



# Virtual action

Jan-Hendrik Heinrichs<sup>1</sup>

Accepted: 13 November 2020 / Published online: 11 December 2020  
© The Author(s) 2020

## Abstract

In the debate about actions in virtual environments two interdependent types of question have been pondered: What is a person doing who acts in a virtual environment? Second, can virtual actions be evaluated morally? These questions have been discussed using examples from morally dubious computer games, which seem to revel in atrocities. The examples were introduced using the terminology of “virtual murder” “virtual rape” and “virtual pedophilia”. The terminological choice had a lasting impact on the debate, on the way action types are assigned and on how moral evaluation is supposed to be conducted. However, this terminology and its theoretical consequences, while sometimes resulting in correct results, lead to absurd results when applied across the board. It will be suggested that these absurd consequences can be avoided by a different answer to the question what people in virtual worlds are doing. Alleged virtual actions are first and foremost the creation and modification of data-structures and the resulting output in computer hardware. Such modifications of data structure and imagery can be performed with different intentions, purposes and styles, which will influence the type and moral evaluation of a user’s actions. This reinterpretation allows for a more complex analysis of the moral reasons for praiseworthiness or blameworthiness of actions in virtual environments. This analysis takes not just harm and effects on character into account but the peculiar ways in which speech acts can be morally wrong: e.g. agitatory, deceptive, bullshitting.

**Keywords** Virtual world · Gamer’s dilemma · Virtual murder · Speech acts

## Issue and question

People increasingly act and interact ‘in’ virtual environments, chatting, dating, trading, as well as mining, crafting, fighting, looting etc. There were virtual environments before the advent of the computer, such as the theater (cf. Chalmers 2017), tabletop games, or maybe even immersive literature. Today, however, single- and multiplayer computer games are the most common version of virtual environments, followed by non-gaming social environments such as Second Life. Recent developments in computer hardware have significantly increased the immersiveness of these environments, adding visual and auditory depth, so-called open worlds, and, last but not least, real 3-D immersion with virtual reality devices (Bailenson 2018).

Actions in virtual environments raise two types of urgent question: First, what is a person doing who, e.g., mines ores in a virtual environment? Is she ‘virtually mining’? Or more generally: Is an action in a virtual environment of an analogous type to its non-virtual counterpart<sup>1</sup>? Second, can virtual actions be evaluated morally? And if so, how can they be evaluated?

These questions have been pondered with a focus on gaming-action<sup>2</sup> and more extreme examples than mining. The

<sup>1</sup> Actually, it should say ‘of the same types’ insofar as any action can be subsumed under several different types. If in the example above the type is “intentional (vs non-intentional)” then virtual and non-virtual mining are of the same type. They are both intentional activities. But if the type is “activity that directly changes the distribution of rock on the surface of the earth” then they are not of the same type. For more detail see below.

<sup>2</sup> That actions in video games stand in the focus of the debate has good systematic and historic reasons. It is the most ubiquitous form of virtual action today and most likely in the foreseeable future, and there already is substantial literature on the philosophy of games prior to video games which gets taken up in literature in contemporary debates about computer games (e.g. Caillois 1961; Huizinga 1949; Juul 2004; Suits 2005). Nevertheless, this article will take into account the more general group of actions in virtual environments of which actions in games are a subgroup.

✉ Jan-Hendrik Heinrichs  
J.Heinrichs@fz-juelich.de

<sup>1</sup> Institute for Neuroscience and Medicine, Forschungszentrum Jülich GmbH, INM-8: Ethics in the Neurosciences, 52425 Jülich, Germany

most frequently used examples come from dubious computer games such as *RapeLay*, *Hatred* or *Custer's Revenge* and the actions in the focus have been described as “virtual murder” “virtual rape” and “virtual pedophilia”.<sup>3</sup> The terminological choice in the introduction of these examples had a lasting impact on the debate, on the way action types are assigned and on how moral evaluation is supposed to be conducted. However, this terminology is far from innocent, as it draws a close but tension-fraught analogy between events depicted on some computer hardware and events between flesh and blood human beings. In the following it will be shown that this analogy—while occasionally resulting in correct results by chance—leads to absurd results when applied across the board.

These absurd consequences can be avoided by a different, possibly more conventional answer to the question what people in virtual worlds are doing. If one accepts that a virtual action is only rarely of an analoguous type to its non-virtual counterpart, it will still be possible to hold that actions in virtual environments can be evaluated morally. However, they will usually be evaluated for different reasons than those that apply to or are relevant in evaluating their non-virtual counterparts.

## The analogical model of virtual action

How have the questions concerning action type and concerning the possibility of moral evaluation of actions in virtual environments been addressed in the debate? While the debate offers a rich variety of very differentiated answers, most can be matched to two general models, the analogical model and the representational model. The first model, the analogical model, is characterized by assigning a type to the user's action on the basis of an analogy to the content of the medium in question, in the case of computer games in analogy to what is depicted on the screen. In a nutshell it says: There is at least one morally salient type *virtual- $\varphi$*  to be assigned to events in a virtual worlds on the basis of similarity of these events to non-virtual behavior of type  $\varphi$ .<sup>4</sup>

<sup>3</sup> The term “Pedophilia” as used in the present context is not intended to refer to the psychiatric diagnosis. It is a common but too euphemistic (if not cynical) term for child sexual abuse, since cases of sexual abuse generally have nothing to do with philia. Perhaps it is better to call a spade a spade. If it is sexual abuse, it should be called that.

<sup>4</sup> One problem in this formulation of the analogical model is the reference to similarity. Similarity has famously been pointed out to be problematic (Goodman, 1972). Similarity claims were thought to be uninformative because everything is similar to everything else in some regard and thus a project of deriving properties from partial similarity judgements—as proposed in Carnap's *Aufbau*—was bound to fail. Goodman's critique was focused on a geometrical model of similarity and prompted a number of alternative models which try to overcome the issues he raised. A full treatment of these models

The moral valence of events of type  $\varphi$  extends to the event of type virtual- $\varphi$  or to the user's bringing them about.

I use valence and value/degree to be terms applicable across ethical theories. The moral valence of an action is the general direction of the final moral judgment. These valences typically come in polar opposite options of moral praise- or blameworthiness such as *according to duty* vs. *counter to duty*, *virtuous* vs. *vicious* or of *positive utility* vs. of *negative utility*, more generally: *permitted* vs. *forbidden* or most blatantly *good* vs. *bad*. Typically, such polar opposite options allow for degrees of praise- or blameworthiness. Even the Kantian distinction between actions according to or counter to duty for which such graduations are often denied allows for degrees of moral praiseworthiness, namely acting from duty and acting according to duty but from a different motive (Baron 2002).

The analogical model captures a very plausible intuition, which has been worked out repeatedly in the literature, namely that what a user of virtual worlds does and how it is to be evaluated must be closely related to the events presented in the medium. It clearly isn't possible to say what a user of a virtual environment is doing without first registering what is depicted on the screen—or rather what happens in the diverse media constituting the environment in question. Repeated pressing of the space button can for example be a case of producing Morse code, of making a simulated spacecraft accelerate, of making an avatar jump, with in turn nearly unlimited different types of semantic content in multiplayer games. What is contested between the analogical model and its alternatives is how the information about the events depicted on screen is to be interpreted, which other information is required to assign an action type, and whether and how this assignment shapes its moral evaluation. In the following, I will outline the analogical model and some challenges before pointing out how the alternative, representational model can handle some of its weaknesses. This alternative model in a nutshell claims that bringing about an event in a virtual world that represents events of

Footnote 4 (continued)

clearly cannot be provided here, for an overview see (Goldstone and Son 2012). I merely take it to be possible—on behalf of the analogical model—to systematically register, compare and correct domain-specific similarity judgments. These judgements will typically claim that the depiction on the screen *look similar* to some event between human beings, or that an event described by a game's text *sounds similar* to a description of events in the non-virtual world. So much is granted by most contemporary authors in this field and this has not even been the target of Goodman's critique, which targeted domain-independent claims of similarity. If even that should turn out to be impossible, so much the worse for the analogical model and it would reinforce the need for an alternative such as the representational model suggested below.

type  $\varphi$  in the non-virtual world is of the type *representing*  $\varphi$ . The moral valence of creating such a representation of  $\varphi$  is in principle independent of that of actions of type  $\varphi$ .

In an early contribution to the debate, Matt McCormick tested whether any of the established theory families in ethics had the resources to cover the intuition that playing violent video games is morally objectionable (McCormick 2001). McCormick looks to virtue theory for a possible answer and seems to think that a person can build his character and its constitutive habits by playing such games without these habits necessarily being action guiding beyond the gaming environment. His description of actions in video games is ambiguous throughout most of the article. While in some passages he refers to specific examples such as “[b] lasting someone into bloody pieces with a rocket launcher” (McCormick 2001, p. 283) most of the time he talks about “pulling the joystick trigger” (McCormick 2001, p. e.g. 285) or quite explicitly “playing a game” (throughout). However, when establishing his core thesis, that such actions are morally problematic from a virtue ethical perspective, he suddenly employs a different description. He introduces—seemingly as a thought experiment—improved gaming-devices at the level of the fictional holodeck from Star Trek and then goes on to talk about holo-crimes and about the perpetrators as e.g. holo-pedophiles or holo-murderers (McCormick 2001, p. 285). Although McCormick does not use an analogical model himself, his re-description of gaming actions as ‘holo-crimes’ paved the way. This style of action description invents an action type *virtual- $\varphi$*  in analogy to real-world events and assigns it to visually similar events depicted in the hardware of the virtual environment (holodeck, screen). This is the core of the analogical model.

One of the most-cited contributions to the debate cemented this way of describing action in virtual environments in its very title, Morgan Luck’s ‘The gamer’s dilemma: An analysis of the arguments for the moral distinction between virtual murder and virtual pedophilia’ (Luck 2009). Luck chose these examples because of the supposedly consensual moral evaluation by gamers and many non-gamers alike, namely the acceptability of virtual murder and the blameworthiness of virtual pedophilia and virtual rape. No empirical evidence is, however, presented for this.<sup>5</sup> Luck devises the following dilemma (actually a trilemma): Virtual murder is not immoral. There is no morally salient difference

between virtual murder and virtual pedophilia. Virtual pedophilia is immoral.

For present purposes, the way of describing the actions in question is decisive: “A player commits an act of virtual murder in those cases where he directs his character to kill another in circumstances such that, were the game environment actual, the actions of his character would constitute actual murder.” (Luck 2009, p. 31) And “A player commits an act of virtual pedophilia in those cases where she directs her character to molest another in circumstances such, were the game environment actual, her character would be deemed a paedophile.” (Luck 2009, p. 32) Luck already points out one consequence of describing actions in virtual environments this way: it applies to other media and their content as well: “this dilemma could be adapted to other types of virtual worlds, such as films, paintings and books” (Luck 2009, p. 35). While he distinguishes the active role played by gamers and the passive role of readers and movie goers, this distinction does not save the writers of novels, directors and actors of movies from moral blame for literary or cinematic murder. Luck seems to be the first to insinuate that there might be a transfer of moral evaluations between virtual and non-virtual versions of an action insofar as he sees this transfer not just for video games but across the media landscape.

Admittedly, Luck’s whole trilemma presupposes that virtual murder and murder require different evaluations, the former being morally neutral, the latter obviously blameworthy. Thus, initially Luck suggests that there is no analogy in evaluation for this case. However, he seems to imply that there is something wrong with the neutral evaluation of virtual murder and one obvious solution to the dilemma—only hinted at in Luck’s text—is that what he calls virtual murder has the same moral valence as the non-virtual action, if not to the same degree.

This way of describing the actions of players and users in virtual environments has become quite common, especially because a number of authors reacted to Luck’s article and tried to solve the gamer’s dilemma. Jeff Dunn for example makes this type of analogical thinking his core thesis, when he asks whether actions performed by a player via his or her virtual character are wrong if the same action would be wrong if the virtual world were real (Dunn 2012).

Rami Ali (Ali 2015) affirms the core strategy of the analogical model in a surprising way. Starting with the accurate observation that the individuation of an act is based on its context, he distinguishes between in-game context and gamer’s context. What makes an act of killing into a murder is the context, motive, means of manslaughter etc. Then he goes on to explain how the in-game context is relevant but not sufficient for act individuation of virtual murder. It is relevant whether the killing depicted on screen was embedded in the same context and motive, e.g. if it happened in

<sup>5</sup> There are data showing that real world murder is considered to be more forgivable than real world child abuse by many. The so called 2019 Forgiveness Poll has been conducted by Savanta ComRes and reported by the BBC (URL: <https://www.bbc.com/news/uk-england-47652050>). According to the data, available at Savanta ComRes, 85% said that they find child abuse impossible to forgive, while 73% said the same about murder.

depicted warfare against legitimate military targets or in depicted stealth killings for gain of virtual money. In addition to this game context, Ali thinks the gamer's motive for making a certain in-game move is defining of the action. Did the gamer fantasize about cruel bloodshed on a battlefield or did he merely try to beat the game and reach a maximal game score? The surprise in his analysis is that he does accept that in-game context and gamer's context are relevant for the decision whether a virtual killing is a virtual murder, but he does not consider which context makes the act in question a virtual killing in the first place. He calls a depiction of killing on the computer hardware a case of virtual killing without checking any criterion. Thus, Ali accepts a part of the analogical model, namely the description of a player's action as virtual killing at the outset of his discussion.<sup>6</sup> He does not however, employ the analogical model for moral evaluation.

Indeed, Ali introduces another important element, namely the intent and style of the gamer's action.<sup>7</sup> He suggests that whether a virtual killing is virtual murder depends on the style of the player's engagement with the game context. It makes a difference whether he or she plays the game without following the narrative at all, with the intent of acting out his or her murderous desires or with the intent of beating a complex video game.<sup>8</sup>

To summarize, the analogical model consists of two parts, which are complementary but of which the first can—and sometimes does—stand without the second.

First, it claims that there is at least one non-trivial type virtual- $\varphi$  to be assigned to events in a virtual worlds on the basis of similarity of the behavior described or depicted in the medium to non-virtual behavior  $\varphi$  (cf. Sheng 2020). This can be, but rarely is, employed for purely taxonomic reasons. Even then, taxonomy is not fully innocent. The taxonomy presupposes that actions of type  $\varphi$  and virtual- $\varphi$  share something that licenses the taxonomic decision. Normal speakers would be perplexed by a claim such as: “This is a case

of virtual- $\varphi$  and it is structurally, morally, aesthetically etc. completely different from cases of  $\varphi$ ”. One would wonder why to call it virtual- $\varphi$  then, and not something completely different and unrelated. Because the analogical model is employed in ethical texts predominantly, the first claim is usually made stronger, namely that the non-trivial type to be assigned to the event in the virtual environment is one, which is relevant for the moral evaluation of the real-world behavior  $\varphi$ .

The second part of the analogical model claims that the valence of the moral evaluation of actions of type  $\varphi$  extends to the event of type virtual- $\varphi$  in the virtual world or to the user's bringing about this event. This does not include the value, i.e. the full force of the moral evaluation. Nobody would claim that murder and its virtual counterpart are morally blameworthy to the same degree.

This second claim presupposes the first, taxonomic claim. The straightforward transfer of the evaluation typically associated with type  $\varphi$  to actions of type virtual- $\varphi$  finds fewer explicit supporters but plays a relevant role because after adopting the first, taxonomic claim of the analogical model it has become the new fallback option. Here is why: Without assigning an action type analogous to some morally salient real-world behavior, the cultural default option for the evaluation of actions in virtual environments is moral neutrality.<sup>9</sup> Our standard reaction to actions in virtual environments is just like our default attitude toward games, which is expressed by the phrase: “It's just a game”. This changes with employing the first, taxonomic part of the analogical model. The intuitive fallback option for the evaluation of the type virtual- $\varphi$  is not that of ‘virtual’ but that of  $\varphi$ . Several of the authors discussed above provide alternative modes of evaluation for actions in virtual environments, such as McCormick's use of virtue ethical methods. But even for those authors, the background of the evaluative landscape has changed significantly. If it turns out that the virtue ethical mode of evaluation comes up empty for a specific action—e.g. that a case of virtual burglary does not result

<sup>6</sup> In some passages, Ali even sounds as if he were leaving the analogical description behind and focusses on what the user does: “The act of using the game controller is obviously not what is at stake, it is rather that the inputs we give allow the representation of pedophilia to unfold.” (Ali 2015, p. 272) He does, however, not stick to this way of describing the user's actions as representing something, but rather to the analogical model.

<sup>7</sup> For another approach of integrating the specific motivation into moral evaluation see (Young 2013).

<sup>8</sup> These different motivations seem to be associated with different game modi, which Ali calls sports, storytelling and simulation games. According to this distinction sports games typically provide a setting, where even the most atrocious on-screen events serve the purpose of beating the game, not of enjoying the scene as such, while in simulation games, there is no competition allegedly justifying playing through immoral events.

<sup>9</sup> This is not to claim that popular or political culture is free from extremely strong moral reactions to computer games, especially games with violent or sexual content. The opposite is true, there is still widespread condemnation of violent or sexually explicit computer games, even if it has probably lost its social dominance in most countries.

The point made here is slightly different. It merely points out, that in general, actions in computer games are considered to be morally neutral, unless one makes the transfer of action types suggested by the analogical model. This is well illustrated by other actions such as hugs, saving the world, feeding the hungry or sheltering the homeless in a computer game. If someone were to brag that he just fed 20 hungry and homeless citizens of Westfall in the WoW quest “Feeding the Hungry and the Hopeless”, the standard reaction would be: “It's just a game!”.



in habituation of vicious behaviors or similar detrimental developments of character—it does not immediately follow anymore that the action is morally neutral. As a case of virtual- $\phi$  it might well have the same moral valence as  $\phi$ .

The analogical model does have its critics, however. Mark Coeckelbergh is probably the first to give an apt description of the analogical model and reject its suitability as a tool for moral evaluation: “A common approach to ethics of computer games considers the content of the games, and the relation between playing the game with that content and behavior in the real world. The content of the games is judged by generally accepted moral norms that forbid certain acts. The metaphor used may be that of contamination: if the content is bad, surely we must prevent it to spill over from the virtual world into the real world.” (Coeckelbergh 2007, p. 223). This description is to the point, but in his ethical analysis Coeckelbergh does not fully exhaust its critical potential. His main worry about immoral virtual actions is that they provide some kind of reinforcement or training for actions of an analogous type. He supports the case for such a training by spelling out the similarities between in-game action and off-game action in terms of immersion and interactivity of modern games. This emphasis on similarities brings him close to the analogical model, which he initially seemed to reject. The worry of immoral training is more than justified, but it can be developed without accepting the analogical model at all.

Schulzke (Schulzke 2010) points out the difference which has been covered up by the talk of virtual killing with a simple description of what gamers do: “Games involve simulated killings, but players do not intend to kill another person when they play. They only mean to destroy an avatar.” (Schulzke 2010, p. 129) Schulzke importantly makes philosophical theory of action available for the debate by drawing the attention towards the gamers’ intention as the central source of information about their action type. He drives a wedge between the action as depicted on the screen and the action of the user. One cannot describe one by analogy to the other.<sup>10</sup>

<sup>10</sup> One might, however, disagree with his description of gamers’ intentions as to “destroy an avatar”. While this is a possible description of what happens, it seems to be slightly misleading. In gamer’s talk, one will often hear of the intention to destroy or kill a specific character. In WoW for example, people will brag of having killed a specific boss, say Argus the Unmaker in Antorus, the burning throne in the Legion Expanse of the game. Talk about killing or destroying that *avatar* are, however, rare. This is because the character name belongs to the fictional framework, in which talk of destroying or killing makes sense. Talk of an avatar belongs in the game-framework, where ‘destroy’ has another meaning. Destroying an avatar would mean to delete or corrupt the relevant file(s) on the game server, which no gamer expects to happen. They are perfectly aware of the fact that the avatar in question still exists, that it has merely been marked as inactive by the game mechanics and looks accordingly on screen.

In a similar vein, Seddon has worked out clearly what the analogical model does, or rather does not do, and in so doing identified the different events occurring during a so-called virtual murder: “Neither is playing a game understood as the creation of pictorial representations, even when the game provides feedback by means of images on a screen; on the contrary, virtual murder is what a screenshot of in-game killing is supposed to depict, not simply to be.” (Seddon 2013, p. 1) Following up on this short remark by Seddon, four different events need to be distinguished: the pictorial representation on screen, the digital events causing this depiction, the user’s activity with computer periphery, and the events depicted, which might as well be fully fictional events.<sup>11</sup> And as Seddon points out, “it is no minor terminological judgment to decide that gaming violence not merely resembles or depicts or represents or models murder, but is ‘virtual’ murder.” (Seddon 2013, p. 2).

Patridge puts more focus on the moral evaluation of actions in virtual environments and turns against the analogical thinking in this regard. She characterizes this model as “a mistaken moral assumption, namely that if our virtual activities are subject to non-harm based moral assessment then they must derive their moral status in a straight-forward way from the status that they would have in the real world.” (Patridge 2013, p. 33) Rather than deriving moral evaluation from what is depicted on the screen, ethical awareness should rest on “the nature of representational detail that we confront in-game and how reasonably it invokes thoughts of our actual moral reality” (Patridge 2013, p. 33). Patridge is probably the first in this debate to point out that the acts of creating on-screen representations are viable targets of moral evaluation themselves (Patridge 2010, 2013).<sup>12</sup> She calls out the enjoyment of game imagery with morally problematic moral meaning as at least insensitive and argues that it expresses or reveals a flaw in the character of the player. Her position has therefore been called *expressivist*.

Based on Patridge’s work, Sebastian Ostritsch has devised what he calls the endorsement view (Ostritsch 2017). The core thesis of the endorsement view is that certain pieces of fiction such as computer games may—under certain conditions—be “not merely fictional, because on a pragmatic level, [they] also endorse[...] a normative view about the real world” (Ostritsch 2017, p. 122). The main target of Ostritsch’s analysis is not the action of an individual within a

<sup>11</sup> Analogue differentiations can be made for any virtual environment, not just modern computer games. It need not be depicting of events on screen, it can be description in written text or spoken word, it need not be digital events and users manipulating computer periphery, it can be more or less any medium and a user’s interaction with it.

<sup>12</sup> Similar arguments against the analogical model have been formulated by Robert Sparrow (Sparrow 2017), if in the context of sex-robots, and John Tillson (Tillson 2018).

virtual environment but the virtual environment itself. Thus, he recognizes that games, and derivatively the actions within games, are carriers of meaning, are predominantly representations together with a certain style and attitudes towards the represented.

Ostritsch remains silent about the relation between virtual actions and their non-virtual counterparts. For him, the metaphysical and moral status of individual action within the virtual environment is derivative to that of the game itself. Nevertheless, he asserts that games—and virtual environments as a whole—are ontologically incomplete without the user's action, the user completes what is only potential in an unplayed game. Thus, if games are carriers of meaning so is the action of the player. According to Ostritsch, the gamer's moral duty is not to enjoy—in a strong sense of having fun—but to have contempt for immoral games, even if he or she comes to play them. Consequently, few acceptable reasons for playing such games remain, among them are scientific or journalist investigation.

## The representational model of action in virtual environments

While I'm sympathetic to the endorsement view, I want to present a thesis with a wider scope, one focusing on virtual action in general and its relation to possible non-virtual counterparts. I take the phrase 'x acts in a virtual world' to already be misleading and reference to his or her action as 'virtual  $\phi$ -ing' even more so. 'Virtual world' mostly refers to computer generated imagery, sound and possibly tactile stimulation, which is called thus for its immersive effect. From the perspective of an external observer—and that of the person herself if she closes her eyes for a moment—it is a form of interaction with computer hardware, (head-mounted) displays, speakerphones etc. in our shared, very non-virtual world. The person allegedly acting in a virtual world first and foremost interacts with a virtual environment, an environment mostly consisting of computer hard- and software and its states. There are virtual worlds not generated by computers, such as the theater, role-playing games, several board games etc. The differences in the underlying technology do not create a principled difference for the philosophical analysis, rather, there are differences in the degree of immersiveness and interactivity.<sup>13</sup> The following will focus on computer generated virtual environments, for reasons of ubiquity and fairly high immersiveness. Thus, if in the following I refer to 'events depicted on screen' the

argument extends to other forms of presenting virtual worlds as well.

In developing an alternative to the analogical model, I would like to use a less drastic action type as an example: lifting an arm. How one makes an avatar lift its arm depends on the system in use. Because of its wide distribution, World of Warcraft (WoW) will have to do as an example here. There is no specific command for lifting an arm but typing in "/wave" will do the trick. It is a bit more than just lifting the arm, it includes waving, but typed in as above, without any target mentioned in the command line, the result will be an output in the console saying "[name of one's avatar] waves" and the avatar making a waving movement with its arm. If one uses the analogical model to describe what a user does in typing "/wave" into the WoW console, one starts with the imagery on the screen showing a person lifting her arm, and from that infers that the user lifts her virtual arm, or that she virtually lifts her arm.

In an alternative, more adequate model of action in virtual environments, the primary focus should lie on what the user is doing in her context by means of the technology at her disposal, not on the images on the screen. The minimalist way of describing the user's action would be: The user types "/wave" into the WoW console. But that will hardly suffice because it takes neither information about the user's intention and context into account, nor information about the events depicted on screen—the virtual-world-context. Both can and one of them must play a role in identifying the type of a user's action, as will be shown in the following.

The virtual-world-context is in most cases important for identifying the type of the user's action, but it can in some cases be ignored by the user and thus in these cases be completely irrelevant to describing his or her action. If such a situation occurs in a gaming environment, the actions in question are typically not considered to be gaming actions anymore. Imagine a situation where a friend asks a gamer "Can you lift that figure's arm?" and the gamer types in "/wave" to show that he can. Even more abstractly: Two game designers talk about technical details of the programming of the avatars and test whether their movement patterns have been implemented correctly by typing in "/wave" and observing the imagery on screen. In order to understand whether and how much influence the events on screen have on deciding the action type of a user's action, it is important to identify the interpretative framework of the user and his possible interaction partners.

A convincing distinction of such frameworks has been introduced by David Chalmers (Chalmers 2017). Chalmers distinguishes between virtual and fictional worlds. Virtual worlds are—primarily—immersive, computer-generated, interactive environments (in the case of virtual worlds generated by other media, they would primarily be environments generated by the medium in question). These

<sup>13</sup> For a philosophical discussion of immersiveness see (Chalmers 2017), for comparisons of the immersiveness of different media (in comparison to virtual reality systems) see (Bailenson 2018).

media-generated virtual worlds can be, often are, but need not be, interpreted by the user as a fictional world. Fictional worlds are the content of stories told in some medium, such as Tolkien's middle-earth, WoW's Azeroth, or the fictional version of Europe in the 1940s narrated in several World War II novels, movies or games. The interpretative relation between virtual and fictional worlds holds not just for specific fictional content such as Frodo's journey in the Lord of the Rings video game or the assassination of Hitler in WW II video games or the movie *Inglorious Basterds*. It holds for generic fictional content such as the interpretation that there is a physical space and physical objects moving in this space, some of which are avatars interacting, that this particular object is a chair, etc. One can interpret a video game this way, but one need not do so.

Virtual versions of actions, such as movements of an avatar's arm, clearly depend on interpretations. With Chalmers, we can take it as a given that actions in virtual environments are actions. However, what type of action a particular move in the virtual world is, will depend on whether there is any fictional interpretation of this move and its respective content. As mentioned, it is quite possible for a user to ignore any type of interpretation of the movement in the virtual environment, consequently this interpretation cannot be sufficient for assigning an action type to the user's action.<sup>14</sup>

In contrast, the gamer's context can never be ignored when evaluating what type of action occurs. The gamer's context allows for an incredible amount of different action types to be realized by typing "/wave",—greeting other users, roleplaying, testing out the controls, showing a friend how one moves an avatar's arm, testing the connection to the game server etc. What action in particular the user realizes depends amongst other things on her or his intentions (Schulzke 2010; Ulbricht 2020), on her or his performance and on the interpretation of the possible audience of this action. There are endless constellations even if one focusses on the "/wave" command only: A user can against her own intention greet someone if this someone interprets the user's action—typically pressing the space bar and thereby making the avatar jump—as a greeting. A user can fail to greet someone by not controlling for the conditions under which she types "/wave", which can result in output such as "[name of the user's avatar] waves at the lamp post", etc.

In addition to the endless options of intention, performance and interpretation, there are some stable elements in the gamer's action, namely the interaction with the hard- and software of the game environment. For the "/wave" command, this stable element can be described thus: By typing

"/wave" into the WoW console, the user causes changes in a data structure, which in turn makes a computer-generated image of some avatar move in a way that looks like it is lifting an arm. This is similar across input devices, even in the case of modern VR equipment where the computer-generated image mimics the person's movement: The person by lifting her arm modifies a data-structure which makes a computer-generated image of an avatar move in a way that looks like lifting an arm.

This description of actions in the virtual environment are often not the descriptions under which a person performs them *actualiter*. The player in WoW waves at another player under exactly this description: wave at another player. However, the means (hard- and software)-based description is one which is always at the user's disposal and which does continuously shape his or her behavior.

Combining the intention-and-performance-based and the means-based component of this description of action will yield something like: The user modifies a data-structure by typing "/wave" into the WoW console, in order to greet another player, which makes a computer-generated image of an avatar move in a way that looks like lifting an arm. This way of describing actions in virtual environments can be called the representational model, because it identifies the user's action as a case of creating or manipulating representations, both electronic and audiovisual for some further purpose.

The representational model has several advantages over the analogical one—if regrettably not its suitedness for snappy terms such as 'virtual murder'. As will be shown later in this article, it is better suited to evaluative purposes, too. But first and foremost, it provides a more adequate metaphysics of action in contact with virtual environments. Unlike the analogical model, the representational one does not have to invent a new action sub-type 'virtual  $\varphi$ -ing' for every action type  $\varphi$ , but can treat action in or in contact with virtual worlds as belonging to a familiar, representation-creating or -manipulating action type. What according to the analogical model is a case of virtual- $\varphi$ , the representational model simply considers representing  $\varphi$ .

Therefore, the representational model can keep the description of virtual reality consistent with the description of production and consumption of other arts and media. When Picasso created *Guernica* he did not commit a canvassy airstrike. He painted an airstrike. And someone who looks at the picture today does not look at an airstrike either.<sup>15</sup> When an ancient Greek actor played Pheidippides hitting his father Strepsiades in Aristophanes' *Clouds* he did

<sup>14</sup> Thi Nguyen's would probably insist that if you're not engaged in the narrative, you're not playing the game at all (Nguyen 2019). However, even if you are not playing the game you are acting in the virtual environment in question.

<sup>15</sup> The case is slightly more complicated for the question of what an observer sees. The famous example is this: An observer looks at a photo of Winston Churchill, but she sees Winston Churchill (as the causal basis of all photos and film of Winston Churchill). The greater the methodological and intentional distance from the documentary

not commit theatrical assault. A theatre goer blaming the actor for assault or even coming to Strepsiades' aid would have seriously misunderstood the events. Similarly, if a programmer generates a battle-scene content for virtual reality, he does not virtually set up a battle, and if a user plays to fight in said battle, he does not virtually kill enemy soldiers.

While the difference between the action of the artist and the action that the artist depicts holds most of the time, it does not throughout. Take for example insults to the audience. An actor in modern theater turning to the audience and clearly addressing the audience as the audience in an insulting style, does not merely depict insults, he also really insults the audience (Handke 1966). David Chalmers in his seminal 'The virtual and the real' (Chalmers 2017) pointed out, that while a virtual library may be a library, a virtual kitten is not a kitten. The same amount of differentiation is required for events and not just objects. It is quite likely that virtual sculpting and sculpting share a non-trivial type such as artistic creation, but the same seems to be dubious for virtual piloting and piloting—in most cases at least, and for virtual murder and murder.<sup>16</sup>

## Whether actions in virtual environments are morally evaluable

The choice between the analogical and the representational model of action regarding virtual worlds has direct repercussions on their moral evaluation. Admittedly, the analogical model does have linguistic flair on its side: Starting for example with the depiction of a murder on the screen and considering the user's computer-input as *virtual murder* makes for a temptingly snappy term. However, the fact that a morally wrong action would happen if the events shown on screen were to occur between flesh and blood human beings does not make participation in the data manipulating and resulting on-screen events wrong. Thus, the linguistic flair of 'virtual murder' comes at the cost of an implausible model of action.

If the description of the events depicted on screen is not sufficient for assigning an action type, neither is it for moral evaluation. However, that does not preclude the possibility that the action in question is wrong, if for other reasons. In

Footnote 15 (continued)

imagery, the more complicated this relation between what is depicted and what the observer sees will get. For the current purpose it will suffice to concentrate on what the observer looks at, and that always is the depiction itself.

<sup>16</sup> Exceptions have been thought up in science fiction literature, e.g. Ender's Game, in which the protagonist finds out that all his virtual training sessions in piloting and space battle have actually not been virtual at all.

some cases, the analogical model will generate correct moral evaluations, but not for systematic reasons, as will be shown in the following.

Why might an action in a virtual environment be morally blameworthy or praiseworthy at all, if it is primarily manipulation of data structures which results in images on screen (and sound etc.)?<sup>17</sup> First and obviously, because certain types of data manipulation hurt others. If one modified data-structures and images on screen in a way that another is constantly confronted with aversive imagery or sound, that is a case of harassment. If someone manipulates a data set in a way that denies access to its original owner without consent and accepted compensation, it is a case of theft. If one were to destroy someone else's data-structures (including their avatar) or modify them in a way not accepted by the owner, it is a case of property damage etc. These latter two ways of acting might be very grave because of the attachment many users have to their avatars.<sup>18</sup> One need not (but can) go as far as claiming that the avatar is a part of the gamer's person, as some extended cognition theories suggest. It is sufficient

<sup>17</sup> Saying that actions in a virtual world are primarily manipulation of data structure does not imply that this description is the only one under which a user performs the action, nor that this description is always fully transparent to her or him during performance. A gamer's behavior in the middle of a battle scene might not look much like what the dry term "manipulation of data structures" seems to imply. But the gamer would surely not act as she or he does, if it were not constantly accessible background knowledge, that the gameplay is virtual. Neither would the gamer continue to shoot, if he lost that belief, nor would he quicksave, watch out for lag (delay in gameplay due to data-transfer limits) or open the quickchat menu to communicate with the team.

<sup>18</sup> The extent to which malicious manipulation of a person's avatar—avatar assault—can harm a user has been discussed in detail in (Huff et al. 2003), and in (Powers 2003). The case in which this phenomenon has first received public attention has been reported in (Dibbell 1993). It occurred in a text-based multi-user platform, in which one participant via a hack took over the control over other member's avatars and had them play out a brutal sexual scene. I take it to be a still open—and in parts empirical—question how harmful such and other manipulations can be for the affected user.

It should, however, be noted that even the case of malicious manipulation of a person's avatar is slightly at odds with the analogical model. If the analogical model were adequate, a case of, say, virtual bodily assault should be wrong, because the real-world counterpart—bodily assault—is wrong. But this is not what harm-based accounts say. Rather, according to harm-based accounts these actions are wrong because of the suffering of the person emotionally invested in the avatar. If the same action occurred but the avatar's owner did not suffer from it (say in a computer game's Player-versus-Player scenario, or in a Live-action-roleplay where one person acts out a non-player role), harm-based accounts would see no wrong while the analogical account would insist that there still is. The representational model can point to the suffering caused by a person's avatar being involved in representations which either are simply against the will of the avatar's author or even themselves aversive and disturbing. One might, admittedly, wonder, whether the harm and suffering caused by such unwanted or aversive representations is equal to the harm reported by the victims of avatar assault.



to point to the significant investment of time, energy and sometimes money gamers put into their avatars, and to the emotional reaction to harm to this special data structure (Schulzke 2010; Wolfendale 2007).<sup>19</sup>

What is at issue in the whole debate about alleged virtual murder and virtual sexual abuse of children is whether in addition to uncontroversial cases of actual harm, there is a type of moral blameworthiness—and possibly praiseworthiness—which does not result from harm to others. Many authors highlight the fact that they are after moral reasons beyond reasons from harm by insisting that actions in virtual environments can be morally blameworthy, even if no other person or that person's avatar is affected. One way to make this limitation salient is by limiting the perspective to actions aimed at mere computer-directed avatars on the perpetrator's own single-user system (Luck 2009). Nobody gets hurt if for example someone generates a data-structure, which in VR looks exactly like a colleague and then—unknown to that colleague—hacks it to pieces in VR (Tillson 2018).

The second fairly common answer, after that referring to harm, draws on the effects on the person's character. From a virtue consequentialist or virtue ethical perspective, the moral damage of virtual brutality is to be sought in the habituation of the user. Someone who regularly engages in gameplay or other virtual action which depicts high levels of physical or psychological brutality will become used to acting in such a way and might well transfer this type of action into non-virtual environments (see e.g. Coeckelbergh 2007).

Whether such effects really occur is an empirical question, which has found broad attention for decades. Whether brutality in the media is cathartic or habituating, whether it has an effect at all, is, and will continue to be, a valuable research topic for psychology and sociology, but cannot be resolved from the armchair (Polman et al. 2008). If such effects occur, they are a valid consideration within moral evaluation. Still, even in conjunction with the abovementioned harm-based reasons, the potential of moral evaluation of actions in virtual worlds seems not to be exhausted. Somehow, there still seems to be something wrong when a person plays at brutally murdering a virtual version of her colleagues on her private computer, even if no habituation or spill-over occurs. Here is another example: in the WoW quest “Maintaining Discipline”, the player walks a slave mine, zapping into action exhausted workers with a

‘disciplining rod’, some obviously torturous device. Repeating the quest often enough will give the player reputation with the local non-player characters and thus allow the player to buy a rare pet. Now imagine a player returning to the quest after having received that reward, just for the sake of playing at striking slaves. Or take an opposed example: there is an increasing number of gamers trying out pacifist playthroughs for several games including *The Elder Scrolls: Skyrim*, a game strongly oriented towards violent conflict. One impressive feature of this pacifist gameplay surely is the challenge of beating the game literally speaking with their hands tied. But another reason for being impressed is the avoidance of scenes of violence and bloodshed for the sake of scenes of persuasion and conflict avoidance.

Having put aside harm-based and character-based arguments for the moral evaluation of action in virtual environments, is there another normative resource for their assessment, even after leaving behind the easy path suggested by the analogical model? The representational model provides such a resource. While it does not allow us to transfer an action type and thereby the moral evaluation from the non-virtual counterparts, it does identify the actual type of actions in a virtual environment and thereby opens up a path for evaluating specific actions. What does a person do who creates and hacks up a VR-version of his colleague, or who whacks WoW's slave avatars with a discipline rod? According to the representational model, she primarily creates audiovisual depictions of violence by means of manipulating data-structures. And while there are good reasons and morally unproblematic ways to create such a depiction of violence, there are very bad ones, too: When Picasso painted *Guernica*, he created a warning against the horrors of war by depicting them in a specific way. When an extreme right-wing journal like the NS-German ‘*Kunst dem Volke*’ (art for the people) depicts war events in the trenches in heroic, glossy images, it tends to glorify war. The one is morally praiseworthy, the other morally blameworthy. The reason for this does not lie in the effects of the imagery alone. Even if nobody ever looked at *Guernica* or those heroic, glossy trench images, the act of creating them seems to be morally evaluable. What is more, the moral reasons for and against such creations do not refer to the creator's interests or conception of the good. They are—as will be shown in the next paragraph—agent-neutral (cf. Parfit 1984, p. 27).

## How the representational model of action influences moral evaluation

One reason why they might indeed be morally evaluable is that such depicting is a speech act. By creating or modifying imagery, one can glorify or incite violence, revel in cruelty, mock weakness, but also praise kindness or decry

<sup>19</sup> One way of approaching the gamer's dilemma via the representational model, pointing out the harms specific representation has been presented by Bartel (2012). He points out that virtual pedophilia is subsumable under child pornography and therefore morally—and legally—culpable. While there is little to be objected to in this analysis, especially if considering the legal codes of several countries (for a discussion of the American, British and Japanese case see chapters 1.3 and 2.1 of Young 2018), it does not generalize easily.

oppression. Thus, these actions are, amongst other things, expressions of attitudes and evaluations themselves; by creating a certain type of imagery, a person endorses certain norms or values (cf. Ostritsch 2017, who focuses on games as interactive systems of symbols, not on the actions of game producers or players). Such an endorsement is an opening move for moral discourse; it invites replies which support or criticize the norms and values and their endorsement.

If indeed actions in virtual environments are moves in moral discourse, they can be wrong in just the same way as other such discursive actions can be wrong. We can make use of the ethical resources employed for the evaluation of discursive actions and speech acts for the evaluation of actions in virtual environments. Admittedly speech acts can cause harm just as any other action can, and they can affect the speaker's and the audience's character just as other action can. As mentioned above, these ways of being wrong surely are relevant and possibly even dominant in the evaluation of actions in virtual environments, but they do not exhaust the potential for such evaluation. There are additional ways of being wrong for a speech act.

Two basic types of being wrong for a speech act are (a) to be false,<sup>20</sup> and (b) not to be adequate to the speaker's intention, i.e. being misleading. These two basic types can take slightly different forms and be combined in various ways, generating several quite specific ways for speech acts to be wrong: they can be mere error, they can be agitation, they can be deceptive, they can be bullshit (in the sense used by Harry Frankfurt). This list is by no means exhaustive but contrasts the most harmless case of simple error<sup>21</sup> to the ethically most salient ways of being wrong: agitation, deception, and bullshit.

I'll only take a short glance at the first before turning to two special forms of bullshit and deception.

<sup>20</sup> An assertive statement being false is the prototype of this way of being wrong. Other speech acts can be wrong in a similar if not exactly the same manner. This point has already been pressed by Kenny: „For if a man, without lying, says something false, then he makes a mistake; but a man who gives a command which is not obeyed, does not therefore make a mistake. I shall indeed argue below that it is possible to be untruthful in giving a command; but a command is untruthful when it is not meant, not when it is not obeyed “. (Kenny 1963, pp. 152, Fn. 151)).

<sup>21</sup> Patridge (2010) shows that even the alleged simple case of mere error might be morally problematic. Some facts ought to be known by members of society. Patridge's example is the fact that imagery showing African-Americans eating watermelons have a racist meaning. She considers a person who, in an American social context, does not have this knowledge and enjoys this type of imagery in a virtual environment. Although this person seems to make a simple error out of ignorance, she takes her to be culpably lacking sensitivity.

For an action in a virtual environment to be a case of agitation it needs to somehow be accessible to—if not accessed by—others. Users need to tell others about their exploits in a virtual world, to brag about gameplay or their actions and the options on a certain platform. They need to invite others to take part or to provide their own modifications to a community. Regrettably, this is not a philosopher's scenario but a ubiquitous phenomenon in the use of virtual environments. Countless cultural subgroups with an agenda of hate create and distribute mods for games which make computer-controlled enemy avatars look like people of a specific ethnicity, gender or social background. There are racist or misogynist modifications for many computer games, in addition to computer games originally produced with a similar agenda. These games and environments are shared within the relevant subgroups and the actions therein communicated either directly via interactive software modes or via other channels.<sup>22</sup> These uses of virtual environments can be treated within the existing moral and legal frameworks for incitement, agitation and hate speech.

An action can only be incitement or agitation if it aims at an audience other than the speaker, but surprisingly enough it need not do so in order to be deceptive. While it is still a matter of discussion in philosophy whether people can in the strong sense of the word deceive themselves, the phenomenon so described in folk-psychological terminology is rarely denied.

A person might by her very actions in a virtual environment express positive attitudes towards some morally corrupt action and in her verbal behavior deny such an attitude. She might say that murder or sexual abuse of children are not something to be enjoyed and surely morally blameworthy, but by her action in the virtual environment express the opposite attitude.

Depending on one's view of self-deception, one will have to provide slightly different explanations of how these two attitudes can occur in the same person at the same—or nearly the same—time. Take the above examples of the depiction of brutally murdering an office colleague or of needlessly beating slaves with a discipline rod. The person expresses both views: that murdering others or that holding and torturing slaves is morally blameworthy, is not to be enjoyed and

<sup>22</sup> Please note, however, that there are positive examples of such mods as well. There are mods to major games, which allow a user to experience the effects of racism or other forms of stigmatization by putting her or him in the place of a group often so treated, be it real (see for example the VR experience 1000 cut journey) or fictional groups (see for example the 'More Racism in Windhelm SSE' mod to Skyrim).

that it is to be enjoyed even if morally blameworthy.<sup>23</sup> This seemingly paradox situation has been explained by positing that he or she holds these views at different times (temporal partitioning, see e.g. Sorensen 1985), in different contexts or in compartmentalized parts of her mind (psychological partitioning, see e.g. Davidson 1982) or by treating some of the states involved as sub-doxastic or motivationally biased. In more or less any version, the person's mental states contain an inconsistency.

This inconsistency can—and in the examples provided does—concern deep moral convictions. Typically, inconsistencies in such deep moral convictions are considered to be at least a character flaw, if not a culpable flaw. Whether a person can be responsible and thus culpable for a case of self-deception depends, again, on the specific account of self-deception. Given an intentional account of self-deception, i.e. one which accepts that individuals can hold two inconsistent doxastic states at the same time (cf. Bermúdez 2000), for which I have sympathies, this person is guilty of an important omission. She omits to attend to her own doxastic states and allows a form of morally problematic inconsistency, which she could—and should—clear up. This form of self-deceptive ignorance of in the right circumstances not knowing that some things are not to be enjoyed goes beyond the lack of sensitivity admonished by Patridge. Furthermore, the moral reason against this form of self-deception do not presuppose that the person in question has a preference or moral conviction against self-deception. There is something objectively amiss with the act of deceiving oneself about such deep moral convictions, which goes beyond harm to others or to herself and beyond habituating problematic behavior or character traits.<sup>24</sup>

<sup>23</sup> Admittedly, what the person expresses in playing at murdering the avatar depicting the colleague depends on her intention and the style of her action. In the very best case, she might just express that make-belief-murdering colleagues is to be enjoyed while real murder is not. In other cases, however, she will not make this distinction and simply express that murdering this colleague *simpliciter* is to be enjoyed. While this seems to be what children do in playing with toy guns, an adult's doing so seems to include a more complicated set of considerations: For children it is *just* play and they tend not to express any endorsement of norms or actions. An adult, however, this innocence will usually not be attainable anymore. Once someone creates an avatar looking like a colleague in order to play at killing it, he usually is quite aware that this behavior does express something, which is at least morally salient. Otherwise we would have many more dartboards with colleagues' images on them openly hanging in offices. Thus, if an adult engages in this type of behavior, he forms an extra meta-attitude about the expressive function of his play. He or she cannot claim that it is just play but will have to hold in addition that this type of play does not entail endorsing violence towards colleagues but merely endorses playing at violence towards colleagues.

<sup>24</sup> If one were to fall back on a non-intentional account of self-deception, the only thing one could hold against the person is a defect of character, but not the self-deceiving action itself.

A person who before and after engaging with the virtual environment endorses some moral norm and in the virtual environment produces representations which express rejection of the same rule need not self-deceive. This behavior need not involve a person's endorsing two different attitudes. Rather it is possible that he simply does not care about the moral norms he professes to endorse. Either in the virtual environment or his verbal behavior or in both he merely touts what he thinks socially desirable. In what sounds like an explicit endorsement as well as in his expressive endorsement in the action in the virtual environment, this person pays mere lip-service and does not engage with the norms in question at all. Such a person is bullshitting himself in the established philosophical sense of the term. While the person who self-deceives at least recognizes that her action calls for moral engagement, the person who self-bullshits lacks even this awareness (cf. Frankfurt 1986).

As in the case of self-deception, bullshitting oneself with regard to important moral topics and to whether one even has deep moral convictions concerning these topics does more than reveal a character flaw. A person who does not care for the moral norms she endorses in her expressive behavior does not invest the level of attention normally required in this type of discourse, no matter whether he or she personally considers this kind of attention worth having. She—as the person who self-deceives—engages in speech acts which do not live up to the relevant normative requirements.<sup>25</sup>

The representational model does more than provide a powerful alternative for the assignments of action types to actions in virtual worlds, even more than to provide a more plausible starting ground for the moral evaluation of such actions. It can explain and unify several previous approaches to classifying and evaluating actions in virtual environments. As shown above it can integrate the expressivist and endorsement view of Patridge and Ostritsch. Both focus on special cases of wrongs in the production of representations in virtual worlds. Games are indeed forms of endorsing certain norms (Ostritsch), because, much like other artwork, games do not simply depict events, but provide interpretations and transport attitudes. And so do players' actions. Moreover, the enjoyment of certain imagery does reveal flaws of sensitivity and sympathy (Patridge), because in such

<sup>25</sup> Even if this level of inattention seems to be a flaw of its own, it might—as Kathi Jenni (Jenni, 2003) argued—result in further moral issues ranging from bad consequences to development of vices and breaches of autonomy.

enjoyment a person expresses something as enjoyable<sup>26</sup> and can err, deceive or bullshit herself.

Moreover, the present version of the representational model explains why the analogical model often generates correct results concerning the wrongness of actions in virtual environments. As a reminder, the analogical model is based on the plausible intuition that what a user in a virtual environment does must be closely related to and explained by the event on the screen. It puts this intuition to use by assigning the user's action a type in analogy to the action type of the events depicted on screen and adopting the moral valence these would deserve if they happened between flesh and blood human beings. This model often generates correct results, because in many cases, the depiction of morally dubious acts happens in an uncritical, sometimes even in an explicitly endorsing way. The depiction of murder in virtual environments is all too often not conducted with an intention and a style which parodies, criticizes, decries violence. But depicting murder without such a style and intention seems to be a problematic move in moral discourse. It withholds endorsement from a norm worth endorsing (murder is wrong) or even endorses a morally corrupt attitude (murder can be fun). What the person does in the misnamed virtual murder is thus not a special type of murder, but an endorsement of a problematic norm concerning murder.

The analogical model does not, however, generate correct results reliably. It does not do so because it is quite possible to engage in depiction of murder in ways, which are not expressing or endorsing morally corrupt attitudes or norms. There is a huge leeway for the attitudes and norms expressed and for the style of expression, leaving open among other things stern criticism, parody, simple abstention on one side and endorsement, cheer, or praise on the other. The representational model can distinguish between a user engaging in a new version of *Guernica* (say a VR-experience depicting the horrors of an airstrike in Yemen today) and a user playing enthusiastically through some battlefield simulator modified to make the imagery extra gory. The analogical model cannot, because it does not account for style of representation and attitude expressed.

The representational model fits our stance towards art and media much better than the alternative. We do tend to take a critical stance towards the positive depiction of war in arts, but we can still distinguish a respectful depiction of military action such as that in Heinlein's books from simplistic glorification of violence. In both cases, our

evaluation of the piece of art does not merely depend on the fact that war and violence are depicted, but on how they are depicted. If the depiction of murder in virtual environments were wrong because of its analogy to murder, the style of depiction would hardly matter. Consequently, the acceptable content for virtual environments, and by analogy that for arts and crafts would be excessively small. We would have to condemn the majority of artists across history as committing narratory, canvassy, theatrical, cinematic, sculptural etc. murder. If, however, actions in virtual environments and other media and arts do not inherit their moral standing from the depicted actions, but from the style of representation and attitude expressed, no comparable reduction on acceptable content of the art, of story-telling, role-playing, video-gaming and VR-systems follows.

### An easy way out of moral critique?

A possible counter-argument points to the fact that including style and intention of the author or recipient will allow too much leeway for depiction of violence. Will it not be possible for authors to make threadbare excuses for their representational actions? This risk is quite real. There are a number of games on the market which fail to meet even the minimal standards of acceptable representation of violence. Some of those are justified by their producers as for example reclaiming video games as “a rebellious medium” or criticism of our culture of political correctness.<sup>27</sup> Gamers can even play tamer games in a most disturbing style, reveling in unnecessary violence—as described above for the case of *WoW*—or even installing modifications which make the whole imagery gorier.

However, the leeway in expression is not unlimited. The representational model is capable of identifying actions in virtual environments as morally blameworthy even against flimsy justifications. Style and intention are not purely subjective. The expression of an attitude and the style of that expression are subject to intersubjective standards of interpretation. As Patridge pointed out: many of the meanings of video game imagery are incorrigibly social. Lame excuses by game producers or gamers notwithstanding, it is quite possible to identify such actions as morally blameworthy, even without referring to them as ‘virtual murder’. They are glorifications of murder.

In addition, we can draw a distinction between creating representations in an interactive virtual environment and

<sup>26</sup> Actually, one does not immediately express that something is enjoyable by enjoying it. If one encounters a certain experience for the very first time, immediately enjoys it, and then makes up one's mind that it is not to be enjoyed, nothing much has been expressed by the enjoyment up to this point. If, however, one keeps returning to this type of experience, the expressive component arises quite fast.

<sup>27</sup> Similar phenomena have been observed for other media: books, movies, board games, roleplay games can and often have failed standards of the depiction of violence and often have been defended by publishers by similar phrases.



in one used by a single user alone. While even in the case of a lone user the depiction of actions in a certain style is a morally evaluable action itself, in shared environments it becomes a speech act by gaining a perlocutionary force, i.e. a social effect (Powers 2003). The latter will aggravate the moral evaluation of blameworthy acts of representation, but not change their moral evaluation in any principled way.

## Results

I have suggested that actions in virtual worlds should be interpreted according to a representational model. Using the representational model, I claimed, generates metaphysical and ethical benefits. The metaphysical benefit is a more adequate assignment of action types to the behavior of users of virtual environments. Their actions are first and foremost the creation and modification of data-structures and the resulting output in computer hardware, or more generally the creation and modification of representational content by means of some medium. Such modifications of data structure and content can be performed with different intentions and different styles, which will influence the type and moral evaluation of a user's actions. Only in rare cases such as 'virtual sculpting' or 'virtual reading' will an action in a virtual environment be of the same type as the action depicted on the screen would be if it happened in our non-virtual environment.

The ethical benefit is generated by allowing a more complex analysis of the moral reasons for praiseworthiness or blameworthiness of actions in virtual environments. While the alternative, the analogical model, simply imports moral evaluation from action types beyond the virtual environment, the representational model allows for an analysis which takes not just harm and effects on character into account but the peculiar ways in which speech acts can be morally wrong: e.g. agitatory, deceptive, bullshitting.

**Acknowledgements** I'd like to thank several colleagues and friends for helpful comments, in particular Sebastian Ostritsch, Anneli Jefferson and my colleagues at the Institute for Ethics in the Neurosciences Bert Heinrichs and Charles Rathkopf as well as two anonymous reviewers for this journal.

**Author contributions** Single author.

**Funding** Open Access funding enabled and organized by Projekt DEAL. No extra funding was received for this research.

**Data availability** Not applicable.

**Code availability** Not applicable.

## Compliance with ethical standards

**Conflict of interest** The authors declare that they have no conflict of interest to disclose.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Ali, R. (2015). A new solution to the gamer's dilemma. *Ethics and Information Technology*, 17(4), 267–274. <https://doi.org/10.1007/s10676-015-9381-x>.
- Bailenson, J. (2018). *Experience on demand. What virtual reality is, how it works, and what it can do*. New York: W. W. Norton.
- Baron, M. (2002). Acting from duty. In A. W. Wood (Ed.), *Immanuel Kant: Groundwork for the metaphysics of morals*. New Haven/London: Yale University Press.
- Bartel, C. (2012). Resolving the gamer's dilemma. *Ethics and Information Technology*, 14(1), 11–16. <https://doi.org/10.1007/s10676-011-9280-8>.
- Bermúdez, J. L. (2000). Self-deception, intentions, and contradictory beliefs. *Analysis*, 60(268), 309–319. <https://doi.org/10.1111/1467-8284.00247>.
- Caillois, R. (1961). *Man, play, and games*. New York: Simon & Schuster.
- Chalmers, D. J. (2017). The virtual and the real. *Disputatio*, 9(46), 309–352.
- Coeckelbergh, M. (2007). Violent computer games, empathy, and cosmopolitanism. *Ethics and Information Technology*, 9(3), 219–231. <https://doi.org/10.1007/s10676-007-9145-3>.
- Davidson, D. (1982). Paradoxes of Irrationality. In R. Wollheim & J. Hopkins (Eds.), *Philosophical essays on freud* (pp. 289–305). Cambridge: Cambridge University Press.
- Dibbell, J. (1993). A rape in cyberspace. *The village voice*. Oxford: Blackwell.
- Dunn, J. (2012). Virtual worlds and moral evaluation. *Ethics and Information Technology*, 14(4), 255–265. <https://doi.org/10.1007/s10676-012-9298-6>.
- Frankfurt, H. (1986). On Bullshit. *Raritan Quarterly Review*, 6(2), 81–100.
- Goldstone, R. L., & Son, J. Y. (2012). Similarity. *The Oxford handbook of thinking and reasoning* (pp. 155–176). New York: Oxford University Press.
- Goodman, N. (1972). *Problems and projects*. Indianapolis/New York: Bobbs-Merrill.
- Handke, P. (1966). *Publikumsbeschimpfung und andere Sprechstücke*. Frankfurt am Main: Suhrkamp.
- Huff, C., Johnson, D. G., & Miller, K. (2003). Virtual harms and real responsibility. *Technology and Society Magazine, IEEE*, 22, 12–19. <https://doi.org/10.1109/MTAS.2003.1216238>.
- Huizinga, J. (1949). *Homo ludens*. London/Boston/Henley: Routledge & K. Paul.

- Jenni, K. (2003). Vices of inattention. *Journal of Applied Philosophy*, 20(3), 279–295. <https://doi.org/10.1046/j.0264-3758.2003.00253.x>.
- Juul, J. (2004). *Half-real: Video games between real rules and fictional worlds*. Cambridge: MIT Press.
- Kenny, A. (1963). *Action, emotion and will*. London and New York: Routledge.
- Luck, M. (2009). The gamer's dilemma: An analysis of the arguments for the moral distinction between virtual murder and virtual paedophilia. *Ethics and Information Technology*, 11(1), 31–36. <https://doi.org/10.1007/s10676-008-9168-4>.
- McCormick, M. (2001). Is it wrong to play violent video games? *Ethics and Information Technology*, 3(4), 277–287. <https://doi.org/10.1023/a:1013802119431>.
- Nguyen, C. T. (2019). Games and the art of agency. *The Philosophical Review*, 128(4), 423–462. <https://doi.org/10.1215/00318108-7697863>.
- Ostritsch, S. (2017). The amoralist challenge to gaming and the gamer's moral obligation. *Ethics and Information Technology*, 19(2), 117–128. <https://doi.org/10.1007/s10676-017-9420-x>.
- Parfit, D. (1984). *Reasons and persons*. Oxford Oxfordshire: Clarendon Press.
- Patridge, S. (2010). The incorrigible social meaning of video game imagery. *Ethics and Information Technology*, 13(4), 303–312. <https://doi.org/10.1007/s10676-010-9250-6>.
- Patridge, S. (2013). Pornography, ethics, and video games. *Ethics and Information Technology*, 15(1), 25–34. <https://doi.org/10.1007/s10676-012-9310-1>.
- Polman, H., de Castro, B. O., & van Aken, M. A. (2008). Experimental study of the differential effects of playing versus watching violent video games on children's aggressive behavior. *Aggressive Behaviour*, 34(3), 256–264. <https://doi.org/10.1002/ab.20245>.
- Powers, T. M. (2003). Real wrongs in virtual communities. *Ethics and Information Technology*, 5(4), 191–198. <https://doi.org/10.1023/b:Etin.0000017737.56971.20>.
- Schulzke, M. (2010). Defending the morality of violent video games. *Ethics and Information Technology*, 12(2), 127–138. <https://doi.org/10.1007/s10676-010-9222-x>.
- Seddon, R. F. J. (2013). Getting 'virtual' wrongs right. *Ethics and Information Technology*, 15(1), 1–11. <https://doi.org/10.1007/s10676-012-9304-z>.
- Sheng, E. (2020). Why Is Virtual Wrongdoing Morally Disquieting, Insofar As It Is? *Practical Ethics*. Retrieved from <http://blog.practicalethics.ox.ac.uk/2020/03/oxford-uehiro-prize-in-practical-ethics-why-is-virtual-wrongdoing-morally-disquieting-insofar-as-it-is/>.
- Sorensen, R. A. (1985). Self-deception and scattered events. *Mind*, 94(373), 64–69.
- Sparrow, R. (2017). Robots, rape, and representation. *International Journal of Social Robotics*, 9(4), 465–477. <https://doi.org/10.1007/s12369-017-0413-z>.
- Suits, B. (2005). *The grasshopper: Games, life and utopia*. Ontario/New York: Broadview Press.
- Tillson, J. (2018). Is it distinctively wrong to simulate doing wrong? *Ethics and Information Technology*, 20(3), 205–217. <https://doi.org/10.1007/s10676-018-9463-7>.
- Ulbricht, S. (2020). *Ethik des Computerspielens*. Stuttgart: Metzler.
- Wolfendale, J. (2007). My avatar, myself: Virtual harm and attachment. *Ethics and Information Technology*, 9(2), 111–119. <https://doi.org/10.1007/s10676-006-9125-z>.
- Young, G. (2013). Enacting taboos as a means to an end; but what end? On the morality of motivations for child murder and paedophilia within gamespace. *Ethics and Information Technology*, 15(1), 13–23. <https://doi.org/10.1007/s10676-012-9306-x>.
- Young, G. (Ed.). (2018). *Resolving the gamers dilemma examining the moral and psychological differences between virtual murder and virtual paedophilia*. Cham: Palgrave Macmillan.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.