimate artefacts, such as artworks, landscapes, and fictional characters (Pinotti, 2017). Susan Lanzoni reports that during the "Cheves Perky's 1910 experiments" conducted at the Cornell psychology

laboratory on the imagination and *Einfühlung*, when a participant was told to imagine a bunch of grapes, the report was of "a cool, juicy feeling all over"; when imagining a parrot, the response was of "a feeling of smoothness and softness all over me"; and the image of fish elicited a "slippery feeling and coolness" (2012, p. 301). Nature is also often identified as an object of aesthetic empathy. As discussed by Peter Hacker (2017), we sometimes tend to *read* psychological attributes *into* nature. This is manifested in utterances such as "*raging* storms, *menacing* rumbles of thunder, *tranquil, ominous*, or *brooding* landscapes" (Hacker, 2017, p. 378). Lipps also considered this early connotation of empathy as a *sine qua non* of aesthetic experience, understanding it as the projection of inner bodily feelings and movement into art forms and nature (Jahoda, 2005; Lanzoni, 2012, 2018).

The utilisation of empathy was also noticeably expanded to the context of empathic responses towards fictional characters, as explored in Susan Keen's theory of narrative empathy. It might come as a surprise that some theorists and philosophers of art argue that empathy toward *fictional* characters and inanimate objects is possible. Nevertheless, when closely examining the etymological history of empathy, it becomes apparent that this connotation of empathy is not only coherent but also considered one of the first utilisations of the term in aesthetic theory. In what follows, I discuss aesthetic empathy with close reference to experiencing emotional contagion and self and other-oriented perspective-taking towards characters in narratives, focusing on the implications of employing the imagination as a vehicle for understanding emotional expressions of characters that may be different from or similar to the reader or viewer.

Coplan argues that "among the implications, of emotional contagion is the fact that films and television shows can generate emotional contagion responses, while literary [text-based] narratives cannot" (2011, p. 46). Prima facie, it might be suggested that emotional contagion can also occur, to a lesser extent, without direct exposure to multisensory stimuli that are often present in 2D, 3D, or VR films, when textually communicated events are vividly created in one's imagination. For example, in the abolitionist novel, *Uncle Tom's Cabin* (1851), Harriet Beecher Stowe attempts to *represent* the cruelty and savagery of slave owners leading to the escalation of the American Civil War, chiefly by describing the excruciating suffering that Uncle Tom (the novel's protagonist) endures, building up to the moment of his imminent death. For instance, in one of the passages describing Uncle Tom's plight, Stowe writes: "the night was damp and close, and the thick air swarmed with myriads of mosquitos, which increased the restless torture of his wounds; whilst a burning thirst – a torture beyond

all others – filled up the uttermost measure of physical anguish….” (p. 279-280). Such vivid expressive language can trigger the reader's visceral response, leading to some degree of implicit emotional contagion, not through direct exposure to a film adaptation of the novel, but to the events mentally and reflexively simulated in the reader's imagination.

Perhaps unbeknownst to some, Stowe also published a follow-up, non-fictional book entitled *A Key to Uncle Tom's Cabin* (1853), revealing the *real* characters and events that inspired the writing of her bestseller novel. Now, let us imagine that after reading the novel, one decides to recommend it to a friend, alongside the book discussing the veridical elements which inspired the novel. After the friend reads both pieces, she reports that she feels empathy for the character of Uncle Tom, anger towards the character of Simon Legree (the brutal slave dealer in the novel), and admiration for the resilience and courage of Josiah Henson, who Stowe says to be the main inspiration behind the character of Uncle Tom. Although the question of whether or not the friend experienced emotional contagion remains contingent on the dynamism of her imagination, emotional engagement, and what de Waal terms visceral "sensitivity", it is undeniable that she at least developed a certain *degree* of embodied responses upon reading the detailed description of Uncle Tom's suffering.

Relevantly, a quite intriguing point is that upon watching an emotionally charged film or reading about distressful events, we can develop heightened emotional responses, even in the absence of a *belief* in the truth of these events. This issue is discussed in the context of "the paradox of fiction". Here, philosophers of art ponder the motives behind our often *realistic* responses towards fictional events, despite knowing that these events are not real. Cognitivist theories of emotions, suggesting that our realistic emotional responses to fiction are often caused by underlying reasons or beliefs, do not provide much help in solving this riddle, for a variety of reasons. For instance, if moviegoers *believe* that a fast train is headed towards them from a screen, they would perform actions in response to this, such as evacuating the movie theatre. This shows that the cognitivist theory of emotions, as previously mentioned, cannot explain our realistic responses to multisensory media or to narratives that can stimulate emotional contagion, at least to some extent.

In *Mimesis as Make-Believe* (1993), Walton suggests that, like children playing in the forest pretending that tree stumps are bears, when engaging with fiction we equally engage in imaginative, pretence games of "make-believe", where we pretend that fictional events and characters are truly existent in a given world of fiction. An alternative interpretation of how

and why viewers interact with fictional content is provided by Alex Neill (1993) who argues that cognitivism does not require the belief that the fictional target of one's emotional response *exists* in the actual world. According to Neill, one can develop emotions such as pity or admiration toward a fictional character, simply because one believes that the character exists in a pitiable or admirable situation *in the world of fiction* (Davies, 2011, p. 136). However, more should be said to support this line of reasoning. I will discuss this issue in more detail in the fourth chapter of this thesis to argue that the paradox of fiction can be extended into "a paradox of virtuality", discussing (realistic) emotionally heightened and even behavioural responses to virtual content.

Returning to the topic at hand on experiencing aesthetic empathy, Susan Keen confirms that empathy plays an essential role in both the artistic processes of creating a narrative (especially its characters) and visual representational artwork, as well as shaping the audience's interpretative responses. Here, viewers and readers are said to mentally place themselves in the perspective of a given character in a narrative to get a better understanding of it as well as themselves. This partially explains why readers and viewers often have heterogeneous responses to works of fiction, either based on resonance or on detecting differences between themselves and depicted characters. Intriguingly, Keen stresses that consuming different forms of fictional narratives that portray intense negative events can sometimes lead to personal distress, as a result of experiencing emotional contagion or purposeful self-oriented perspective-taking, where a reader or viewer places themselves in the shoes of a given character in a narrative. Keen explains that this sometimes pushes the reader or viewer to divert their attention away from the source of distress, preventing full empathic understanding from successfully occurring. Keen adds that this is easier to do in the context of fiction or watching films as opposed to interpersonal, real-life events.

In the third chapter, I argue that VR, as an artistic medium for storytelling, psychophysiologically and visually "transports" VR users (so to speak) *into* the space of a given narrative. Through new technological affordances, VR offers viewers the opportunity to be active participants in the unfolding of a given narrative, hence immersing them more effectively than other non-interactive mediums (such as written narratives and 2D films). This is reported to facilitate and broaden imaginative perspective-taking and emotional contagion which I have argued to be preconditions for experiencing empathy.

5.   On the Limitations of "Perfect" Interpersonal Empathy

In the context of interpersonal empathy, errors of interpretation and misattribution may often occur, either due to lack of clarity, failure to consider things from others' viewpoints, or simply due to overattachment to aspects of one's worldview. Scheler (1954) as well as other philosophers in the phenomenological tradition, maintained that there are certain limitations to "accessibility" to others' emotional states through direct perception, cogitative and reflective imagination, and finite communicative channels of emotional expressions; including linguistic and facial expressions and other manifestations of body language. For instance, gustatory sensations and physiological pain can only be experienced first-personally. Although we can do our best to third-personally imagine the physical agony resulting from a soldier's bullet wound or the pain that pregnant ladies endure during labour, our imagination-based experiential understanding will always be missing a piece when attempting to empathise. This sometimes leaves us with an intellectual understanding of others' emotional experiences and an inability to perform what Coplan refers to as affective matching through other-oriented perspective-taking. Perhaps this means that restricted emotional contagion is the best we can do when attempting to understand some physiological layers that often characterise another's lived emotional experience.

As pointed out by Hacker, and phenomenologists such as Husserl and Edith Stein (1989), empathy is not an emotion at all, instead, it is a *sui generis* form of intentional understanding of objects or persons that are external to an empathiser, with reliance on cogitative imagination. Stein compares some aspects of direct sensory perception to interpersonal empathy, by arguing that others' emotional expressions and corresponding connotations are, to some degree, "given" to us in the same way that sensory (visual, auditory, haptic, and gustatory) experiences are. That is to say, in interpersonal interactions with others, we do not merely see arbitrary muscle distortion and behavioural adjustments, instead, we see those emotional expressions as meaningful indications of a given emotional state. For instance, consider Kris McDaniel's example; "I don't simply see faces. I see angry faces, faces transfixed with wonder, or expressions of grief. I don't simply see physical bodies as mere physical things, but rather as embodying the lived experiences of the people in front of me" (2016, p. 7). At the same time, we cannot introspectively experience or understand others' emotions with an identical direct givenness to that of our own emotions. Contradicting Stein's intuitions, it can then be argued that multidimensional properties of others' complex emotional experiences are

not as available nor as *verifiable* as sensory experiences*,* through sense perception alone. For example, I cannot take my friend's sarcastic grin as an indication of happiness, or someone else's happy tears as an indication of sadness. This shows that a communicative effort is required for attaining and enhancing what is called "empathic accuracy" in clinical psychology (Ta & Ickes, 2017). It also demarcates the limitation of achieving full empathic understanding by means of sense perception and imagination alone.

Helpfully, Husserl distinguishes two different attitudes that we can take towards other individuals when attempting to be empathic; a naturalistic attitude and a personalistic one. The naturalistic attitude depicts others as composed of a twofold materialistic self that is functional and observable, in addition to an underlying experiential layer of self. As for the personalistic attitude, it describes the person as a unit rather than a twofold, causally related entity. Husserl explains that when adopting the personalistic attitude, instead of seeing a conjunction of two realities, we experience "one expressive unity" (Zahavi, 2014, p. 128). Similar to Stein, Scheler and Zahavi, Husserl also emphasises the importance of experiencing others directly as others and not through mental mirroring or projection. He explains that:

> Just as what is past can be originally given as past only through memory, and what is to come in the future can as such only be originally given through expectation, the "foreign" can only be originally given as foreign through empathy. Original givenness in this sense is the same as experience. (Husserl 1959, p. 176) as quoted in (Zahavi, 2014).

Nevertheless, instead of Husserl's contention that third-person givenness is "the same as" direct experience, it may be more fitting to emphasise that the first and third-person modes of understanding are *not* the same per se, and yet they are both epistemically *reliable* sources for obtaining a degree of understanding about oneself and others. With that said, an important limitation persists, as discussed by Ronald de Sousa (1990), in that we can only recognise and hence make sense of emotional expressions that fall within the range of our emotional repertoires. This means that some aspects of others' emotional experiences and their signification are bound to remain elusive, or at least reliant on the explanatory powers that others are willing to exert (either propositionally or representationally) to render the signification of their emotions more lucid.

In this vein, Karsten R. Stueber (2006) discusses this limitation of empathy in the context of cultural, historical, and biological interpersonal differences. He explains that

achieving empathy is always contingent on the preconceptions and experiential paradigms of the empathiser. Apart from basic affects, one cannot make sense of complex emotional experiences or their underlying reasons without having a prior grasp or manifestations of these experiences at some point in the past. Coplan similarly suggests that individuals often tend to assume more similarities than typically exists, leading to errors, egocentric biases, and misinterpretations when utilising imaginative, self-oriented perspective-taking to empathically understand others, On the other end of the spectrum, Stueber discusses another error which occurs due to "non-projectionism" (p. 205), which takes place when an empathiser presupposes radical *differences* between themselves and another target, characterising the latter as completely "foreign". This leads to missing out on epistemically beneficial cues and common features that can support a more accurate empathic understanding. All in all, it is evident and widely agreed upon that the underlying mechanisms of empathy are epistemologically and phenomenologically fallible and restrictive. That is to say, retrieving accurate, supportive justifications for understanding either the reasons triggering someone's emotional experience or the phenomenological contents (i.e. how it feels like) to undergo said experience is often restricted due to the restrictiveness of imaginative perspective-taking.

Within the overarching context of this research, these limitations highlight a possible need for technologically-mediated empathy. Through VR representational narratives, which provide propositional and representational evidence-based input for supporting the experiential understanding that is necessary for empathy, one can obtain a more epistemically enriching background against which imaginative perspective-taking can be extended, hence leading to a higher accuracy of visceral empathic understanding. However, these claims are not completely unchallenged, as I will seek to show later in this thesis.

6.   Chapter Conclusion

In an attempt to contribute to mitigating the widespread conceptual vagueness and confusion surrounding the notion of empathy across the literature, this chapter has proposed that it is more functionally and pragmatically productive to speak of empathy as a phenomenon that is the result of the inner workings of developmentally interconnected systems of emotional and cognitive (experiential) understanding of others' states. Such understanding is widely said to heavily rely on the imaginative process of perspective-taking and emotional contagion. Many insist that empathy is not reducible to mere mimicry or involuntary emotional contagion unless we are referring to a basic level of empathy that is often noticed in non-human primates

and human infants. This highlights the crucial involvement of critical and interpretative (cognitive) mechanisms in practising higher-level (human) interpersonal empathy, which I will continue to focus on throughout this thesis. Due to etymological reasons and the widespread lack of conceptual consensus on empathy, I have sought to provide a rationale for discussing empathy from the lens of the pluralistic approach, as discussed by de Waal, Preston, and Fagiano.

I proposed to thematically identify empathy vis-à-vis its underlying mechanisms and its object-directedness. The upshot of this characterisation is that it respects the polysemic nature of the term, and accounts for the etymological shift that the term has undergone from the aesthetic concept of *Einfühlung* towards interpersonal empathy as it is discussed in Anglophone contemporary writings on the topic. In this chapter, I have also discussed emotional contagion and self and other-oriented perspective-taking, while making sure to shed light on some philosophical issues raised concerning both of these lower-level and higher-level processes, respectively.

I have also argued that emotional contagion and self and other-oriented perspective-taking can be coherently experienced towards fictitious characters, portrayed in films and novels, often producing heightened emotional experiences that can lead to personal distress, which can in turn push readers or viewers away from the recognised source of distress, further obstructing the imaginative, experiential understanding necessary for empathy.

The final part of this chapter briefly discussed some of the inevitable limitations in experientially understanding some facets of others' emotional experiences, namely including physiological aspects that may not be directly given to us through self or other-oriented perspective-taking. Nevertheless, I have suggested that we can predict that technological affordances provided by VR experiences, can, at least to some extent, aid our imaginative and experiential understanding of others' otherwise concealed emotional experiences, by providing second-hand, evidence-based representations. However, this claim is still widely contested across disciplines as I will discuss in the fifth chapter of this thesis. In the next chapter, I seek to disambiguate some important dimensions of virtuality and VR technology, as a prerequisite for evaluating claims about VR's utilisation as a so-called "empathy machine".

Chapter 2: On Virtuality and Virtual Reality (VR) Technology

1. Chapter Overview

In academic discourse, there seems to be a tendency to juxtapose the "real" with the "virtual". For instance, Webster's New Universal Unabridged Dictionary (1989) defines the word virtual as "being in essence or effect, but *not in fact*" and reality as "the state or quality of being real. Something that exists independently of ideas concerning it". The term "virtual reality" is hence sometimes described as an oxymoronic term (Jerald, 2016) and (Stanovsky, 2008). A similar juxtaposition is reflected in philosophy of technology where there exist two dominant positions; namely "virtual realism" and "virtual irrealism". Proponents of both positions often draw upon their metaphysical accounts of virtuality to respond to questions about the possible types of value we can (or cannot) attribute to virtual experiences. On the one hand, virtual realists, most noticeably David Chalmers, argue that: "virtual reality is a sort of genuine reality, virtual objects are real objects, and what goes on in virtual reality is truly real" (2017, p. 1). On the other hand, virtual irrealists, either argue that virtuality can be equated to an *illusion* or a type of fiction, as virtual fictionalists claim (McDonnell & Wildman, 2019).

In this thesis, I subscribe to a third view, which characterises virtuality as a *sui generis* mode of actualisation where both psychophysiological illusions and veridical components coexist. This view is inspired by writings on virtuality by Pierre Lévy (1998), Fiona Macpherson (2020), Jasper Juul (2019), Peter Ludlow (2019), and Philip Brey (2003, 2014). The crux of the view is that even though virtuality stimulates psychophysiological illusions, namely of *virtual embodiment* and *virtual presence*, these do not cause a defect in users' rational judgment and understanding of the metaphysical underpinning of virtual "objects" and "events". In short, I strive to show that describing an experience as virtual does not mean that it is "unreal" tout court, as virtuality can be a host to virtually existent social objects and other virtual objects that have referents in the physical world. As I will explain, virtual media (i.e. virtual objects and events) is often utilised to meet teleologically significant ends; such as navigation, visualisation, and interaction. These would not be possible had VR been purely illusory or fictional. I suggest that both (absolute) virtual realism and irrealism stem from misleading and sometimes hyperbolic conceptualisations of virtuality that do not adequately account for the status quo or prospects of VR technological development. As philosopher Lucy Osler writes, some writings on "techno philosophy" "present philosophy as primarily occupied

with ivory-tower thought experiments about the nature of reality in imagined digital worlds, rather than an engaged and applied practice that can have a societal impact" (2022, p. 4).

Alternatively, post-phenomenologists such as Don Ihde (1975, 1979, 1990, 2002) and Peter-Paul Verbeek (2015) prefer to speak of human-technology *praxis* and relations, constantly redefined by our goals and intentions as well as technological "specs" that a given technology can afford. Such a research approach does not ascribe a fixed "essence" to technological instruments but instead highlights the dynamism with which they are constantly readapted for serving problem-solving functionalities. When it comes to VR, Fiona Macpherson (2020) asserts that this technology involves an amalgamation of both illusory and veridical (functional) features, simultaneously present in technological interfaces. She explains that multisensory illusions in VR can be understood in analogy to optical illusions that nevertheless *mediate* legible fictional or non-fictional representations.

This chapter will, therefore, be divided into three parts. The first part will be divided into two sections. The first section of which will aim toward putting forth a clear understanding of the notion of virtuality. Making use of Pierre Lévy's writings on the topic, I define virtuality as "a technological mode of *actualisation*", sometimes consisting of the transformation of some properties – of a given entity in the "real world" – into a virtual form, for pragmatically serving new functionalities. Virtual actualisation can also be utilised to create virtual media that is purely referent to fictional and imaginary objects that need not exist in the real world. The myriad of VR video games available in the market is perhaps the most common paradigm of this category of virtual media. As Lévy explains, virtual technologies need not be restrictively understood with a fixation on ontological "derealisation" (i.e. the move away from the physically actual towards the virtually actualised) but instead, they should be identified with emphasis on their *virtually existent* constituents, which are ontologically fluid and pragmatically shaped to solve real "continuously renewed problems" (1998, p. 28). Common examples of virtually actualised media that have physical counterparts in the world include virtual texts, virtual cooperations, virtual libraries…etc, that users *truly* perceive and virtually interact with.

In the second section of the first part, I narrow down my focus from the broad notion of virtuality to discuss VR technology, with attention to VR-specific technological hardware which generates multisensory stimuli that set VR technology apart from earlier mediums, such as 2D non-interactive films. In this vein, Heim boils down VR's basic constituents to three

"Is"; "immersion, interactivity, and information intensity" (1998, pp. 6–7). Considering the objective of this section to highlight the novelty introduced by VR technology, I will only be focusing on VR immersion and interactivity.

As explained by Slater and Wilbur (1997) immersion is the technological system's capacity to generate vivid, multisensory stimuli. More thoroughly, it is "the technical capability of the system to deliver a *surrounding* [emphasis added] and convincing environment with which the participant can interact" (p. 333). Immersion occurs in degrees and it is possible to objectively measure it against the quality of technological modalities and the physical hardware involved in generating a given VR simulation. For instance, the quality of numerous features, such as latency bandwidth, audio-visuals, frame rates, stereoscopy, and field of view, is considered the chief criterion for determining the extent to which a given VR system is immersive. Across the literature, immersion is sometimes interchangeably used with "virtual presence" (Oh et al., 2018). Nevertheless, as some suggest, it is more useful to terminologically differentiate between immersion and the subjective experience of virtual presence, because the former doesn't always entail the latter. In other words, it has been empirically shown that just because a VR system consists of optimal VR gear, this does not guarantee that the user will feel virtually "present" in a given VE. More will be said about the subjective experience of virtual presence throughout this chapter.

Next, interactivity in the context of VR technology is, at least, twofold. The first layer of interactivity in VR refers to physical, embodied interaction causally occurring between the user and a given model of VR gear, often involving hardware such as VR goggles, controllers, haptic gloves and, most recently introduced, full haptic body suits. The second layer of interactivity is a virtual, or more clearly, representational kind of interaction occurring between the user's virtual representation and virtual objects. This layer of interaction has produced a great deal of confusion and dispute in the philosophical literature on VR. Here, I make sure to highlight that representational, virtual interactions in VEs only entail pseudo-causality, in Neil McDonnell and Nathan Wildman's terms (2019). I contend that it is only at the level of the first layer of interaction that physical causality occurs. However, this observation need not be taken as evidence for undermining the pragmatic value of virtual interactions occurring at the level of the second layer, because this pseudo-causality remains crucial for successfully navigating VEs and making use of affordances that often intersect with teleologically significant aims in the physical world.

Having explained the preliminary structural features of virtuality and VR technology, the second part critically discusses two psychological *illusions* often stimulated in VR; namely "the illusion of place" (commonly known as virtual presence) and "the illusion of embodiment" (also referred to as the illusion of virtual body ownership). As mentioned above, I argue that acknowledging these VR-induced illusions upfront is crucial for addressing any questions revolving around the types of value that we can legitimately ascribe to virtual representations in themselves. As prolifically discussed by researcher Mel Slater and his colleagues (Slater, 2003, 2009, 2018; Slater et al., 1998; Slater & Sanchez-Vives, 2016; Slater & Wilbur, 1997), using VR technology is often observed to trigger a variety of psychological and sensory illusions, which in turn stimulate heightened (realistic) emotional and reflexive responses to virtual stimuli.

Virtual presence is, hence, defined across the literature as the *subjective experience* of situatedness or "being there" in an immersive VE. Some attempt to argue that virtual presence occurs mainly due to a defect in users' rational judgment, in that users are said to be "tricked" into believing that their virtual environment is no longer mediated by technological hardware (Mania & Chalmers, 2001). Another view maintains that in order to experience virtual presence, users need to engage in an imaginative exercise of "making- believe" that virtually represented events are either fictionally or non-fictionally truly occurring (Studt, 2021). Others argue that experiencing virtual presence requires users to "willingly suspend their disbeliefs" about VEs being virtually actualised, as an attempt to account for users' reflexive, sometimes emotionally heightened responses in immersive VEs. Alternatively, I contend, drawing upon Don Ihde's theory of human-technology "transparency relation", that users do not need to resort to conscious (in the sense of intentional or purposeful) mental or imaginative processes as a means for experiencing virtual presence (or virtual embodiment). Instead, similarly to emotional contagion discussed in the previous chapter, I identify virtual presence as a visceral, non-intentional, phenomenological by-product of technological immersion.

Then, I discuss another intriguing VR-induced illusion; namely virtual embodiment. Numerous experiments in psychology such as the well-known "rubber hand illusion" and "chameleon effect" (Bertrand et al., 2018; Botvinick & Cohen, 1998; Ehrsson et al., 2004, 2005) report that humans experience an illusory sensation of body or limb ownership upon exposure to special techniques of "visuotactile" stimulation. To trigger the rubber hand illusion, for instance, a scientist needs to stroke a participant's real, covered hand with a brush or feather while simultaneously performing the same motion on a visible rubber hand. Gradually, subjects

47

develop a sense of body ownership over the rubber hand, which is often followed by behavioural consequences. For instance, participants demonstrate an intense fear and shock reflex when the scientist subsequently strikes the rubber hand with a hammer.

This fascinating psychological illusion has inspired the creation of "virtual bodies" standing for the user's representation in virtual reality. Through sensorimotor synchronisation between the user's physical and virtual body, the body ownership illusion is reported to also occur in VEs. Drawing upon the transparency relation illustrated by Merleau-Ponty's famous example of "the blind man's cane" (1962), we can establish a clear phenomenological description of the embodiment illusion in VR. More thoroughly, provided that VR technology is sufficiently immersive, VR hardware becomes "ready-to-hand" instead of "present-at-hand", where VR users' *attention* would be directed toward the experience of functionality and interactive opportunities instead of the VR hardware itself. Here, similar to a carpenter's embodied experience of using a hammer and to how a visually impaired individual uses a cane to compile epistemically valuable substitute cues, we phenomenologically experience VEs and stimuli *through* VR hardware, making the latter "ready-to-hand". The hardware, and more generally technological tools and instruments, only become "present-at-hand" (i.e. at the centre of attention in and for itself) in case technical issues or malfunctions arise. In these cases, our attention gets shifted back *to* the technological device, severing both the illusion of embodiment and virtual presence. The transparency relation discussed in the phenomenological tradition hence provides rich theoretical insights for understanding virtual embodiment and virtual presence. Nevertheless, the application of this theory sheds light on important limitations in the ability of virtual experiences to proprioceptively or physically *extend* human functions.

Finally, in the third section, I show that we can equally encounter "social objects" (sometimes known as "institutional facts") *veridically* generated through representational virtual media, building upon John Searle (1995) and Philip Brey's characterisation of virtual social objects (2003, 2014). The ontological independence of social objects from fixed physical objects and properties in the physical world makes them a fitting candidate for being fully generated virtuality. At the end of this chapter, my characterisation of virtuality as a medium where illusory and veridical components coexist will hopefully be made more lucid.

2. Disambiguating Virtuality: The Transformation from the Physically Actual to the Virtually Actualised

It is important to acknowledge right from the onset that elements of virtuality, i.e. virtual media (consisting of virtual objects and events) are representations which expressively project amplifications or simplifications of features and functions either corresponding to objects in the physical world or in imagined, fictional worlds (or a combination of both) (Brey, 2003, 2014; Grant, 2019; Juul, 2019). Within the constraints of hardware and software specifications, and the intended objectives, which steer interface design, we can, at best, encounter virtual representations of physical laws in the way that virtual entities seem to causally behave. As eloquently put by Grant Tavinor: "many philosophical accounts of VR begin with Cartesian considerations of 'perfect matrix-like' virtual realities… it could be best not to speak of virtual worlds or realities at all: rather, we might restrict ourselves to talk of *virtual media*, which are capable of depicting fictions or reality" (2022, p. 6).

Here, it can be safely assumed that physical properties cannot be fully replicated in virtuality. Perhaps a compelling counterargument would suggest that technological innovation will eventually succeed in creating a full-blown counterpart of physical properties in virtuality. However, as explained by Jesper Juul: "…this is neither desirable nor feasible. If we consider the complexity of simulating anything down to the subatomic level (where the science is uncertain anyway)" (2019, p. 8). Juul describes the nature of virtual experiences as inevitably (ontologically) "incomplete", using an example of a virtual simulation of a laser beam. He concludes that:

> It [the Unity3d VR tutorial] has a laser beam, but it lacks the photons that comprise an actual laser beam, and even if it did include photons, these would lack the interactions that photons have with air particles, and even if they did interact with air particles, they would lack relativistic and quantum effects, and even if they did include those effects, and so on. (ibid,.)

In their paper "There's no there there" (2001), Peter Fisher and David Unwin assert that "all principles of real space may be violated in cyberspace [i.e. in virtuality] and the characteristics and constraints are only determined by the specifications that define the particular virtual space" (2001, p. 343). Fisher and Unwin articulately explain that we can identifying at least two categories of virtual media. The first type is virtual media conveying information about a referent that already exists in the physical world. Virtualisation of this type

is often intended for the visualisation of remote, condensed, or complex information or characteristics of a physical object or location. The resulting virtual version would then be made more accessible and navigable than the original one, while also preserving – and sometimes pragmatically amplifying – key features of the latter. Paradigms and examples of this type of virtual media include Global Positioning Systems (GPS), Geographical Information Systems (GIS), audio cues, and digital spatial data that produce improved navigation tools. In these contexts, authenticity, coherence with reality, and sometimes high realism, are often expected and required. For instance, using enhanced photogrammetry[7] technology and laser scanning, Oculus has created an immersive VR piece on *The Nefatari Tomb,* where users can have a very detailed view of Queen Nefatari's tomb in Egypt from the comfort of their living rooms, through an interactive VR experience. Here, "state-of-the-art technology has made it possible to digitally scan Nefatari's tomb with millimetre accuracy" (Oculus, 2018). These kinds of VR projects are sometimes pitched as a pro-environmental alternative for physical tourism as discussed in Jeremy Bailenson's book *Experience on Demand* (2019).

The second category of virtual media, as outlined by Fisher and Unwin, refers to data with no situated referents in nature. This includes files, docs, webpages, and other virtual objects that do not referentially have spatial attributes, yet still make sense to computer users. The non-illusory character of these virtual items is mostly anchored by the way they display real functionality; if a user clicks on a folder, it "opens", showing the files contained in the folder. This does not stand for something in the physical world, but instead, a virtual file is virtually (instead of physically) real in its own right.

As mentioned before, we can also easily distinguish a third category of virtual media which represents fictional and imagined objects that are neither existent in the real world nor are specific to a digital environment (as explained in the previous paragraph). For example, In *Resident Evil 7: Biohazard* (2017), a PlayStation 4 VR survival horror game, the player virtually takes the perspective of a fictional character called Ethan Winters, who is on a mission to find his missing wife Mia. Throughout the game, the player has several encounters with a group of crazed, infected cannibals known as "the Baker Family". While playing the game,

---

[7] The American Society of Photogrammetry defines the term as: "the art, science and technology of obtaining reliable information about physical objects and the environment through processes of recording measuring and interpreting images and patterns of electromagnetic radiant energy…". See Nebel, Beege, Schneider, & Rey (2020) for more details.

players often exhibit heightened emotional responses of fear and anxiety. Nevertheless, these responses do not result from a sudden defect in players' judgment about the ontological quality of their virtual experience, nor from the belief that there is an actual danger to their safety. In other words, players do not take their virtual experience at face value. For instance, they do not believe that a three-dimensional being is about to cut them in half using a chainsaw. Instead, they understand, to some technical extent, that the 3D virtual objects being fed into their perceptual experience are a product of an optical illusion, rooted in the way that the brain neurologically fuses 2D displays from the HMD binocular lenses to create the virtual experience of a 3D object.

Keeping these categories of virtual media in mind, French philosopher Pierre Lévy broadly defines virtuality as "a technological mode of actualisation". Lévy suggests that virtuality should, hence, not be restrictively understood as purely antithetical or separate from reality, because although it remains valid that virtual media has a *distinct* underlying ontological structure in being instantiated or existent *qua* "bits and bytes" and computational processes, it still, nevertheless, fulfils real functionalities that are perhaps indispensable in nowadays modern lives. That is to say, the straightforward and concrete referential relationship that some virtual representations have with the real world attests to the genuine functionality that virtuality serves, which cannot be otherwise achieved if virtual content was not genuinely existent. Some of these functionalities include but are not limited to, indirect perception of objects in the real world through VR genres such as immersive (non-fictional) journalism. Another important feature that grounds the reality of virtual content is that virtualisation cannot occur without a human operator and certain technological tools which make acts of virtual actualisation possible. I will return to the significance of this aspect later in this section.

Concerning the process of transformation from the physically actual to the virtually actualised, Lévy explains that this is done by, first, steering technological design towards a continuous problem-solving functionality. For instance, the virtualisation of a business corporation can be shaped vis-à-vis the problem of geographic boundaries which Lévy describes as a "contingent constraint" (1998, p. 31). Virtualisation can hence be partially understood as the reshaping of physical actualities into virtually existent, dynamic actualisations through technological tools that mediate multisensory output, hence enabling numerous functions, such as reliably transmitting sensory input for facilitating communication.

Some may still insist that the virtual is still not "real" in the sense that it is not spatiotemporally "there". Following Michel Serres (1994), Lévy acknowledges that there are germs of truth in the view that recognises the virtual, in some regards, as unreal, namely due to its detachment from the here and now. To overcome this explanatory challenge, Lévy pushes back by suggesting that despite its spatiotemporal detachment, virtual media unproblematically generates one-to-one perceptual experiences and opportunities for virtual interaction, setting it apart from fully imagined experiences, such as dreams or drug-induced hallucinations.

As a concrete example, Lévy refers to virtual texts. This text, for instance, does not currently "exist" on paper nor does it occupy physical space at a given time. Instead, it is a virtual actualisation of a digital hypertext that also "has no place" geographically (1998, p. 28). In other words, the original – physical – manuscript of the text (if existed) is not present in the flesh, so to speak, in every point in the network connected to the digital memory in which the text is encoded, nor is it present in every terminal from which the text can be copied. Nevertheless, as Lévy explains, the spatiotemporally detached hypertext still generates real events of textual actualisation, navigation, and reading (ibid., p. 30). Therefore, Lévy grounds the reality of virtual media on a *sui generis* mode of actualisation. Virtual actualisation is then not only real in the sense that it is experienced in real-time by several users, but also in its pragmatic impact and functionality, as previously stated.

However, there is still a worry that following these somewhat loose criteria for defining reality and virtuality, acts of memory, thoughts, dreams, and other spatiotemporally detached states may also fit to be referred to as "a special mode of reality", which further blurs the lines between the real, the imaginary, and the virtual. This worry can be avoided by emphasising that dreamlike states lack virtual actualisation that can only occur through *technological mediation*, essentially involving the use of technological hardware supervised (or at least pre-engineered) by human operators, consequently simulating virtual objects of perception and enabling opportunities for pragmatic and collective virtual interaction. This type of concrete mediation, offered by technologies such as VR, and even by simpler technological equipment, hence makes a clear separation between virtual experiences and fully imaginary, non-mediated experiences such as dreams, fantasies, and drug trips.

In what follows, and following Michael Heim, I shift my attention toward discussing VR technology as a subset of the broader notion of virtuality, with a focus on the novelties that VR introduces in comparison to earlier mediums such as 2D non-interactive films. Among the three

essential elements that Heim sets for VR technology, I discuss immersion and interaction as theoretical prerequisites for critically tackling some psychological and perceptual illusions that VR often triggers, in the following section of this chapter.

2.1. On VR Technology

Heim writes: "if we are to truly understand virtual reality as part of the dynamics of cultural evolution, we have to focus on what exactly it is, and we have to put aside for a moment the loosely associated meanings and spin-offs" (1994, p. 28). Here, Heim advises against understanding VR with exaggerated reference to conceptualisations most prevailing in sci-fi and popular culture, often defining this complex technology in terms of a "consensual hallucination" as described by William Gibson in his dystopian novel *Neuromancer* (1984). Conceptualisations inspired by science fiction hence tend to be detached from the actual status quo of technological development. Such do not only breed confusion and ambiguity about what VR technology really is but also produce awkward philosophical questions such as whether or not it would be possible to live our whole lives, permanently plugged into a VR experience machine, and how can we tell the difference between reality and virtuality in the far future. Such questions tend to miss the point about VR, and other technologies, being extensions and dynamic instruments that we strategically tame for a myriad of non-illusory, pragmatic, utilisations that correspond with fulfilling our needs in the real world.

Philosopher of technology Don Ihde, for instance, explains that both human subjects and technology have a dynamic relationship that is shaped reciprocally through embodied interaction. Contrary to phenomenologists like Martin Heidegger, Ihde does not draw any conclusions about a fixed metaphysical essence of either humans or technological tools, and he also acknowledges the instrumental role that technology plays for human users. He adds that technology is not neutral, in that its usage constantly shapes human praxis through mediation, which in turn guides and influences actions and acts of thinking and decision-making. Technological tools hence influence and are influenced by human praxis. To Ihde, this is the core of human-technology relations.

To further illustrate, Ihde uses the example of scientific instrumentalization of technology in the use of the telescope and the microscope, which display objects in more nearness and detail, in comparison to when observed with the naked eye, hence unravelling more potentials for scientific inquiry. This is similar to the case of mediated perception through teleoperation systems. Using technological (robotic) instruments, the human operator perceives

the target environment through mediation, and the acquired sensory input directs her decision-making and mediated actions in the target environment, which would otherwise be too risky or complicated to perform first-hand.

Returning to VR technology, I adopt Heim's framework which precisely defines it as "an immersive, interactive system based on computable information. These defining characteristics boil down to the "three I's" of VR: immersion, interactivity, and information intensity." (Heim, 1998, pp. 6–7). More thoroughly, VR is a technology that fosters the simulation of computer-generated, virtual (representational) environments that can be experienced and interacted with in ways that sometimes mimic our interactions in the physical world. Multisensory experiences, as Heim stresses, are a fundamental aspect of VR technology. The most popular hardware associated with VR is the head-mounted display (HMD) which represents the observer with a 2D perspective projection, or images, for each eye, from which the brain infers a 3D property of simulated "virtual objects".

In addition to sense perception, motor movement in VR is very crucial in terms of acquiring multisensory information represented in a given VE. In addition to visually immersing users in computer-generated multisensory worlds, VR usually incorporates devices for tracking the user's movement in the physical world and updating their real-time body representation in the VE accordingly. Manufacturers often either integrate movement tracking devices in the HMD or through non-optical trackers that attach microscopic, electromechanical sensors to other pieces of VR hardware. Put simply, a change in the user's physical body movement would lead to a change in their visual field in VR, mimicking our sensorimotor channels and mechanisms in the physical world when acquiring sensory information from our environment. In what follows, I will briefly elaborate on two of VR's key features, distinguishing this technology from non-VR mediums such as 2D non-interactive films.

2.1.1. Immersion

Philosophers of technology have different understandings of immersion in VR, in that some use this term interchangeably with "virtual presence" (Oh et al., 2018). Alternatively, I contend that while immersion is closely related to virtual presence, it is nevertheless useful to distinguish the latter from the former, namely because one does not always entail the other. Empirical studies have shown that high levels of immersion do not guarantee the generation of the subjective experience of virtual presence (Amin et al., 2016; Slater, 2018). Immersion can, therefore, be separately defined as a system's technological capacity to create a sensory, vivid

illusion of inhabiting a VE. This is achieved when footage from the HMD *blocks* the user's field of view of their immediate physical environment.

Filmmaker Chris Milk (2015) describes VR immersion as centralised on taking the user "out of the frame". Usually, 2D non-interactive films are created as a compilation of shots, producing scenes which we view, as Milk describes, within clearly defined rectangles or frames. Through the use of 360-degree footage, VR filmmakers succeed in creating the illusion of generating a navigable "open world". When feeling "surrounded" in a VE, users often report a visceral sense of being perceptually transported into said VE. A crucial defining feature of immersion is that it can be objectively measured against the technological modalities and the physical hardware of the VR simulation in question. For instance, the *quality* of latency bandwidth, audio-visuals, frame rate, stereoscopy, and field of view, determine the extent to which a VR system is immersive.

In his new book *Reality+* (2022), David Chalmers explains immersion in virtuality as experiencing a VE as "a world all around us, with ourselves present at the centre" (p. 193). Chalmers adds that immersion occurs in degrees. Across the literature, immersion is commonly described as the unique property that distinguishes VR from more traditional mediums (Shin & Biocca, 2018; Slater, 2003; Steuer, 1992). In other words, the use of tools such as VR goggles, haptic gloves and motion detectors provides a straightforward distinction between VR and non-VR. However, in their widely cited paper (1997), Slater and Wilbur suggest that reference to the tools involved in building or having a VR experience does not thoroughly capture all unique aspects of a VR user's experience, in that by strictly focusing on immersion we tend to marginalise the user's subjective sense of "being" in another space that is different from their physical environment. Prima facie, earlier mediums do not seem to generate this kind of experience as VR does, at least not to the degree that VR is reported to. I will discuss the subjective experience of virtual presence in more detail in the next section of this chapter.

2.1.2. Interactivity

As mentioned in the overview of this chapter, we can distinguish, at least, two layers of interaction integrated into almost all VR experiences. The first layer involves an unproblematic kind of causal – physical – interaction, occurring directly between the user and the VR hardware, manifested in users' utilisation of VR controllers and other VR-specific devices. When discussing the second layer of interaction, however, things tend to get a bit (philosophically) tricky. The second layer refers to an interactive loop, involving virtual media

and the user's avatar (or designated body representation). In addition, virtual objects in VR are also sometimes designed to interact with one another without the need for users' interference. There are two main views in the literature attempting to accurately put forth a causal description of this layer of (pseudo) interaction.

Before discussing these two views, we first need to note that it is widely agreed upon, by virtual realists and irrealists, that virtual representations are generated *qua* digital objects. The phrase "digital object" refers to the plethora of computational processes, elements (e.g. bits and bytes, algorithms, servers…etc), and raw data structures which developers use for generating virtual media, which in turn generate virtual media (or objects, in virtual realists' term). Virtual media refer to multisensory data such as sounds, images, and other representational features of virtually *animated* content, generated as a product of hardware and software dynamics.

As mentioned above, we can distinguish two main views seeking to capture the kind of interaction that occurs at the representational, second layer of virtual experiences. The first view is defended by Chalmers (2017, 2019). As a virtual realist, Chalmers seeks to ascribe a genuine sort of causality to virtual representations by *identifying* such representations *with* digital objects. To iterate, Chalmers argues that virtual representations (or media) are *identical* to their underlying computational processes; i.e. digital objects.

Contrarily, the second view defended by *virtual fictionalists*, such as Neil McDonnell and Nathan Wildman, equally allows that for every virtual object there are underlying digital objects, i.e. computational processes, making the generation of virtual objects possible. However, they reject the identification relation that Chalmers postulates between virtual and digital objects. With that said, they conclude that interactions between digital objects are genuinely (causally) real, while they deem interactions between virtual representations as a manifestation of pseudo-causality at best.

In an attempt to settle this disagreement on the right kind of interaction that occurs between virtual objects and the users' virtual representation in VEs, we can refer to Chalmers' observation that "virtual objects *do not look like* (emphasis added) digital objects, at least to the naive user" (D. J. Chalmers, 2017, pp. 319–320). This undermines the identification relation that Chalmers himself constitutes between virtual and digital objects, making it more coherent to adopt virtual fictionalists' description of the pseudo-causality that occurs at the representation level of virtual experiences. In other words, virtual and digital objects cannot be identical and different at the same time.

Borrowing Peter Ludlow's example, if one is invited to a virtual sushi dinner, "there will be no rice or fish entering their physical body. And the virtual sushi may be represented on a virtual table, but we needn't worry about bruising our shins on the virtual table" (2019, p. 2). There are, nevertheless, digital objects instantiating virtual media and representations, made possible by the existence of technological hardware and human operators. Therefore, it is more reasonable to not commit a categorical conflation between these elements and concede that the seeming interaction between multisensory output in virtuality only entails pseudo-causality at best. From this perspective, virtual fictionalists (another appellation for virtual irrealists) are right in suggesting that all virtual objects are, so to speak, ontologically fictive; i.e. virtual objects are certainly not causally equal to their full-blown, physical counterparts. At best, we can often encounter representations of properties that are unique to the physical world in the way that virtual objects seem to behave. In a similar vein, the next section discusses common, subjective, psychophysiological illusions that VR users often experience when using a VR immersive and interactive system.

3. VR-Induced Psychophysiological Illusions

Having an immersive, interactive, VR experience, essentially implies exposure to the optical illusion of inferring 3D images from receiving output from *two* 2D screens on a given HMD. This binocular structure – mimicking left-eye and right-eye functions – creates the impression of depth perception by "fusing" both images together upon utilising VR HMDs. Moreover, most VR systems also tend to involve a technical mapping of the user's sensorimotor functions. This mapping is subsequently synchronised with the representational displays shown on the 2D screens. As mentioned before, undergoing a VR experience entails immersion through the utilisation of VR-specific tools, which strive to create an open-world view, perceptually placing the VR user at the centre of the virtual environment. Immersion is sometimes interchangeably used with the notion of virtual presence, however, the latter can be more specifically defined as the *subjective experience of spatially occupying another place*, while maintaining an "offline" awareness of one's immediate physical environment. Some assume that virtual presence requires the "willing suspension of disbelief" about one's physical spatiotemporal location, reinforced by an attentional shift toward virtual stimuli.

Another illusion that is often felt when using VR is that of virtual embodiment, also known as the illusion of body ownership. Building upon research on psychophysiological phenomena such as the "rubber hand illusion" (Botvinick & Cohen, 1998; Ehrsson et al., 2005)

and insights on the topic of "embodied cognition" most noticeably discussed by phenomenologist Merleau-Ponty (1962) and others, we can understand VR's ability to stimulate these illusions vis-à-vis the technological enhancement of instantaneous, uninterrupted synchrony between the user's movement in the physical world and their body representation in the virtual environment. This section will hence critically elaborate on these two types of illusions and some of the philosophical questions they inspire.

3.1. The Illusion of Place: On Virtual Presence

In the 1980s, Marvin Minsky introduced the term "telepresence" to describe the subjective experience triggered during the use of human teleoperation systems, which are designed for creating safer, more cost-effective work conditions for employees, especially in dangerous work environments (e.g., military training, mining, and construction) (Held & Durlach, 1993, 1992). Teleoperation allows human operators to remotely control artificial operators for receiving quantity and quality feedback that is harnessed for task completion in designated, physical, target environments. The more the human operator experiences an embodied sense of being in the target worksite, the better their performance becomes. Since the term has been coined, telepresence has been closely associated with the sense of embodied *transportation*, in that maximising the feeling of "being there" is considered key to dexterity (Schloerb, 1995).

Since Minsky, the concept of telepresence has undergone major conceptual expansions, surpassing workplace interaction and teleoperation to include *virtual presence*. First coined by Thomas B. Sheridan in 1992, virtual presence refers to the subjective, psychophysiological experience (or illusion) of non-mediated situatedness in virtually simulated environments that are virtually actualised, instead of physically, remotely existent. According to Waterworth & Riva's description (2014), VR users *simultaneously* develop an experience of presence both in the physical environment where their bodies and technological hardware are situated, as well as a feeling of presence in VEs toward which their attention and sensorimotor functions are directed. The subjective experience of virtual presence in highly immersive VEs, according to Lévy, is only different in degree, but not in kind, from the sense of presence developed when having a telephone conversation or when interacting with text-based content on virtual platforms like websites and word-processors (1998, p. 28). Sheridan bases his conceptualisation of virtual presence on the success of perceiving sensory cues and then employing this sensory input for virtual interaction and task completion in a designated virtual environment. Here, we can deduce that virtual presence relies on the successful acquisition of

virtual, sensory stimuli as well as synchronised embodied interaction. When discussing virtual presence, Sheridan still metaphorically draws upon the illusory sense of transportation experienced during teleoperations. It goes without saying that VR users are not literally physically transported to a different location. However, due to the vividness of sensory stimuli and engaging virtual interaction, they often experience an impression of transportation.

A helpful feature distinguishing between telepresence and virtual presence is that in the latter the artificial operator, controlled by the human operator, is *digitalised* and virtually represented most commonly as an avatar, or a cursor in cases of less representationally loaded virtual environments, meaning that the user does not remotely manipulate or alter any objects in a remote physical environment. Instead, technological hardware simulates perceivable content and affords mediated perception and manipulation of virtual objects (although there exists some hybrid training simulations that involve both teleoperation and virtual representation).

When describing the subjective experience of virtual presence, some theorists, such as Baudrillard, interpret it as essentially caused by a defect in users' rational judgments about the quality of their virtual experience. In other words, some assume that virtual presence and users' often heightened, realistic responses towards virtual stimuli are only generated iff users *mistake* or are "tricked" into believing that they are actually transported into another environment.

There are two watered-down versions of this characterisation of virtual presence. First, building upon Walton's fictional theory of "make-believe", where he states that our engagements with fictional events *mandate* imagining said events to be "fictional truths" (1990, p. 39), philosopher Eric Studt postulates that in order to effectively experience virtual presence in VEs, one is mandated to imagine virtual events as either fictionally occurring or non-fictionally so (2021). Even in response to non-fictional VR experiences (such as VR documentaries), Studt contends that there is a requirement to employ imagination, in making believe that portrayed, non-fictional events, have been really documented and are not instead staged fictional movies. If we do not use imagination as a vehicle to make sense of VR experiences, Studt insists that we would not be able to experience virtual presence nor display heightened emotional responses toward virtual stimuli.

Secondly, Another, perhaps less implausible, characterisation of potential criteria for virtual presence is cashed out in terms of "*willingness to suspend belief*" as discussed by Marie-Laure Ryan (2001) and VR and AR (artificial reality) developer Jamie McRoberts (2017).

Here, VR users are claimed to intentionally sanction their beliefs about their virtual experience being ontologically virtual, as discussed in the first section, as a condition to successfully experience virtual presence. Nevertheless, it must be acknowledged that both Ryan and McRoberts do not argue that suspension of belief is the only criterion for developing the subjective experience of presence. For instance, in his framework to study virtual presence, McRoberts proposes four necessary dimensions that affect the subjective experience of situatedness in VEs, namely immersion, the positionality of the user, interactivity and narrative agency.

With that said, I contend that we need not associate belief suspension or make-believe with the subjective, psychophysiological, experience of virtual presence. Mainly because these mental processes do not effectively capture virtual presence as a visceral, reflexive, experience, which is more akin to emotional contagion (discussed in the first chapter) than to intentional make-believe and disbelief suspension. Furthermore, attempts to describe virtual presence and users' realistic responses in VR vis-à-vis cognitively demanding, attentional processes do not account for the immediacy of such embodied experiences. Just like emotional contagion, we can characterise virtual presence as a contentless reflex that can be described as neurologically *involuntary* to a large extent, similar to the rubber hand illusion commonly discussed in psychology. That is to say, when using VR, we do not experience virtual presence thanks to intellectually and cognitively informed thought processes of belief suspension or make-believe. Instead, we *directly* encounter virtual media and make use of the pseudo-causal interactive affordances it offers, without the need for our visceral responses to depend on an extra layer of imaginative thinking.

To iterate, I argue that the claim that VR users need to employ imagination in order to experience virtual presence does not capture instantaneous, visceral, responses that users display when faced with virtual stimuli. Even more, it seems redundant to assume that we need to resort to imagination to make sense of already, directly, perceived stimuli in virtuality. Studt's contention that we also need to employ imagination to experience virtual presence in response to non-fictional documentaries seems paradoxical; why should we imagine non-fictional (and even fictional) events as being truly fictional or non-fictional while we can already sensorily and contextually perceive them as such? I, therefore, argue that this view creates more explanatory gaps and ambiguity instead of unequivocally explaining what virtual presence entails.

3.2. The Illusion of Embodiment (Experiencing Body Ownership in VR)

In his paper, "The Cyborg's Dilemma: Progressive Embodiment in Virtual Environments", Frank Biocca writes: "before paper, wires, and silicon, the primordial communication medium is the body" (2006, p. 13). Biocca explains that the body plays an important role as an information acquisition device, which has the ability, not only to receive but also to react to acquired information, in turn producing more information that can be received by other bodies generating more responses, and so on. Building upon this empirical observation, Biocca explains how several computer interfaces are seeking to integrate what he terms *progressive embodiment* within their framework of development (ibid., p.14). In their early humble beginnings, most computer interfaces representationally indicated the user's locus of control within a virtual environment through virtual cursors, sometimes portrayed as a "hand". This tacit anthropomorphic design tendency eventually evolved into full-body virtual representations. As Jacquelyn Ford Morie further explains: "VEs, for example, engage the body as kinaesthetic input via the specialised interface devices that not only permit but *require* bodily actions to be performed kinaesthetically – within a full 3D virtual construct" (2007, p. 126).

With accelerated technological innovations, bodily senses such as vision, hearing, and touch have been mapped into virtuality through HMD, quality sound systems, motion trackers and haptic bodysuits, respectively. This is often done alongside the isomorphic use of "avatars" that can be representative of the user's body, while also functioning as semantic markers for sociocultural identity. Biocca explains that: "implicit and explicit *social norms* that may be partially idiosyncratic to the virtual environment and imported from the user's social environment finalize the social-semiotic role and identity of the avatar. Issues of class, gender, occupational role, body type, etc. are raised when considering this aspect of embodiment" (Biocca, 2006, p. 23, emphasis added). Therefore, virtual body mapping includes the replication of sensory information channels as well as social, often stereotypical, representations attributed to avatars. In the context of virtual spaces dedicated to online communication and socialising, users can select their virtual "skinsuits", so to speak, and make use of interactive possibilities afforded by their virtual environment, creating multi-layered communication channels through which they express themselves and respond to each other within virtual spaces. Famous examples of these types of virtual spaces include video games like *The Sims* and *Second Life*. Therefore, through virtual embodiment channels, users can

"read"[8] sensory feedback and alterations of virtual media that, in turn, correspond to their physical movements and are designed to instantiate continuously updated alterations in users' collective, perceptual, field of view in VEs intended for social interaction.

As mentioned at the beginning of this chapter, the transparency relation that happens through embodied (i.e. practical and instrumental) use of technology provides a reliable theoretical foundation for explaining the sensory illusion of virtual embodiment. In his book, *the Phenomenology of Perception* (1962), Merleau-Ponty describes the transparency relation that is created when visually impaired individuals use a stick to gather sensory input about their environments as follows:

> The blind man's stick has ceased to be an object for him and is no longer perceived for itself; its point has become an area of sensitivity, extending the scope and active radius of touch, and providing a parallel to sight. In the exploration of things, the length of the stick does not enter expressly as a middle term: the blind man is rather aware of it through the position of objects. (1962, p. 165)

Analogously, in her article, "Living a 'Phantom Limb'" (2010), Vivian Sobchack describes her own experience of undergoing amputation and consequently implementing a prosthetic limb. Sobchack discusses the notion of a "phenomenal lived body", experienced as a result of having a given anatomical structure. She writes that through our phenomenal bodies (i.e. the subjective experience of our bodies) "we perceive and express our bodily comportment and movement but not our anatomy, our gestures and thoughts but not the firing of our neurons" (2010, p. 52). This observation is made not in order to discredit physical and neurological underpinnings of perceptual experiences but to simply point out that "material causes and processes of these sensations are not identical to their experiential effects" (ibid,.). What is under scrutiny here is, therefore, not only the anatomical boundaries of the user's physical body, but also the lived sensory experiences that lead to grasping parts of the body as a unified whole, separating it from the rest of the world, and characterising it as one's own body.

Looking closer at Sobchack's description of the embodied experience of using a prosthetic leg, it can be highlighted that the proprioceptive feedback provided by the leg carries several affinities with Merleau-Ponty's example of the blind man's cane, which functions as

---

[8] Ihde also explains another type of human-technology relation, referred to as "hermeneutic relation", where humans perceive and interpret data as displayed by technology in reference to the world (reading the temperature off the thermostat is an example). See Verbeek (2001) p. 124.

an "area of sensitivity" that is used to locate the blind man's body vis-à-vis his environment. This embodied use of such tools offers transformed sensory input that is acquired with reliance on haptic (instead of visual) feedback to guide actions and decision-making. In parallel, Sobchack states the following:

> Now, having incorporated the prosthetic, I primarily sense my leg as an active, quasi-absent 'part' of my *whole body*…I do not focus on or feel my leg as 'some thing'. I do not feel the objective 'place' where the flesh of my stump ends and the material of my prosthesis begins. (2010, p. 62)

She further explains that the embodied use of the prosthetic is characterised by a general "seamlessness" (Ibid.), as accounted for in Ihde's theoretical framework on transparent, human-machine relations. This seamlessness, is hence, anchored by the embodied functionality that underpins a transformed perceptual experience of the body, leading to a recognition of the leg as a tool for mediated haptic (and motor) functionality, that works in harmony with other body parts.

Even though the transparency relation theory accounts for the seamlessness effect that occurs through embodied virtual interactions, where implemented VR hardware fades into the background of the user's experience, creating an illusion of non-mediation, virtual embodiment, prima facie, does not seem to afford the same extension of somatic and bodily capabilities in the way that a prosthetic limb or the blind man's cane does. That is to say, if a user has a type of physical impairment, her embodied representation in VR cannot amend or contribute to enhancing or modifying any functions of her body in the physical world. Focusing on the aspect of proprioception, Kirk Besmer (2015) shows that virtual embodiment does not provide any relevant sensory feedback that would lead to an alteration in the grasping of physical body boundaries. He explains: "We should avoid thinking of remote robotic arms and avatar 'bodies' as technological extensions of embodiment in the same order as bodily co-located tools such as the blind man's cane or the carpenter's hammer. Doing so implies the abolition of distance and integration into the body schema that does not, in fact, occur"(2015, p. 68). This observation is also valid because, as explained by Heim, the technological hardware associated with virtual embodiment blocks the user's reception of sensory input from the physical world. What remains is the virtual, sometimes socio-culturally significant representation, which can only resonate with the user at best, yet cannot be recognised as a literal extension of one's body. Besmer better explains this by stating that: "while one might identify with one's avatar, one does so in a self-referential manner indicative of an intentional

body image. In no sense can it be understood that one's avatar comes to be integrated into one's pre-personal body schema but, then, neither do the remote 'arms' of the telerobotic system" (2015, p. 67).

4. Social (Veridical) Objects in Virtual Reality

In an attempt to provide an alternative view to the dichotomously constructed realist and irrealists characterisations of virtuality in philosophy of technology, I will now argue that virtual environments involve both the existence of illusory effects and veridical components, together constituting a sui generis mode of virtual actualisation, which reflect coherent and legible representations and affords different layers of causal and pseudo-causal interaction with virtual content. So far, I have elaborated on virtual presence and virtual embodiment, identified as psychophysiological illusions, typically experienced when interacting with virtual media. Concerning paradigms of veridical components, the first section already foreshadowed the existence of referential relations that sometimes ties virtual media to geographic locations and other aspects of the physical world, enabling the fulfilment of teleologically significant functions such as navigation, visualisation, communication, and reading. These paradigms underline the functional realism that can be legibly and coherently mediated by virtuality, unhampered by the psychophysiological illusions outlined above. To support my contention for the existence of hybrid illusory and veridical constituents of virtuality, this section elaborates on the purely veridical generation of *social objects* (also known as *institutional facts*), building upon John Searle and Philip Brey's writings.

I suggest that what Searle terms "social objects", can also be recreated in virtuality, beyond illusions. To unpack this claim, this section will first briefly explain Searle's distinction between physical "brute" facts and socially constructed facts. The latter are said to be conditionally held by *collective intentionality*, establishing that object (x) counts as (y) in context (c), not by virtue of the physical properties *inherently* existent in (x), but because of a function that is imposed on (x) through a collective agreement among members of a given society. After laying down the foundations of Searle's theory of social ontology, I will argue that virtuality can also be a host to *some* institutional facts, such as currency, property, and other forms of socially-established facts. I will be supporting these claims with reference to Brey's paper, "The Social Ontology of Virtual Environments" (2003). By the end of this section, I will highlight some of the limitations and determinant factors for the virtualisation of social objects.

64

In his book, *The Construction of Social Reality* (1995), Searle suggests that there are features of our world which exist independently of our existence, attitudes, and judgements (p. 9). These features are intrinsic to physical nature, including, for instance, mass and the chemical composition of objects. That water is compounded of two hydrogen atoms and one oxygen atom, and that it exists in gaseous, liquid, and solid states, are examples of facts that are intrinsic to the nature of water. That is to say, water will maintain these states and properties even in the absence of an observer.

Next, Searle distinguishes another category of objects – more accurately functions assigned to objects – which he calls institutional facts (ibid., p. 48). Such facts are commonly assigned through collective intentionality, or simply through convention, where an object (x) counts as (y) in context (c), not necessarily by virtue of anything substantial in the physical form of (x). In this vein, Searle distinguishes *two* main occasions where a function is agentively assigned to an object that may be initially devoid of it. The first is attributed due to a supplementary, usually embodied, i.e. practical, function that object (x) is *designed* to fulfil. A recurrent example of these kinds of objects is screwdrivers (ibid., p. 20), in that they are intentionally designed to drive screws. This tool (and many others) is often made of materials like wood and metal that do not separately, intrinsically possess the functionality of driving screws, unless manufactured and assembled in a specific way, creating an object that is identified as a screwdriver.

The second context, closely relevant to the topic at hand, refers to functions assigned to certain objects not thanks to any practical functionality or properties inherently present or designed in such objects, but instead by a sheer collective, social agreement. Searle illustrates that: "…bits of paper, are objects like any others, but the imposition of *status-functions* [emphasis added] on these objects creates a level of description of the object where it is an institutional object, e.g., a twenty-dollar bill" (ibid., p. 48). The imposition of a status function is, therefore, *less dependent* on the intrinsic features of the object in question and more so on collective intentionality, *context*, and sometimes the practicality with which a given object fulfils the required institutional function. For instance, a twenty-dollar bill may not be an accepted currency of trade in a tribal community, underlining the constitutive importance of the social context for defining a given institutional fact. Searle adds that social objects are often distinguished by agentively added "identity markers", such as using signs that read "Stop" to organise traffic or placing a sign that reads "sold" to semantically indicate a status of a property.

As mentioned in the first section, some core functions of physical objects can be represented in virtuality, such as calculation, guiding navigation...etc. Drawing upon Searle's theory of social ontology, Brey argues that social status functions can also be genuinely established through virtual media. He uses as an example that there can be an agreement that yellow rooms are women-only chat rooms in a given virtual environment. "For a user in this environment to recognize a virtual entity as a women-only chat room, he or she must recognize that virtual yellow rooms (x) *count as* [emphasis added] women-only chat rooms (y) in the context of that virtual world (c)" (Brey, 2003). As virtual irrealist may insist, things like "yellow" and "rooms" cannot physically exist in virtuality beyond virtual display, in the sense that such display can only be a projection of computational processes afforded by certain types of hardware. This need not be problematic for defending the reality of social status function in virtuality, namely because what we identify as virtual entities, can also be treated as real, social tokens of meaning that users recognise via collective intentionality beyond the confines of illusions. These tokens function as abstract symbols that organise *genuine* communication and interaction in virtuality.

This account explains why activities such as buying, selling, voting, chatting, playing a game of chess, gambling, or stealing can fully exist in virtuality. Following Brey, I argue that such things are possible because these flexible activities are not ontologically reliant on a fixed or finite set of physical properties or objects to be created in reality in the first place. For instance, a virtual voting poll on social media exists in full capacity as a non-fictional and non-illusory institutional object and its results are identified as a true representation of the public's opinion, without the need for a ballot box. In other words, the functionality of a ballot box is replicated by its virtual counterpart. The latter is not ontologically equal to its physical counterpart, and yet it is still suitable for preserving the function of collecting votes.

To recapitulate, virtual objects can be designed to symbolically represent institutional objects that stand for functions we socially assign to certain objects despite the fact that said objects naturally lack these functions. As mentioned in Brey's example, yellow rooms do not naturally carry the social significance of being a women-only chat room, yet they can be rendered so in reality, as in virtuality, provided there is sufficient agreement on that. In such situations, users can often encounter virtual representations that reflect social norms they are already familiar with. A widely reported outcome of such familiarity is described in terms of lowering users' learning curve, allowing them to interact in virtuality *as if* they were in reality, as shown in (Jerald, 2015, p. 227).

Virtual fictionalists may contend that grasping and interacting with social objects in virtuality is also reliant on make-believe, where it is claimed that in order for a user to successfully follow the rules made to organise a given virtual environment, she needs to *pretend* that there are yellow rooms that also happen to be women-only chat rooms. In response, it is evident that to naïve users, the intricacy of generating virtual objects through computational processes might be ambiguous, in comparison to more experienced users or programmers (D. J. Chalmers, 2019; Juul, 2019). However, even if I have limited technical knowledge of how virtual display is generated, it does not follow that I *need* to pretend that said display is "genuinely" real in some vague, metaphysical sense for me to be able to make use of the interactive possibilities afforded by a virtual environment.

As previously discussed, it is sound to maintain that virtual objects are not perceptually identical to the raw codes and data through which they are generated. Some virtual fictionalists, such as McDonnell and Wildman, also allow that there is a pseudo-causality between virtual entities in the way they seem to behave in virtuality. At a bare minimum, users would know that the displays they experience are a result of a computer generation, which perceptually affords sensory display and *virtual* interactivity. Instead of pretending that virtual yellow rooms are genuine yellow rooms, as children pretend that stumps stand for bears in the forest in McDonnel and Wildman's example, it is sufficient to simply recognise the virtual as virtual (as pointed out by Chalmers). This involves *reading* and perceptually recognising the principles of generation specific to a given virtual experience, which is in turn undeniably grounded in digital objects. This recognition can be done without committing to any sort of pretence that entails that virtual entities are real beyond the confines of virtuality.

Juul explains that: "As users, we maintain a double consciousness where we think of an object in virtual reality as a full *fictional* entity, as well as something we can interact with in the specific way determined by the *rules* of the virtual world (Juul, 2019, p. 9). "Fictional" here refers both to the consistently virtually grounded ontology of virtual display, as they do not themselves bear any causality or any other physical properties, as well as the possibility of using virtuality to deliver fictional content. Therefore, as mentioned before, I believe that successful virtual interaction does not require make-believe, instead, it is sufficient for the user to be made aware of the principles of generation organising the virtual environment one experiences.

Understanding virtual media from the scope of Searle's theory of social ontology reveals an additional layer of pragmatic functionality in virtuality, where a virtual object (v) is recognised as (y) in context (c). As is the case with virtual maps, institutional facts, such as currency, can equally be present in fiction. Brey sensibly addresses this by maintaining that: "for virtual institutional objects and actions to be fully real, they must be part of an institution in the real world, rather than a simulated one. For example, virtual money is only real money if it can be transferred to one's bank account or be used to make real purchases" (Brey, 2014, p. 6). Here, the contextual and institutional intersection between virtuality and reality is essential for distinguishing non-fictional, virtual social objects from their fictional counterparts.

Another issue that virtual fictionalists may raise is that in the same way a coin may be used as a screwdriver and not count as one, virtual objects can function as social objects when it would still be insufficient to consider them as such. Cryptocurrency, for instance, can be used as money to purchase goods, yet due to its restrictions, it cannot be used as effectively as hard currency in the physical world, perhaps discrediting it as a fully real currency. One way to address this worry is to emphasise that these limitations only stress the importance of intended use and design, social context, and collective intentionality as chief markers of the boundaries between fictional and non-fictional social objects. This observation also sets reasonable limitations on the creation of social objects in virtuality, wherein these virtual objects need to contextually intersect with real-life usage. Perhaps in the future, virtual marriage ceremonies can be considered legal and cryptocurrency might replace paper money (although this might be problematic due to the limited amount of cryptocurrency available and the demanding process of mining it). However, if this indeed occurs, it will essentially depend on collective intentionality and the rules of the age that would socially assign such status functions to virtual entities through societal agreement. This assignment will definitely be relevant to the practicality that this use would harbour. However, building upon Searle's insights, the ontologically sui generis nature of virtual objects will not disqualify them from obtaining such status functions, because in theory, as previously explained, the imposition of a status function is not dependent on the intrinsic (metaphysical) properties of the object in question.

5. Chapter Conclusion

VR technology pioneer Jaron Lanier wrote in the opening of his 2017 book, *Dawn of the New Everything* that "VR is one of the scientific, philosophical, and technological frontiers of

our era. It is a means for creating *comprehensive illusions* that you're in a different place…" (qtd. in Chalmers 2022, p. 206, emphasis added). Echoing these intuitions, this chapter has identified virtuality as a sui generis mode of actualisation, through which virtual representational media (i.e. images, sounds, animation, simulations…etc) are generated *qua* computational processes and data constructs (or digital objects for short). Virtual media can function as a referent to things in the physical world or pure fiction, but both of these implementations are contextually comprehensive and generated by human operators and concrete technological hardware.

As alluded to by Lanier, VR experiences are also, in some sense, illusory. I have suggested that it is crucial to clearly explain what makes VR illusory insofar as it mitigates hyperbole about the psychological and behavioural effects that virtual experiences are sometimes said to induce in real-time and on longer terms. At a bare minimum, when using VR HMD, we go through the optical illusion of inferring a 3D object from the two 2D screens. As discussed in this chapter, some of the other psychophysiological illusions that VR is reported to trigger include virtual presence and virtual embodiment, referring to the subjective experiences of non-mediated situatedness in a different place and feeling bodily agency over a virtual representation. Here, it is important to note that these so-called illusions do not cause a defect in users' rational judgment, nor necessitate cognitively demanding processes such as imaginative make-believe or the willing suspension of belief about the ontology of virtual experiences. With that said, I have, therefore, rejected Studt's claim that in order to experience these illusions in VR, one needs to engage in a playful act of intentional make-believe, where they imagine represented events as either fictionally or non-fictionally true. I have rejected these intuitions because they fall short of explaining the immediacy of reflexive responses to virtual stimuli. I have also argued that the make-believe account does not accommodate the availability of direct, collectively experienced perception of virtual media, rendering imaginative make-believe a redundant, unnecessary step.

Another, less ambiguous, interpretation argues that psychophysiological illusions necessitate a willing suspension of disbelief. I have argued that both of these interpretations which assume the occurrence or necessity of a belief formation process for experiencing illusions seem counterintuitive and do not cohere with the phenomenology of virtual experiences. Alternatively, the post-phenomenological interpretation of the two illusions in question emphasises that the embodied use of technological instruments reliably and uninterruptedly builds up a transparency relationship between humans and tools. Here, the tool

fades into the background of the user's experience, constituting a direct relation between the user and the world (in the case, the virtual environment). It is worth mentioning that technological immersion is never completely seamless. This is what Don Ihde refers to as "echo focus", asserting that tools can never be fully absorbed into our subjective experiences of "the self". Macpherson also underestimates the permanent sustainability of the psychophysiological illusions in VR. She writes: "virtual reality [is] experienced as less realistic than it once did. The power of the technology wanes!" (2020, p. 32). This further emphasises that experiencing these illusions does not hamper the mediation of coherent and legible information that does not require make-believe and does not cause a defect in rational judgment, separating virtuality from hallucinations and dream-like states.

In the next chapter, I begin by highlighting some important disanalogies between the current implementations of VR technology and Nozick's "experiment machine", introduced in his influential thought experiment (1974). The purpose of this comparison is to support the contention that engagement with standard VR technology entails real (as opposed to illusory) harm and benefits. As a paradigm of the concrete benefits that can arise as a result of using VR technology, I discuss know-how knowledge transfer and the experimental use of VR as an empathy machine for reinforcing interpersonal understanding. As for the potential harm that VR users can experience when engaging in a virtual experience, I discuss the intriguing case of "virtual theft". Refocusing on the use of VR as an empathy machine, this chapter ends by a discussion of users' idiosyncratic responses to VR experiences, specifically designed for reinforcing communication and interpersonal understanding.

Chapter 3: VR as an "Experience Machine" Versus an "Empathy Machine"

1. Chapter Overview

In his book *Anarchy, State, and Utopia* (1974), Nozick proposed his famous "experience machine" thought experiment, where a user can choose to plug into a device that generates "virtual" simulations that are sensorily indistinguishable from physical reality. What is special about Nozick's experience machine is that it generates *pleasurable* experiences, more specifically whatever the agent inside the machine desires, also erasing their memories of ever entering the machine. Such an option, in Nozick's thought experiment, would also be available to other members of the agent's society. More recently, in a series of articles for Forbes magazine (2000), Nozick attempted to build an analogy between current VR technology, which he dubbed an "environment machine" and his hypothetical experience machine, eventually concluding that the creation of environment machines runs the risk of blurring the lines between reality and illusions, spreading paranoias about the dangers of this technology, and encouraging escapism as a result of users' pursuit of addictive pleasures.

If Nozick is right in holding current VR technology on par with his experience machine, it would follow that: (1) VR cannot produces significant effects that extend beyond the confines of the virtual environment into the physical world, and (2) no genuine value can be acquired from embarking on virtual experiences, at least not beyond stimulating pleasure and escapism. To argue against these claim, I attempt to highlight important disanalogies between Nozick's description of VR as an experience machine and VR's current utilisations. Throughout this chapter, I paradigmatically discuss a few teleological implementations of virtuality to emphasise the real-life effects of virtual interactions. I then argue that these effects entail concrete harms and benefits that surpass the confines of virtual environments.

Hereafter, this chapter will be composed of four sections. The first section, as mentioned above, aims to highlight disanalogies between Nozick's experience machine and current VR technology. In the second section, I explore VR's implementation as a so-called "empathy machine", advertised to function as a supportive extension of interpersonal empathy, specifically through enriching imaginative perspective-taking through evidence-based, representational narratives. I tackle two distinctive elements of VR narratives; namely *visual storytelling* and *interactive narration*. Initially, narration or storytelling (used interchangeably) comes in different forms. The most basic characterisation of which can be made in terms of

"diegesis" and "mimesis" narratives (Bordwell, 1986), respectively referring to delivering stories through either verbally "telling" a story (through speech acts) or indirectly "showing" elements of the story through (e.g. audio-visual) representations. Making use of VR technological affordances, which often allows the stimulating of virtual presence and embodiment as discussed in the previous chapter, producers and storytellers are offered the opportunity to visually construct immersive environments (or 'worlds') which engulf the user's attention in the narrative, generating the sensation of being "inside the story". Virtual objects can also function as explanatory cues for efficient, representational depictions of important narrative elements, such as the setting, the plot, and represented characters. As discussed by Henry Jenkins (2003), visual (also known as spatial or environmental) storytelling offers practical and engaging methods for delivering immersive stories, which compress and convert propositional and representational elements into *lived* experiences. Describing the feature of visual storytelling in VR, AI researchers Ruth Aylett and Sandy Louchard affirm that users' narrative experiences are transformed from storytelling into what is called "story living" in immersive virtual environments. This narrative structure is often reliant on users' interactive participation as an essential element for unfolding a given narrative in VR (2003; 2005).

To a great extent, VR interactive narratives differ from the traditional Aristotelean narrative structure in terms of plot linearity and reassigning the roles of narrators (or producers) and spectators, who become more active and involved participants without which the narrative cannot unfold. In the traditional narrative format, narratives have a beginning, middle, and end, often integrating a climax and conflict around which the story evolves towards a resolution. In VR storytelling, this is not necessarily the case. Due to the interactive freedom that users are granted, the progression of a story depends on the users' movement and decisions in the virtual environment. It is important to note that VR affordances do not only come with possibilities for improvement and innovation but also with challenges, chiefly involving what is known across the literature as "the interaction paradox" (Aylett & Louchart, 2003; Louchart & Aylett, 2005).

As discussed in the previous chapter, interactivity is one of VR's defining features. Despite the advantages interactivity offers, especially in relation to engaging users' attention and enhancing the memorability of the narrative, this feature poses a challenge for storytellers to sustain a balance between users' interactive freedom and the coherent unfolding of a given narrative *as the narrator envisions it*. Scholar Marie-Laure Ryan suggests that, for the time being, "there is no solution to the paradox but only acceptable compromises" (2019, p. 94).

Ryan proposes that it is required to either limit the users' interactive freedom to create a consistent narrative or expand it at the expense of maintaining a coherent narrative form. The second outcome, in the context of non-fictional narratives, threatens to collapse VR non-fictional narratives into a game-like experience, jeopardising the "seriousness" with which VR narratives and their underlying meaningful messages are taken, and hence hampering the development of interpersonal empathy.

After providing an overview of the characteristics and challenges of storytelling (or story living) in VR, the third section will then exhibit some VR non-fictional narratives, designed for enhancing empathy and perspective-taking through representationally portraying aspects of other targets' lived experiences. These paradigms of VR narratives include: *Waves of Grace* (2015), *Clouds over Sidra* (2015), *Defying the Nazi in VR* (2017), *Home After War* (2019), and *Authentically Us – She Flies By Her Own Wings* (Part Two) (2019). As for examples of fictional VR narratives, I explore *Kobold* (2018) and *Awake (2018)*. These VR narratives are created with varying degrees of interaction, attempting to balance propositional, representational, and interactive elements to deliver coherent narratives meant for steering viewers in the direction of developing empathy by conveniently providing them with evidence-based input for easing imaginative resistance and to some extent bridging the gap in experientially understanding others' experiences which would otherwise be concealed or inaccessible. As VR pioneer Jeremy Bailenson puts it, the value of virtual representations is made most apparent through "easing the cognitive [imaginative] burden of starting from scratch" (2019, p. 83).

Finally, the fourth section explores some behavioural consequences of experiencing embodied VR narratives. Following Suzanne Keen's theory of narrative empathy, emotionally-charged, negative experiences, vividly portraying the pain or suffering of other individuals, can produce an undesired effect of "personal distress", steering viewers (as well as readers) away from developing an empathic understanding of others' experiences. In other instances, according to the empirical research done in Stanford Virtual Human Interaction Laboratory, viewing representations of others' emotional plight in virtuality can either prompt viewers to be more conscious of their locus of control, consequently stimulating prosocial behaviour, or steer them further away from taking action. That is to say, virtual experiences' success in depicting heightened phenomenal and epistemically informative representations of others' experiences can have a double-edged effect when it comes to behavioural responses. This shows that behavioural effects triggered by virtual experiences are not only dependent on the

73

quality and content of virtual narratives but also on individual users' characteristics, narrative interpretations, personal ethical appraisals and value systems, perhaps among other criteria. Despite mixed empirical results on the behavioural consequences that VR narratives produce, it is still found that embarking on a virtual experience generates concrete effects that stretch beyond the confines of the virtual environment. Now more than ever, with accelerating technological development, there is a pressing demand for the construction of ethical frameworks that assess and guide important dimensions of VR design, instead of dismissing the technology as a mere escapist and pleasure-inducing equivalent of Nozick's experience machine.

2. "The Experience Machine" and VR Technology: Some Disanalogies

With emerging VR technologies that focus on increasing the level of spatiotemporal immersion, there is a growing worry that virtual experiences are not valuable in the general sense in comparison to experiences in the physical environment. Following Cogburn and Silcox's paper (2014), VR's value is often undermined on an axiological basis. In what follows, I will discuss axiological scepticism against VR, traced back to Nozick's thought experiment against hedonism, where he plausibly argues that experiences fed into the brain by an experience machine are deemed invaluable mainly because they are merely illusory.

Criticising the value of life inside the hypothetical experience machine, Nozick famously states that "we want to do certain things, and not just have the experience of doing them" (1974, p. 42). This is why, according to Nozick, choosing to plug into an experience machine would not be worthwhile. For instance, if someone imaginatively experiences writing a musical piece in the experience machine, this would not be a genuine achievement, instead, it would be a mere illusion and a product of the machine. More recently, Nozick suggested that current VR technology should be considered a type of experience machine as he describes. Contrarily, and as pointed out by Cogburn and Silcox (2014, p. 5) and Chalmers (2022, p. 305), Nozick falsely equated *passive* experiences in his experience machine with *active* VR experiences. In the former case, the user is a mere observer of sense data being fed into their brain, while in the latter case, the user consciously performs intentional, virtual actions that produce effects in the virtual environment in real-time.

The fact that there is a finite number of decisions and interactions that users can engage with in VR highlights another difference between VR technology and Nozick's experience machine. More thoroughly, the latter allows the user to simulate any sensory experience they

desire, where the resulting experiences are illusionary and only bound by the user's imagination and not real-time interactions, unlike current VR technology. Therefore, the experience machine does not foster any embodied relations between the user and the technological tool, instead, it only delivers false beliefs about having undergone a given experience which is not occurring in physical actuality nor virtuality, but merely in a spatiotemporally detached dream-like state. This is not the case for current VR technology, which permits the user to perform actions in the physical world (through VR hardware), further producing instantaneous effects on virtual media, as explained in the previous chapter in the section on VR quasi-causality.

Next, Nozick worries that using VR technology paints a vivid, convincing illusion about "being a certain way and being a certain sort of person" (1974, p. 42). This puts the user's identifying features and self-perception when using VR into question. For example, a coward in real life could be convinced that she is brave by undergoing certain virtual representations depicting her as such. Surprisingly, an empirical experiment has shown that participants who were assigned more attractive or taller avatars in VR tended to interact more confidently and have a stronger tendency to openly negotiate. This is referred to as the "Proteus effect"[9] where "an individual's behaviour conforms to their digital self-representation independent of how others perceive them" (Yee & Bailenson, 2007, p. 217). Nozick emphasises that this effect can distort one's self-perception and image similar to illusory experiences in the experience machine.

In contrast, it has been shown that virtual embodiment, in general, can be harnessed for fulfilling more positive (or at least non-illusory) objectives. Here, it is shown that virtual embodiment often leads to either an experiential "coupling" or "decoupling" between the user's physical body (and phenomenal body image) and their virtual representation in VR. Research in social sciences sometimes seeks to establish that the decoupling effect can facilitate the experiential understanding of depicted characters in VR narratives. The main premise is that if a VR user succeeds in being immersed in an other-oriented, embodied perspective-taking in VR, they are more likely to experientially understand the world from an alternative viewpoint they may not have otherwise considered before undergoing a VR experience. This links back to Coplan's contention that other-oriented perspective-taking is enhanced by sanctioning one's egocentric biases and self-oriented viewpoint. This dimension of virtual embodiment which aims towards a decoupling between the user's physical body image and a given virtual

---

[9] See Domna Banakou et al (2013)

representation is often utilised for simulating not only narrative and empathic understanding but also prosocial and pro-environmental behaviour. However, this endeavour remains an open empirical question that is subject to ongoing interdisciplinary research and scrutiny. As for virtual embodied "coupling", i.e. creating virtual representations with close affinities to the user's real body and identity, it is often harnessed for "know-how" knowledge and skills transfer and training (which I will discuss in the next part of this section). More importantly, as I have shown in the previous chapter, undergoing psychophysiological illusions in VR has not been shown to cause a defect in users' rational judgment. This means that VR users' real self-image remains untouched, for they are not obliviated to it as is the case in Nozick's experience machine.

Subsequently, in the case of active VR, there are situations where actions performed inside VR have genuine concrete value. For instance, if a VR user is playing *Guitar Hero* (2015) using VR HMD and controllers, they are actually producing music which can be assessed as good or bad. Nevertheless, when this reasoning is applied to "virtual actions" that do not include physical activities or produce concrete outcomes, the value of virtual interactions may not be as straightforwardly apparent. For example, if I demonstrate pro-environmental or prosocial behaviour inside a virtual environment like *The Sims* (2000) or *Second Life* (2003), would these acts have real value? According to common sense, the answer would be no. However, VR has been shown to sometimes have a long-standing effect on the users' real-life behaviours and habit formation. An example of this is shown in an experiment, explained in (Ahn et al., 2014) where, after completing a VR simulation designed for raising awareness about saving water, one of the researchers pretends to knock a 2 Oz glass of water (in real life) and asks the participants to help using an already counted number of napkins that was placed on the table. Participants who had a VR experience that seeks to stimulate pro-environmental behaviour were reported to use 20% fewer napkins than the participants in the print condition, where participants were asked to read a text describing the negative consequence of wasting water. More will be said about this in the last section of this chapter. Therefore, it may be initially suggested that some positive (and perhaps also negative) habits acquired or practised in VR may sometimes be genuinely extended to the real world. The same cannot be said for Nozick's experience machine because the thought experiment states that participants who choose to plug into the machine would be choosing to do so for life, completely detached from the physical world. Also, VR does not only produce pleasurable experiences, as I will explain when discussing VR narratives, designed to reinforce perspective-taking.

Another complaint that Nozick raises about VR goes as follows: "a man-made reality, a world no deeper or more important than that which people can construct" is ultimately unsatisfying. He adds that: "There is no actual contact with any deeper reality, though the experience of it can be simulated" (1974, p. 43). This worry is rooted in the belief that VR (especially video games) is man-made and only consists of a limited set of scenarios and possibilities that are pre-determined by hardware and software limitations. Nozick stresses that VR experiences are artificial and excessively embarking on them separates us from the natural world. As Chalmers paraphrases Nozick's worry: "At best, we're in contact with a simulation of the natural world. The simulation is not itself natural; it's artificial" (2022, p. 305). In this critique of VR experiences, there are two issues under scrutiny; namely that VR is pre-programmed and that it is artificial, leading Nozick to contend that it is hence less valuable than physical (natural) reality.

As a response, Chalmers explains that VR scenarios are not *fully* programmed, instead, they are *open-ended* and flexibly determined by users' decisions. "even in a simple video game like *Pac-Man*, the user chooses which direction to go in" (ibid.,). Cogburn and Silcox also state that: "part of the lusory joy of playing such games is beating the algorithm, finding strategies that the designers and play testers did not foresee" (2014, p. 12). This shows that, at least in the case of contemporary video games and some other VR representations, the content is not completely pre-determined by the designers or the computer code.

As for the issue of VR being artificial, Chalmers contends that several non-virtual experiences are also artificial, and still considered valuable. "So artificiality of an environment is no bar to value" (ibid., p. 307). Moreover, some artificial artefacts are often even more valuable. For instance, consider the example of prosthetic limbs discussed in the previous chapter and scientific instrumentalization. These synthetic artefacts are indeed artificial but far from being valueless, they enhance the quality of life and provide means for overcoming difficult obstacles. That is to say, it is hasty to juxtapose the artificial with the natural, favouring one element over the other when it comes to complex questions about value.

As I mentioned in this chapter's overview, Nozick's analogy between standard VR and the hypothetical experience machine, as well as his characterisation of the latter as a mere escapist, paranoia spreading or pleasure-inducing device does not accurately capture the concrete benefits and harm that can occur *through* (and not necessarily due to) this novel technology. In what follows, I explore two cases of VR utilisation that, respectively, reflect concrete instances

of real benefits and harm. Namely, know-how knowledge transfer and the puzzling case of virtual theft. This exploration is intended for providing a broader characterisation of status quo VR devices.

2.1. "Know-how" Knowledge Transfer in VR

VR was first remarkably used as a flight simulator by the US military and NASA to train pilots (Stanovsky, 2008, p. 168). More recently, The British Army has benefited from The Virtual Reality in Land Training (VRLT) created by Bohemia Interactive Simulations (BiSim), a global developer of military simulation and training software. The training program was designed to enhance the Army's military skills in a less compromised environment. Ever since, a large number of training simulators have been made available in other fields such as sports, surgery, and disaster preparation. Some examples of these VR projects are detailed in Jeremy Bailenson's book, *Experience on Demand* (2018), the most emphasised of which is perhaps the use of VR by the Stanford (American) football team. Using a VR simulation of the game (made by the STRIVR lab founded by Bailenson himself) has reportedly enhanced players' performance and helped the placekicker, Ukropina, secure a position as a semi-finalist for the nation's best placekicker award.[10] Considering that VR simulators intended for skill transfer are getting more and more common, it would be worthwhile to wonder how valuable this kind of VR physical training is in this context.

When discussing the value of VR simulations for training purposes, Nozick concedes that VR technology, which he dubs an environment machine, can be considered teleologically valuable in this regard. However, he quickly dismisses this possibility, worrying that without real-life application, these simulations remain illusory and invaluable. In response, it can be suggested that these activities do not need to be "real" or related to reality in order to carry a sort of value. For instance, virtual poker might not include physical cards, yet it can still count as poker. VR pieces intended for entertainment purposes (such as video games) would still carry a hedonistic value despite including fantasised worlds. However, is there another kind of value in VR that surpasses pleasure? Especially when taking into consideration the physical skills that some VR projects promise to enhance.

---

[10] See https://www.strivr.com/resources/customers/stanford/ for more on immersive learning with specific reference to football.

In his paper, "Virtual Reality, Ontology and Value" (2006), Norman Mooradian shows that visually mimicking reality is not enough for ascribing an instrumental value to virtual activities. He closely examines virtual martial arts (more specifically, virtual karate) to how much it compares to real karate training. Mooradian concludes that the execution of techniques in virtual karate is inaccurate in comparison to real karate. He explains that "even though made with similar motions, the strikes and blocks [in VR] would not be done with the proper force and in the proper form. The players might not be moving properly even if it appeared that they were since the precise movements of the body and muscles would be absent" (Mooradian, 2006, p. 682). In addition, Mooradian mentions that players' endurance would not be tested with the same intensity as in real karate. Pain and other possible risks would also be eliminated in virtual karate. He concludes that VR does not sufficiently mimic real-life karate training, undermining the overall value we can ascribe to it when it comes to physical skills transfer, because it omits most of the "essential" properties of the sport. To Mooradian, the value of a VR experience that simulates physical activities will always be grounded in how compliant it is with how these activities are practised in real-life. In most cases, such implementations will only provide a fun 3D computer game, carrying the name of the activity in question (Mooradian, 2006, p. 689).

Building upon Mooradian's useful insights, we can concede that VR projects designed for transferring "know-how" knowledge do have their value grounded in how accurate they are vis-à-vis reality. In this case, fictive as opposed to real content would grant the experience less value, while potentially preserving the hedonistic value of the experience. In other words, this genre of VR projects would need to maintain strong ties with reality in terms of the physical execution of movements, techniques, and instructions rather than metaphysical content. This point sheds light on important points for improvement, especially considering the monetary value of investments directed towards the overall development of this technology for these purposes. Ongoing empirical and technological research is still actively working on resolving physical, haptic and sensory deficiencies of VR (such as those pointed out by Mooradian) so that more valuable training and more reliable know-how knowledge transfer can be provided by VR. Until then, discussions on the value of VR for skill transfer remain inconclusive. However, dramatic future development that would make VR "more realistic" would, as a result, spark ethical questions about inflicting physical pain or harm upon users during a VR simulation for training. VR proponents would be faced with a dilemma; for VR simulations to be made more valuable, they would have to be more compliant with physical reality, which

would need to comply with a certain threshold of physiologically negative experiences. This perhaps draws some reasonable restrictions on the value that we can ascribe to a virtual experience in this context after all.

Mooradian finishes by suggesting that other genres of VR, designed with different contents and intentions may have what he refers to as "intellectual value" that lies in delivering theoretical rather than practical knowledge. In this chapter, I seek to further discuss Mooradian's suggestion by examining a few VR paradigms, intended for fictional and non-fictional (e.g. journalistic) usage of VR technology, which I argue stimulates real emotional effects and epistemically valuable content which supports a more profound understanding of some underpinning features others' emotional states through VR narratives.

2.2. Real Harm in Virtually Real Environments

After exploring the paradigm of know-how knowledge transfer in virtual reality, which entails epistemic benefits, this section will attempt to account for actual inflicted harm occurring in virtuality and escalating into real-world consequences. From the onset, it is clear that there are permissible and even encouraged (virtual) actions the "real" (i.e. non-virtual) equivalent of which is usually considered substantially wrong, harmful, and sometimes punishable by law. The most well-known example of this is *virtual murder*. Several video games are designed in a way that necessitates virtually killing other avatars in order to make progress in the game. Notably, no matter how devastating and nerve-wracking it may be for one to lose a "death match" in games like *Call of Duty* or *Apex Legends*, the death of one's avatar would never be morally nor legally treated as a real murder, simply because no one is *actually* murdered.

Nevertheless, some wrongdoings in virtuality can result in actual harm to the victim both within and beyond the confines of the virtual environment. For instance, an example of harmful behaviour occurring within a virtual environment is often inflicted by "griefers". A griefer is an individual who enjoys spoiling the game experience for other gamers, by looting, harassing or sabotaging others within the game world. Griefers with ill intentions can use the symbolism afforded by virtual objects – admittedly constituted by bits, bytes, and pixels – to communicate and direct offensive content toward other gamers, ruining the enjoyment and experience of the gameplay and causing psychological harm to others. Among other reasons, griefers' actions are considered wrong because they entail a violation of the terms and conditions agreed upon prior to starting the game.

Apart from roleplaying video games in which gamers are *supposed to* misbehave (such as *Grand Theft Auto*) unscripted theft and sabotage, especially in multiplayer virtual environments, are usually frowned upon within gaming communities. Individuals who commit such violations within the game world are often punished by banishment from the game after issued warnings. However, there are crimes committed in virtuality which call for the involvement of higher authorities (such as the legal system). The determining factors for such involvement are worth discussing and clarifying, for they make the difference between a mere roleplay game and a serious crime that is punishable by law.

On 6 September 2007, a 15 and a 14-year-old boy used physical violence to force a 13-year-old to log into his *RuneScape* account and give up virtual objects deemed valuable, including "a magic amulet and an enchanted mask", in addition to game coins that can be sold for monetary value. The victim was forced to "drop" the virtual items in the game for one of the offenders to "pick up" and keep in their account. After settling issues on the value and ownership status of the virtual objects in question, the Court decided that the two offenders were guilty of theft accompanied by violence. However, some believe that it is more accurate to refer to the case as a case of extortion.[11]

This virtual theft, if we may call it so, can be puzzling to virtual irrealists, who contend that virtual objects "do not exist" beyond bits and bytes. In other words, it is argued that no real items were dropped or picked up and the only crime that can be clearly identified is that of physical violence against the victim. Neil McDonnell and Nathan Wildman further explain that "the objects that were stolen are certainly digital objects (the 'bits and bytes' encoded on silicone chips). These were *broadly physical* objects [emphasis added], so there is no puzzle as to their existence" (2020, p. 498). Therefore, McDonnell and Wildman attempt to disambiguate the underlying technicalities behind the theft in question. In other words, they suggest that it is inaccurate to claim that virtual objects were stolen, instead, it is the underlying code (digital objects) that was concretely subject to the theft. The issue of the items' intangibility will be further discussed in this section, where I will seek to clarify that the intangibility of a given "object" does not make such object immune to theft.

Relevant explanatory questions arise on virtual objects being owned, exchanged, and used in a variety of practices that inevitably entail real harm and benefit to the involved parties. For instance, the two thieves have gained items of some *value* and the victim of the theft has been

---

[11] See Lodder (2010)

harmed. In addition to physical harm, the victim has suffered a loss of valuable items, the value of which can be interpreted monetarily and also in consideration of the time and effort dedicated to retrieving the items in the first place. For example, in games such as *World of Warcraft* (*WoW*), the central goal of the game is to collect in-game money called "*WoW* gold", which is required for advancing from level to level in the game. Retrieving *WoW* gold often requires effort and repetitive, tedious in-game activities. Therefore, some gamers opt for paying someone else real money to do it for them. As Susan Brenner explains: "Gold farmers are individuals who play *WoW* and other MMORPGs to earn *WoW* gold or the currency applicable in another similar virtual world; they earn real-world wages by the hour and work for businesses that sell the gold (or other currency) they generate online" (2008, p. 26). Similar to *WoW* gold, the mask and amulet in *RuneScape* are not only (intangible) bytes and pixels but also a testament to the effort and time spent in overcoming several challenges in the game world and they are therefore of personal value to the victim and can, as previously mentioned, also be exchanged against real currency. This supports the Court order and renders the case less puzzling; if the items in question were not economically valuable, the dispute would not have reached the Supreme Court.

Returning to the previous issue, the defence in the *RuneScape* case attempted to argue that the virtual items in question are intangible and therefore the case of theft cannot be established. The Court, then, based its judgment on two precedented case rulings: a case concerning the theft of electricity and another on the theft of electronic, transferrable money. It can be argued that although both of these objects of theft are intangible they remain physical *phenomena* with an undisputable value, and are hence unproblematically owned and protected by law against theft. To the Court, the economic value and ownership status of the items in question mattered a great deal more than their ontological description. This leads us to the next question on the ownership status of virtual objects.

As reported by Arno Lodder (2010), the defence also attempted to argue that the producer of the game (Jagex), who is the lawful owner of all components of the game, did not lose possession of the disputed items. Therefore, no case of theft should be established. However, the counterargument was more compelling. The victim was argued to be the rightful *holder* of the items of which he was robbed and denied access. Hence, just because the theft did not violate the rights of the game owners, the same cannot be said about the holder's rights. The example that was used to support this claim is that of passports, commonly owned by the government and held by citizens. In case of theft, the inconvenient consequences are endured

by the passport holder who loses access to the passport until a new passport is issued. Therefore, the Court decided that the victim is the rightful holder of the mask and amulet and by losing possession of and access to the items, the incident can count as a case of theft, at least from the Court's perspective.

Next, it is important to note that virtual objects can either have status functions within the *fictional*, imaginary boundaries of a given virtual environment or they can surpass such boundaries to be integrated into real-world context through, for instance, being assigned a real monetary value. For example, in the online game *Happy Farm,* users can *own* virtual crops and animals but only within the imaginary confines of the game. Whereas, in *RuneScape* virtual "goods" are exchangeable against real monetary value. The crops in *Happy Farm* and the enchanted mask and amulet in *RuneScape* are all generated by digital objects and presented in the format of images and sounds, i.e. virtual objects. However, what sets the former virtual items apart from the latter is not their material structure, instead, it is the conventionally and institutionally assigned status (and value) imposed on each category of virtual objects. Such (sometimes arbitrary) imposition is primarily done by the producers and programmers of the game, negotiated by users themselves, and, as in the case discussed above, it can sometimes be regulated by real-life institutions such as the judiciary system. Furthermore, as mentioned above, virtual reality as a communicative medium can be used to offend and psychologically harm other users. Approaching the issue from a virtual irrealism point of view, claiming that virtual objects "do not exist" can be misleading and dismissive of the concrete damage that can be done within virtual environments. Instead, despite the fact that virtual entities are not physical, they still exist as bytes and pixels and they, more importantly, constitute a social space and an institutional reality that users collectively inhabit.

All in all, when attempting to explain the reality of virtual entities, complex layers are invoked beyond the material description of such entities. In this section, I have discussed how the intangible nature of virtual objects does not prevent them from playing an integral role in our social lives, highlighting their rich and fluid social ontology. Dismissing virtual objects as non-existent tout court may run the risk of underestimating the degree of psychological and economic harm that can be inflicted within virtual environments. Therefore, in order to seriously address the ethical and legal issues that can arise in virtual environments, more emphasis should be put on the institutional and not only the material reality of virtual entities.

In what follows, I discuss another paradigm of using VR technology, namely as an "empathy machine". More specifically, in highlighting its use to communicate fictional and

non-fictional narratives, we can spot further disanalogies between Nozick's experience machine and current VR implementation. I also suggest that the value of narrative representations in VR can be assessed vis-à-vis the extent to which they extend the imaginative perspective-taking necessary for developing empathy, taking it for granted that empathy is in itself valuable for enhancing interpersonal communication, negotiations and conflict resolution. In other words, understanding others' experiences from their perspective provides epistemically reliable foundations for successful communicative exchange. I argue that these social benefits can be more easily reached by overcoming the *limitations* of imaginative perspective-taking discussed in the first chapter.

3. VR as an "Empathy Machine": From Storytelling to "Story living"

VR used for storytelling and immersive journalism is often designed to serve the purpose of facilitating human connections wherein immersive narratives showing what it is like to be someone else, or portraying otherwise inaccessible or distant events in time or space, are often meant to generate a line of affectivity that bridges the divide in interhuman understanding. Such a divide might be the result of differences in subjective, experiential knowledge or paradigm scenarios[12] based on which we build our emotional responses. Further obstacles in attaining an accurate interpersonal understanding of others individuals may occur due to social and cultural differences or a misguided clinging to false stereotypical identifications of other individuals. In an attempt to overcome these obstacles and divides, VR has witnessed the rise of a new genre, namely VR (fictional and non-fictional) storytelling. This genre makes use of a wide range of multimedia formats and materials to deliver immersive narratives, characterised by interactivity and an audio-visually induced impression of non-mediated "situatedness" in a perceived virtual environment.

In the journalistic context, VR technology has been claimed to provide ground-breaking innovations and affordances that approach delivering what Martha Gellhorn termed "a view from the ground". For instance, commenting on her VR film entitled *Project Syria* (2014), immersive journalism pioneer Nonny de la Peña claims that through VR technological affordances, which often trigger an illusory sense of presence and virtual embodiment in the represented target environment, viewers can experience a first-person simulation of Aleppo's bombing events, through the transmission of original (and spatially immersive) footage taken

---

[12] See de Sousa (1990)

from the sites where the bombings occurred. It might be worth noting that these bombings lasted 55 days before the conflict de-escalated. Hence, de la Peña's VR film and many others only offer a fraction of the experience, which nevertheless communicates a glimpse of the original event. De la Peña enthusiastically affirms that VR is a powerful medium for fostering more authentically expressive news reporting that, to some extent, viscerally places viewers "in the scene" where the narrative unfolds.

VR has been famously dubbed "an empathy machine" that transmits perceptual experiences from one agent to another via a first-person or third-person perspective. Prima facie, the characterisation of VR technology as an "empathy machine" (de la Peña et al., 2010; Milk, 2015) is juxtaposed to Nozick's description of VR as an experience machine. Far from being a tool for escapism and inducing pleasure, VR can also be photographically utilised to enable *indirect perception* of objects and events taking place in the physical world. These events are not necessarily about inducing pleasurable experiences. As shown in the example above about *Project Syria,* VR simulations can be about horrific and traumatising events, especially in the context of immersive journalism.

As Macpherson explains, representational, photographic VR experiences enable indirect perception of real-life objects and events by making it possible to perceive one thing by virtue of seeing another. "You indirectly see Donald Trump in virtue of directly seeing the photograph. You indirectly see the jellyfish migration in virtue of directly seeing the television screen" (Macpherson, 2020, p. 24). This is not to say that veridical indirect perception is the sole function of VR narratives, for VR can be used to create fictional worlds and narratives that do not have a referent in the physical world. Fictional narratives as well, according to Keen's theory of narrative empathy, can enforce an empathic understanding of fictional instead of veridical characters through cultivating imaginative perspective-taking of fully imagined instead of truly existent characters.

Next, non-representational, text-based mediums also attempt to mediate "seeing" certain events through the use of expressive language that "paints a picture" of a given scene or events' sequence in the reader's imagination. However, what differentiates VR from earlier mediums is that perceived objects appear in a 3D format, "fusing" 2D images received from binocular HMD lenses. As discussed by Henry Jenkins (2003), VR affordances facilitate visual (also known as spatial or environmental) storytelling and immersive world-making, practically compressing and converting propositional and representational elements into lived

experiences. AI researchers Ruth Aylett and Sandy Louchard affirm that users' narrative experiences are transformed from storytelling into what is called "story living" in immersive virtual environments. Virtual objects can also function as explanatory cues for economic (compacted) representational depictions of important narrative elements, such as the setting, the plot, and represented characters. As previously mentioned, VR narrative structure essentially necessitates users' interactive participation in unfolding the events of a given narrative in VR.

To a great extent, VR interactive narratives differ from the classical Aristotelean narrative structure which involves chronological plots and clear designation of narrator and audience roles. That is to say, storytelling delivered through earlier mediums assigns a more passive role to the spectators, who receive the story in a non-interactive manner. In VR, however, users are typically granted more interactive freedom and responsibility for deciding which event is viewed first. Considering the early mentioned "interaction paradox" (Aylett & Louchart, 2003; Louchart & Aylett, 2005), VR content creators (especially in the non-fictional genre) are sometimes obliged to limit users' interaction to allow the coherent progression of a story in the sequence they initially envisioned. Another challenge is to not interrupt the immersive experience of virtual presence due to issues of either insufficient or excessive interactivity. One of the techniques employed for mitigating this issue is the subtle integration of "spatialised sounds" to guide users' attention to important points of interest in the narrative, as a substitute for using animated indicators such as arrows, which can interfere with the user's field of view, interrupting the flow and, "photorealism" of the virtual experience. Hence, despite the advantages interactivity offers for maintaining users' engagement, attention, and memorability of the narrative, the challenge remains in preventing VR non-fictional narratives from collapsing into a game-like experience. Ryan suggests that, for the time being, "there is no solution to the paradox but only acceptable compromises" (2019, p. 94). Since Plato, where *The Republic* itself is formed in the shape of a story, narratives are indisputably said to fulfil several functions, ranging from education and entertainment to broadening the boundaries of imaginative and hypothetical thinking, which is often crucial for problem-solving and promoting critical thinking. Although VR is still a fairly new medium, it is predicted that it is suitable for harnessing enacted story living, bringing the long-pursued ideal of mimetic "showing" of narrative aspects, indispensably shaped and organised by contextual and propositional elements.

In continuation of discussing the utilisation of VR as an empathy machine for supporting perspective-taking, the following section outlines some of the paradigms of VR narratives in both fictional and non-fictional genres.

4. Paradigms of Utilising VR as an "Empathy Machine"

Waves of Grace (non-fiction) (2015): A VR that follows Ebola survivor Decontee Davis in exploring the damages that Ebola caused in her town. From disserted schools, hospitals, abandoned buildings and burial grounds, VR viewers are given a first-hand look into Decontee's journey and struggle in the post-epidemic environment. Taking a third-person perspective, the viewer virtually accompanies Davis (in a character-based storytelling structure) to burial ceremonies, providing an experiential understanding, supported by statistical and propositional information about the casualties. The VR also integrates comparative (visual) glimpses about life before and after Ebola.

Clouds over Sidra (non-fiction) (2015) This VR film captures the quotidian life, characterised by resilience and hope in the face of compelling struggles in the Syrian Za'atari Refugee Camp, which is home to 80,000 Syrians fleeing war, half of which are children. The film is shot through a 360-degree customised camera which switches scenes between the desert, a classroom, and the living quarters of individual families in the camp. The user has the option of manipulating the camera, and scrolling left or right, up or down within a flexible field of view. The documentary's main character is a 12-year-old girl named Sidra, who narrates her story throughout the VR documentary. Her speech is then translated to English by a "voice over" which is a commonly used translation technique in news reports. Ryan describes the utility of identifying Sidra in a detailed manner and on a personal level as follows: "by giving an identity to the narrator, even if this identity is not genuine but acted out, the film strengthens its emotional impact, because we are more inclined to be affected by the experience of one particular individual than by the experiences of thousands of anonymous people" (2019, p. 103). This highlights another limitation of (experiential) empathic understanding in VR as well as in non-mediated empathy, which is limited to one individual (or group) at a time. More will be said about this in the fifth chapter of this thesis.

Defying the Nazi in VR (non-fiction) (2017) In this short animated (instead of photographic) VR film, the narrative portrays the experience of child refugees who were fleeing Nazi France in the 1940s towards the United States. Most of the events in the story take place on a virtual representation of "the Excambion", which was the ship carrying the survivors. When using

HMD to view the film, viewers hear veridically recorded testimonies from the actual survivors after they have grown up, describing how their trip went back then, providing vivid descriptions and matching VR animation of the emotionally loaded lived experience. Amélie Diamant-Holmstrom, who was one of the 29 children who arrived in the United States on board the ship represented in the VR piece, was pleasantly surprised and emotionally overwhelmed upon seeing her own experience embodied in VR, hoping that more VR pieces of this sort are made to provide a communicative medium for an experiential understanding of at least some degree and facets of others' lived experiences.

Authentically Us – She Flies By Her Own Wings (Part Two) (non-fiction) (2019) This is one of three parts VR films made as a part of Oculus' "VR for Change" global initiative. Part 2 captures Shannon Scott's fight for justice, representing veteran transgender individuals who were threatened by Donald Trump's 2017 decision of discharging them from military service if they continue to transition, due to "high medical costs and other distractions", in Trump's words. After almost 12 years of military service, Shannon's career in the U.S. Air Force has been negatively affected due to injustice committed against her gender identity. In this VR film, Shannon narrates her struggles and current achievements as an LGBT+ representative after federal judges ruled in her favour. The VR accompanies Shannon inside the courtroom while defending her case before the federal judges, and also shows important parts of her inspirational speech, advocating for social justice and equality.

Home After War (non-fiction) (2019) Returning residents who were previously displaced from the Iraqi city of Fallujah due to war and ISIS domination, guide VR viewers through what is left of their homes, still surrounded by the life-threatening dangers and striking signs of destruction after continuous bombing events. This VR is created using photogrammetry, which consists of taking several pictures and footage (about 4000 photos) to recreate a precise 3D replica of the real location in VR. This VR documentary also attempts to create virtual representations of some post-traumatic symptoms, integrating original audio testimonies for a full cinematic (authentic) experience. This VR experience is somewhat limited to a few locations due to the persisting difficulty of passing through the checkpoints in the Iraqi city, as explained and captured in the VR documentary.

Awake (fiction) (2018) This is an animated VR video game where the player follows the experience of the story's protagonist Harry, who is confined in a wheelchair and appears to be stuck in a lucid dream. Using techniques like frame narration (a story within a story), scene

transition, and audio-visual effects, this VR strives to generate effects that seem like a hallucination. There is also an omniscient narrative voice which gives the players some clues for progressing in the game/narrative. Interaction and control in this VR also seem a bit disorienting, where the viewer is bombarded with sensory cues. Intriguingly, this VR integrates impressive mapping of facial expressions, adding to the photorealism of the experience.

Kobold (Fiction) (2018) In this VR horror video game, viewers step into the shoes of an investigator who is trying to solve a mystery about a little boy's disappearance. The story includes sinister creatures inspired by old, German mythology. What is special about this VR experience is that the real location where the story sets (a villa in Brandenburg, Germany), as well as real-life actors, were meticulously scanned with photogrammetry and advanced facial capture technology. This VR experience has been observed to trigger heightened emotional responses in players, namely reflexive fear, anxiety and anticipation.

The plot, setting, and main characters of these VR films may differ, but what they all have in common is that they are immersive, presenting the viewer with a 360-degree navigable, interactive field of view that is not granted by earlier mediums. This creates a sense of place and proximity to either existing places, or imaginative fictional spaces. Ryan points out an important dilemma that is often present in VR storytelling where a great deal of rich sensory data is compressed into narratives over a short span of time. VR users often have an instinctive urge to explore their interactive agency within virtual spaces. When this is continuously done in random directions, there is a risk of missing out on important elements and key points of interest in the narrative, unless the virtual environment (virtual setting) itself is an integral part of the narrative, as is the case in *Clouds over Sidra*. As mentioned before, the challenge that VR content creators face is allowing users to be interactively free within a virtual environment, while also making sure to coherently guide them to points of interest in the narrative, allowing the latter to successfully unfold as envisioned by a given producer. As Ryan eloquently explains:

> The art of VR narration thus requires a compromise between the user's instinctive tendency to focus on the heart of the narrative action and his desire to exercise his agency by exploring the scene. Or, to put this differently, the art of VR narrative must find the right balance between temporal immersion, which relies on interest in the evolution of the story world, and spatial immersion, which relies on interest in the environment. (2019, p. 102)

Ryan confirms that in optimal conditions, empathy is experienced in VR narratives when a user *identifies* with a given character, especially on an embodied level. This can also be done on social and cultural levels, depending on the context of the VR narrative and the user's imaginative flexibility, as well as several other criteria. As mentioned before, experiencing emotionally and visually engaging scenarios eases imaginative resistance and cognitive effort of picturing a given narrative from scratch. Through VR, users are said to acquire experiential input which enforces their other-oriented perspective-taking, resulting in an evidence-based understanding of what it is like being in the shoes of another character, be it fictional or non-fictional. In the fifth chapter, I critically engage in discussing some important limitations and (normative and technical) lines of criticism of using VR to stimulate empathy. For the time being, I turn to discuss some of the behavioural consequences of being exposed to VR narratives, aimed to enhance empathy.

5. Users' Behavioural Responses to VR Narratives: Between Personal Distress, Egocentric Bias and Prosocial Behaviour

Does embarking on a VR experience that promotes empathy and perspective-taking necessarily lead to prosocial behaviour? Finding responses to this empirical question has been one of the main focuses of researchers at Stanford's Virtual Human Interaction Laboratory. Their line of dynamic research has shown different, sometimes conflicting, results which demonstrate that stimulating prosocial behaviour and what we may consider "sound" ethical judgment may not be as easily accomplished as one may think. From the onset, it might be too idealistic to argue that mere exposure to emotionally engaging VR narratives would produce a radical change in users' behaviour, especially in cases when users purposefully cling to a certain pattern of behaviour out of habit instead of clear ethical justification. Nevertheless, VR narratives have been shown to *sometimes* stimulate positive behaviours, directed to serving an altruistic "greater good", provided that the absence of such behaviour was initially caused by an epistemic deficit in understanding certain facts about the world, or unawareness about one's locus of control.

First introduced by J.B Rotter (1960), the expression "internal locus of control" refers to an agent's belief that their behaviours are causally and directly connected to other events or outcomes, as opposed to believing that things occur due to uncontrollable, external influences that are beyond the individual's control (Ahn et al., 2014, p. 2). Having a good grasp of one's internal locus of control in mind often serves as a strong motivation for committing to

behaviours that lead to desirable, prosocial outcomes. In other words, it is more likely for an individual to act prosocially if they believe their actions have a direct effect on external things and events in the world as opposed to when they do not have these beliefs. It is sometimes argued that individuals fail to grasp the internal locus of control due to an *epistemic gap* in understanding that they have a prominent, causal effect in the chain leading to unfortunate events such as environmental damage and social injustice.

This claim can be challenged by the fact that we are surrounded by information about the environmental and social status quo, highlighting the potential risks of not taking precautions and not having behavioural responses towards environmental or social issues. For example, in the environmental context, a study has shown that there is a plethora of research (about 12000 papers) that attributed 97% of the rising gravity of climate change and global warming to human-related causes. However, only 41% of the American public acknowledges this information (ibid.). Despite this abundance of theoretical knowledge (in text-based, written and spoken forms), individuals remain unalarmed towards future risks.

Some argue that the knowledge (or understanding) that effectively drives prosocial behaviours is of an *experiential* kind. Supporting this intuition, Elke U. Weber (2006)traces the lack of pro-environmental behaviours to a lack of recent personal experiences where global warming may manifest itself. Weber acknowledges that first-hand experiences of environmental damage are not as common as mediated experiences portraying such damage. She adds: "The time-delayed, abstract, and often statistical nature of the risks of global warming does not evoke strong visceral reactions. These results suggest that we should find ways to evoke visceral reactions towards the risk of global warming, *perhaps by simulations* of its concrete future consequences for people's home or other regions they visit or value" (p. 2, emphasis added). This foreshadows the potential that VR may have in guiding prosocial responses by putting agents in a vivid, highly immersive situation where potential risks are, to some extent, brought to life.

To investigate the legitimacy of these claims, two studies measured the long-term behavioural changes after undergoing VR experiences, specifically designed to enhance positive behaviour. What is interesting about these two studies is that they were carried out over extended periods; over one week (Ahn et al., 2014) and eight weeks (Herrera et al., 2018). The first study compared the effects of having a VR experience of cutting a tree (dramatised by enhanced spatial sound effects, intense haptic feedback and textual cues) to the effects of

being presented with written descriptions of the consequences of anti-environmental behaviours; namely how deforestation occurs as a result of failing to recycle paper. The main purpose of the study was to figure out what may lead participants to more effectively grasp an environmental, internal locus of control that would consequently lead to a better understanding of their direct, causal influence on environmental outcomes.

In the VR condition, participants stood facing a large tree holding the handle of a virtual chainsaw. Before being instructed to do any cutting, they were asked to look around the virtual forest they were in with all its details to have a sense of presence and to take notice of the virtually animated living beings and other trees in the forest. Then, using a haptic joystick, participants were instructed to cut the tree they were facing. The movement of the joystick in the physical world was synced to the movement of the virtual arms moving the chainsaw. After two minutes of cutting the virtual tree, participants saw and heard the tree crash down to the ground. After the tree fell, the virtual forest got completely quiet, with no signs of life, to emphasize the damage inflicted upon the forest. In comparison, in the print condition, participants were instructed to create a mental simulation of cutting a tree in the forest on the basis of a written description that depicted in detail what happened in the virtual experience. Both groups were then led to a computer where they filled in a questionnaire. Participants were then guided to participate in another irrelevant 30 minutes experiment so that the effect of the first experiment would wear off.

The most interesting part of this experiment is what happened next. The researcher pretended to have knocked a 2 Oz glass of water and asked the participants to help using an already counted number of napkins that were placed on the table. Participants from the IVR condition were reported to use 20% fewer napkins than the participants in the print condition, which shows that VR had a more powerful effect on bridging the knowledge-to-action gap, hence stimulating higher pro-environmental behaviour. After a week, it has been found that these effects persisted for participants in the VR condition but deteriorated in participants of the other condition (at least according to the methodological standards of measurements employed by the study).

The second study involved watching a VR animation entitled *Becoming Homeless: A Human Experience* (2017) and also compared the impact of virtual perspective-taking to non-mediated perspective-taking using mental simulation alone or relying on written materials. Here, participants were invited to undergo a vivid experience of becoming homeless in VR;

from losing their homes to having to live in their cars, to losing everything and having to protect themselves from potential thieves on the bus. Interestingly, the effect of the virtual experience endured longer than the print condition and perspective-taking by means of mental simulation alone and so did the resulting prosocial behaviour, where participants who took part in the VR condition were more likely to sign a petition for supporting the homeless than participants in other conditions.

However, as mentioned before, studies on the direct effect of technologically-mediated perspective-taking on enhancing prosocial behaviour do not always confirm the relationship between these two aspects. For example, in a study conducted by Groom and colleagues (2009), participants virtually embodied either an African American or white avatar from the perspective of which they completed a mock job interview. Surprisingly, participants who took the perspective of an African American avatar exhibited higher levels of racial prejudice than their white counterparts. This was said to be due to the competitive environment the participants were placed in. Another study also showed that mediated perspective-taking can lead to anti-social behaviours (Vorauer, 2013). Here, after experiencing a VR that detailed some of the manifestations of racial injustice, a few participants felt villainised and cast as an outgroup, which hampered their ability to engage in perspective-taking and prosocial behaviour. Thus, despite the fact that VR experiences transmitted experiential knowledge, this did not lead to increasing prosocial behaviour by definition, instead, it even led to more hostility and prejudice.

Another consequence, emphasised by Coplan and Keen, is where representations of heightened, distressful events, cause individuals to experience personal distress rather than empathy. Surprisingly, these instances can sometimes lead to prosocial behaviour, but, as mentioned before, such prosocial behaviour would only be intended for the elimination of the source of distress, rather than being the direct consequence of experiencing empathic understanding. As concluded by most of the empirical studies referenced in this section, more interdisciplinary research needs to be carried out to unveil the relationship between perspective-taking and prosocial behaviour. Such relation has been shown to involve an interplay among several criteria that do not solely involve the content or quality of VR immersive narratives, but also subjective characteristics that are specific to the VR user.

6. Chapter Conclusion

Returning to the main question that this chapter tackles, it is now hopefully evident that the status quo of VR technology is pragmatically disanalogous to Nozick's procrustean description of VR as an escapist, pleasure-inducing experience machine. This chapter has thus far sought to highlight some of the real-life implementations of VR mediums, entailing concrete (non-illusory) harm and benefit which surpasses the capacities of Nozick's experience machine. This calls for a need to take VR seriously by asking deeper ethical questions about its various, continuously evolving utilisations beyond metaphysical concerns.

This chapter has sought to highlight disanalogies between Nozick's experience machine and real-life implications of using VR technology to mediate and enhance know-how knowledge transfer. I have also attempted to disentangle the puzzling case of virtual theft, arguing that virtual objects can be considered property, the violation of which can result in real-life harm and repercussions. More central to this research, this chapter has chiefly focused on elaborating on novel usages of VR technology as a storytelling medium. I have also explored issues relating to the interactive paradox, shedding light on some of the recurrent challenges that VR affordances create in balancing narrative coherence and the progression of the story in the virtual environment with the user's interactive freedom. After outlining a few paradigms of VR narratives, designed for extending and enhancing users' imaginative perspective-taking capacities, I have discussed a few behavioural consequences that are empirically shown to result from exposure to VR experiences. Here, I have emphasised that embarking on a virtual experience of some kind does not, by definition, lead to radical paradigm shifts or changes in the user's behaviour or ethical judgment. In continuation of the discussion started in this chapter, the next chapter delves deeper into the philosophical problem of the value of VR experiences. As I proposed in this chapter, it might not be coherent to assign an overarching all-or-nothing value to all VR paradigms, considering that these are designed with different objectives and intentions in mind. It is also equally important to clarify which type of value is being investigated.

With that said, the next chapter investigates the status and significance of users' heightened emotional responses to virtual events they do not *believe* to actually occur in their spatiotemporal environment. I attempt to rationalise these responses and their value to research projects on human emotional and behavioural responses through the lens of Kendall Walton's theory of "make-believe". I also draw upon Neil McDonnell and Nathan Wildman's view,

referred to as "virtual fictionalism" (2019) to better account for VR users' engagement with virtual experiences. The next chapter also discusses some of the psychological and ethical concerns, as well as some benefits, of VR's ability to induce emotional responses in VR users.

Chapter 4: The Intricacy of Users' Responses to Virtually Real Experiences

1. Chapter Overview

Consuming certain types of VR content is reported to stimulate heightened bodily sensations and physiological changes, associated with different emotions.[13] The resulting VR-stimulated emotions are often accompanied by actions, performed within the virtual environment. These psychophysiological and behavioural responses are intriguingly elicited without the belief that virtual objects and events exist beyond pixels, sounds, bits and bytes. If a gamer displays signs of "fear" while playing a VR survival horror game, their response and attempts to escape or fight off "monsters" may seem *inconsistent* with their disbelief that collections of sensory inputs are dangerous, or even actually (instead of virtually) existent. The question that follows is how can we fittingly rationalise VR users' emotions and behaviours when engaging with virtual content.

A possible interpretation suggests that users lose their "grip on reality" when using VR, and are *tricked into falsely believing* that virtual events actually transpire in their immediate (physical) environment (Slater, 2009). The resulting false beliefs, then, push the user to either "run", "hide", or "attack" to avoid or eliminate the threat. From the onset, I suggest that this interpretation is flawed. For, it involves a mischaracterisation of VR users' commonly reported phenomenological experiences, which do not involve trickery and deceit. In addition, technological limitations make it easy to differentiate between virtual and non-virtual environments. It is, therefore, crucial to highlight that VR users are perfectly aware of the causal limitations of interacting with virtual media, at all times. With that said, the question still stands; How can we rationalise VR users' responses in consistency with their disbelief in the full-blown "reality" of virtual experiences (without resorting to the argument from illusion)?

As an alternative to the interpretation mentioned above, this chapter will explore "virtual fictionalism" (VF), a view defended by Neil McDonnell and Nathan Wildman (2019). VF explains users' engagement with virtual content in terms of participation in a game of make-believe, whereby VR users are prescribed to *imagine* that virtual objects and events *truly* exist and transpire *in a fictional world.* VF is based on Kendall Walton's theory of make-believe,

---

[13] For example, see Wang et al (2019) for a study on heightened emotional responses to an immersive virtual simulation of height exposure.

designed to account for audience engagement with representational artworks. Walton's theory suggests that emotions that arise as a result of engaging with imagined scenarios and states of affairs are called "fictional emotions", empirically identified as natural and reflexive responses that are instrumental in problem-solving and decision-making (Van Leeuwen, 2016). In *Mimesis as Make-Believe* (1993), Walton claims that fictional emotions are mainly constituted by *quasi-emotions,* defined as phenomenological sensations that typically characterise ordinary emotions (Walton, 1993, p. 251).

Through extrapolating Walton's description of fictional emotions and accepting McDonnell and Wildman's claim that proper engagement with virtual content requires make-believe, we can establish that VR experiences can trigger quasi-emotions when users believe that objects, events, and states of affairs depicted in VR are the case in a fictional (or non-fictional) world of a given narrative. This interpretation is consistent with the assertion that VR users do not believe virtual objects and events actually exist in their immediate, spatiotemporal environment. Nevertheless, as Walton acknowledges, it is sometimes difficult to identify quasi-emotions, because not all emotions have clear and well-defined phenomenological constituents. Apart from this issue, this chapter suggests that we can effectively use the notion of make-believe to explain what motivates VR users to emotionally and behaviourally engage with virtual content.

After further discussing the nature of VR-stimulated emotions and behaviours vis-à-vis make-believe, I argue that we have no strong reasons to label these responses as unfitting or irrational unless they violate the norms and standards of using VR technology. A fitting response can be understood as one that is suitable and appropriate for the situation in which it is elicited and it also should be backed by sound reasons. For instance, experiencing (quasi)fear mixed with amusement is an expected and fitting response to make-believing that a (fictional) monster chases us in a VR video game. Making use of the interactive features in the game is also a fitting and expected response. All these responses are required for the full appreciation of a VR experience. Without them, the user's experience would be severely impoverished. In contrast, high levels of terror, leading one to use a real-life weapon to eliminate zombies in the VR video game would be an unfitting and irrational response, for it violates the norms of engaging with a VR experience, *as a VR experience*. That is to say, one of the norms of fittingly using VR is to act and respond within the causal boundaries of the virtual content. Failing to do so can lead to producing unfitting responses, perhaps caused by "forgetting" that VR content is not real, but only virtually so. After discussing the (un)fittingness of VR-stimulated

responses in fuller terms, this chapter tackles some of the ethical concerns and potential benefits of experiencing heightened responses in virtuality.

2. Emotional and Behavioural Responses in Virtual Reality

In this thesis, emotions are understood as multifaceted, complex phenomena that typically involve voluntary and "automatic" bodily sensations and expressions. They are usually expected to be accompanied by "cognitive" elements (such as judgments and beliefs) which enable individuals to recognise things in the world as characterised by certain features that *merit* or call for eliciting a given emotion. For instance, if someone claims to be afraid, it is reasonably expected for their fear to be accompanied by a judgment or belief that something is dangerous. In the context of his influential theory of fiction, Walton also proposed that emotions should essentially involve *a disposition to action.* He writes, "Fear emasculated by subtracting its distinctive motivational force is not fear at all" (Walton, 1993, p. 202). A typical (but not the only) *paradigm* of emotional experiences, hence, involves physiological, evaluative, and motivational elements. [14]

With this simplified conceptualisation of emotions in mind, let us take a look at an online review written by a gamer, let us call Charlene, about *Resident Evil 7: Biohazard* (2017), a VR horror game:

> …Using my actual body, I lean forward on my couch, craning my neck around the in-game corner to my left, trying to see if Jack, the hulking mass of seemingly invincible mutated flesh, has passed by yet. Just before I take my step around the corner, much to my horror, Jack grabs me from behind and for a moment it's as if I can feel his breath on my face. He throws me to the ground, cursing at me, and raises his shovel high above his head. With a loud thud, he slams it down into my shin, slicing my leg in half. For a split second, there's the illusion of pain — a sudden twitch of realism — before Ethan [the game's main character] bleeds out and dies not just before my eyes, but within myself. (Jagneaux, 2017)

In addition to using VR content for entertainment, a growing number of researchers in psychology use VR simulations to study human emotions more closely. For instance, in a recent study, Kirill A. Fadeev, Alexey S. Smirnov and their colleagues used four VR simulations to expose participants to scenarios that involve stressful events. Subsequently, participants' activity, their autonomic nervous systems, and their avoidance responses were closely monitored. This study concluded that "the use of stressful VR content can cause high

---

[14] See Lazarus (1991) for an overview on the definitional problem of emotions.

emotional stress to a user and restrictions should be considered" (Fadeev et al., 2020). Another study (Boccignone et al., 2022) was specifically focused on assessing participants' fear of heights through a virtual simulation. The study findings were then considered reliable for discerning prospective workers' suitability for positions requiring working at high altitudes. In his paper entitled "Virtually Real Emotions and the Paradox of Fiction" (2010), Garry Young compiles an insightful survey of earlier psychological studies that rely on virtual simulations for investigating phenomena such as arachnophobe's fear response to virtual representations of spiders. These pieces of evidence attest to VR's ability to stimulate heightened emotional and behavioural responses.

As previously mentioned, VR-stimulated emotional and behavioural responses can appear inconsistent with the beliefs that users hold about the causal limitations of the virtual environment. Charlene, as well as the participants in VR psychological experiments, are all well aware that characters in horror games and virtual representations of all sorts are nothing more than animated images and sounds (bits, bytes and pixels). Nevertheless, VR users act *as if* the virtual stimuli they perceive actually exist. In his influential paper "How Can We Be Moved by the Fate of Anna Karenina?" (1975), Colin Radford formulates a thought experiment to better highlight the explanatory problem at hand. Suppose that someone told you a "harrowing story" about their sister, causing you to feel a range of emotions such as pity, sadness or anger, but then revealed that they have no sister and that the story is all made up. Consequently, according to Radford, it would be incoherent for one's feelings to persist after learning that the story is only fabricated. Nevertheless, these feelings can be replaced by other feelings (such as anger or amusement upon knowing the events are not real). Another recurrent example goes: if a hiker is terrified of a bear in the forest at night, her fear is grounded and justified by the belief that her life is in danger. If the hiker discovers that she mistook a steel statue for a real bear, it would be inconsistent for her fear to persist.

Radford draws an analogy between emotional responses produced in these cases and our (emotional) engagement with fiction. Normatively, is it more fitting for our emotional responses to die out upon knowing that the content of our experience is a "mere" fiction? Radford affirms and concludes that while our responses to fiction and imagined states of affairs are natural, they are still incoherent and inconsistent with our disbelief in their real-time, actual existence (1975, p. 78). Just like the imaginary sister in the harrowing story, events and objects in VR do not *actually* exist, and according to Radford's reasoning, only actually and concretely existing stimuli can *coherently* stimulate emotional responses. This formulation of the problem

(concluding emotional and behavioural inconsistency from disbelief in the spatiotemporal existence of a stimulus) implies that a great deal of our (usually unproblematic) emotional responses are incoherent. This includes emotions produced while reliving past events, speculating the occurrence of possible events, and entertaining improbable or impossible scenarios in fantasy. This means that the issue at hand is not restricted to emotions triggered by fictional and virtual content but also touches on some of the emotions we experience in everyday life. Radford's conclusion also poses a threat to the validity and generalisability of psychological experiments implementing VR technology. As Gary Young (2010) points out, if Radford is correct, then such experiments only succeed in capturing inconsistent participants' responses.

In response, Peter Langland-Hassan calls the thought experiments introduced by Radford the "pull rug" cases. Langland-Hassan argues that these cases are disanalogous to our engagement with fiction. Radford claims that responding emotionally to fiction is as incoherent as continuing to respond emotionally to events we believed to be true but discovered to be otherwise. But it can be suggested that, unlike the pull rugs cases, we are aware of engaging with fiction at all times, and we willingly engage with it accordingly, *as* fiction (2020, p. 242). That is to say, we respond a certain way and not another due to knowing that the content of our experience is fictional. Equally, when using VR, we are perfectly aware that the content of our experience is only multisensory media, depicting either fictional or non-fictional events. This awareness guides us to feel and act a certain way and not another. For instance, because I do not mistake the content of my experience for real events while using VR, I do not attempt to barricade my house when I see that zombies are approaching me in a VR video game. Instead, I use the interactive features of the video game to avoid or fight off zombies only *virtually and fictionally*. However, highlighting the disanalogy between the pull rug cases and our guided, conscious engagement with virtual content does not make VR-stimulated emotions and behaviours any less intriguing; users still respond emotionally and behaviourally to VR content they don't believe to actually exist.

To better understand the nature of these responses, I propose that it is paramount to examine users' physiological reactions and biologically "wired" reflexes when interacting with virtual content. More thoroughly, phenomena such as the "rubber hand illusion" can provide a good background for discussing the topic at hand. The rubber hand experiment is conducted by stroking a participant's arm with a feather (or a similar object) simultaneously with a rubber arm, in a synchronised motion, while keeping the participant's real arm covered. The more

time elapses, the more participants develop a sense of agency over the rubber arm. Although they know for a fact that the rubber arm is not a part of their bodies, participants display heightened emotional responses when a researcher suddenly strikes the rubber arm with a hammer. By responding "realistically", participants' reflexive feedback contradicts their awareness of the disconnectedness and artificial nature of the rubber arm. In VR, something similar transpires. VR users experience a sense of (virtual) embodiment, which leads them to reflexively react to stimuli as if they were truly present inside a given virtual environment *as their avatar.*

So far, it is clear that certain emotions and actions *can* (reflexively) arise without belief in the existence of their corresponding states of affairs. Relevantly, Richard Moran (1994) explains that there are certain paradigms of emotions, such as nostalgia, regret, and relief, that can only be about states of affairs that no longer exist. These responses can be considered unproblematic. But can the same be said about responses to virtual content (and fiction)? I suggest that if we expand our paradigmatic repertoire and understanding of what constitutes a coherent emotional response, we can unproblematically consider "emotions triggered by stimuli that do not spatiotemporally exist" as a salient category of emotional responses, particularly characterised by not necessitating an existence belief. However, the issue remains that disbelief (in the actual existence of the virtual stimuli) cannot, on its own, explain how VR users are *motivated* to respond to virtual stimuli in one way and not the other. As a solution, McDonnell and Wildman propose to explain VR users' (proper) engagement with virtual content in terms of participation in a game of *make-believe*. In the next section, I attempt to show that this interpretation does not entail any incoherence between users' beliefs, disbeliefs and resulting responses. However, it raises new questions on the nature and fittingness of emotions and actions produced as a result of make-believe in VR.

3. The Phenomenology of Virtual Experiences From the Lens of Virtual Fictionalism

In their paper (2019), McDonnell and Wildman apply Walton's theory of make-believe to address issues on the ontological status of virtual objects and events. Using Walton's terminology, they suggest that the sounds, images, and other forms of multisensory media in virtual environments are *props,* which make playing games of make-believe possible. They affirm that props genuinely exist in reality*,* while virtual objects (properties assigned to props within games of make-believe) only exist fictionally. Props are said to be organised by "principles of generation", i.e. rules that determine how a game of make-believe ought to be

played. These rules also dictate the characteristics of the props and determine what is (to be imagined as) true in the (fictional or non-fictional) virtual environment. For example, a "zombie" is a virtual object, constituted by (real) images and sounds alongside (fictional) indicators that prescribe gamers to imagine that such images and sounds truly possess "zombie-like" properties (such as running fast and launching acid attacks). Virtual objects can, therefore, only *fully* exist within the world of users' make-believe and imagination. Walton uses these two terms interchangeably. Notably, McDonnell and Wildman do not discuss VR users' emotions and behaviours from the lens of Walton's theory of make-believe. I will attempt to do so in what follows.

In his famous thought experiment, Walton introduces Charles, a moviegoer watching a horror movie displaying "terrifying", vivid scenes of a green slime. "The slime, picking up speed, oozes on a new course straight toward the viewers. Charles emits a shriek and clutches desperately at his chair. Afterwards, still shaken, he confesses that he was 'terrified' of the slime" (Walton, 1993, p. 196). Despite Charles' confession, Walton insists that the latter cannot be afraid of the slime. He argues that although Charles experiences a genuine emotional state, it cannot count as one of fear. Walton explains that if Charles was truly afraid, he would be disposed to act in a way that reflects real fear of the slime. Contrarily, Charles did not, nor was he even inclined, to flee the cinema or call the police on a dangerous slime being on the loose. Similarly, Charlene's emotional response would not be considered one of true fear either. If she was truly afraid, she would, at least, abstain from playing the game. But for most VR users, such abstention does not usually occur.

Walton describes Charles' (and hence Charlene's) reactions as "quasi-fear", defined as the psychophysiological symptoms, such as muscles tensing, accelerating heart pace, and adrenaline flow, which typically characterise fear elicited in ordinary circumstances. That is to say, quasi-emotions refer to the phenomenological and physiological components that characterise *both* fictional and ordinary emotions. Walton's distinction between fictional and ordinary emotions resonates with Jean-Paul Sartre's differentiation between "genuine feelings" and "imaginary feelings" ("sentiments vrais" and "sentiments imaginaires") (2004, p. 145). Both Walton and Sartre affirm that fictional/imaginary emotions are *genuine* emotional states, differentiated from ordinary emotions by their directedness towards imagined states of affairs and their stimulation of different behavioural outcomes. We can plausibly assume that the quasi-emotions that arise when engaging with virtual content are a result of forming certain beliefs and judgments about events and states of affairs depicted in VR. Users, hence, evaluate

the multisensory media they perceive in virtual environments and may sometimes find a given depiction worthy of eliciting a given emotional response. This plays an important role in how VR users come to fully grasp and (aesthetically and intellectually) appreciate and engage with a virtual experience.

VR users' cognitive and affective evaluations are performed within the frame and boundaries of make-believe and imagination. Charles and Charlene and other consumers of representational (art)works, therefore, imagine that the states of affairs they perceive (through a given multisensory medium) actually occur in a fictional (or non-fictional but still spatiotemporally distant) world. In response, they naturally and even involuntarily produce quasi-emotions. This interpretation coheres with the fact that VR users have disbelief in the spatiotemporal existence of virtual objects and events and it also explains what can motivate VR users to react one way and not another when using VR. At this point, we can wonder whether emotional and behavioural responses prompted by make-believe are coherent or not. Following a flexible theory of emotions, I propose that responses stimulated by imagination can be coherent *if* they do not violate the norms of engaging with imagined (virtual) states of affairs. That is to say, responses that arise as a result of entertaining the occurrence of certain scenarios in imagination are subject to certain norms of fittingness. More will be said about this in the next section. For the time being, it does not seem sufficient to deem these responses as unfitting or incoherent (in Radford's terms) only by virtue of their being directed to or stimulated by imagination.

Problematically, however, several connotations and functions have been attached to the term "imagination" throughout history. For instance, Leslie Stevenson counts twelve different (but overlapping) conceptualisations of imagination (2002). Pursuing similar goals, Annis Flew (1953) presented perhaps a more concise alternative by distinguishing three possible conceptualisations we can attribute to "imagination". Both Stevenson and Flew identify one of the meanings of imagination as "the ability to entertain mental images" (Flew, 1953, p. 246; Stevenson, 2003, p. 243). This connotation may not be useful for our present purposes. For there might not be any obvious reasons why VR users would need to build mental images on top of the virtual multisensory content they perceive in their virtual environment. The challenge, therefore, is to select a conceptualisation of imagination that productively highlights and justifies emotional and behavioural responses to virtual stimuli. Going back to Walton's theory of make-believe, he suggests that imagination entails a sort of pretence that certain things "count as" other things within the confines of a fictional world (1993, p. 37). This

understanding of imagination need not involve the construction of mental imagery, and it aligns with Flew's second understanding of imagination as "supposing" something to be the case or "thinking of what would happen" if something is the case. He calls this type of imagining "propositional entertainment" (1953, p. 247).

To further disambiguate the concepts of imagination and make-believe as utilised in this context, Walton provides the example of children playing a game of make-believe, where they pretend or allow stumps in the forest to count as bears. Every encounter with a stump qualifies as a – fictional – encounter with a bear. By engaging in this game, the children create a fictional world that is constituted not only by fictional bears but also by their in-game behaviours and emotional responses. Without make-believe, the stumps cannot be (fictionally) ascribed the property of bears and hence cannot trigger children's responses of quasi-fear and quasi-excitement. Similarly, if a VR user is fixated on the fact that virtual objects are merely constituted by pixels, sounds, and programming codes, they cannot engage with the virtual content as it is intended by its creators. As McDonnell and Wildman assert, proper engagement with virtual experiences necessitates employing imagination and make-believe. This kind of engagement often triggers users' quasi-emotional and behavioural responses.

With that said, there might be some difficulty in identifying the constituent(s) of quasi-emotions. Walton acknowledges this minor issue and writes: "It is not hard to specify sensations characteristic of fear (intense fear, anyway). But other quasi-emotions are more elusive. Quasi-admiration may seem especially hard to put one's finger on. What does it feel like to admire someone?" (1993, p. 251). While emotions can sometimes be *expressed* in a certain way and can entail distinct physiological responses, this is not always the case. For the time being, this need not be an issue considering that research is often steered toward focusing on "basic emotions" (Ekman, 2005) that are associated with clearly defined facial expressions and distinct physiological symptoms. More elusive or "complex" emotions can always be better clarified through introspection and verbal reports.

To summarise, in this section, I have attempted to show that Walton's theory of make-believe is suitable for explaining VR users' emotional responses and behaviours in reaction to content they do not believe (but only imagine) to occur within the confines of the virtual environment. When a VR experience is built properly, it immerses and transports the users to another world of a given fictional or non-fictional narrative (Green, 2021). During engagement and interaction with virtual content, users form beliefs and judgments about the content of their

virtual experience and naturally and reflexively experience quasi-emotions, which can also be accompanied by certain behavioural responses. Following Flew and Walton, I have defined imagination and make-believe as the ability to assume or accept that certain things count as other things in an imaginary or spatiotemporally distant world. This conceptualisation need not (but can) involve forming mental images.

Concerning Radford's worry, we do not have sufficient evidence to consider responses stimulated by virtual content, fiction, and probably other categories of stimuli we encounter in everyday life, as incoherent *only* by virtue of being direct and stimulated by states of affairs that do not spatiotemporally exist, yet still exist in some capacity (fictionally, virtually, mentally…etc). The coherence of these responses is anchored by the fact that users do not mistake the content of their experience for actual events and objects. There is also no inconsistency between users' existence disbeliefs and their make-believe that certain things truly occur in imagination or virtuality. It might be worth mentioning that VR users' responses are instrumental in the pursuit of enjoyment and intellectual or aesthetic appreciation of a given VR experience. Such responses remain coherent and fitting results of make-believe unless they violate certain norms. In the next section, I discuss the norms of fittingness of VR emotional and behavioural responses more thoroughly. The main contention is that if these norms are violated, engaging with VR can result in incoherent, unfitting and even irrational emotional and behavioural responses.

4. A Fittingness Analysis of VR-Stimulated Responses

The fittingness of emotional and behavioural responses can be determined based on different grounds. A given response can be *instrumentally* fitting if it leads to beneficial outcomes, or *morally* fitting if such a response is compliant with standards of virtue and ethical values. Commonly, fittingness is also ascribed to a given response when such a response correctly coheres with evaluative features of a given situation or state of affairs in the world. For instance, joy is a fitting response to hearing great news and anger is a fitting response to unjustified aggression and so on. This last connotation of fittingness presupposes that all emotions are directed towards particular objects. For a given response to be fitting, it must therefore reflect that the individual correctly recognises certain features in an object of emotion.

In this section, I explain a fitting emotional or behavioural response as one that is supported by appropriate and sound *reasons.* For instance, being angry at a messenger for delivering bad news is an unfitting emotional response (D'Arms, 2022, p. 1) not because this response is

harmful or unethical per se, but because no good reasons can be provided to justify it. It is safe to assume that the reasons behind emotionally and behaviourally engaging with virtual experiences are various. In what follows, I also discuss the role that *users' interest* plays in rationalising their heightened responses to VR experiences (and perhaps other states of affairs in the world). Then, I explore how the *context (or genre)* of a given VR experience acts as a factor for evaluating the fittingness of VR-stimulated emotions and behaviours.

4.1. Evaluating the reasons behind VR-stimulated responses

Previously in this chapter, I argued that proper engagement with VR content requires making believe that objects and events in virtual environments are truly the case in a world of a narrative. Taking it for granted that emotional and behavioural responses are a vital aspect of users' "proper" engagement, one can wonder if (propositional) make-believe is a strong enough reason for causing users to be emotionally and behaviourally moved. Surely, we can imagine a given scenario to be occurring or even witness its depiction in VR without being inclined to feel or do anything in response. To expand this idea, it is reasonable to affirm that not every event or state of affairs in virtuality (or imagination) *calls for* an emotional response. For instance, it is common for video games and other types of virtual content to contain interactive login and saving screens that can be labelled as "neutral" content. This type of content is not *intended* to stimulate any special emotions or actions. Hence, dispassionate interaction with mundane depictions in VR is not problematic.

However, there are other instances where emotional or behavioural unresponsiveness is a result of quality or content issues in what we imagine to be true in a virtual environment. Peter Langland-Hassan provides a parallel example of a similar case involving fictional content. He writes: "We can easily imagine a terribly written and poorly acted tragedy—something concocted by a group of fifth graders over the course of a few hours. We ought not to be moved by it, even if it contains fictional truths to which a congruent response would be considerable sadness" (2020, p. 253). Boredom and even disorientation would be the equivalent responses to VR experiences that are not *well-crafted*. Therefore, being prescribed to make-believe is necessary but often not sufficient for stimulating VR users' heightened responses. Users' cognitive and affective evaluations of objects, events, and other states of affairs in virtual environments are performed with consideration to what truly occurs in these environments and how it is (technologically and aesthetically) presented. Supporting this contention, Paisley Livingston and Alfred Mele (1997) state that only if "the aspect of a work

to which one responds is of at least moderate quality," is one "justified in responding emotionally" (p. 173).

This is not to say that low-quality virtual content cannot stimulate any intense emotional or behavioural responses. For example, games like *Tank* (1990), *Pac-man* (1982), and even *Minecraft* (2001) do not depend on high-quality graphics or soundtracks. And yet, these games still attract a large audience that actively engages with such content. Are users' responses to such (low-quality) content unjustified or unfitting? I believe that this is not necessarily the case. When interacting with a relatively low-quality game or virtual experience, users' interest can still be captured by the challenges depicted in these games, or the ease of use and interaction with in-game objects. Or perhaps playing these games is associated with a memory of personal value to the user or acquiring epistemically valuable input, despite the low quality of the game. In any case, various reasons can enable users' engagement with virtual content. Langland-Hassan explains that "once our interest in a fiction (well-crafted or not) [or a VR content] is made reasonable, our emotional reactions will come along as reasonable for free" (2020, p. 257).

Therefore, our *interest in* certain aspects in a make-believe, virtual environment gives us reasons to emotionally and behaviourally engage with VR content. We may take an interest in VR experiences for various reasons, including experiential vividness, media richness, compelling narratives, and reliable hardware and software performance and quality (Gall et al., 2021; Kisker et al., 2021; Otondo et al., 2008). More importantly, we respond a certain way and not another as a result of making cognitive and affective evaluations of what is depicted in a virtual environment. As long as our responses are coherently aligned with sound reasons and evaluations, such responses remain (aesthetically) fitting. However, it is crucial to note that our interest in a VR experience (for whatever reasons) cannot justify *misplaced* emotional and behavioural responses, for instance, elicited with an exaggerated degree of intensity. Such responses remain unfitting and not properly justified. Next, I discuss context relevance as another helpful factor for assessing the fittingness of given (VR-stimulated) emotional and behavioural responses.

4.2. Context relevance to the fittingness of VR responses

It is reasonable to assume that responses towards virtual content are elicited alongside the awareness that virtual objects and events do not spatiotemporally exist in one's environment. This fact can warrant certain emotional and behavioural responses in VR that would be otherwise considered inappropriate if produced in real-life, everyday life. For instance, amusement upon witnessing another team's avatar's "death" in a video game is fitting in this context but unfitting if elicited in response to witnessing a real-life death or devastating events (either in VR or in unmediated reality). Such a response can even be considered immoral. There is, hence, an expected asymmetry between our real-life and VR (make-believe) responses. Since each category of these responses is prompted with different considerations in mind, it can be suggested that reality-directed, "ordinary" emotions ought to be subject to different norms and measures of fittingness from those governing emotions directed toward virtual states of affairs. Some may disagree with this intuition and counterargue that responses to VR (fictional and non-fictional) content are nothing but an extension of our real-life, everyday responses. And hence, they should be assessed against the same norms of fittingness and appropriateness. As Jonathan Gilmore explains in a relevant discussion (2011), there are arguments in favour of both positions.

Regardless of whether or not we decide to apply the norms of fittingness usually used to assess everyday responses to analyse VR-stimulated responses, deciding what is a fitting response for a certain situation (real or virtual) depends on moral, social, cultural and perhaps even aesthetic norms and considerations. For instance, when exposed to realistic (interactive) VR content such as *Richie's Plank Experience* (2016)), users tend to react as if virtual events are truly transpiring. Is (quasi)fear in this case an appropriate response to being on a skyscraper, walking on a wooden plank *in VR?* It can be argued that it is indeed a fitting response, resulting from evaluating a situation as being make-believedly dangerous or fearsome in the VR context. It is important to note that the degree of intensity (and perhaps even the *kind*) of responses elicited in reaction to playing *Richie's Plank Experience* noticeably differ from the responses that can be produced if the user was truly experiencing the same situation in real life. So far, quasi-emotions, devoid of the belief that virtual events spatiotemporally transpire, remain an appropriate response in the context of VR content.

However, what if a user experiences pleasure while engaging with (socially, culturally and morally) inappropriate content in VR? For example, *Custer's Revenge* (1982) is an adult

action game, published by American Multiple Industries for the Atari 2600, where players are required to sexually assault a Native American Woman in order to advance in the game. With cases like these, we are pushed to resort to the more robust norms of fittingness that we usually utilise to assess the appropriateness of responses in everyday life. Pleasure and amusement experienced in response to immoral representations in VR would, therefore, be considered inappropriate even if the content is not spatiotemporally actualised. Eventually, it might be more productive to concede that VR-stimulated responses are often *congruent* with those elicited in actuality. If this is the case, we can unproblematically apply the same rules of fittingness to analyse the fittingness of VR-stimulated responses. As Livingston and Mele explain: "If anger is the appropriate response to a certain kind of unjustified aggression in actuality, anger (though not necessarily of the same intensity) is also the congruent response to such events in fiction [and imagined domains]" (1997, p. 171). However, as a medium for artistic expression, VR can put users in virtual situations where it is artistically fitting to engage with virtual events in a way that can be considered inappropriate in real life, such as enjoying – virtually – stealing cars and destroying properties in the *Grand Theft Auto* franchise. The spectrum of (contextual) fittingness of emotions and behaviour, hence, remains to be decided by the everchanging social and cultural norms. It might be best to investigate the fittingness of VR responses on case to case basis, for VR experiences are too varied, in context, genre, content, and many other features, for us to apply a unified notion of appropriateness. Also, as we have seen, there are various ways in which a response can be fitting or unfitting. Considerations of the context of the experience and the reasons causing a VR user to respond a certain way and not another are nevertheless always helpful in determining the fittingness of a given response, understood vis-à-vis the fittingness of reasons and appropriateness of a response in a certain situation.

5. Ethical Concerns and Benefits of VR-Stimulated Emotional States

In *The Republic*, Plato criticises poets and painters of his age for making inferior imitations; "a copy of a copy", of truth (Harris, 1929). To Plato, when artists do not faithfully present reality, they produce artworks that are corruptive of the soul, leading viewers to emotionally respond to characters' overdramatised (deceptive) plight, suffering, and other emotional experiences. Some of Plato's concerns are echoed by philosophers and researchers who worry that VR's ability to trigger heightened emotional responses runs the risk of manipulating and distorting VR users' rational judgment and decision-making mechanisms. Therefore, the structure of highly immersive experiences in VR, coupled with certain types of content that are

classified as inappropriate or misleading, led some to describe some VR paradigms as "unethical". For instance, Erick Jose Ramirez and Scott LaBarge worry that virtual (fictional) experiences that stimulate "excitative" emotions and events can negatively influence users' real-life behaviours. They write: "Many examples of ethically problematic VR content come quickly to mind: we might be concerned about violent or pornographic content, or about harmful stereotypes of various groups that might be encoded in a simulation" (2018, p. 4).

The main issue that Ramirez and LaBarge underline is that these depictions normalise unethical behaviours in the virtual environment and can be taken as encouraging users to commit said behaviours in real life. However, this line of reasoning merely rests on several correlations between how emotions affect judgment and decision mechanisms and assumptions about VR's power to control and manipulate users into mimicking depicted events without any prior rational reflection. Elsewhere, Ramirez and his colleagues (2021) also worry that the content and immersiveness of some virtual experiences pose a threat to users' autonomy if they are not accompanied by full transparency and accuracy of representations that should comply with reality. They go on to assume that some virtual experiences are deceptive, in that they lean toward "nudging" users towards unconsciously making certain decisions or forming certain beliefs and judgments.

Presupposing there is a "likelihood of deception" in being exposed to virtual experiences, Ramirez and his colleagues write: "…Another possibility is that designers are producing VR simulations without understanding how they are misleading their users. On this view, designers are not necessarily engaged in manipulative trickery, but they are nonetheless guilty of a type of culpable wrongdoing" (ibid., p. 18). Even when giving VR content creators the benefit of the doubt by suggesting they do not intend to misguide VR users into forming false beliefs or committing unethical or irrational behaviours, such outcomes may still occur. However, one wonders, is there enough evidence to support a causal correlation between heightened emotions and fictional scenarios elicited in virtuality and the production of such undesirable outcomes? Moreover, can VR content creators fully foresee and be held accountable for the alleged effects of virtual experiences? Initially, we may be able to reject both of these suppositions. But this does not mean that these issues are not worthy of deeper scrutiny.

Studies in psychology (e.g. Lerner et al., 2015) indeed confirm that emotions have a great impact on our decision-making mechanisms, this impact can either be positive or negative, based on an interconnected web of conditions. Furthermore, as discussed in this chapter, the

same type of emotion can yield different motivational responses. Guilt, for instance, can be seen as a motive for acting morally in the future. And fear can inhibit approaching dangerous situations. As Gregory Currie argues (2020, p. 128), emotions also tend to interact with other elements of our mental economy, and vice versa, leading to dynamic *updates* in emotional, behavioural, and cognitive responses. The interplay between emotions and other conditions steering judgment and decision-making, hence, undermines the claim that emotions (elicited in VR or elsewhere) *essentially* blind and obstruct sound (rational) judgment and decision-making. As this association between emotions, irrationality and manipulation cannot empirically stand, VR experiences can perhaps be examined from an alternative angle as a tool for the *assessment* of moral judgment and decision-making.

For instance, philosophers, psychologists, and programmers are already working hand in hand to create representations of moral dilemmas in virtuality, such as the famous "trolley problem" and "mad bomber dilemma" (Niforatos et al., 2020; Pak et al., 2019). These virtual experiences indeed trigger heightened emotional responses, but this need not be taken as a hazard or a tool for manipulation, instead, it seems to be an extension of the aforementioned notion of "affective forecasting", offering users an interactive, risk-free generation of hypothetical scenarios for practising and assessing their moral intuitions and decision-making. Hence, I argue that exposure to events that stimulate emotions cannot be by definition taken as evidence for content creators' intentions or accountability for users' irrational or unethical behaviour in the real world, for such behaviour is less likely to occur unless there was already a disposition for a given user to act unethically or irrationally.

Further, we can identify some benefits of experiencing emotional responses to virtual stimuli, especially in the context of psychotherapy. For instance, exposure therapy is said to be done via imaginative activities, where patients are guided into recollecting and summoning past traumatic memories and emotions they intentionally suppressed to avoid mental agony. The goal behind such imaginative recollection is to release these emotions with the help of a practitioner. However, in some cases, patients are unwilling or simply unable to effectively engage with the said imaginative activity. JoAnn Difede and Hunter Hoffman, therefore, attempted to use VR for facilitating exposure therapy for a survivor of the World Trade Center attack of 9-11-01, who had developed acute Post-traumatic Stress Disorder (PTSD) and was not responding to the traditional methods of exposure therapy. The VR recreation of the event included "jets crashing into the World Trade Center with animated explosions and sound effects, virtual people jumping to their deaths from the burning buildings, towers collapsing,

and dust clouds" (Difede & Hoffman, 2002). The results of the study are reported to be dramatically successful, encouraging Difede and Hoffman to use VR for more beneficial aims, such as pain mitigation and distraction for severe and mild cases of burn victims. [15]

However, there is another recurring worry concerning the negative psychological effects resulting from the overstimulation of stressful and negative emotions. Ramirez and LaBarge report that some studies show that undergoing excessively violent or traumatising events in VR, or other communication mediums, can cause considerable, and potentially lasting, psychological damage. Therefore, they recommend that transparency is a crucial requirement for designing and advertising VR experiences that stimulate heightened emotional responses, especially for entertainment purposes. They also propose the "equivalence principle", which advises that "if it would be wrong to allow subjects to have a certain experience in reality, then it would be wrong to allow subjects to have that experience in a virtually real setting" (p. 18). This plea is specifically directed towards VR content creators and psychologists who use VR to put participants through extremely stressful conditions to obtain an in-depth understanding of their neurological and emotional responses.

It can be agreed that transparency and providing fair disclaimers and instructions for engaging with virtual experiences are perfectly reasonable requirements that are already being applied, at least in the fields using VR in video games and entertainment. In psychology, ethical questions are still raised and debated on the limitations of using VR for experimentation. There isn't yet a recognisable threshold or distinction between ethical and unethical requirements in this domain. However, it can be argued that generalising the equivalence principle can be too limiting to the hedonistic, aesthetic, and as discussed above, the therapeutic value of virtual experiences. A less restrictive and perhaps less practical and more demanding approach would be to ethically and psychologically assess VR paradigms on a case-to-case basis or to develop an ethical framework that takes into consideration the ethical and psychological "passes" often granted to fictional and non-fictional representations as tools for transmitting nuanced emotional expressions for the benefits listed above, instead of putting VR on an equal footing to real-life experiences.

It is also more reasonable to advocate for a shared sense of accountability between VR users and content creators, instead of holding the latter fully responsible for how users respond to virtual experiences. This takes us back to Ihde's account of human-technology relations,

---

[15] For more on this see http://www.vrpain.com/

112

where he states that both the technological tool, as well as human intentions, constitute a dynamic relationship where neither element is neutral. Hence, managing and accurately predicting all facets of technological effects remains an ongoing, experimental process, in need of the construction of balanced and reliable, ethical and psychological frameworks to steer it towards maximising the benefits and mitigating the negative effects.

6. Chapter Conclusion

This chapter attempted to put forth an expanded version of virtual fictionalism, as presented by McDonnell and Wildman, by investigating the possibility and fittingness of users' emotional and behavioural responses in VR. I argued that what Walton terms "fictional emotions" and their corresponding (proper) behavioural outcomes in VR can be rationalised using the concepts of imagination and make-believe. Following rich insights in the literature, I have understood imagination, in this context, to entail supposition or simulation of certain events or states of affairs without necessarily constructing mental imageries. I also argued that make-believe is not always intended to trigger heightened emotional responses. We can unproblematically allow that certain events in virtuality are truly the case without being moved to respond emotionally or behaviourally. Such responses are only produced when users' evaluate a state of affairs in virtuality as deserving of a given response.

With that said, I also claimed that we have no strong reasons to consider VR-stimulated emotions as incoherent or inconsistent, unless they violate the norms of using VR as a VR, respecting the causal limitations of virtual content. The fittingness and appropriateness of VR-stimulated responses can also be anchored by the users' reasons for engaging with a virtual experience. Some of these reasons can be spelled out with reference to users' interest in the virtual content and the context and content of the virtual experience itself. At the end of this chapter, I discussed some positive and negative, ethical and psychological ramifications of VR's ability to elicit emotions. This discussion will be further pursued in the next and final chapter of this thesis in relation to using VR as an emotionally-charged empathy machine.

Chapter 5: Addressing some Criticism of Using VR as an "Empathy Machine"

1. Chapter Overview

As I foreshadowed in the previous chapters, using VR as a medium for storytelling, specifically dedicated to reinforcing perspective-taking and empathic understanding of others' emotional experiences, comes with the potential for innovative developments in this genre, but also with many challenges and limitations. Scholars from a variety of disciplines, most noticeably philosophy, media studies, and XR design, have expressed some scepticism about whether or not VR *should* be used as an "empathy machine". Accordingly, in this chapter, I explore normative objections raised by Paul Bloom (2016, 2017), Joshua A Fisher (2017), and Grant Bollmer (2017). These objections question both the use of VR as a medium for stimulating empathy and the very significance of empathy itself, especially in the context of guiding moral judgment and good behaviour.

First, psychology professor Paul Bloom (2016) provides a long list of grievances against what he terms "emotional empathy" and he explicitly advises against pursuing technologically-mediated empathy. To Bloom, this attempt is futile because empathy is ethically biased. Bloom argues that we are more inclined to empathise with those with whom we share more similarities and those closest to us than with complete strangers or individuals with whom we have conflicts or disagreements. To Bloom, empathy is like a spotlight, which restrictedly illuminates the spot on which it is directed, leaving the rest in the dark. That is to say, he suggests that we tend to selectively empathise with certain individuals and not with others, based on purely personal considerations. Due to this, he equates empathy to prejudice and claims that "empathy can distort our *rational* judgment [emphasis added]" (2016, p. 31). Therefore, the "spotlight metaphor" that Bloom uses is meant to show the narrow range of empathy, as well as its directedness towards narrow vantage points, which are, to Bloom, only determined by biased, self-interested, irrational reasons.

Bloom confesses that he "finds it exhausting to spend even a short time with someone who is depressed or anxious" (ibid., p.131) because absorbing others' negative feelings is both overwhelming and preventative from performing helping behaviours. He claims that a good therapist, for instance, needs to restrain her empathy mechanisms in order to provide more effective help to her patients, as he sees no use in "feeling what others feel", a feature that

Bloom deems central to emotional empathy. He also suggests that empathy, especially that acquired through technological mediation and different news and communication mediums, is characterised by innumeracy and blindness to statistical facts. In one of his interviews, he suggests that we are more likely to feel empathy for a little girl stuck down a well, shown on the news than with hundreds and thousands of children whose suffering we are not aware of on a deep emotional level. He also rejects the ability to empathise with more than one or two targets of empathy at once. To Bloom, the "design" of empathy mechanisms is not adapted for undertaking such an enormous and overwhelming task.

Bloom also discusses "cognitive empathy"; the ability to understand others' experiences and perspectives from their point of view. He seems to criticise cognitive empathy less than emotional empathy. However, he still contends that the former is morally *neutral* and not particularly necessary (or reliable) for guiding moral behaviour. For instance, Bloom argues that we do not need to imagine how it must feel like for a child that is drowning or to imaginatively put ourselves in their situation, as an incentive to save them. For, it is reasonable to rush to save them without resorting to empathy or perspective-taking. In his paper, "Is Empathy Important for Morality?" (2011), Jesse Prinz makes similar arguments about empathy not being necessary, and even harmful, for moral judgment and good behaviour. To Prinz, more reliable motivations for acting morally include rational assessment of the *consequences* of a given act, and disapprobation or strong disapproval of an act such as lying, killing, or stealing without consideration of the consequences, but only in virtue of deeming such act as inherently immoral. Prinz adds that individuals should engage in moral judgment based on evaluations of actions and situations, separate from biased considerations about the person involved in such actions and situations. Since empathy is centralised around understanding a target's emotions, both Bloom and Prinz, hence, deem it an unreliable foundation for morality.

Bloom also puts forth other objections specifically targeting the use of VR as an empathy machine. His main two concerns are with the safety and control that VR experiences integrate when representing horrific events. For instance, criticising *Clouds over Sidra* and other VRs which seek to represent war events, Bloom argues that these films turn other people's suffering into a fun experience. He adds that experiencing war is not necessarily about the sounds and sights that VR communicates, but more about the fear, illness, and anxiety that people experience, which VR can never faithfully fully deliver. Finally, one more disadvantage that Bloom highlights is that the short duration of VR films, making them "consumable" to the

public, obstructs the authenticity with which the original event is communicated. In an article for *the Atlantic,* Bloom writes:

> It's not hard to try out certain short-term experiences, such as dealing with a crying baby for a few minutes, sitting alone in a closet, or having strangers gawk at you on the street. But you can't extrapolate from these to learn what it's like to be a single parent, a prisoner in solitary confinement, or a famous movie star. You can't take an event of minutes and hours and generalize it to months and years. (2017)

After elaborating on what Bloom means by emotional empathy in more detail, I seek to address some of his objections about the significance of technologically-mediated empathy, while conceding that *some* of these objections draw reasonable limitations on what VR for storytelling can accomplish when striving to stimulate interpersonal empathy. With that said, I suggest, following Rueda and Lara (2020), that VR experiences depicting socio-political and cultural representations or traumatic events, should be treated as complementary to other epistemic sources, instead of an absolute replacement.

In addition, I seek to show that Bloom's understanding of emotional empathy in complete separation from cognitive empathy, faces some conceptual issues. As discussed in the first chapter, it is much less likely for an empathiser to merely mimic or absorb the emotions of another individual, separate from any contextual interpretations, except in cases of reflexive emotional contagion. I, therefore, suggest that Bloom seems to be conflating reflexive emotional contagion with intentionally guided empathy. Here, his arguments about emotional contagion being an unreliable mechanism for moral judgment may stand. However, if Bloom picks up a gloomy mood after being around a depressed person for an extended period of time, it is unclear how 1) the target of empathy's sadness is (fully) contextually appropriated by Bloom, and 2) how emotional contagion (or emotional empathy) functions as a sufficient justification for hampering prosocial and supportive behaviour towards the depressed person. Therefore, it is still not well established how interpersonal empathy has definite behavioural consequences in misleading decision-making and moral deliberation**,** as Bloom attempts to argue in some parts of his book.

Alternatively, following Masto (2015), I contend that the very often interconnected cognitive and emotional empathy mechanisms have an *epistemic worth* in guiding better-informed decision-making and moral judgement. The latter can be assessed against several criteria, separately from empathy itself. I agree with Bloom that acquiring information about

others' experiences through empathy cannot be straightforwardly tied to things like kindness and compassion, as some would assume, especially in non-academic parlance. Even more, empathically understanding what others are thinking and emoting can indeed be used as a tool for manipulation and immoral behaviour, as discussed in the first chapter. This shows that the acquisition of an *understanding* (be it cognitive or emotional) of other individuals or states of affairs in the world, is itself a raw material or a tool, which can be steered towards either moral or immoral ends. However, this is not sufficient for discarding empathy altogether as inert, insignificant, and harmful tout court as Bloom and Prinz do.

Subsequently, if empathy is indeed empirically shown to be biased and more likely to be directed towards those we care for than those we care less for or share fewer similarities with, I argue that this highlights, even more, the importance and value of VR storytelling, which gives those interested in acquiring a visceral and evidence-based understanding of others' "foreign consciousness", as Edith Stein describes targets of empathy, a chance to attain a certain degree of such understanding. In other words, our empathy mechanisms are inherently limited. This observation should be taken as a motivation for seeking more epistemically reliable sources for extending our empathy capacities, instead of abiding by our evolutionary restrictions. I propose that the limitation in "natural" empathy gives storytellers and VR content creators good reasons for creating more VR narratives for offsetting this propensity. I also highlight some contradictions in Bloom's account of the implications of empathy. In some passages, he describes empathy as ethically neutral, but in others, he strongly suggests that it is harmful and we could do better without it for it is suggested to lead to irrational judgment.

Grant Bollmer also voices his scepticism on the effectiveness of using VR as an empathy machine. In his paper "Empathy Machines" (2017) he argues that: "technologies intended to foster empathy merely presume to acknowledge the experience of another, but fail to do so in any meaningful way" (p. 63). Bollmer's main worry is that this implementation of VR technology tends to "efface" others' personal experiences by reducing them to mere fragments, the construction of which heavily relies on what can be made sensible to a target audience, inevitably marginalising certain aspects of the original experience. Bollmer believes that developing genuine empathy requires a "universal transmissibility" and full access to all aspects of others' emotional experiences, which he deems impossible due to communicative restrictions due to cultural differences, language barriers, and diversity in emotional expression.

Problematically, Bollmer also suggests that genuine empathy necessitates isomorphic *assimilation* with other targets' personal, political, and ethical stances. Since satisfying these conditions is often neither possible nor desired considering how nuanced and different these stances are, Bollmer denies VR's ability to stimulate genuine empathy and calls VR users to adopt *radical compassion* instead, which in Bollmer's words: "refers to an ethical stance that refuses any attempt to experience, or to completely understand, the experience of another, but instead embraces an openness to understanding and refuses assimilation into one's own self" (p.71). As this quote shows, Bollmer's intuitions seem contradictory, switching between denying the very possibility of interpersonal (experiential) understanding and contending that radical compassion necessitates and entails it.

Bollmer makes the supposition that all VR experiences labelled under the heading of "empathy machine" are meant for leading viewers to acknowledge, which Bollmer uses synonymously with assimilating, all dimensions of others' experiences, including ethical and political stances. In addition to challenges in the universal transmissibility of aspects of others' lived experiences, Bollmer also worries that users in VR can be excessively immersed in their first-person perspectives, preventing them from grasping a given experience as belonging to another. He writes that viewers "hastily *absorb* the other's experience into their own experience" (p. 64). Also, similar to Bloom, Bollmer presupposes that VR can only trigger "empathy circuits", i.e. "nonconscious reflex response that mirrors the experience of another in one's brain" (2017, p. 64). Building upon the discussion in the first chapter, Bollmer seems to be referring to emotional contagion; the reflexive, contentless, psychophysiological tendency to mimic other people's expressed emotions or "catch" the prevailing mood or ambience of a given situation or place.

In response to Bollmer, I concede that it is true that our mirror systems are shown to be activated while undergoing a virtual experience, as they would be in ordinary circumstances. However, it is important to point out that this is not all VR narratives offer. The complementary propositional and non-propositional (i.e. representational) constituents of VR narratives also foster an intellectual understanding of others' experiences, instead of merely triggering certain neurological circuits. Furthermore, as discussed above, empathy is an ethically neutral, epistemically valuable tool that involves several consciously employed mechanisms (namely other-oriented perspective-taking) and thinking processes. To say that empathy is morally neutral, I mean that the very *acquisition* of evidence-based input or other forms of understanding (e.g. through imaginative perspective-taking) does not in itself determine the

course of ethical judgment or behaviour that an agent opts for. Although it can be argued that empathy can be characterised as an epistemic good, which can be employed for informing moral (or immoral) judgment and decision-making. However, whether or not it does so remains contingent on the agent, their heterogenous rational judgment, emotions, value and belief systems and other criteria that shape decision-making and (subjective) moral deliberation.

I also seek to show that it is farfetched to presuppose that VR-mediated empathy necessitates full assimilation. That is to say, one can preserve their identifying features and belief systems while also being able to resort to perspective-taking to acquire an experiential understanding of what characterises given emotional experiences, without radically endorsing political or ethical views of a target of empathy. As mentioned previously, to empathically understand a psychopath or serial killer's perspective or worldview does not necessitate us to also *identically share* their moral and political stances. Bollmer's conceptual suppositions about empathy, when extended to the context of VR, would characterise the perfect VR empathy machines as automated "brainwashing machines", which are highly (luckily) implausible. It is more accurate to describe empathic understanding as occurring *in degrees* in a way that may or may not influence one's values and decision-making, depending on an interplay among several criteria discussed in psychology and many other disciplines. Therefore, I suggest that Bollmer is at fault for considering empathy an all-or-nothing phenomenon that necessitates full assimilation. Following Fagiano, for instance, it is more helpful to think of empathy as a relational process of "feeling with" another agent than fully "feeling as" another, for the latter outcome might be unnecessarily demanding. Coplan also confirms the coherence (and importance) of self and other distinctions when feeling empathy for another person, as discussed in the first chapter.

The next objection I tackle in this chapter is presented by Joshua A. Fisher (2017), and it resonates with some of Bollmer's claims. Fisher worries that through VR narratives, we can only empathise with the VR content creators' egocentrically-shaped understanding and representation of other targets' experiences. He claims that VR "aesthetics" often disrupt (especially non-fictional) narratives. Following Grierson, he terms this view "truth made beautiful" (p. 236). This approach, to Fisher, runs the risk of dooming aspects of original features of a target's experiences to be lost in transit when portrayed in VR, either due to misinterpretations or reshaping the represented narrative to fit the experiential paradigms of the viewer. Fisher reminds us that VR narratives are representational, i.e. only experienced second-hand, either through animated virtual media or photographic content. He adds that

attempting to capture and recreate targets' emotional experiences will always be determined, and often limited, by the content creator's degree of experiential understanding of the target experience in question. Continuing to use *Clouds over Sidra* as an example, according to Fisher, we only get to empathise with Chris Milk's interpretation of Sidra's emotional experience, and not Sidra's herself. The worry is that it is never a target's emotional experience being represented in the VR narrative, but only the VR designer's interpretation that is often liable to misinterpretations and mischaracterisations. Fisher's critique speaks directly to the epistemic reliability of virtual representations in transmitting aspects of other targets' emotional experiences.

In an attempt to respond to Fisher's worries, it can be argued that even in the case of face-to-face communication, emotional expressions are not always fully or even correctly interpreted and understood among individuals. As I mention in the first chapter, there are limits to "perfect empathy"; complete understanding of others' experiences from their viewpoint. Nevertheless, I concede that the issue Fisher raises on the central role that VR storytellers and content creators play in the interpretation and representation of target emotional experiences is a legitimate one. Evaluative questions on the "process reliabilism" of how virtual representations are created are beyond the scope of this research. But at least it can be suggested that Fisher's worry further emphasises the significance of empathic understanding in the stages preceding the production of a given VR narrative. This is because the degree to which a content creator empathises with a target's emotional experience partially determines the degree to which viewers can experience empathy for the represented character. Therefore, the complex topic of empathically understanding others' emotional experiences is not only relevant to philosophers and scholars looking into it, but also to VR designers who aspire to create narratives pursuing these objectives.

With that said, this chapter will be divided into three main sections, respectively discussing Bloom, Bollmer, and Fisher's lines of criticism against using VR as an empathy machine. The significance of this discussion can be illuminated in terms of setting realistic expectations for VR-mediated empathy as well as recognising the key role that VR producers, together with VR technological affordances, play in mediating empathy. Issues such as misrepresentation, short duration of VR films in comparison to the original experience, and elusiveness of authentic framing of VR narratives are considered. Similar to earlier communicative mediums, VR as well is prone to be utilised for communicating fake news and false representations. Coupled with the level of emotional engagement and memorability of

VR experiences, content created in this medium should indeed be created with caution, especially in the non-fictional genre. Nevertheless, adopting a more optimistic outlook than Bloom's, I content that the powerful impact that VR technology entails can also be harnessed as a tool for promoting human connectedness and bridging the epistemic gap in understanding what it is like to undergo a given experience from another's perspective, albeit for a short amount of time. When successful, VR can be an epistemically valuable gateway medium for enhancing viewers' appetite for more reliable information from other sources, which can shape better-informed decision-making.

I agree with Bloom that it is conceptually inaccurate to associate empathy with sound moral judgement (however heterogenous that may look), kindness and prosocial behaviour. But just because VR does not condition users to do "the right thing" by definition, it still provides them with invaluable forms of information that would be, in turn, autonomously interpreted vis-à-vis users' subjectivities. Hence, Bloom's contention that mediated and non-mediated empathy lead to immoral judgment is not well-supported. After a viewer experiences a VR representation, their responses will inevitably vary from being emotionally moved, to wanting to do more for a given cause, or remaining completely unmoved or criticising the still imperfect glitches in the technology. What is evident is that the implementation of VR in communicating emotional experiences is still in its infancy. Hence, more interdisciplinary research is still dynamically being carried out to shape and guide the construction of virtual worlds and narratives. Therefore, highlighting existing deficiencies and limitations remains crucial to addressing them and for transitioning to the next stage of VR development as an empathy machine, not for personal appropriation or brainwashing, but for supporting human connectedness and evidence-based interpersonal understanding, hopefully eventually leading to disambiguation of a distant "other".

2. Paul Bloom's Criticism of Empathy

In the 2006 Commencement Speech to Northwestern University graduates, former US president Barack Obama famously said that the country suffers from an "empathy deficit". He explains the term as "the ability to put ourselves in someone else's shoes; to see the world through those who are different from us – the child who's hungry, the laid-off steelworker, the immigrant woman cleaning your dorm room… After all, if they are like us, then their struggles are our own. If we fail to help, we diminish ourselves". As mentioned above, Paul Bloom vehemently objects to associating empathy with good behaviour and sound moral judgment. In

his thought-provoking book *Against Empathy: The Case for Rational Compassion* (2016), Bloom describes empathy as "biased and parochial". To further unpack and support this claim, Bloom presents several pieces of evidence from neuroscientific studies and a plethora of real-life examples.

To say that empathy is biased and parochial, Bloom means that individuals can only selectively and narrowly empathise with the few over the many. Bloom also argues that bound by our evolutionary prejudice, we tend to empathise with members of our "tribe", sometimes at the expense of others who might have more merits. For instance, Bloom insightfully borrows Peter Singer's example which is meant to highlight arguably misplaced sentiments and prosocial behaviours in the selection of charitable work. The example goes that the Make-A-Wish foundation has helped Miles Scott, a five-year-old with leukaemia, to fulfil his dream of spending a full day as a superhero "bat kid". The activity cost a total of at least $7,500. Singer contends that this money could have been better used in saving the life of 3 children living in areas infected with Malaria, by providing bed nets and other supplies. Singer argues that if Scott's parents had an option between saving their kid's life and fulfilling his dream to be a superhero for a day, they would have chosen the former option. Singer then states: "When more than one child's life can be saved, the choice is even clearer. Why then do so many people give to Make-A-Wish when they could do more good by donating to the Against Malaria Foundation, which is a highly effective provider of bed nets to families in malaria-prone regions?" (qtd. in Bloom, 2016, p. 90).

Another intriguing case that Bloom discusses concerns what is known as "the identifiable victim effect", referring to intensified and emotionally driven responses that the public often displays in reaction to well-identified victims, especially in news outlets, and not other cases that the "empathy spotlight" is not directed toward. This, as Bloom argues, shows that empathy is innumerate. Echoing Paul Slovic's complaints, Bloom for instance explains that "each day more than ten times the number of people who died in Hurricane Katrina die because of preventable diseases, and more than thirteen times as many die from malnutrition" (p. 84). However, responses to the devastating Hurricane Katrina remained stronger, amplified by broad media coverage. Here, Bloom sheds light on the power of novelists, journalists, and content creators in selecting and drawing attention to certain calamities and not others. This selection and coverage are not necessarily guided by statistical or concrete considerations, which, to Bloom, constitute a more reliable rational foundation for decision-making and moral deliberation than emotions. At this point, it is sensible to inquire whether these shortcomings

necessarily occur due to empathy per se. In some passages, Bloom responds to this question in the negative, but in others, he still insists that empathy is one of the causes of harming and distorting rational judgment because it mainly involves biased reasoning that is often indifferent to broader, long-term consequences.

One of the examples Bloom uses to highlight the alleged negative effect of practising empathy concerns parenting. He suggests that parents who empathically choose to favour "short-term buzz" which makes their children momentarily happy risk causing negative long-term consequences. Bloom also draws upon examples in the context of policy-making. If a ruler or a policy-maker empathises with all parties, to Bloom this may lead to making irrational decisions that might not be beneficial in the longer run. Another one of many examples that Bloom provides concerns practising empathy, which ends up leading to helping behaviours toward child beggars in the developing world. Bloom suggests that "by giving, you make the world worse" (p. 93) because the funds offered to these children can end up promoting criminal activity of gangs enslaving them. As a remedy, Bloom advocates for compassion guided by rationality, which consists of clear justifications of actions, led by the concrete analysis of costs, long-term effects, and the scale of influence that certain behaviour may trigger. Then Bloom suggests this should be coupled with an assessment of the happiness and the well-being of the majority, as long as it does not clash with rational considerations.

In what follows, I will attempt to respond to some of Bloom's objections to empathy and technology-mediated empathy in VR by highlighting a few contradictions and unfounded associations of negative *behavioural* consequences with empathy. I first begin by disambiguating Bloom's conceptualisation of "emotional empathy", characterised as sharing and feeling what others feel and its effect on moral deliberation and decision-making.

2.1. The Myth of "Emotional Empathy"

Bloom does not provide a precise definition of what he means by emotional empathy but throughout his account, this notion seems to chiefly revolve around a strong sense of *sharing,* perhaps even replicating, what another is emoting. Bloom's main premise is that emotional empathy is exhausting and neither necessary nor reliable for moral judgement, as there are more reliable routes for accomplishing such ends. Bloom writes: "I can worry about a child who is afraid of a thunderstorm and pick her up and comfort her without experiencing her fear in the slightest. I can be concerned about starving people and try to support them without having any vicarious experience of starving" (p. 128).

Here, Bloom's understanding of emotional empathy seems to be more akin to emotional contagion, which is reflexive and unintentional. Bloom attempts to differentiate between emotional empathy and emotional contagion by suggesting that the former is more than just a reflex; it is "an act of will", in that we can choose to either amplify or restrain our empathic responses based on contextual factors such as beliefs, expectations, motivations, and judgments. This is well-established by several neuroscientific studies, which show that individuals are less likely to feel empathic if their target of empathy is a member of an opposing political party or even a supporter of their least favourite soccer team. For instance, As shown by Jamil Zaki's extensive review of studies on empathy in psychology (2014), empathy has been confirmed to be a consciously (in the sense of intentionally) triggered phenomenon. Even at the level of its underlying *automatic* (neurological) mechanisms (such as emotional contagion), it has been shown that empathic responses can be either *avoided* (i.e. suppressed) or *approached*, depending on one's relationship with the target of empathy as well as several other criteria; such as "pain avoidance, material costs, and interference with competition" (Jamil, 2014, p. 5). For instance, it has been observed in numerous studies that people often intentionally avoid being empathic towards others who are experiencing distress, in an attempt to not "catch" negative emotions. Overall, employing empathy is often modulated and heavily context-dependent.

As mentioned before, echoing Prinz, Bloom argues that since empathy is flawed by being biased and innumerate, there are more reliable means for making moral judgments, namely rationality; which is grounded in providing reasons and justifications based on concrete evidence that would be convincing to an interested third party. More importantly, Bloom juxtaposes rationality with "gut feelings" which he deems empathy to be one of. This categorisation shows one of several issues with Bloom's account. The main problem is that Bloom insists on making a clear distinction between emotions and cognitive components in conceptualising empathy. For instance, "catching" a feeling of sadness from another individual is often not done independently of contextual considerations, as Bloom himself acknowledges. My contention here is, as I argued in the first chapter, that practising emotional empathy is rarely separate from cognitive empathy unless we're using emotional empathy synonymously with emotional contagion. Now, in practising cognitive empathy, which aims toward acquiring a better understanding of others' experiences or emotional states, it is hard to trace irrationality in the plain act of the pursuit of understanding. We can even go a step further to argue that it is more rational than not to seek out a multidimensional understanding as a guide for informed

moral judgment and actions. What would be unfitting is perhaps making decisions and moral deliberation *only* by virtue of knowing how another person is feeling. For instance, imagine that a judge would reduce a serial killer's sentence because a more severe verdict would hurt the killer's feelings. Here, I tend to agree with Bloom that empathy alone might not be sufficient for decision-making. However, I disagree with the parts of his argument that insist that empathy is inherently irrational and leads to harmful consequences for relying solely on gut feelings, for empathy can also be a foundation for good behaviours. Indeed, emotions are often traditionally characterised as antithetical to rationality. As the famous Latin phrase "ira furor brevis est" suggests, anger is a brief episode of madness. Emotions may lead people to perform irrational acts. Nevertheless, as discussed in the first chapter, empathy is not an emotion at all. It is instantiated by complex dimensions and components that do not simply comprise gut feelings, but also contextual and rational considerations, as sometimes acknowledged by Bloom himself. Whether or not the behavioural consequences and quality of reasoning resulting from empathically engaging with others are sound is a completely different question from determining whether or not empathy itself is good or bad. Instead, empathy is simply instrumental in making sense of others' emotional experiences, but it does not inherently entail a decisive inclination to one moral position over the other. Contra Bloom's intuition, this is also not to say that empathy is a cause of harm, instead, as mentioned before, it is more coherent to characterise it as neutral to morality.

Next, and more problematically, Bloom suggests that absolute emotional empathy is possible and that it leads to appropriating another's emotions to oneself, consequently paralysing prosocial behaviour. He writes:

> What gives empathy its power, after all, is that we appreciate that we are feeling what another feels. If I feel your pain but don't know that it's your pain—if I think that it's *my* [emphasis in the original text] pain—then I'm not going to help you. (p. 156)

Here, Bloom is too quick in assuming that through empathy, individuals acquire a full sense of others' emotions and are paralysed from performing helping behaviours as a result. This is especially problematic in the case of distress. As Zahavi explains:

> Just as we ought to respect the difference between thinking about a lion, imagining a lion, and seeing a lion, we ought also to respect the difference between thinking about Adam's distress or embarrassment, imagining what it must be like for him to be distressed or embarrassed, and being empathically acquainted with his distress and embarrassment in the direct face-to-face encounter. (2014, p. 150)

Some philosophers, such as Fagiano, instead prefer to speak of empathy in terms of "feeling with" instead of "feeling as", without exaggerating the ability to experience another's emotion completely as another. Coplan also makes sure to emphasise that the ability to maintain a clear self-other distinction is a necessary feature for experiencing empathy, considering that we are only acquainted with others' emotions third-personally and hence with visceral restrictions. It is not apparent how, even in first-person VR, a person would readily appropriate another person's distress to themselves. For example, if Sally catches a gloomy mood from spending too much time with her friend Chris who is clinically depressed, it would be ludicrous to suggest that Sally is now depressed too.

Next, concerning sound moral judgment and prosocial behaviour, it is not immediately clear how experiencing another's pain would paralyse these functions. Surely, when Sally has a feel of Chris' sadness, albeit restricted, she would also have a motivation to act in favour of alleviating some of his mental pain, unless she is a psychopath or someone who particularly takes pleasure in watching other people suffer, or simply disinterested in embarking on such task, or judges that giving Chris space would be more beneficial for him. Bloom additionally argues that even when empathy successfully facilitates prosocial outcomes, this would be egocentrically motivated because it is reached as a result of experiencing others' feelings as *one's own* through projection. In other words, through imaginative perspective-taking, one would pose the question "what would it have been like for me to be in another individual's challenging situation?" and *then* proceed to act prosocially. More extremely, in other cases, Bloom explains that empathy can be implemented immorally. This "dark side" of empathy is also discussed by Frans de Waal (2009) and Mark Fagiano (Fagiano, 2016) and can be spotted in examples of torturers, bullies, and psychopaths who sometimes utilise empathic understanding merely as a tool for inflicting more pain and suffering on their victims. That is to say, in the context of morality, empathy can be a hit or miss. According to Bloom, it does more harm than good when it falls "into the wrong hands" (2016, p. 37).

The claim that the knowledge or understanding that can be acquired through empathy can be used either for good or bad is legitimate. Even more, we can also concede that such epistemic input can be completely inert in cases where the agent chooses not to act upon that knowledge or understanding. Nevertheless, I argue that suggesting that empathy, as a tool for acquiring these epistemic goods, is particularly hampering good behaviour, is not well supported. It is instead more helpful to assess behavioural outcomes separate from empathy, which can be

more accurately described as a morally neutral epistemic tool for acquiring information. In other words, empathy's mechanisms are not essentially constituted of (im)morality. If an individual behaves immorally, it is not plausible to blame this on the process of interpersonal knowledge acquisition, but instead, it is more coherent to independently assess individual choices of moral deliberation and behaviours apart from the means of acquiring the information itself. Even more, Meghan Masto (2015) argues that empathic understanding is an epistemically valuable tool not only for working out the right course of action but also because it is important for morally assigning blame or praise to individuals' behaviours. For instance, if empathy enables me to understand that my friend is bothered by the scent of tobacco and despite this understanding, I nevertheless continue to smoke around her, this makes me even more accountable for not acting upon my empathic knowledge. Therefore, Bloom's supposition that empathy is not significant and sometimes harmful to moral judgment might be too hasty.

2.2. Addressing a Critique on the Use of VR as an "Empathy Machine"

More relevant to this research, Bloom argues that *VR-mediated empathy* runs the risk of producing irrational beliefs and misguided decision-making. These contentions stem from Bloom's supposition that emotionally charged experiences shown in VR are essentially obstructive of rationality. Put simply, Bloom believes that VR narratives which confrontationally depict emotionally loaded experiences, especially in association with political and ethical stances that may contradict the user's experiential paradigms or belief systems, often threaten to hinder rationality and contaminate our decision-making mechanisms. More explicitly, Bloom argues that by blinding rational judgment, empathy-inducing, emotionally-charged VR narratives tend to selectively amplify a given perspective at the expense of another and can hence be used as a tool for *manipulation*. Bloom also prolifically writes about the limitations of VR *sensory* affordances.

In an attempt to address these worries, it is crucial to point out that the integral ability to consciously either avoid or enhance empathic responses highlights the existence of underlying cognitive processes that steer empathy. This means that empathic responses to VR narratives are not standardised nor automatic. This goes to show that it is very less likely for irrational beliefs or behaviours, outside of the viewer's behavioural patterns and character to emerge merely as a result of exposure to VR narratives that are directed at enhancing empathic understanding. For instance, if one watches a VR film representing a serial killer's experience, from their perspective, perhaps even integrating the killer's "heartfelt" testimony in the

storylines, this would indeed trigger a variety of emotional responses. However, it would be counterintuitive to believe that only by virtue of being exposed to this VR experience, would we consider the killer's actions as warranted, or even comprehensible, unless we initially endorse such unethical behaviour prior to embarking on the VR experience. There are of course more complex scenarios. For instance, if someone does not support a given political party and views a VR narrative that seemingly supports certain beliefs of that party, one possible outcome is that they would acquire a new perspective or arguments that may change aspects of their worldview. But this remains fully contingent on the subject, the strength of their inclination to a given position, and how convincing the VR narrative is, as there aren't enough grounds for causally explaining a "change of heart" with reference to only one of these factors.

It is hence important to note that VR-mediated empathy, similar to non-mediated empathy, is not automatic and it can only occur through an interplay among several factors among which are the user's past experiences, beliefs, and experiential paradigms. This ought to be seen as an advantage rather than a drawback, because otherwise, VR would go from an epistemically valuable "empathy machine" to a "brainwashing machine", which is neither plausible nor desired. Emphasising Bloom's assertion that empathy is a morally neutral tool shows that irrational decisions and behavioural responses are not a plausible outcome of having empathy per se, but simply of mere irrationality and flawed moral judgment. It is hence more productive to assess the effects of VR narratives on a case-by-case basis, in parallel with users' real-life interpretations and reactions, instead of condemning the whole genre as causing irrationality and immorality.

Nevertheless, Bloom's critique of the sensory limitations characterising VR narratives is a legitimate one. For instance, there exist many VR simulations that claim to have captured an accurate depiction of mental illnesses such as schizophrenia, integrating exaggerated audio-visual effects that are more akin to the ones used in horror movies or video games. According to the feedback provided by individuals diagnosed with schizophrenia, these kinds of VR films advance erroneous and negative misrepresentations of the original emotional experience in question. This type of stereotypical and not well-researched characterisations might deliver a false image of individuals who struggle with mental illness or of other represented targets. As Arielle Michal Silverman points out, these simulations "give the mistaken impression that the entirety of being disabled is marked by loss, frustration, and incompetence." This highlights the need to proceed with caution and recognise the limitations and intricacies of communicating a target's emotional experiences in VR. Luckily, there is always room for creative

improvements in constructing VR narratives for stimulating empathy, based on informed research and ongoing feedback.

Bloom, therefore, highlights limitations in VR representations, especially with regard to restrictions in delivering a replica of others' experiences. However, these limitations are an inevitable characteristic of technological mediation and representation in broad terms. It would not be an exaggeration to identify such restrictions as a defining feature of second-hand representation. However, do these imperfections mean that we should give up the endeavour of communicating people's experiences and voicing their concerns altogether? This position seems to be a bit too extreme. As shown by Rueda and Lara (2020), idealistic expectations should be managed when it comes to stating the role that VR can play in driving radical change. In Bloom's words, we cannot expect that simply strapping a VR helmet to a user's head would immediately generate a full empathic understanding of others' innermost feelings. Rueda and Lara use the example of VR implementation as an "add-on" for raising more awareness about domestic abuse. This VR implementation is also sometimes integrated into rehabilitation programmes but is in no way sufficient on its own for reaching desirable change. With that said, it is highly recommended to treat VR as a complementary tool for stimulating empathy and interpersonal connectedness. Therefore, I suggest that VR's significance for empathy need not be articulated with a focus on the technological hype about this new technology. Instead, it is important to recognise the realistic limitations of this endeavour. However, it remains true that state-of-the-art VR narratives are useful when it comes to broadening imaginative perspective-taking and empathy, which should be understood as an ongoing, constantly renewed process. In what follows, I explore more criticism presented by Grant Bollmer on the use of VR as an empathy machine.

3. Grant Bollmer on the Impossibility of "Universal Transmissibility" of Experiences

Grant Bollmer argues that VR *cannot* (and hence should not attempt to) trigger empathy in viewers because he believes that in the process of creating representations and simulations of others' emotional experiences, a serious reductionism and amputation of many lived features of others' personal experiences occurs. This reduction, to Bollmer, is inevitable so that the final product; i.e. the VR film, would be tailor-made to fit what de Sousa terms a target viewer's "paradigm scenario". In the context of emotional education, de Sousa explains that paradigm scenarios are *stories* and past experiences based on which we learn to associate an emotional

response or expression with a given stimulus or situation.[16] Indeed, in one of her interviews, prolific immersive journalism pioneer Nonny de la Peña suggests that one of the reasons inspiring her to create VR films for immersive journalism is to trigger American youth's curiosity about news around the world, perhaps also increasing their desire to consult richer, epistemically reliable sources.

Bollmer worries that through the reductionism and reshaping of the "other" into digestible narratives that can be comprehended from an alternative viewpoint, VR which claims to be an empathy machine hastily absorbs the other's experience into the viewers' own experience. For instance in *Clouds over Sidra,* the original audio and speech of Sidra are translated into English for it to be comprehended by a wider audience, and the children's movement in the refugee camp is, to some extent, *directed* by Chris Milk and his team. Hence, Bollmer worries that the dramatic effect present in almost all VR films risks negatively affecting the authenticity of the experience. Bollmer also claims that others' emotional experiences cannot be universally transmittable. By emotional experiences, Bollmer also means others' political and ethical stances, which he assumes that content creators and storytellers communicate with the subtle intention to "convince" the audience to change their pre-existing views and adopt new ones as a prerequisite for experiencing empathy. In simpler words, Bollmer is arguing that in order to accredit empathy to a viewer, they must be willing to personally indulge and subscribe to all represented views in a VR representation.

Clearly, this requirement is not only too demanding but also not plausible, especially considering how nuanced and sometimes conflicting VR representations are, in reflection of people's real-life various stances. As Bloom discusses, it is not physically possible to empathise with every single person, as empathy mechanisms are often either restrained or enhanced on a context-dependent basis. This should be seen as a positive aspect, attesting to viewers' mindfulness and capability to empathise (imaginatively shift perspective with another) without radically accepting every single stance they are presented with. Hence, that political and ethical stances cannot be universally engraved in all users poses no threat to the implementation of VR technology as an empathy machine for the purposes outlined at the beginning of this chapter. Empathy, as discussed in the first chapter, occurs in *degrees*. Therefore, it should not be identified as an all-or-nothing phenomenon. Additionally, as I mentioned before, we can successfully develop an experiential understanding of the gruesome perspective of a

---

[16] See de Sousa (1990) pp. 434–446.

psychopath or serial killer, without the need to radically share their ethical stance. Assimilating others' views should, therefore, not be seen as a requirement for experiencing interpersonal empathy.

Bollmer writes: "empathy, I claim, while seeming to be about *positive relations* [emphasis added], seeming to be about acknowledging another's experience in a full and politically productive way, is a negative annihilation of the Other as their otherness becomes nothing beyond what can be absorbed and experienced by oneself" (2017, p. 72). In other words, Bollmer suggests that VR content creators cannot authentically experience, understand or communicate others' experiences. Moreover, the main priority would always be to create high-selling VR films, which comply with users' standard experiences. Prima facie, this supposition seems to be dismissive of the ability to understand new experiences beyond our usual paradigms. As discussed by Scheler (2017) for instance, we come to learn that if a dog is wagging its tail, that means it is happy, despite our inability to physically experience such a thing. Examples are numerous about inter-species interaction and grasping the significance of new experiences that are beyond the confines of our experiential paradigms. Moreover, according to Ekman's research on basic emotions, humans have been shown to share several manifestations of emotional expression, making interpersonal understanding possible to a great extent. Nevertheless, editions and omissions are an inevitable step in post-production to make a given VR representation *accessible*, which is the central aim behind (technological) mediation; to deliver representations of experiences that would otherwise be out of reach.

Another recurrent, and often exaggerated, criticism of VR's usage as an empathy machine concerns the so-called problem of over-immersion. As Bollmer proposes, VR experiences can be so immersive that the viewer absorbs and adopts another's experience as their own, losing sight of the represented "other". This claim seems to be overly magnifying the effect of virtual embodiment. For instance, in the context of video games, gamers might be immersed in an embodied manner with the avatar they play as. This simply refers to the synchrony and psychophysiological illusion that VR is often reported to generate. However, in the context of non-fiction, what would it be for a user to mistake their own experience for the experience of a refugee, a homeless person, or a prisoner in solitary confinement? I believe this assumption does not provide an accurate description of users' phenomenology when embarking on virtual experiences. Moreover, users are not empty vessels who are completely oblivious to the contextual and propositional cues that almost always frame and organise a VR narrative or

documentary. Hence, there will always be a certain distance that allows experiencing empathy *with* and *for* another, instead of mere appropriation.

A more plausible suggestion is that users may experience *resonance* with another's experience, perhaps leading to connectedness and compassion. Alternatively, they may identify a given character in VR as highly antagonistic. In any case, it is not empirically established to suggest that the user's *self* simply fuses or dissolves due to experiencing a highly immersive VR experience. This is not only due to technological imperfections, distinguished by issues like motion sickness, high latency, and the heavy weight of VR gear, just to name a few deficiencies, but also because users' awareness and cognitive capacities are not shut down during a VR experience if anything their critical capacities are said to be. As discussed in the second chapter, VR experiences are not equivalent to hallucinations or mind-altering states.

In what follows, I tackle Fisher's criticism. Similar to Bollmer, Fisher also suspects that VR entails misrepresentation and generation of what he terms "empathic actualities", which involve intensified dramatic effects that do not fully reflect the reality of represented targets in VR.

4. "Truth made Beautiful": Discussing J.A. Fisher's View on Using VR for Empathy

In his paper, "Empathic Actualities: Toward a Taxonomy of Empathy in Virtual Reality" (2017), Fisher does not necessarily object to using VR as an empathy machine. Instead, he highlights the ambiguity in the way that VR producers use the term. This observation was shared by a great number of filmmakers attending the 2017 Tribeca Film Festival – an annual event supporting the innovative integration of technology in filmmaking – who told Adi Robertson that empathy (as a term) has been overused in the industry to the point of confusion and that sometimes it is merely utilised as a publicity stunt, without any profound or clear meaning (Robertson, 2017). Fisher then suggests that there is a need to create clearer rhetoric when discussing this implementation of VR technology.

Fisher explains that VR producers create "empathic actualities", which he identifies as dramatic, emotionally engaging representations of reality. He then emphasises that these representations are "creatively treated" to encourage users to participate in roleplay, in "almost real life" scenarios, arguing that when it comes to empathy, viewers' empathic intentionality is only directed towards the VR producers or designers instead of the represented subject, whose experience remains elusive and concealed by dramatic effects and modification done in post-

production. Fisher assumes that in VR empathic actualities, we are not presented with a fully authentic view of the world, but instead, we are restricted by the producer's "vision", which can involve techniques and strategies that steer viewers towards producing a given emotional response and not another. Fisher then importantly states that: "it is critically important for those working in the VR documentary space to do justice to those they represent in their experiences if claims of empathy are to be made with any legitimacy" (p. 238).

What might be a bit confusing in Fisher's account is that he suggests that viewers' empathy is directed towards the VR designer, while their sympathy (in the sense of care and compassion) is directed towards the represented subject. But how can that be? If users can experience sympathy towards a subject in VR, this implies that aspects of the subject's experience are successfully communicated through this medium, perhaps enabling the user to also acquire a degree of empathic understanding. Furthermore, it is ambiguous to claim that viewers come to empathise with the VR designer themselves. One way to interpret this claim can be made with reference to "cinematic grammar"; the use of wide, middle, and close-up shots, using music, voice-over, and other techniques to make the experience more lucid. Here, the viewer can be said to empathise with the content creator's cinematic tone, but surely not their experience, because the latter is not personally represented in the VR; it is instead the experience of another subject. On a marginal note, Fisher contends that it is challenging to perform some of these features in VR interactive storytelling due to the interaction paradox explained in the third chapter.

Next, describing interactions in VR in terms of roleplay undermines the user's genuine experience of both empathy and sympathy, because it implies pretence. This might be unproblematic in the context of fictional VR experiences, but in VR non-fictional documentaries, presupposing roleplay just by virtue of the existence of interaction may not be well grounded. Instead, direct interactions and users' intellectual understanding of the context of a given narrative in VR hardly necessitates any kind of roleplay or pretence. For instance, the user does not need to pretend they are Sidra in order to empathically understand her depicted experience on a deeper level. All in all, I agree with Fisher's first two claims that 1) it remains valid that VR producers have an important role to play in mediating empathic experiences that may indeed include emotionally charged expressions, and 2) there is indeed a need for creating a clearer discourse when discussing the use of VR as an empathy machine. The second endeavour is harder due to the lack of consensus in using the term "empathy". However, it might be confusing to resort to terms such as "roleplay" and "empathic actualities", because

these terms restrictively overemphasise the artificial nature of virtual experiences while leaving out the genuine effect they are often reported to induce in viewers, especially in the non-fictional (veridical) genre. Without this, it might not be convincing to invite VR producers to represent their target subjects and experiences with justice, authenticity and care. That is to say, if VR empathy machines are interpreted as nothing more than a theatre where we engage in amusing roleplay, why should VR producers bother with making these experiences authentic? Hence, it is crucial to not equate the virtual with the unreal, as discussed in the second chapter. And even though we might be inclined to discuss the "illusory" effects of virtuality, it is equally important to discuss its genuine effects with accuracy so that the real implementations of this medium would not be understated. Virtual representations of sensitive topics such as the social and economic struggle of under-represented minorities, the subjective experience of individuals with mental illnesses, solitary confinement, and many more are a reflection of real-life occurrences. The fact that they are communicated through an artificial medium that incorporates psychophysiological sensations such as virtual embodiment and presence does not make these experiences any less real. Therefore, unequivocally highlighting the veridical and non-veridical aspects of virtual experiences and the intentionality of users' responses remains crucial for determining the standards and ethical frameworks for governing VR narratives.

5. Chapter Conclusion

In conclusion, some of the issues discussed above on procrustean misrepresentation and the sensory limitations of using VR as an empathy machine are indeed valid. They also shed light on avenues of potential development in this genre of VR implementation. For instance, there exists an accelerating number of VR simulations that attempt to capture distressful experiences in VR, such as becoming homeless, becoming visually impaired, or experiencing mental illnesses such as schizophrenia. At times, these virtual experiences can be epistemically informative by developing a better understanding of a certain human condition. However, more often than not, VR "aesthetics" tend to implement exaggerated audio-visual effects that distort the original experience. This might be unproblematic for fiction, but when dealing with real-life, sensitive issues, these representations need to be created with accuracy. Moreover, the short duration of these VR narratives leaves the user with a fragment of the original experience and at times, with erroneous and misleading representations. It is sometimes noted that VR users' *attention* often gets more drawn to technological features rather than to the message that is being transmitted by the VR narrative. In this case, the overall user experience becomes more

akin to a video game than an emotionally stimulating narrative that is meant for enhancing empathic understanding. This means that a great deal of research must be conducted not only on the technicalities of constructing virtual experiences but also on narrative contextualisation, accuracy and authenticity. These endeavours would be beneficial for mitigating procrustean misrepresentations and balancing between the user's interactive "freedom" in the VR experience and the sensitive content delivery that a given VR narrative seeks to communicate.

However, we must not forget that viewers' response is not solely the product of how accurate and immersive VR narratives representing others' experiences are. As pointed out by Bloom, empathy remains a context-dependent phenomenon and this applies to technologically-mediated empathy as well. The takeaway from the discussion in this chapter is that empathy need not be associated with benevolence and prosocial behaviour or with immorality and manipulation, for it is more accurate to characterise it as a morally neutral, epistemically valuable source of acquiring a degree of understanding of others' emotional states. In extension, the role of VR as an empathy machine was discussed with consideration of realistic limitations and avoiding discussing this implementation with the prevailing hype that is often present in conversations on VR technology. On the other hand, it is undeniable that VR offers new affordances and interactive possibility which is said to add to the memorability and emotional engagement of viewers in VR narratives. This ascribes a considerable degree of power to VR storytelling and hence calls for raising and tackling more profound research questions in this field of research.

General Conclusion

This thesis has explored conceptual and ethical issues, promising potentials and limitations surrounding the utilisation of VR technology as an empathy machine, starting with conceptual problems revolving around the notions of empathy and virtuality. Following the pluralistic approach, I have distinguished two varieties of empathy; namely aesthetic and interpersonal empathy and continued to focus on the second variety throughout this thesis due to its close relevance to the topic at hand. I have also discussed emotional contagion and self and other-oriented perspective-taking as essential underlying mechanisms making empathy possible. Perspective-taking is often said to be reliant and hence restricted by the confines of the imagination. Therefore, I have attempted to show that technologically-mediated empathy, through the use of VR multisensory affordances, can act as an extension to perspective-taking, offering evidence-based, experiential and memorable input for enhancing empathic understanding of others' narratives that surpasses imaginative reflection.

After highlighting the disanalogies between VR's current implementation and Nozick's hedonistic experience machine, I elaborated on the novelties and challenges introduced by VR's visual, interactive storytelling. I also provided some examples of fictional and non-fictional paradigms, reported to successfully enhance viewers' empathy or at least willingness to acquire a more informed understanding of others' emotional experiences. I have highlighted that the technological, especially interactive, features offered by VR created more potentials and challenges for spatial storytelling. Immersion in virtual environments is often supposed to reinforce VR's engagement with a virtual experience and strengthen the memorability of said experience. However, users behavioural and emotional responses remain idiosyncratic and influenced by a variety of technological and subject-specific factors.

Expanding the discussion of VR users' often heightened and realistic responses to virtual stimuli, I have attempted to rationalise such responses from the lens of Walton's theory of make-believe. Countra Radford's contentions, I have argued that responses to spatiotemporally inexistent states of affairs, that nevertheless exist in some capacity (e.g. imaginatively, virtually…etc.), cannot be deemed unfitting only by virtue of their ontological status alone. I defined fitting emotional and behavioural responses as ones that are supported by appropriate reasons, vis-à-vis the situation in which they are elicited. As long as VR-stimulated responses are not elicited with an exaggerated degree of intensity and as long as they do violate any measure of fittingness, (quasi) emotions and certain types of behaviours

136

remain fitting and even necessary for the full appreciated of a virtual experience. In the remainder of the thesis, I have recentred my focus around addressing some criticism of using VR as an empathy machine.

Overall, human-technology relations, as Don Ihde affirms, are constantly being reshared and redefined into new actualities. In turn, these are shaped accordingly to the capacities of the technological tool in question and the intentions and utilisation of human users and programmers. Taking an optimistic outlook on the issue, I have argued that VR technology offers artists and content creators rich, multisensory channels for expression that can be exploited for several usages, among which is the representation of aspects of other targets' emotional experiences, which otherwise may remain elusive due to a great number of restrictions, including imaginative resistance, temporal and geographic boundaries, biological and sociocultural differences, just to name a few. Therefore, I have argued that using VR as an empathy machine is worth pursuing and investigating in clearer terms, to more effectively disentangle the conceptual and ethical issues and limitations surrounding the endeavour.

References

Ahn, S. J. (Grace), Bailenson, J. N., & Park, D. (2014). Short- and long-term effects of embodied experiences in immersive virtual environments on environmental locus of control and behavior. *Computers in Human Behavior*, *39*, 235–245. https://doi.org/10.1016/j.chb.2014.07.025

Amin, A., Gromal, D., Tong, X., & Shaw, C. (2016). Immersion in cardboard VR compared to a traditional head-mounted display. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *9740*, 269–276. https://doi.org/10.1007/978-3-319-39907-2_25

Ammann, J., Hartmann, C., Peterhans, V., Ropelato, S., & Siegrist, M. (2020). The relationship between disgust sensitivity and behaviour: A virtual reality study on food disgust. *Food Quality and Preference*, *80*, 103833. https://doi.org/10.1016/J.FOODQUAL.2019.103833

Aylett, R., & Louchart, S. (2003). Towards a Narrative Theory of Virtual Reality. *Virtual Reality*, *7*(1), 2–9. https://doi.org/10.1007/S10055-003-0114-9

Bailenson, J. (2019). *Experience on demand : what virtual reality is, how it works, and what it can do*. W. W. Norton & Company, Inc.

Baudrillard, J. (1981). Simulacra and Simulation. In *Simulacra and Simulation*. University of Michigan Press. https://doi.org/10.3998/MPUB.9904

Beisbart, C. (2019). Virtual Realism: Really Realism or only Virtually so? A Comment on D. J. Chalmers's Petrus Hispanus Lectures. *Disputatio*, *11*(55), 297–331. https://doi.org/10.2478/DISP-2019-0008

Bertrand, P., Guegan, J., Robieux, L., McCall, C. A., & Zenasni, F. (2018). Learning empathy through virtual reality: Multiple strategies for training empathy-related abilities using body ownership illusions in embodied virtual reality. In *Frontiers Robotics AI* (Vol. 5, Issue MAR, p. 26). Frontiers Media S.A. https://doi.org/10.3389/frobt.2018.00026

Besmer, K. M. (2015). What robotic re-embodiment reveals about virtual re-embodiment. In *Postphenomenological Investigations*. Lexington Books.

Biocca, F. (2006). The Cyborg's Dilemma: Progressive Embodiment in Virtual Environments [1]. *Journal of Computer-Mediated Communication*, *3*(2), 0–0. https://doi.org/10.1111/j.1083-6101.1997.tb00070.x

Blair, R. J. R., & Blair, K. S. (2013). Empathy, Morality, and Social Convention: Evidence from the Study of Psychopathy and Other Psychiatric Disorders. *The Social Neuroscience of Empathy*, 139–152. https://doi.org/10.7551/MITPRESS/9780262012973.003.0012

Bloom, P. (2016). *Against Empathy: The Case for Rational Compassion*. Ecco.

Bloom, P. (2017, February 3). It's Ridiculous to Use Virtual Reality to Empathize With Refugees . *The Atlantic*. https://www.theatlantic.com/technology/archive/2017/02/virtual-reality-wont-make-you-more-empathetic/515511/

Boccignone, G., Gadia, D., Maggiorini, D., Ripamonti, L. A., & Tosto, V. (2022). Wuthering heights: gauging fear at altitude in virtual reality. *Multimedia Tools and Applications*, 1–22. https://doi.org/10.1007/S11042-022-13366-1/FIGURES/5

Bollmer, G. (2017). Empathy machines. *Media International Australia*, *165*(1), 63–76. https://doi.org/10.1177/1329878X17726794

Bordwell, David. (1986). *Narration in the fiction film*. Routledge. https://www.routledge.com/Narration-in-the-Fiction-Film/Bordwell/p/book/9780415018777

Botvinick, M., & Cohen, J. (1998). Rubber hands 'feel' touch that eyes see. *Nature 1998 391:6669*, *391*(6669), 756–756. https://doi.org/10.1038/35784

Brey, P. (2003). The Social Ontology of Virtual Environments. *American Journal of Economics and Sociology*, *62*(1), 269–282. https://doi.org/10.1111/1536-7150.T01-1-00011

Brey, P. (2014). The physical and social reality of virtual worlds. In M. Grimshaw (Ed.), *The Oxford handbook of virtual reality* (pp. 42–54). Oxford University Press. https://doi.org/10.1093/OXFORDHB/9780199826162.013.029

Chalmers, D. J. (2016). The Matrix as Metaphysics. In *Science Fiction and Philosophy* (pp. 35–54). John Wiley & Sons, Inc. https://doi.org/10.1002/9781118922590.ch5

Chalmers, D. J. (2017). The Virtual and the Real. *Disputatio*, *9*(46), 309–352. https://doi.org/10.1515/DISP-2017-0009

Chalmers, D. J. (2019). The Virtual as the Digital. *Disputatio*, *11*(55), 453–486. https://doi.org/10.2478/DISP-2019-0022

Chalmers, D. J. (2022). *Reality+: Virtual Worlds and the Problems of Philosophy*. W. W. Norton.

Cogburn, J., & Silcox, M. (2014). Against Brain-in-a-Vatism: On the Value of Virtual Reality. *Philosophy and Technology*, *27*(4), 561–579. https://doi.org/10.1007/s13347-013-0137-4

Coplan, A. (2011). Will the real empathy please stand up? A case for a narrow conceptualization. *Southern Journal of Philosophy*, *49*(SUPPL. 1), 40–65. https://doi.org/10.1111/j.2041-6962.2011.00056.x

Coplan, A. (2012). Understanding Empathy: Its Features and Effects. In A. Coplan & P. Goldie (Eds.), *Empathy: Philosophical and Psychological Perspectives* (pp. 3–30). Oxford University Press.

Cuff, B. M. P., Brown, S. J., Taylor, L., & Howat, D. J. (2016). Empathy: A review of the concept. *Emotion Review*, *8*(2), 144–153. https://doi.org/10.1177/1754073914558466

Currie, G. (2020). *Imagining and Knowing: The Shape of Fiction*. Oxford University Press.

Davies, S. (2011). Infectious Music: Music-Listener Emotional Contagion. In A. Coplan & P. Goldie (Eds.), *Empathy Philosophical and Psychological Perspectives* (pp. 134–148). Oxford University Press .

Davies, S. (2014). Emotional contagion. In W. F. Thompson (Ed.), *Music in the Social and Behavioral Sciences: An Encyclopedia* (Vol. 1, pp. 376–380). Sage Reference.

Davis, M. H. (2006). Empathy. *Handbooks of Sociology and Social Research*, 443–466. https://doi.org/10.1007/978-0-387-30715-2_20/COVER

de la Peña, N., Weil, P., Llobera, J., Giannopoulos, E., Pomés, A., Spanlang, B., Friedman, D., Sanchez-Vives, M. v., & Slater, M. (2010). Immersive journalism: Immersive virtual reality for the first-person experience of news. *Presence: Teleoperators and Virtual Environments*, *19*(4), 291–301. https://doi.org/10.1162/PRES_a_00005

de Sousa, R. (1990). Emotions, Education and Time. In *Metaphilosophy* (Vol. 21, pp. 434–446). Wiley. https://doi.org/10.2307/24436885

de Vignemont, F., & Singer, T. (2006). The Empathic Brain: How, When and Why? *Trends in Cognitive Sciences*, *10*(10), 435–441. https://doi.org/10.1016/j.tics.2006.08.008

de Waal, F. (2009). *The Age of Empathy: Nature's Lessons for a Kinder Society*. Crown Publishing Group.

de Waal, F. (2021). Empathy, the Umbrella Term. *Neuroscience & Biobehavioral Reviews*, *129*, 180–181. https://doi.org/10.1016/J.NEUBIOREV.2021.07.034

Debes, R. (2009). Which empathy? Limitations in the mirrored "understanding" of emotion. *Synthese 2009 175:2*, *175*(2), 219–239. https://doi.org/10.1007/S11229-009-9499-7

Difede, J., & Hoffman, H. G. (2002). Virtual Reality Exposure Therapy for World Trade Center Post-traumatic Stress Disorder: A Case Report. *Cyberpsychology and Behaviour*, *5*(6), 529–535. https://doi.org/10.1089/109493102321018169

Ehrsson, H. H., Holmes, N. P., & Passingham, R. E. (2005). Touching a rubber hand: Feeling of body ownership is associated with activity in multisensory brain areas. *Journal of Neuroscience*, *25*(45), 10564–10573. https://doi.org/10.1523/JNEUROSCI.0800-05.2005

Ehrsson, H. H., Spence, C., & Passingham, R. E. (2004). That's My Hand! Activity in Premotor Cortex Reflects Feeling of Ownership of a Limb. *Science*, *305*(5685), 875–877. https://doi.org/10.1126/SCIENCE.1097011

Ekman, P. (1992). An Argument for Basic Emotions. *Cognition and Emotion*, *6*(3–4), 169–200. https://doi.org/10.1080/02699939208411068

Ekman, P. (2005). Basic Emotions. In *Handbook of Cognition and Emotion* (pp. 45–60). John Wiley & Sons, Ltd. https://doi.org/10.1002/0470013494.ch3

Fadeev, K. A., Smirnov, A. S., Zhigalova, O. P., Bazhina, P. S., Tumialis, A. v., & Golokhvast, K. S. (2020). Too Real to Be Virtual: Autonomic and EEG Responses to Extreme Stress Scenarios in Virtual Reality. *Behavioural Neurology*, *2020*. https://doi.org/10.1155/2020/5758038

Fagiano, M. (2016). Pluralistic conceptualizations of empathy. In *Journal of Speculative Philosophy* (Vol. 30, Issue 1, pp. 27–44). Penn State University Press. https://doi.org/10.5325/jspecphil.30.1.0027

Fagiano, M. (2019). Relational Empathy. *International Journal of Philosophical Studies*, *27*(2), 162–179. https://doi.org/10.1080/09672559.2019.1598086

Fisher, J. A. (2017). Empathic Actualities: Toward a Taxonomy of Empathy in Virtual Reality. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *10690 LNCS*, 233–244. https://doi.org/10.1007/978-3-319-71027-3_19

Friend, S. (2020). Fiction and Emotion: The Puzzle of Divergent Norms. *The British Journal of Aesthetics*, *60*(4), 403–418. https://doi.org/10.1093/AESTHJ/AYAA010

Friend, S. (2022). Emotion in Fiction: State of the Art. *The British Journal of Aesthetics*, *62*(2), 257–271. https://doi.org/10.1093/AESTHJ/AYAB060

Grant, T. (2019). *Towards an Analysis of Virtual Realism*. http://www.digra.org/wp-content/uploads/digital-library/DiGRA_2019_paper_24.pdf

Gromer, D., Reinke, M., Christner, I., & Pauli, P. (2019). Causal Interactive Links Between Presence and Fear in Virtual Reality Height Exposure. *Frontiers in Psychology*, *10*(JAN). https://doi.org/10.3389/FPSYG.2019.00141

Groom, V., Bailenson, J. N., & Nass, C. (2009). The influence of racial embodiment on racial bias in immersive virtual environments. *Social Influence*, *4*(3), 231–248. https://doi.org/10.1080/15534510802643750

Hacker, Peter. Michael. S. (2017). *The Passions: A Study of Human Nature*. Wiley-Blackwell.

Harris, M. S. (1929). Two Postulates of Expressionism. *The Journal of Philosophy*, *26*(8), 210. https://doi.org/10.2307/2014043

Heim, M. (1994). The Metaphysics of Virtual Reality. *The Metaphysics of Virtual Reality*, 1–200. https://doi.org/10.1093/ACPROF:OSO/9780195092585.001.0001

Heim, M. (1998). *Virtual Realism*. Oxford University Press, Inc.

Held, R. M., & Durlach, N. (1993). Telepresence, Time Delay and Adaptation. In *Spatial Displays and Spatial Instruments* (Proceeding).

Held, R. M., & Durlach, N. I. (1992). Telepresence. *Presence: Teleoperators and Virtual Environments*, *1*(1), 109–112. https://doi.org/10.1162/PRES.1992.1.1.109

Herrera, F., Bailenson, J., Weisz, E., Ogle, E., & Zaki, J. (2018). Building long-term empathy: A large-scale comparison of traditional and virtual reality perspective-taking. *PLOS ONE*, *13*(10).

Humbert-Droz, S., Garcia, A. L., Sennwald, V., Teroni, F., Deonna, J., Sander, D., & Cova, F. (2020). Lost in intensity: is there an empirical solution to the quasi-emotions debate? *Aesthetic Investigations*, *4*(1), 54–76.

Ihde, D. (1975). The Experience of Technology: Human-Machine Relations: *Cultural Hermeneutics*, *2*(3), 267–279. https://doi.org/10.1177/019145377500200304

Ihde, D. (1979). *Technics and Praxis: A Philosophy of Technology*. Springer Netherlands. https://doi.org/10.1007/978-94-009-9900-8_1

Ihde, D. (1990). *Technology and the lifeworld : from garden to earth*. Indiana University Press. https://books.google.com/books/about/Technology_and_the_Lifeworld.html?id=s3u1QgAAC AAJ

Ihde, D. (2002). *Bodies in technology*. University of Minnesota Press.

Jagneaux, D. (2017, January 23). *Resident Evil 7 VR Review - Bringing The Survival Back to Horror Games*. https://uploadvr.com/resident-evil-7-vr-review/

Jahoda, G. (2005). Theodor Lipps and the Shift from "Sympathy" to "Empathy." *Journal of the History of the Behavioral Sciences*, *41*(2), 151–163. https://doi.org/10.1002/JHBS.20080

James, W. (1884). What is an Emotion? In *Mind* (Vol. 9, pp. 188–205). Oxford University PressMind Association. https://doi.org/10.2307/2246769

Jamil, Z. (2014). Empathy: A motivated account. In *Psychological Bulletin* (Vol. 140, Issue 6, pp. 1608–1647). American Psychological Association Inc. https://doi.org/10.1037/a0037679

Jenkins, H. (2003). *Game Design as Narrative Architecture*.

Jerald, J. (2015). The VR Book. In *The VR Book*. Association for Computing Machinery. https://doi.org/10.1145/2792790

Juul, J. (2019). Virtual Reality: Fictional all the Way down (and that's OK). *Disputatio*, *11*(55), 333–343. https://doi.org/10.2478/DISP-2019-0010

Keen, S. (2006). A Theory of Narrative Empathy. *Narrative*, *14*(3), 207–236. https://doi.org/10.1353/nar.2006.0015

Kisker, J., Gruber, T., & Schöne, B. (2021). Behavioral realism and lifelike psychophysiological responses in virtual reality by the example of a height exposure. *Psychological Research*, *85*(1), 68–81. https://doi.org/10.1007/S00426-019-01244-9

Kisker, J., Lange, L., Flinkenflügel, K., Kaup, M., Labersweiler, N., Tetenborg, F., Ott, P., Gundler, C., Gruber, T., Osinsky, R., & Schöne, B. (2021a). Authentic Fear Responses in Virtual Reality: A Mobile EEG Study on Affective, Behavioral and Electrophysiological Correlates of Fear. *Frontiers in Virtual Reality*, *2*, 106. https://doi.org/10.3389/FRVIR.2021.716318/BIBTEX

Kitchin, R., & Dodge, M. (2001). 'There's no there there' : Virtual reality, space and geographic visualization. In *Virtual Reality in Geography* (pp. 353–373). CRC Press. https://doi.org/10.1201/9780203305850-30

Kvanvig, J. (2009). The Value of Understanding . In D. Pritchard, A. Haddock, & A. Millar (Eds.), *Epistemic Value* (pp. 95–112). Oxford University Press. https://philpapers.org/rec/KVATVO-5

Lamarque, P. (1981). How Can We Fear and Pity Fictions? *The British Journal of Aesthetics*, *21*(4), 291–304. https://doi.org/10.1093/BJAESTHETICS/21.4.291

Lanzoni, S. (2012). Empathy in Translation: Movement and Image in the Psychological Laboratory. *Science in Context*, *25*(3), 301–327. https://doi.org/10.1017/S0269889712000154

Lanzoni, S. (2018). Empathy: A History. In *Empathy: A History*. Yale University Press. https://doi.org/10.1177/00030651211035518

Lazarus, R. S. (1991). Progress on a cognitive-motivational-relational theory of emotion. *The American Psychologist*, *46*(8), 819–834. https://doi.org/10.1037//0003-066x.46.8.819

Lerner, J. S., Li, Y., Valdesolo, P., & Kassam, K. S. (2015). Emotion and Decision Making. *Annual Review of Psychology*, *66*, 799–823. https://doi.org/10.1146/ANNUREV-PSYCH-010213-115043

Lévy, P. (1998). *Becoming Virtual : Reality in the Digital Age* (R. Bononno, Ed.). Plenum Trade. https://catalogue.nla.gov.au/Record/1834644

Louchart, S., & Aylett, R. (2005). Managing a non-linear scenario - A narrative evolution. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *3805 LNCS*, 148–157. https://doi.org/10.1007/11590361_17/COVER

Ludlow, P. (2019). The Social Furniture of Virtual Worlds. *Disputatio*, *11*(55), 345–369. https://doi.org/10.2478/DISP-2019-0009

Macpherson, F. (2020). Is Virtual Reality Experience Veridical, Illusory or Hallucinatory? A Complex Answer Based on a New Theory of Illusion and Hallucination and the Nature of the Technology Used to Create Virtual Reality. *The University of Glasgow*. https://eprints.gla.ac.uk/226457/

Mania, K., & Chalmers, A. (2001). The effects of levels of immersion on memory and presence in virtual environments: A reality centered approach. *Cyberpsychology and Behavior*, *4*(2), 247–264. https://doi.org/10.1089/109493101300117938

Masto, M. (2015). Empathy and Its Role in Morality. *The Southern Journal of Philosophy*, *53*(1), 74–96. https://doi.org/10.1111/SJP.12097

Matravers, D. (2014). The (so-called) 'Paradox of Fiction.' *Fiction and Narrative*, 102–117. https://doi.org/10.1093/ACPROF:OSO/9780199647019.003.0008

McDaniel, K. (2016). Edith Stein. In *Ten Neglected Classics of Philosophy* (pp. 195–221). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199928903.003.0008

McDonnell, N., & Wildman, N. (2019). Virtual Reality: Digital or Fictional? *Disputatio*. https://doi.org/10.2478/disp-2019-0004

McRoberts, J. (2017). *Are we there yet? Media content and sense of presence in non-fiction virtual reality*. *12*(2), 101–118. https://doi.org/10.1080/17503280.2017.1344924

Merleau-ponty, M. (1962). *The Phenomenology of Perception, translated from French* (C. Smith, Ed.). Routledge.

Milk, C. (2015). *How virtual reality can create the ultimate empathy machine /*. TED Talk. https://www.ted.com/talks/chris_milk_how_virtual_reality_can_create_the_ultimate_empathy_machine?language=en

Mooradian, N. (2006). Virtual Reality, Ontology, and Value. *Metaphilosophy*, *37*(5), 673–690. https://doi.org/10.1111/j.1467-9973.2006.00460.x

Moran, R. (1994). The Expression of Feeling in Imagination. *The Philosophical Review*, *103*(1), 75. https://doi.org/10.2307/2185873

Morie, J. F. (2007). Performing in (Virtual) spaces: Embodiment and being in virtual environments. *International Journal of Performance Arts and Digital Media*, *3*(2–3), 123–138. https://doi.org/10.1386/PADM.3.2-3.123_1

Mosselaer, N. van de. (2018). How Can We Be Moved to Shoot Zombies? A Paradox of Fictional Emotions and Actions in Interactive Fiction. *Journal of Literary Theory*, *12*(2), 279–299. https://doi.org/10.1515/JLT-2018-0016

Neill, A. (1991). Fear, Fiction and Make-Believe. *The Journal of Aesthetics and Art Criticism*, *49*(1), 47. https://doi.org/10.2307/431648

Neill, A. (1995). Emotional Responses to Fiction: Reply to Radford. *The Journal of Aesthetics and Art Criticism*, *53*(1), 75. https://doi.org/10.2307/431739

Niforatos, E., Palma, A., Gluszny, R., Vourvopoulos, A., & Liarokapis, F. (2020). Would you do it?: Enacting Moral Dilemmas in Virtual Reality for Understanding Ethical Decision-Making. *Conference on Human Factors in Computing Systems - Proceedings*. https://doi.org/10.1145/3313831.3376788

Nozick, R. (1974). *Anarchy, State and Utopia*. Basic Books.

Nozick, R. (2000, February 10). The Pursuit of Happiness. *Forbes*.

Oculus. (2018). *Nefertari: Journey to Eternity on Oculus Rift | Oculus*. https://www.oculus.com/experiences/rift/1491802884282318/?locale=en_GB

Oh, C. S., Bailenson, J. N., & Welch, G. F. (2018). A Systematic Review of Social Presence: Definition, Antecedents, and Implications. *Frontiers in Robotics and AI*, *5*(OCT). https://doi.org/10.3389/frobt.2018.00114

Pak, J., Guirguis, A., Mirchandani, N., Cummings, S., & Maoz, U. (2019). The Trolley Problem in Virtual Reality. *Student Scholar Symposium Abstracts and Posters*. https://digitalcommons.chapman.edu/cusrd_abstracts

Pinotti, A. (2017). A Question of Character: Analogy and the Empathic Life of Things. *Empathy*, 245–269. https://doi.org/10.1057/978-1-137-51299-4_10

Preston, S. D., & de Waal, F. B. M. (2002). Empathy: Its ultimate and proximate bases. *Behavioral and Brain Sciences*, *25*(1), 1–20. https://doi.org/10.1017/S0140525X02000018

Prinz, J. J. (2004). *Gut Reactions: A Perceptual Theory of the Emotions*. Oxford University Press.

Prinz, J. J. (2011). Is Empathy Necessary for Morality? In A. Coplan & P. Goldie (Eds.), *Empathy: Philosophical and Psychological Perspectives*.

Radford, C., & Weston, M. (1975). How Can We Be Moved by the Fate of Anna Karenina? *Aristotelian Society Supplementary Volume*, *49*(1), 67–94. https://doi.org/10.1093/ARISTOTELIANSUPP/49.1.67

Ramirez, E. J., Elliott, M., & Milam, E. (2021). What it's like to be a _____: Why it's (often) unethical to use VR as an empathy nudging tool. *Ethics and Information Technology*, 527–542.

Ramirez, E. J., & LaBarge, S. (2018). Real Moral Problems in the Use of Virtual Reality. *Ethics and Information Technology*, 249–263. https://doi.org/10.1007/s10676-018-9473-5

Rizzolatti, G., & Craighero, L. (2004). The Mirror-Neuron System. *Annual Review of Neuroscience*, *27*, 169–192.

Robertson, A. (2017). VR was sold as an 'empathy machine' — but some artists are getting sick of it. *The Verge*. https://www.theverge.com/2017/5/3/15524404/tribeca-film-festival-2017-vr-empathy-machine-backlash

Rotter, J. B. (1960). Some implications of a social learning theory for the prediction of goal directed behavior from testing procedures. *Psychological Review*, *67*(5), 301–316. https://doi.org/10.1037/h0039601

Rueda, J., & Lara, F. (2020). Virtual Reality and Empathy Enhancement: Ethical Aspects. *Frontiers in Robotics and AI*, *7*, 160. https://doi.org/10.3389/FROBT.2020.506984/BIBTEX

Ryan, M.-L. (2001). *Narrative as Virtual Reality: Immersion and Interactivity in Literature and Electronic Media*. The Johns Hopkins University Press.

Ryan, M.-L. (2019). *Narrative in Virtual Reality? Anatomy of a Dream Reborn*. https://doi.org/10.5281/ZENODO.3515106

Scarantino, A., & de Sousa, R. (2011). Emotions . *The Stanford Encyclopedia of Philosophy* . https://plato.stanford.edu/archives/sum2021/entries/emotion

Scheler, M. (2017). *The Nature of Sympathy*. Routledge. https://doi.org/10.4324/9781315133348

Schloerb, D. W. (1995). A Quantitative Measure of Telepresence. *Presence: Teleoperators and Virtual Environments*, *4*(1), 64–80. https://doi.org/10.1162/pres.1995.4.1.64

Searle, J. R. (1995). The Construction of Social Reality. *Ethics*, *108*(1), 208–210. https://doi.org/10.1086/233794

Serres, Michel. (1994). *Atlas*. Flammarion.

Sheridan, T. B. (1992). Musings on Telepresence and Virtual Presence. *Presence: Teleoperators and Virtual Environments*, *1*(1), 120–126. https://doi.org/10.1162/PRES.1992.1.1.120

Shin, D., & Biocca, F. (2018). Exploring immersive experience in journalism. *New Media & Society*, *20*(8), 2800–2823. https://doi.org/10.1177/1461444817733133

Siegel, D., & Dray, S. (2019). The Map is not the Territory: Empathy in Design . *Interactions*, *26*(2), 82–85. https://doi.org/10.1145/3308647

Slater, M. (2003). *A Note on Presence Terminology*. www.cs.ucl.ac.uk/staff/m.slater

Slater, M. (2009). Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1535), 3549–3557. https://doi.org/10.1098/rstb.2009.0138

Slater, M. (2018). Immersion and the illusion of presence in virtual reality. *British Journal of Psychology*, *109*, 431–433. https://doi.org/10.1111/bjop.12305

Slater, M., & Sanchez-Vives, M. v. (2016). Enhancing our lives with immersive virtual reality. In *Frontiers Robotics AI* (Vol. 3, Issue DEC, p. 74). Frontiers Media S.A. https://doi.org/10.3389/frobt.2016.00074

Slater, M., Steed, A., McCarthy, J., & Maringelli, F. (1998). The influence of body movement on subjective presence in virtual environments. *Human Factors*, *40*(3), 469–477. https://doi.org/10.1518/001872098779591368

Slater, M., & Wilbur, S. (1997). A framework for immersive virtual environments (FIVE): Speculations on the role of presence in virtual environments. *Presence: Teleoperators and Virtual Environments*, *6*(6), 603–616. https://doi.org/10.1162/pres.1997.6.6.603

Slote, M. (2010). *Moral sentimentalism*. Oxford University Press.

Sobchack, V. (2010). Living a "Phantom Limb": On the Phenomenology of Bodily Integrity. *Body and Society*, *16*(3), 51–67. https://doi.org/10.1177/1357034X10373407

Sofronieva, D. (2018). The Epistemic and Moral Value of Empathy [University of Leeds]. In *Empathy*. https://doi.org/10.1057/978-1-137-51299-4_12

Solomon, R. C. (1973). Emotions and Choice. *The Review of Metaphysics*, *27*(1), 20–41.

Spaulding, S. (2017). Cognitive Empathy. In H. Maibom (Ed.), *The Routledge Handbook of Philosophy of Empathy* (1st Edition, pp. 13–21). Routledge. https://doi.org/10.4324/9781315282015-2

Stanovsky, D. (2008). Virtual Reality. In *The Blackwell Guide to the Philosophy of Computing and Information* (pp. 167–177). Blackwell Publishing Ltd. https://doi.org/10.1002/9780470757017.ch12

Stecker, R. (2011). Should we Still Care about the Paradox of Fiction? *British Journal of Aesthetics*, *51*(3), 295–308. https://doi.org/10.1093/AESTHJ/AYR019

Stein, E. (1989). *On the problem of empathy*. ICS Publications.

Steuer, J. (1992). Defining Virtual Reality: Dimensions Determining Telepresence. *Journal of Communication*, *42*(4), 73–93. https://doi.org/10.1111/j.1460-2466.1992.tb00812.x

Studt, E. (2021). Virtual reality documentaries and the illusion of presence. *Https://Doi.Org/10.1080/17503280.2021.1923147*, *15*(2), 175–185. https://doi.org/10.1080/17503280.2021.1923147

Stueber, R. K. (2006). *Rediscovering Empathy: Agency, Folk Psychology, and the Human Sciences*. The MIT Press.

Ta, V. P., & Ickes, W. (2017). Empathic Accuracy. *The Routledge Handbook of Philosophy of Empathy*, 353–363. https://doi.org/10.4324/9781315282015-32

Tavinor, G. (2022). Virtual Reality Media and Aesthetics. *The Oxford Handbook of Philosophy of Technology*, 416–433. https://doi.org/10.1093/OXFORDHB/9780190851187.013.25

Verbeek, P. P. (2015). Beyond Interaction: a short introduction to mediation theory. *Interactions*, *22*(3), 26–31. https://doi.org/10.1145/2751314

Vorauer, J. (2013). The Case For and Against Perspective-Taking. In *Advances in Experimental Social Psychology* (Vol. 48, pp. 59–115). Academic Press Inc. https://doi.org/10.1016/B978-0-12-407188-9.00002-8

Walton, K. L. (1978). Fearing Fictions. *The Journal of Philosophy*, *75*(1), 5. https://doi.org/10.2307/2025831

Walton, K. L. (1993). Mimesis as Make-Believe: On the Foundations of the Representational Arts. In *Mimesis as Make-Believe: On the Foundations of the Representational Arts* (Reprint). Harvard University Press.

Waterworth, J., & Riva, G. (2014). Feeling Present in the Physical World and in Computer-Mediated Environments. In *Feeling Present in the Physical World and in Computer-Mediated Environments*. Palgrave Macmillan UK. https://doi.org/10.1057/9781137431677

Weber, E. U. (2006). Experience-based and description-based perceptions of long-term risk: Why global warming does not scare us (yet). *Climatic Change*, *77*(1–2), 103–120. https://doi.org/10.1007/s10584-006-9060-3

Wildman, N., & McDonnell, N. (2020). The puzzle of virtual theft. *Analysis*, *80*(3), 493–499. https://doi.org/10.1093/ANALYS/ANAA005

Wispé, L. (1986). The Distinction Between Sympathy and Empathy. To Call Forth a Concept, A Word Is Needed. *Journal of Personality and Social Psychology*, *50*(2), 314–321. https://doi.org/10.1037/0022-3514.50.2.314

Yee, N., & Bailenson, J. (2007). The Proteus Effect: The Effect of Transformed Self-Representation on Behavior. *Human Communication Research*, *33*(3), 271–290. https://doi.org/10.1111/j.1468-2958.2007.00299.x

Young, G. (2010). Virtually real emotions and the paradox of fiction: Implications for the use of virtual environments in psychological research. *Philosophical Psychology* , *23*(1), 1–21. https://doi.org/10.1080/09515080903532274

Zahavi, D. (2014). *Self and Other: Exploring Subjectivity, Empathy, and Shame*. Oxford University Press.