Virtual Killing

Carl David Mildenberger - cd@mildenberger.ch

Abstract. Debates that revolve around the topic of morality and fiction rarely explicitly treat virtual worlds like, for example, *Second Life*. The reason for this disregard cannot be that all users of virtual worlds only do the right thing while online – for they sometimes even virtually kill each other. *Is* it wrong to kill other people in a virtual world? It depends. This essay analyzes on what it depends, why it is that killing people in a virtual world sometimes is wrong, and how different virtual killings are wrong in different ways. I argue that killing people online is wrong if it is an instance of deliberately and non-consensually evoking disagreeable emotions in others. Establishing this conclusion requires substantial conceptual work, as virtual worlds feature new kinds of fictional agency, particular emotional responses to fiction, and unique ways in which the fiction of the virtual world relates to the wrongness of the killing.

Keywords: fictional agency; killing; morality and fiction; virtual worlds; wrongdoing

1 Introduction

Every day, millions of people log into virtual worlds like *Second Life* or *World of Warcraft*. Virtual worlds are computer-created environments that visually mimic complex physical spaces, where people interact with each other and with virtual objects (Bainbridge 2007). The users do all kinds

of things online. They explore the environment, they trade objects, they make friends. And sometimes they kill each other. Is it wrong to kill other people in a virtual world? It depends. The purpose of this essay is to analyze on what it depends, why it is that killing people in a virtual world sometimes is wrong, and how different virtual killings are wrong in different ways.

I shall argue that killing people online is wrong, if it is an instance of deliberately and non-consensually evoking disagreeable emotions in others; which it sometimes is. This conclusion might seem trivially correct. But it is surprisingly hard to show that virtual killings indeed sometimes are wrong. This is because of the problems posed by (i) agency in fictional worlds, (ii) emotional response to fiction, and (iii) justifications of harmful actions in playful or fictional contexts.

Establishing this conclusion requires analyzing the new forms of agency and emotional responses to fiction virtual worlds feature. I shall proceed by first giving an example of a virtual killing (Section 2), before looking at the perpetrator's action and the victim's response in detail (Sections 3 and 4, respectively). In addition, we need to have a close look at how the fictional nature of the killing and considerations of consent might justify virtual killings. The challenge is to show how virtual killings may be wrong when we intuitively believe other fictional killings not to be wrong; say an author "killing" one of her characters in a novel. Section 5 examines the abundant justifications for evoking disagreeable emotions in others by means of a virtual killing.

At the end of Section 5 we will have established that it indeed can be wrong to kill people online – and we will have mapped the domain of wrongdoing in virtual worlds in the course of doing so. This allows us to show, in addition, how different virtual killings are wrong in different ways (Section 6). Virtual worlds feature unique ways in which the fiction of the virtual world relates to the wrongness of the killing. For some kinds of virtual killings, virtual worlds only mediate the wrongdoing, whereas in other cases the wrong is based on the fictional aspect of the virtual world, and in yet other cases the wrongdoing shatters the fiction.

Now why should we care about morality in virtual worlds? As recently as 20 years ago, virtual worlds basically did not exist. Nowadays, estimates are that more than 500 million people are regularly participating in them (Watters 2010). Thus, analyzing wrongdoing in virtual worlds updates our picture of the relationship between morality and fiction. As killing is a paradigm case of wrongdoing, analyzing virtual killing teaches us about morality in this increasingly important arena for social interaction more generally. It does not so much tell us about the morality *of* virtual worlds, i.e. for example about how the issue of imaginative resistance affects which moral codes may fictionally hold. It tells us about morality *in* virtual worlds, i.e. about what we ought and ought not do when we are logged in.

2 A virtual killing

Consider the following scenario.

> VIRTUAL KILLING: Sitting in front of her computer in Oxford, Gillian logs into a virtual world. Here she controls the actions of her avatar, called Gillie14. Using keyboard and mouse, Gillian makes Gillie14 walk around to explore the virtual world. After some time, Gillie14 bumps into SeanBean. SeanBean is an avatar controlled by Sean, who is sitting in front of his computer in Cambridge. Without much ado, Gillian right-clicks on SeanBean, which is interpreted by the software as the command for Gillie14 to attack SeanBean. Gillie14 draws a gun and shoots SeanBean, who drops dead.

This is an example of what I call a virtual killing. It features two real persons controlling their avatars. Gillian is not killing a *non-player character*, i.e. a fictional character whose actions are controlled by the software, but Sean's avatar.

When Gillian kills SeanBean in the virtual world she is not playing a role in the way an actress fictionally killing somebody on stage does. It is Gillian's own and free choice to kill SeanBean, as virtual worlds are not scripted. I also assume that the background story of the virtual world is not such that it holds that Gillian has to kill SeanBean in order to win the game. Virtual worlds,

unlike classic single-player computer games, rarely feature such clear, predetermined goals. Typically, users cannot "win" a virtual world at all. All they can do is to continuously participate.

3 The perpetrator's action

How it is even possible to kill somebody online? Without doubt, real actions performed by real persons can make things fictional.[1] If an actor on stage is holding a fake gun, making it touch the chest of another actor, and then pulls the trigger (with the fake gun emitting a loud bang), his real actions make it fictional that a fictional character has just been shot. Yet, the exact modes of agency according to which the real actions performed by real persons make things fictional differ among fictional worlds (Walton 1978, 13–16).

Velleman (2013) argues that virtual worlds allow for a special mode of agency: *virtual agency*. With respect to virtual worlds, he states, it is appropriate to say that real people can literally commit fictional actions (2013, 5–15). Two central ideas underlie Velleman's claim. First, Velleman highlights that it is *Gillian's* beliefs and desires which motivate Gillie14's behavior. Unlike children playing a game of make-believe, in virtual worlds "participants do not generally attribute attitudes to their avatars …; they simply have thoughts and feelings about the world of the game, and they act on that world through their avatars but under the motivational force of their own attitudes" (2013, 11). Say Gillian sees the virtual representation of a mountain on her screen. In this case it is Gillian's own curiosity to see what lies behind the virtual mountain which motivates her avatar's climbing it. Gillian is curious to explore the virtual world, not Gillie14.

Second, Velleman stresses how the intentions of a user at some point stop being directed at mastering keyboard and mouse, i.e. at the real world actions needed to make things fictional in a virtual world. Slowly but surely the user's intentions start to become intentions of making things with his avatar, or as his avatar, or simply intentions to perform fictional actions (2013, 12–15).

---

[1] Following Walton (1990, 35), I will say that something is *fictional* if it is true in a fictional world, and that something is *true* if it is true in the actual world.

For Velleman, avatars are a prosthetic extension of our body. They are tools, like a tennis racket is. Just like skilled tennis players do not intend to maneuver their racket so that it hits the ball over the net, but use the racket as a part of their body, so do skilled users use their avatar as if it were under their direct control. They start intending "to perform avatar-eye-coordinated actions in the virtual world, not real-world actions of controlling the avatar" (2013, 14).

Velleman concludes that, "[w]hen engaged in virtual play … a person really has a fictional body" (2013, 14). This is what virtual agency amounts to: real persons performing fictional actions by means of their fictional bodies. Velleman's account of virtual agency is convincing in that it is able to account for the curious fact that users of virtual worlds frequently attribute to themselves the actions of their avatars. Nobody says "My avatar went there and killed him". Everybody just says "I went there and killed him". But Velleman's account of how users virtually act (*v-act*), is not exhaustive with respect to how users make things fictional in virtual worlds. We can distinguish at least two more modes of agency that users frequently adopt: *playing* and *acting*. I will treat them in turn.

Many users do what Velleman (2013, 11) says they rarely do; namely to invent characters who possess their own beliefs and desires. That is, many users choose to *play* (or: *role-play*). They engage in a game of make-believe within the virtual world, creating an avatar that is visibly and psychologically distinct from themselves. As they engage in *belief-like* and *desire-like imagining* (cf. Velleman 2000), the avatar's imagined beliefs and desires can be very close to the real person's beliefs and desires, but they do not have to be. Users might be closer to creating a fictional self, or they might be creating a truly other.[2] When users (role-)play, they do not use their avatars as

---

[2] It is useful to distinguish between these two opposite ends of a continuous spectrum. We can do so by saying that Gillian is *playing* when she creates an avatar that is very close to her real self yet distinct, and that she is *role-playing* when her avatar is drastically different. But even if Gillian is role-playing, she is not playing a role in the same way an actress does. Gillian is not *enacting* a role. That is, she is not role-playing on the basis of the imagined beliefs and

tools. Instead, the avatar is an independent fictional character, whose imagined beliefs and desires are the motivational force underlying the fictional actions.

Acting is a yet again different mode of agency. The distinction between v-acting and acting is best understood with respect to the direction of fit of the desires motivating the fictional action. For both acting and v-acting, it is Gillian's desires which constitute the motivating force for the fictional action. However, while in the case of v-acting Gillian's desires have a direction of fit which is best described as being *virtual-world-to-mind*, in the case of acting, the direction of fit is best captured by saying that it is *real-world-to-mind*. Let me clarify.

A v-acting user wants to achieve something within the virtual world, e.g. to see what lies behind a virtual mountain. Similarly, a v-acting user's virtual-world-directed desire might be to kill a fictional character, say because he is preventing her from climbing the mountain. If Gillian desires to kill SeanBean, then she is v-acting, and her desire has a virtual-world-to-mind direction of fit. But Gillian might also desire to annoy Sean. That is, she might decide to kill SeanBean not because she wants to accomplish something in the virtual world, but because she wants to accomplish something in the real world; namely, to annoy Sean. In this latter case, it seems appropriate to say that Gillian's desire underlying the fictional action has a real-world-to-mind direction of fit. Gillian is acting.

If acting, Gillian is not only using her *avatar* as a tool. Rather, she is using the *entire virtual world* as a medium that allows her to satisfy a real-world-directed desire. Gillian does not care about what is fictional as long as the fictional action satisfies her real-world-directed desire. If the virtual killing is an instance of Gillian acting, then it is very close to things like "cyber-bullying", in that both actions simply use modern communication technology as a new means to satisfy all too well-known desires.

---

desires that somebody else (notably the author of the play) imagined. Gillian herself engages in belief- and desire-like imagination.

These are then the different modes of agency that Gillian might have adopted when she virtually killed SeanBean. I think there is no reason to say that there can be only one mode of agency with respect to making things fictional in virtual worlds.[3] Probably, much depends on the individual user. Some users will experience their avatars as tools enabling them within a virtual world but which have "no life of their own". They are primarily v-acting or acting. Others will draw a clear line between their actual self and their virtual self (play) or their actual self and the virtual other (role-play). These individual perceptions might also change over time or as regards different avatars of the same user.

Thus, I'm adopting a pluralist stance with respect to how it is possible for real perpetrators to perform virtual killings. Distinguishing different modes of agency will not matter for analyzing why virtual killings are sometimes wrong. But it becomes important with respect to showing that virtual killings differ in how the fiction relates to the wrongness of the deed.

4 The victim's response

Sean's most natural response to the virtual killing seems to be anger.[4] Because SeanBean was killed, Sean is likely do things like hitting the table or uttering some swearwords ("Gillie, you bastard!"). He will probably also show a physiological reaction. His blood pressure, heart rate, and level of adrenaline are likely to rise. Given that either Sean's fictional body, or his fictional

---

[3] The advent of virtual worlds puts pressure on Walton's claim that what I call (role-)playing is the only possible mode of agency with respect to fictional worlds (1990, 191–5). Walton's claim rests on the idea that a person has to exist in a given world in order to act in it, and that real people "usually" (1990, 194) do not exist in fictional worlds. But arguably, and following Velleman's analysis of v-acting, it is quite usual for real people to exist in virtual worlds in the form of a virtual body.

[4] I focus on Sean's response, leaving the consequences for SeanBean aside, as what it means to die varies greatly among virtual worlds. Note, however, that virtual death typically is not permanent, but the avatar is instantly reborn. This grants Sean the possibility to (role-)play SeanBean's reaction to his own death.

self, or some fictional character he holds dear have just fallen victim to an unprovoked killing, this seems to be a *prima facie* plausible reaction.

But what kind of anger are we dealing with here? How can Sean be genuinely angry about something that never really happened (cf. Radford 1975, 77)? Let me outline how, according to the two currently most widely endorsed theories of emotional response to fiction, *Pretend Theory* and *Thought Theory*, we can explain Sean's emotional response. I will then add a new way of accounting for Sean's reaction that is particularly relevant when dealing with virtual worlds.

Following Walton (1978), Sean does not experience real anger but *quasi-anger*. According to Pretend Theory, it is not literally true that Sean is angry about the virtual killing. This is because in order to be angry about the virtual killing, Sean would have to believe that the killing really happened. But this is not what Sean believes. Sean only believes that the killing fictionally happened. While the belief that somebody has been killed may evoke real anger, the belief that somebody fictionally has been killed evokes quasi-anger. Pretend theorists argue that Sean "plays along" with the fiction of the virtual world when logging into it. Because Sean plays along, he pretends to be angry when SeanBean is killed. His actions (yelling etc.) as well as his physiological reactions (higher blood pressure etc.) are consequences of his pretending to be angry.[5]

By contrast, Thought theorists like Carroll hold that what underlies Sean's *real* anger is Sean's "entertaining in thought" that SeanBean has been killed (1990, 74). Thought theorists deny the idea that belief is a necessary condition for a rational emotional response. We can react emotionally to thoughts we entertain or events we imagine. As Schneider puts it: "At the heart of the Thought Theory lies the view that, although our emotional responses to *actual* characters and events may require beliefs in their existence, there is no good reason to hold up this particular

---

[5] If Sean is playing, he experiences quasi-anger about his own fictional death. If he is role-playing, he experiences quasi-anger about the fictional death of some character he holds dear.

type of emotional response as the model for understanding emotional response *in general*" (2006), and emotional response to fiction in particular.

Here is the third way to account for Sean's emotional response. Sean might be angry because he believes that he himself, the actual Sean, has been provoked by Gillian. Quite generally speaking, the belief or thought which mediates our anger is the belief or thought that somebody provoked us, e.g. by threatening us or by violating a certain bodily or psychological boundary of ours (Thomson 1990, 250). Following Pretend Theory, we might say that the belief that aroused Sean's pretended anger is that SeanBean fictionally has been provoked. Following Thought Theory, we might say that the thought which aroused Sean's anger is the entertained thought that SeanBean has been provoked. But Sean might also be angry because he believes that he himself has been provoked. Sean might believe that he has been provoked in the real world and by Gillian, although by means of a fictional action.

This is because Sean is aware that Gillian might be *acting*. If Sean is forming the belief that Gillian provoked him, this is the natural counterpart to the possibility of this mode of agency for Gillian.[6] It is an understandable reaction on Sean's part to form the belief that it was not SeanBean who has the target of the provocation, but he himself. Just as Gillian might ignore the fiction – seeing the virtual world merely as a piece of communication technology allowing her to interact with distant people – so might Sean when he speculates about what the real person controlling her avatar-tool is thinking and feeling.

This kind of emotional response is not limited to virtual worlds. Say two children engage in a make-believe game of being pirates. The stronger child plays at torturing the weaker, by pressing her head under water for an exceedingly long time. In such a situation, the weaker child might

---

[6] Empirically speaking, *grief play* is a very prevalent phenomenon in virtual worlds. Grief play (or: *griefing*), is the practice of deliberately irritating and harassing other users. Griefers do not engage with the fiction of the virtual world, as all they are trying to do is to annoy others. Consider that around 25 per cent of customer support calls to the developers of virtual worlds are griefing related (Davies 2006).

well get angry and complain that the stronger is "not really playing". The weaker child might complain that the stronger only pretends to pretend to torture, when in fact there is a real desire to upset the weaker underlying his actions. That is, she might accuse him of only using the game of make-believe as a pretext and a medium for mistreating her.

These are then the ways in which Sean might emotionally respond to the virtual killing. Again I propose to take a pluralist stance that allows for differences among individuals. I think that some actual Seans out there experience *quasi-anger*, whereas others are really angry – either because they entertain in thought that SeanBean has been provoked (let us call this *thought-mediated anger*), or because they form the belief that they themselves have been provoked (*real-world-directed anger*). We do not have to opt for one theory over the other in order to proceed.[7]

5 Virtual killings: consent, fictionality, and wrongness

Let us turn to analyzing *why* and *when* it is wrong to kill people in a virtual world. Given that we now know precisely what a virtual killing amounts to, i.e. in which ways Gillian might perform the virtual killing and how Sean might react to it, one answer to the former question actually suggests itself. This answer is that it is wrong to virtually kill people because it is wrong to deliberately evoke disagreeable emotions in others. When Gillian in some way makes it fictional that SeanBean is killed, she thereby makes Sean feel some kind of anger he dislikes experiencing.[8] This is *why* it is wrong for Gillian to perform the virtual killing.

---

[7] I am aware that Pretend Theory and Thought Theory typically are conceived of as competing theories rather than complementary ones. But this is not the appropriate place to judge whether Pretend or Thought Theory is ultimately right. And as we shall see, for our purpose of analyzing on what it depends, and why it is that killing people in a virtual world sometimes is wrong, it actually makes no difference which theory is right and whether there only is one correct theory. The result is the same for all suggested theories.

[8] Note that this idea of why it is wrong to virtually kill people neither depends on Gillian's precise mode of agency, nor on the exact type of Sean's emotional response.

As I said at the very beginning, this conclusion on its own almost seems trivially correct. Surely, if Gillian deliberately evokes disagreeable emotions in Sean this is *pro tanto* wrong. But only on the basis of our deepened understanding of how users can evoke disagreeable emotions in others by means of a virtual killing can we see how this is what holds true for the case of virtual killings. We now see how this thought naturally applies to and accounts for the wrongness of virtual killings.

There certainly are other possible ways to argue for why it is wrong to kill people in virtual worlds. Consider, for example, the alternative proposal that virtual killings are wrong because they often feature negative material consequences. In the course of many virtual killings virtual objects possessing real world value are destroyed; and virtual killings might be wrong for this reason. Also, if continuously engaging in virtual killings has negative effects on our moral character (e.g. that it increases real-life aggression), this might be a very important reason why virtual killings are wrong.

But I think an argument in terms of evoking disagreeable emotions is the approach that best captures what is *idiosyncratically* wrong with virtual killings. If one were asked what specifically differentiates virtual killings from other wrongdoings, I think it would be most appropriate to highlight how virtual killings actualize our well-known capacity to emotionally respond to fiction (rather than to highlight potential material consequences). We often have the impression that fictional worlds are metaphysically isolated from the actual world, but psychologically connected (Walton 1990, 191). This is precisely why focusing on emotional responses captures the idiosyncrasy of virtual killings. Also, I think that my focus on disagreeable emotions is the *philosophically most interesting* one. I would be happy to admit that if virtual killings negatively affect our character, they also are wrong for this reason – just like other activities having such effects are. But whether virtual killings have such an effect is an empirical question.[9]

---

[9] Those interested in it should see Anderson and Bushman (2001) for a survey.

I take it as given that, generally speaking, it is wrong to deliberately evoke disagreeable emotions in others.[10] But it also seems like this is not always the case. There are considerations that can justify or excuse such behavior; considerations that make the experienced emotional harm a non-wrongful one. This, then, brings us to the question of *when* it is wrong to virtually kill others, i.e. on what it depends whether a specific virtual killing is wrong. The rest of this section is devoted to this question.

In the particular case of deliberately evoking disagreeable emotions in others *by means of a virtual killing,* there are two salient considerations for why it might not be wrong to do so. Consent and fictionality. One might think that, in the case of virtual killings, it is not wrong to evoke disagreeable emotions in others, because the victim has somehow consented to the experiencing them. Similarly, one might think that the fictionality of virtual killings suffices to found an exception. Gillian evokes disagreeable emotions by means of a fictional action. Is not it in some sense "irrational" on the victim's part to feel anger because of a fictional event – and thus not wrong for Gillian to evoke such emotions?

I will address both considerations starting with the former. The goal is to show that consent and fictionality bar us from concluding that all virtual killings are wrongful. But even in light of these important justifying considerations we will see that at least sometimes it *is* wrong to kill others in virtual worlds.

5.1 Consent

It seems as if some form of the maxim *volenti non fit iniuria* governs virtual world interactions. The maxim holds that if someone consents to a certain form of conduct, then the consequences of that conduct, even if disagreeable, are not such that they wrong the consenter. The crucial question in dealing with this potentially justifying consideration is: what kind of conduct has the

---

[10] For a good impression of which issues are at stake when making such a claim, see for instance Thomson (1990, 249–57) and Ellis's (1995) reply to Thomson.

victim consented to? There are two ways of defining the consented-to conduct which strike me as natural or promising.

First, one might say that the victim has consented to *playing the game* the virtual world allows to play. The idea is that because the users are essentially playing a game, and because play is voluntary, the emotions evoked while playing are consented to. I do not think that this reasoning withstands scrutiny. Even though it is a classic idea that all play is voluntary (e.g. Huizinga 1949, 7–8; Caillois 1961, 6), it does not seem to be true. Juul is right to ask the rhetorical question of whether something is not play "if social pressure forces the player to play" (2005, 31).

Besides these definitional issues, the usual caveats of evoking the *Volenti* maxim loom large when we conceive of the consented-to conduct this broadly. That is, we can doubt the genuineness, validity, and extent of the victim's consent. When deciding to play the game, there probably were many aspects of playing the victim did not foresee. Would Sean still have decided to create SeanBean if he had been aware that SeanBean will be killed? Or, if he was aware, what effects the virtual death might have on him? Such doubts seem sufficient to argue that Sean might not have genuinely consented to experience disagreeable emotions upon the death of SeanBean. It is very likely that at least some victims of virtual killings are genuinely surprised about the disagreeable emotions they experience because their avatar has been killed.[11] Thus, if we are defining the consented-to conduct in this broad way, there still will be some non-consensual (and thus wrongful) virtual killings around.

Here is a second way. One might argue that the consented-to conduct is the *competition among users* within virtual worlds. This seems to be promising. First, since we know at least since Mill's *On Liberty* (1991, chap. V.3) that we should evaluate differently those harms that arise in a competitive environment. Competitive harms, e.g. the emotional harm the frustrated loser incurs, are not wrongful harms. This strategy furthermore seems to be promising because the

---

[11] Compare Dibbell (1999), who describes how a virtual world user was traumatized by her avatar being virtually raped.

consented-to conduct is much narrower in nature. While users can hardly overlook all the potential experiences they might undergo when logging into a virtual world, it is easier to have a clear understanding of the rules of competition for a certain virtual world. Users can be expected to know that in virtual world V, the rules explicitly allow for killing as a means of competing.[12] Consequently, consent to competition is more likely to be fully informed and thus valid.

It is true that because of these considerations many instances of virtual killings are not wrongful. Killing is an officially allowed means of competing in many virtual worlds. Thus, the disagreeable emotions experienced when being killed often can be considered a competitive harm. But wrongful virtual killings still happen. The reason for this is, quite simply, that not all virtual killings take place in a consented-to competitive context. Users griefing others, e.g. by killing them over and over again "just for fun", is the most salient example of such non-competitive virtual killings. And whereas the rules of V might allow for competitive killings, they may proscribe non-competitive ones. In fact, because griefing is such a prevalent phenomenon, there typically is an anti-griefing clause in the official *Terms of Use* all users have to agree to before creating an account.[13]

Could we maybe say that all interactions in virtual worlds are inherently competitive in nature (and thus even are competitive if fictionally they are not)? No. There are many virtual worlds which are not inherently competitive in nature, but in which it is still possible to kill other users. The lack of competitiveness in many virtual worlds even is one characteristic that separates them

---

[12] What we see here is the influence of the official rules of a virtual world. Quite plausibly, our judgment as to whether virtual killings are wrong is affected by the rules that govern the virtual world. Whether it also is affected by the potentially deviating fictional moral rules directly leads us to the puzzle of imaginative resistance, which I will no go into here.

[13] It would be wrong to hold that, since griefing is widespread, every user at least tacitly consents to being griefed. Quite the opposite. Far from consenting to being griefed when participating in virtual worlds, every user explicitly accepts not to grief fellow users.

from many more traditional video games. In virtual worlds, we typically are not told what to do. We are just presented with a virtual world that allows us to do whatever we fancy.

To summarize, consent will make sure that some instances of virtual killings are not wrongful even if they evoke disagreeable emotions. But considerations of consent are not sufficient to make all virtual killings non-wrongful instances of evoking disagreeable emotions.

5.2 Fictionality

Let me now consider two arguments as to why considerations of fictionality might be able to turn many instances of virtual killings into non-wrongful cases of evoking disagreeable emotions. The first relies on the idea that, because a virtual killing is a fictional event, it is in some sense irrational to emotionally respond to it in the way Sean does. Because Sean's response is irrational, so the proposed argument goes, Gillian is not wronging him in evoking it.

Here is one sense in which one might say that Sean's anger is irrational. It follows Thomson (1990, 253–4). Thomson holds that anger is irrational, first, if the belief which mediates our anger is false. If I get angry because I believe that you are staring at me, but in fact you are just day-dreaming, then my anger is irrational. Second, my responding with anger to a certain belief or thought might be inappropriate. For example, my anger might be excessive. Or, rather than becoming angry, I ought to experience sorrow, or no emotion at all. On the basis of these thoughts, it seems plausible to hold, as Thomson does, that it is not necessarily wrong to evoke *irrational* anger in others. Consider an example Thomson gives.

> You sneeze, and I feel a mix of horror, fear, and rage. How so? I think you are sneezing *at* me, I think you have cancer and are trying to spread your cancer germs to me. (Thomson 1990, 253)

Intuitively Thomson is right to say that this act of sneezing is not an instance of wrongdoing, because it only evokes irrational anger.[14] Consequently, one might wonder whether, in the case of virtual killings, Sean holds false beliefs, or shows inappropriate emotional reactions to his beliefs or thoughts. This might make his anger irrational and the virtual killing non-wrongful.

Here is a second sense in which one might say that Sean's anger is irrational. It follows Radford (1975, 75). For Radford, an emotional response to fiction is irrational because it is inconsistent. If Sean endorses the idea that we can only be angry about events that we believe actually happened, and if he is not deceived into believing that the virtual killing actually happened, then Sean cannot be both angry and consistent. This is Radford's famous paradox of emotional response to fiction.

Now, does Sean's emotional response rest on an inconsistent set of beliefs, or on a false belief, or is somehow inappropriate – and thus irrational in either a Radfordian or Thomsonian sense? It is easy to see that Sean's anger does not rest on an inconsistent set of beliefs. This is because, on the one hand, Pretend Theory and Thought Theory are developed precisely in order to counter Radford's allegation that responding emotionally to fiction is irrational. If Sean is quasi-angry he is not inconsistent, because he does not experience real anger but quasi-anger. If Sean's anger is thought-mediated, he is not inconsistent because he (tacitly) denies Radford's first premise. On the other hand, if Sean's anger is real-world-directed, Radford's worry of inconsistency does not even arise. For what he is angry about is Gillian's perceived provocation by means of a fictional action, rather than about the fictional event as such.

---

[14] I do not think that is never wrong to evoke this kind of irrational emotions. Consider this example. If you suffer from arachnophobia, and I – being aware of this – place a tiny spider on the table next to you and then point it out to you, I have evoked irrational fear in you. Your belief that the spider presents a threat to you is false. But I take it that, unlike in Thomson's sneezing case, it is plausible that I wronged you by placing the spider. Still, we can assume for the sake of the argument that it is never wrong to evoke irrational anger. Doing so will only make our goal of arguing that some virtual killings are wrong harder.

Also, Sean's reaction does not seem to be inappropriate in Thomson's sense. Yes, Sean's anger *can* be excessive. For example if it makes Sean track down Gillian in real life and punch her.[15] But there is no reason to think that Sean's experiencing anger would necessarily be excessive or misguided. Instead, it seems to be an entirely natural reaction to your fictional body being harmed – or your fictional-but-distinct self being killed, or else some fictional character you are attached to. Even more so, if we compare it to our no-less-natural emotional reactions to past events (things that happened to us) or modal events (things that could have happened to us) (Moran 1994). In all of these cases, one might say that it is mysterious that we are moved. But one could hardly say that these reactions are inappropriate in Thomson's sense. Consider this example of what Thomson deems an inappropriate response. "You draw my attention to the fact that there are three bits of lint on my carpet, and I therefore feel profound grief. How so? I cannot bear that the condition of my carpet be anything less than perfection." (1990, 254) In contrast to this example, to react in the way Sean does just is how we naturally seem to function.

Finally, it is not necessarily the case that a false belief underlies Sean's anger. If Sean is experiencing quasi-anger or thought-mediated anger, it is possible but rather unlikely that he falsely believes ("falsely" entertains in thought?) that the virtual killing happened. By contrast, in the case of real-world-directed anger, the chances that Sean's anger is mediated by a false belief seem to be higher. Sean cannot know whether what underlies the virtual killing is Gillian's desire to annoy him, the actual Sean. Sometimes Sean will be right. In other cases in which Sean suspects that Gillian just wants to annoy him, she might actually be role-playing rather than acting.[16] But even if Sean all too often suspects that everybody just wants to get to him, the odds are that sometimes Sean's real-world-directed anger will be appropriate. In these cases, the potential justification that Sean's response to the fictional event is irrational because being

---

[15] Compare a case in which a virtual sword theft resulted in real-life retribution (Krotoski 2005).

[16] Often there are situational clues Sean can base his judgment on (the exact situation in which the killing happens, text-based communication between perpetrator and victim, etc.).

mediated by a false belief fails. To conclude, anger in response to virtual killings often is rational, according to different senses of rational.

Let us turn to the second argument relying on considerations of fictionality to argue that virtual killings are non-wrongful instances of evoking disagreeable emotions. It goes like this. Intuitively speaking, fictional killings do not seem to be wrong. If an author makes it fictional that his widely-beloved main character dies, or if a villain kills a heroine on stage, we might get angry just like Sean. But it seems implausible to suggest that we have been wronged. Thus, in arguing that virtual killings (i.e. a sub-class of fictional killings) are wrong, we seemingly are running the risk of committing ourselves to a counterintuitive conclusion. If we want to say that virtual killings are sometimes wrong, does not this imply that authors or actors, just as children playing make-believe games, are also frequently wronging their audiences or playmates? Am I suggesting that there are many unnoticed and unpunished wrongs around?

When replying to this argument, it is crucial to emphasize that I am only claiming that *some* virtual killings are wrong. This does not commit us to the (indeed implausible) claim that *all* fictional killings are wrong. What it does imply is that some fictional killings are wrong. First, because virtual killings are a sub-class of fictional killings. Second, because it seems highly likely that if some virtual killings are wrong, some other non-virtual fictional killings will also be wrong for the exact same reasons we are offering in the case of virtual killings. We can then expect that at least some largely unnoticed wrongs are around.

Now it does not strike me as counterintuitive to hold that some non-virtual fictional killings are wrong for the same reasons that some virtual killings are. Consider the following case.

> FICTIONAL KILLING: A father and his daughter are playing with two dolls. One doll, Barbara, has fictionally invited the other doll, Ann, to have tea. After some time, the father who is playing Ann says: "Barbara, there is something I have to tell you. Actually, I'm not Ann. I am an evil alien from outer space who is wearing an Ann-costume. I made the costume by skinning Ann alive earlier today. And

what I'll do next is to kill *you*, Barbara." Then the father laughs a vile laugh and uses his doll to rip off the other doll's head. The daughter screams and runs away frightened.

Did the father wrong his daughter when deliberately evoking disagreeable emotions in her by means of a fictional action? I think it is obvious that he did. This seems to be a case in which the daughter was not fully informed about what kind of emotions the game of make-believe could arouse in her, and thus did not genuinely consent to experiencing disagreeable (quasi-)fear.

I think that the intuition that all members of the family of non-virtual fictional killings are non-wrongful simply is an illusion. Probably, what upholds this illusion again is the prominent role consent plays in many engagements with fictional worlds. If somebody consents to watching a zombie movie, then I think it is typically fair to say that neither the producer of the movie nor the actors wronged him, even if he cannot sleep for days (and if he quite severely dislikes this experience).

Furthermore, it might be true that virtual worlds are more prone than other fictional worlds to highlight the existence of non-consensual fictional killings. And that because virtual worlds only recently emerged as an important cultural phenomenon, our intuitions about the non-wrongness of fictional killings are the way they are. What *is* true is that there are fictional worlds which by their very nature do not permit the non-consensual evoking of disagreeable emotions. Many fictional worlds of children playing games of make-believe are such. Consider an example.

If two children play a game of make-believe, the fictional world in which they interact is a very open one. The children are free in their choice of props that help to imagine the fictional world (Walton 1990, 37–8). In a game of cops and robbers, they may stipulate that all sticks are guns. Also, they typically have equal rights with respect to designing the fictional world. And given that such playing is a collaborative activity, they have to agree on the principles of generation (like "all sticks are guns") for the fictional world (Walton 1990, 40–1). Now, in an analogous way to jointly deciding on props and principles of generation, they have to jointly

agree on the ways in which their real actions affect the fictional world. Say one of the children grabs a stick, holds it like a gun, makes it touch the other kid's chest, and says "Bang!". In this example, the children actually have to agree that holding a stick to the other's chest and saying "Bang!" means to shoot the other. Velleman highlights that, unlike in virtual worlds in which specific actions have predefined meanings (e.g. a right-click on another avatar means that my avatar shall attack it), the fictional worlds of children's make-believe games are not as determinate and "recalcitrant" (2013, 9).

> In the typical make-believe game of pirates, if one player pretends to stab another, there is no fact as to how much damage has been done until one of them makes the requisite stipulation or takes a relevant action, such as pretending to die. The difference between a graze and a fatal wound is not determined by the physical enactment of the blow. (Velleman 2013, 9)

That is to say, if the children start to argue over whether someone has been killed in the fictional shooting, there is no objective way to decide. Often it is not even objectively determined which kinds of bodily movements exactly represent which fictional actions.

This is why there can hardly be any kind of non-consensual evoking of disagreeable emotions in such worlds. The child who was "shot" could simply decide that, in the fictional world, he wore a bulletproof vest, if he does not consent to being killed. There is very little that the fictionally shooting child can do in order to fictionally kill the other against his will – if it is at all possible. Thus, even if children's games of make-believe evoke disagreeable emotions in one of the players, the odds are that these emotions are consented-to.

To summarize, the intuition that fictional killings in general are not wrongful seems to be grounded in how more traditional fictional worlds function and, again, in considerations of

consent.[17] In any case an argument that virtual killings cannot be wrongful in light of this intuition fails. First, because virtual worlds represent a new kind of fictional world and work in their own way. Second, because some non-virtual fictional killings indeed are wrongful for the exact same reason we have suggested for virtual killings; they deliberately and non-consensually evoke disagreeable emotions. Whether a certain virtual killing is wrong depends. But even when we take into account consent and fictionality as the two most salient considerations as to why virtual killings might be justified cases of evoking disagreeable emotions, we find that sometimes it is wrong to virtually kill others.

6 Kinds of virtual killings

Across different modes of agency and different types of emotional responses, the reason why virtual killings are wrong stays the same. The perpetrator can make the killing fictional in different ways, but what is crucial is that she deliberately makes it fictional. Similarly, the victim can emotionally respond to the killing in different ways, but what is important is that the emotion is disagreeable and non-consensually experienced.

This does not mean that the different modes of agency and types of emotional response we have carved out do not matter with respect to our topic of virtual killing. Consider three distinct

---

[17] Here is another route that one might want to take in order to argue that, while virtual killings are wrong, other kinds of fictional killings in general are not. One might emphasize that when Gillian is making it fictional that SeanBean is killed, she is either acting, v-acting, or (role-)playing. That is to say, Gillian is not making things fictional by *fictionalizing* (this is what authors of fictional worlds do), or by *enacting* a role (this is what actresses do). By thoroughly working out the differences between these modes of agency, it is probably possible to defend a claim of the following kind. While it is wrong to deliberately evoke disagreeable emotions by means of acting, v-acting, or (role-)playing, it is not wrong to do so by means of fictionalizing or enacting. Given that the two latter modes of agency are the dominant ones for books, pictures, movies, and plays, a difference in the typical modes of agency between virtual worlds and more traditional fictional worlds might be able to explain the clash of intuitions about fictional killings not being wrong while virtual killings are.

cases of virtual killings to see that differences in modes of agency and types of emotional response lead to differences in how the fiction of the virtual world relates to the wrongness of the killing.

> TRAITOR: In virtual world *W*, Sean is role-playing and so is Gillian. As their joint story has it, SeanBean and Gillie14 are inseparable friends who together fight the forces of evil in *W*. But suddenly, Gillie14 draws a gun and shoots SeanBean. Directly after the incident, Gillie14 says: "It was all a lie, SeanBean. I've always been a servant of the dark side. You're a fool never to have noticed." Sean is startled as this turn of events was not agreed on. Then he gets angry (an instance of quasi-anger) and makes SeanBean shout: "Gillie14, you filthy traitor".

> GRIEFER: In virtual world *W*, Sean is role-playing his avatar SeanBean. Suddenly, Gillie14 approaches and, without further ado, shoots SeanBean – despite there being no virtual-world-related reason to do so. After the killing, Gillian (who is acting with the desire to annoy other users) makes Gillie14 say: "You're a loser and you suck at playing this game!". Sean gets angry (an instance of real-world-directed anger) and makes SeanBean say "People like you destroy this game".

> REALISTS: Both Gillian and Sean are acting with respect to virtual world *W*. Gillian is acting based on the desire to annoy other users, Sean is looking for new friends online in order to meet them offline. When Gillie14 bumps into SeanBean, she draws a gun and kills him without further ado. Gillian makes Gillie14 say: "You're a loser and you suck at playing this game!". Sean gets angry (an instance of real-world-directed anger) and types "You're such a failure".

Intuitively, these three virtual killings are quite different.[18] I think the best way to capture their differences is to say that they differ in how the fictional aspect of the virtual world relates to the

---

[18] The three cases are intended to exemplify three important kinds of virtual killings without implying that these are the only ones. One might feel like some of these killings are worse than others. For example, one might hold that since annoying people for no other purpose than to annoy them is a particularly bad thing to do, GRIEFER features the most serious wrongdoing. But my focus here shall be on differences in kind rather than on differences in degree.

wrongness of the killing.[19] In REALISTS, the virtual world only *mediates* the wrongdoing, whereas in TRAITOR the wrong is *based* on the fictional aspect of the virtual world, and in GRIEFER the virtual world's fictional aspect is *shattered* by the wrongdoing. Let me explain this terminology.

In REALISTS, two acting users clash, and the victim experiences real-world-directed anger. For both users W is only a medium through which they try to satisfy real-world-directed desires. This makes the virtual killing rather similar to Gillian posting "You're a loser" on Sean's social network page. Because both users are acting, to neither of them it is of importance what the virtual killing means within W, how it is represented, or how it relates to the fictional background story. Neither of them cares about W as such. W is but an interchangeable medium for their interaction.

In TRAITOR, two role-players interact, and the victim experiences quasi-anger. The fictional truth that SeanBean has been killed matters to both Gillian and Sean. This is because Gillian and Sean are engaging in an act of joint storytelling in the framework provided by W. Sean is trying to create a certain kind of story with his avatar as the protagonist and other avatars (including Gillian's) as supporting characters. Gillian is trying to create a certain kind of story with her avatar as the protagonist and other avatars (including Sean's) as supporting characters. The virtual killing is based on the fictional aspect of W and inseparably tied to it. W is not an interchangeable medium of interaction. Even as regards his emotional response to the virtual killing, Sean keeps playing along with the fiction of W, experiencing quasi-anger.

In GRIEFER, Gillian (who is acting) kills Sean's character (with Sean role-playing up to this point). As Sean perceives the killing to be an act of griefing, he responds with real-world-directed

---

[19] All those virtual worlds I am interested in are at the same time fictional worlds. But it seems possible, at least for analytical purposes, to separate the computer-mediated communication aspect of virtual worlds (that largely has to do with hardware and software aspects) from their fictional aspects.

anger.[20] Unlike in REALISTS in which the fiction of W is mutually bypassed, and unlike in TRAITOR in which the virtual killing is directly based on W's fictional aspect, in GRIEFER one user pulls another out of the fictional world. Because of Gillian's manifest acting, Sean drops his role-playing and reacts out of character by explicitly acknowledging that his playing experience has been ruined. Gillian is, in Huizinga's terms, a spoil-sport who shatters the fictional world and robs others of its illusion by "revealing the relativity and fragility of the play-world in which [s]he had temporarily shut himself with others" (1949, 11). In the aftermath of the killing, W is (temporarily) reduced to nothing but computer-mediated communication for both users.

If we distinguish different kinds of virtual killings in this way, we are able to highlight what is special about some of them, and how virtual killings are unlike the fictional killings we know. REALISTS, for example, neither is an uncommon scenario in virtual worlds, nor is it limited to them. Using some mediating entity to indirectly provoke somebody instead of, say, outright provoking him with physical threats is an age-old technique. But GRIEFER to some extent is an idiosyncratic scenario closely tied to the kind of interactive fiction virtual worlds feature.

Consider that participating in virtual worlds is not only about escaping to a fictional world, but for many users about joint storytelling. In GRIEFER, Gillian is not just killing a character in a fictional world whom Sean cares about. Because Gillian pulls Sean out of the fictional world, she also puts an end to his contribution to the joint storytelling. At least momentarily, SeanBean's story does not continue, as Sean acts out of character.

---

[20] Generally speaking, there is no definite causal connection between the mode of agency of the perpetrator and the type of emotional response of the victim. That is to say that it is not the case that every time a perpetrator *acts* the victim will respond with experiencing *real-world-directed anger*. For example, the victim might misperceive the perpetrator's acting as an instance of role-playing. Even if the victim realizes that the perpetrator is acting, it seems possible to take no particular offense at such acting and to respond with quasi-anger to an instance of griefing. In turn, the perpetrator can never be sure what type of emotional response his actions will evoke. This also means that we can only judge *ex post* what kind of virtual killing we are dealing with in dependence on the modes of agency and type of emotional response actually present.

On the one hand, GRIEFER describes a phenomenon that is not exclusive to virtual worlds. As mentioned above, two children engaging in a make-believe game of being pirates could face the same scenario, where the victim is pulled out of the fictional world and complains that the perpetrator is "not really playing". But on the other hand, virtual worlds seem to be particularly susceptible to attract people willing to shatter the fictional world from within. First, because they are largely anonymous environments. This means that the negative consequences of spoiling other people's joint storytelling are greatly reduced. Second, all a would-be griefer needs to do in order to become a genuine part of the fictional world is to create an account – which everyone can do at will. By contrast, as far as movies and novels are concerned, it is basically impossible for outsiders to gain access to the fictional world in a way so as to be able to shatter it for others. Similarly, one rarely hears of people joining theater companies in order to then act out of character as soon as they are on stage.

The ease with which griefers enter and linger in virtual worlds is important because potentially GRIEFER is not only an instance of joint storytelling gone wrong, but poses a threat to the entire fictional aspect of the virtual world. Somebody pulling people out of a fictional world *from within* typically seems to affect the fictional world in more drastic ways than outside interruption. If the children playing at being pirates are interrupted by their parents, the fictional world simply pauses. But the griefer from within might enduringly damage the fictional aspect of a virtual world, potentially leading to its overall demise.[21]

TRAITOR is another instance of joint storytelling gone wrong, but in a different way – and it is even more idiosyncratic to virtual worlds. Unlike in GRIEFER, it is not an uninvolved party which puts an end to the joint storytelling. Instead, it is the dedicated role-players themselves who do not agree on the continuation of their joint story. The story continues, but in a way disagreeable for some. For Sean the virtual killing affects the story he was trying to tell with his avatar – which now ends in a very different way from what he was planning (and is not merely

---

[21] Empirically speaking, griefers in the virtual world of *Ultima Online* had this effect (Pham 2002).

interrupted as in GRIEFER). This way in which a virtual killing may be wrong is not to be found in the fictional killings that happen, say, in novels.

A passively receiving reader cannot be wronged by a fictionalizing author in the same way that Sean is wronged by Gillian. What is unique about many virtual worlds is that they break down the sharp distinction between author and reader. They allow for specific kinds of fictional wrongdoing at the same time they allow for specific modes of agency. The role-players in TRAITOR are jointly creating a story (author aspect) but also are responding to pre-defined aspects of W, as well as to events in it brought about by other users over which they have no control (reader aspect). This is the special mix virtual worlds provide. Although the role-players have considerable freedom in telling the stories they want in W, W also features non-negotiable, recalcitrant aspects. Also, every story told is less easily malleable in that its non-negotiable parts are immediately visible to an entire community of users, and possibly affect their actions. Thus, TRAITOR is a kind of virtual killing that highlights in which way a globally successful effort of joint storytelling may wrong some individual users. Namely, the individual storytellers might disagree with respect to how the story is supposed to continue – and the actually occurring plot twist might leave angry users behind.

Analyzing these different kinds of virtual killings we realize how our picture of the relationship between morality and fiction needs to be updated in light of the advent of virtual worlds. It is not only the case that morality limits what kinds of fictional worlds we can imagine, as the puzzle of imaginative resistance shows. The kinds of fictional world we imagine in an effort of joint storytelling within virtual worlds may also subject us to new kinds of wrongs.

References

Anderson, Craig A., and Brad J. Bushman. 2001. "Effects of Violent Video Games on Aggressive Behavior, Aggressive Cognition, Aggressive Affect, Physiological Arousal, and Prosocial Behavior: A Meta-Analytic Review of the Scientific Literature." *Psychological Science* 12 (5): 353–59.

Bainbridge, William Sims. 2007. "The Scientific Research Potential of Virtual Worlds." *Science* 317: 472–76.

Caillois, Roger. 1961. *Man, Play and Games.* Translated by Meyer Barash. Urbana and Chicago: University of Illinois Press.

Carroll, Noel. 1990. *The Philosophy of Horror; Or, Paradoxes of the Heart.* New York: Routledge.

Davies, Martin. 2006. "Gamers Don't Want Any More Grief." *The Guardian.* June 15. http://www.theguardian.com/technology/2006/jun/15/games.guardianweeklytechnologysection2.

Dibbell, Julian. 1999. *My Tiny Life: Crime and Passion in a Virtual World.* London: Fourth Estate.

Ellis, Anthony. 1995. "Thomson on Distress." *Ethics* 106 (1): 112–19.

Huizinga, Johan. 1949. *Homo Ludens: A Study of the Play-Element in Culture.* London: Routledge & Kegan Paul.

Juul, Jesper. 2005. *Half-Real: Video Games Between Real Rules and Fictional Worlds.* Cambridge, MA: MIT Press.

Krotoski, Aleks. 2005. "Virtual Sword Theft Results in Real-Life Retribution." *The Guardian*, March 30. http://www.theguardian.com/technology/gamesblog/2005/mar/30/virtualswordt.

Mill, John Stuart. 1991. *On Liberty and Other Essays.* Edited by John Gray. Oxford: Oxford University Press.

Moran, Richard. 1994. "The Expression of Feeling in Imagination." *Philosophical Review* 103 (1): 75–106.

Pham, Alex. 2002. "Online Bullies Give Grief to Gamers." *Los Angeles Times*, September 2. http://articles.latimes.com/2002/sep/02/business/fi-grief2.

Radford, Colin. 1975. "How Can We Be Moved by the Fate of Anna Karenina?" *Proceedings of the Aristotelian Society, Supplementary Volumes* 49: 67–80.

Schneider, Steven. 2006. "The Paradox of Fiction." *The Internet Encyclopedia of Philosophy.* http://www.iep.utm.edu/f/fict-par.htm.

Thomson, Judith Jarvis. 1990. *The Realm of Rights.* Cambridge, MA: Harvard University Press.

Velleman, J. David. 2000. "On the Aim of Belief." In *The Possibility of Practical Reason*, 244–81. Oxford: Clarendon Press.

———. 2013. "Virtual Selves." In *Foundations for Moral Relativism*, 5–21. Cambridge: Open Book Publishers.

Walton, Kendall L. 1978. "Fearing Fictions." *The Journal of Philosophy* 75 (1): 5–27.

———. 1990. *Mimesis as Make-Believe: On the Foundations of the Representational Arts.* Cambridge, MA: Harvard University Press.

Watters, Audrey. 2010. "Number of Virtual World Users Breaks 1 Billion." *ReadWrite*. October 1. http://readwrite.com/2010/10/01/number_of_virtual_world_users_breaks_the_1_billion.