http://www.eludamos.org

A Conceptual Critique of the Use of Moral Disengagement Theory in Research on Violent Video Games
Jens Kjeldgaard-Christiansen

A Conceptual Critique of the Use of Moral Disengagement Theory in Research on Violent Video Games

JENS KJELDGAARD-CHRISTIANSEN

I am playing the first-person shooter (FPS) video game *Left 4 Dead* (Valve South 2008) on my Xbox 360 video game console. In *Left 4 Dead*, the player takes the role of a survivor of a zombie pandemic who, together with three other player or non-player character (NPC) survivors, must escape the rabid undead. The first level of the game places me in a deserted urban environment at night. Traversing the streets and buildings, my fellow survivors and I encounter hordes of rotting zombies. All of them are mindlessly aggressive, hungering for human flesh and thirsting for blood.

I am heavily armed and could blast my way through them. In fact, I will need to do that to stand a chance of making it to the Safe Zone. But should I? After all, these are living creatures. Had I conscientiously introspected about my situation, I would have known and felt it in myself that it is wrong to kill them. The game, however, does not invite such moral scrutiny. In fact, it actively seeks to disengage the player's moral concern to allow for guilt-free fun: I seem to be fighting zombies, not people. The dehumanized others do not seem to have a strong claim on my sympathy. In addition, my fellow survivors are untroubled by the mayhem. Killing the creatures cannot be bad if they are doing it, too. Perhaps they are unconcerned because the zombies often attack first and have thus demonstrated their murderous intent. Finally, the game's simple narrative premise—escape or die trying—provides a ready excuse for my immoral behavior. I should know better, but the manipulation is successful. I mow down the creatures without feeling as much as a pang of guilt.

If this account of my first experience with *Left 4 Dead* sounds eccentric or even unrealistic, then that, I suggest, is because it is. Within the game's virtual world, *Left 4 Dead* does not deviously disengage my sympathy for my antagonists by *representing* or *dehumanizing* them as unfeeling zombies. They just are unfeeling zombies. It does not *construe* the action as justified self-defense. It just is justified self-defense. The game puts the player in a scenario that demands violent, but morally straightforward, means. There is no devious manipulation or reinterpretation involved. My moral concern for the zombie antagonists of the game is never engaged to begin with, and hence it cannot be disengaged.

Despite its unrealism, my imagined gameplay experience with *Left 4 Dead* is assumed to be typical in an emerging literature on moral disengagement factors in violent video games. This line of research began in the mid-2000s with explorative work on how players might disengage from virtual immorality (Klimmt et al. 2006) and has recently broadened to more systematic work on the specific "cues" embedded in violent video games to morally disengage players, enabling them to perform guilt-free virtual violence (e.g., Hartmann et al. 2014; Hartmann and Vorderer 2010; Klimmt et al. 2008; Moore 2015). Work in this tradition has positioned itself as explanatory input

to the pervasive debates about the consequences of playing violent video games. Contributing researchers suggest that moral disengagement in video games may train the player also to morally disengage from immoral conduct in real life. For example, Gabbiadini et al. (2013) propose that "individuals high in moral disengagement might view reprehensible behaviors in a video game as acceptable or justifiable and that this might leak over to the real world" (p.660).

This article presents a conceptual critique of the emerging literature on moral disengagement in violent video games. I will start by briefly outlining the theoretical foundation of the literature: Albert Bandura's psychological notion of moral disengagement. I will then discuss how the notion has been problematically co-opted by media researchers investigating violence in video games. My argument will be that the literature misapplies moral disengagement theory because the games typically considered in the literature do not *represent* immoral actions as justified, which the theory requires. Instead, they just *present* morally justified violent actions. A faulty game ontology leads the researchers to see the setting, narrative, and concrete situations of violence in video games as misrepresentative of a deeper, underlying reality in which players enjoy violent video games simply because violence is fun. This misconception is made the theoretical fulcrum of empirical claims about the functions and effects of specific "disengagement cues" in video games. When the misconception is replaced with a more psychologically realistic conception, these derivative claims, too, are called into question. I conclude that these issues render many of the findings on moral disengagement in violent video games inconclusive and, in many cases, uninterpretable. The article therefore also challenges the literature's claim to meaningfully inform the debates surrounding violent video games.

My argument can be located within what Egenfeldt-Nielsen et al. (2016) term the "active user" perspective on video games and their risk factors. This perspective emphasizes players' involved meaning-making in actual gameplay scenarios. It is contrasted with the "active media," or "media effects," perspective, according to which players are passively receptive and vulnerable recipients of media messaging. I hope this article strengthens the general case for the "active user" perspective by exposing the fallacious reasoning that leads to its absence in the moral disengagement literature.

Moral Disengagement

Bandura's influential notion of moral disengagement proceeds from his social-cognitive theory of moral thought and action (1991). This theory posits that purely cognitive accounts of moral agency (e.g., Kohlberg 1971, 1981) are limited because the fact that someone thinks they are acting morally does not mean they in fact are acting morally. As Bandura notes, many, if not most, immoral actions are morally motivated or at least feel morally justified to the person committing them (e.g., Fiske and Rai 2015). By contrast, Bandura's social-cognitive theory emphasizes the complex interplay between thought, action, and social situation. It assumes that actions can be immoral even if the perpetrator feels morally justified in performing them. Thus, if people are motivated toward some immoral action, they may *reinterpret* or *reconstrue* the situation in a morally forgiving light. For example, if I feel

like stealing a new iPhone, I could tell myself that Apple has so much money that the act would not put a noticeable dent in the company's revenue, but that having the iPhone would be of real value to me. By focusing on the relative value of the iPhone to me and to Apple respectively, I am able to override my basic moral conviction that it is wrong to steal. In Bandura's dense terminology, I deactivate my inhibitive self-sanction toward my instrumental immoral conduct.

It is not always solely up to the offender to self-justify. As Bandura (1991) explains, other people and even specific situations can facilitate the deactivation of the individual's self-sanctions by representing the immoral conduct in a morally acceptable guise. For example, genocidal tyrants may dehumanize their victims by circulating propaganda that depicts them as uniformly bad, such as was done against the Jews in Nazi Germany. If otherwise decent people are influenced by the misrepresentation, they will find it much easier to squelch the inner moral voice that would normally prevent them from hurting others. They will come to act in ways that they would never endorse if given a full and unbiased account of their actions and the consequences of their actions.

Bandura (1991, 1999, 2002) terms the deactivation of held moral standards *moral disengagement*. He discusses eight different mechanisms of moral disengagement that may be activated from within or facilitated from without: *moral justification*, *euphemistic labelling*, *advantageous comparison*, *displacement of responsibility*, *diffusion of responsibility*, *disregard or distortion of consequences*, *dehumanization*, and *attribution of blame*. I will discuss several of these in what follows.

Video Game Research and Moral Disengagement

Like researchers of violence in film and television (e.g., Krakowiak and Tsay-Vogel 2013), researchers of violent video games have co-opted the notion of moral disengagement to explain how people can enjoy media violence. The explanation runs as follows: People play violent video games because they want enjoyment from these games' involved and highly stimulating gameplay (e.g., Hartmann and Vorderer 2010; Klimmt et al. 2006). However, the violence may also lower enjoyment by challenging players' moral standards (Hartmann and Vorderer 2010; Raney and Bryant 2002). We sense that it is wrong to injure or kill our virtual opponents, and this produces feelings of "distressful concern" and guilt (Hartmann and Vorderer 2010, p.97). Our enjoyment of the game is therefore in conflict with our moral standards. Players of violent video games may seek to mitigate this conflict by employing moral disengagement strategies. Game designers are also aware of the conflict and look for ways by which to disengage players' moral inhibitions (Hartmann and Vorderer 2010, p.99). They make games that maximize players' enjoyment of virtual violence while, and by way of, minimizing players' moral qualms about such violence. Existing research has thus approached the topic from the perspectives of both the players and the creators of violent video games. Recent research has tended to focus on the latter, investigating how video games may be designed to morally disengage players. Though my critique will often straddle both strands, it will primarily target the design perspective.

The most studied moral disengagement factor is moral justification, by which "an otherwise reproachable violent act is interpreted as serving some higher purpose (e.g., killing for a greater good or religious purpose)" (Hartmann 2017, n.p.). My critique will focus on this disengagement factor because it is directly and fully dependent on the fallacy I am highlighting in this article. To give the reader some sense of what is meant by moral justification, as well as the type of research that characterizes the field in general, I will briefly recount a key experimental manipulation from the most cited study (Hartmann and Vorderer 2010) on moral disengagement in violent video games. Participants in the study played a modified level from *Operation Flashpoint: Cold War Crisis* (Bohemia Interactive Studio 2001), a popular FPS game.

> The audiovisual movie-sequence showed a torture camp in the (fictional) Oka region where innocent people were murdered by paramilitary forces. In the remainder of the introduction and depending on the experimental condition, subjects either learned that they would play a soldier of the United Nations (UN), about to attack the torture camp to restore humanity (justified action), or to play a soldier of the paramilitary forces that would continue their cruelty and defend the camp (unjustified action). (Hartmann and Vorderer 2010, p.107)

Unsurprisingly, participants in the "justified" condition—those who attempted to save the innocent victims of torture—felt less guilty about the play session than those who defended the torture camp. The authors concluded that this is because the diegetic setup seemed to provide a justification for the violence. Thus, "the game creates a situation that automatically leads to cognitive disengagement from inner moral standards" (Hartmann and Vorderer 2010, p.98), which would normally bar players from aggressing against real and virtual beings. This example, together with my opening example of *Left 4 Dead*, will serve as reference points for the following critique.

A Conceptual Critique

The core of my critique of the literature on moral disengagement and violent video games is that it ignores the fundamental distinction between reality and representation that moral disengagement theory presupposes. Moral disengagement theory deals with immoral conduct that has been cognitively represented or reconstrued to appear and feel moral. If you falsely represent someone as a cold-blooded killer to justify violence against him or her as self-defense, then you have successfully morally disengaged from your actions. However, if that someone just *is* a coldblooded killer and you are responding accordingly, then, by definition, you have not morally disengaged: there is no misleading representation in play. Bandura himself (Bandura et al. 1996) notes this basic distinction: "People do not ordinarily engage in reprehensible conduct until they have justified to themselves the rightness of their actions. *What is culpable* can be made righteous through cognitive reconstrual" (p.365; my emphasis). The justification must mask a real transgression. "Moral disengagement" names the cognitive process of minimizing or reinterpreting the transgression.

How is the representation/reality distinction supposed to work in a violent video game like *Left 4 Dead*? From which underlying reality is the player disengaging and thereby violating his or her own inner moral standards? The events of *Left 4 Dead* do not happen, and presumably could not happen, in the real world. The real world has no zombies. Instead, the game constitutes a self-enclosed virtual world—a "magic circle" (Huizinga 1970)—in which braindead, rabid zombies actually exist. There is no sense, therefore, in which the game reinterprets the zombies' bloodlust or the player's (or player character's) actions. They just are what they are. This observation would seem to block the application of moral disengagement theory: the term presupposes a distinction between representation and reality that does not apply to games like *Left 4 Dead*.

Remarkably, this fundamental conceptual problem is nowhere acknowledged in the literature. The representational conception is instead uncritically assumed (emphases mine):

> The findings of previous studies suggest that violence in video games is often *portrayed* as justified. (Hartmann et al. 2014, p.315)

> Harming quasi-social characters could be perceived as wrongdoing. But, according to the moral disengagement perspective, features of the game ensure that it is not. Instead, as in other contexts that spur violence in real life, because of cues within video games, "behavior that is ordinarily viewed as unacceptable (killing social beings) is *redefined* as justified and desirable" (Klass, 1990, p.403). (Hartmann and Vorderer 2010, pp.98-99)

> Moral disengagement cues, such as a good reason to fight (e.g., to save the world), particularly against nonanthropomorphic creatures (e.g., aliens), may *frame* violence against game characters as acceptable. (Hartmann and Vorderer 2010, p.99)

> While playing violent video games, the *player typically sees him- or herself* "as a human perpetrator engaging in repeated acts of justified violence" (Smith, Lachlan, & Tamborini, 2003, p. 60). (Greitemeyer and McLatchie 2011, p.659)

> In the process of moral disengagement, individuals reframe reprehensible acts against others in a way that makes them *appear* worthy, just, necessary, or inconsequential. (Hartmann et al. 2014, p.312)

Invariably, the suggestion is that these "framings," "representations," "appearances," "portrayals," "redefinitions," and "construals" work to disengage players from their own moral standards. The game world is a polygonal cover-up of an underlying reality in which players kill virtual beings because killing is fun. The problem with this view, again, is that these games do not *re*present their events and stories. Instead, they just present to the player an encompassing diegesis in which the aliens really do want to destroy humanity, or in which you really are fighting immoral torturers. As Grodal (2003) points out, video games are typically experienced as unfolding in a "progressive present" (p.134) from the player's first-person perspective. They are not experienced as retellings or representations of some distant scenario, but as an involved experience dynamically unfolding around and through the player in the here and now. (I shall later consider the possible counterexamples of historical games.)

In one respect, however, the studies do acknowledge players' direct and involved engagement with the fictional worlds of video games. Most studies advance preliminary arguments to the effect that players relate to virtual characters as though these characters were real, or at least "quasi-social," citing studies into humans' automatic attributions of agency and sociality (e.g., Hartmann and Vorderer 2010, pp.95-97). They use this observation to argue that the games may sometimes activate players' moral intuitions. This immersed mode of fictional engagement, however, presumably has the consequence that players' in-game actions must be morally justified *against the background of the fictional world itself*. Why should only the virtual act, but not the virtual context of the act, matter? Is there really no material moral difference between "shooting a benevolent virtual being" and "shooting a malicious virtual being that is out to kill innocent people"? From the perspective of the moral disengagement tradition, these characters and their actions are somehow assessed completely devoid of context. Nowhere is it explained how this is supposed to work. In fact, the literature on agency and sociality attribution to fictional characters does not support the disengagement perspective. Research starting with Heider and Simmel's (1944) classical paper shows that our social understanding and moral judgments of characters target their motives and actions in the story world (Mar and Oatley 2008 present an overview). If the supposition is that we relate idiosyncratically to the digitized story worlds of video games, then arguments are needed to support that alternative conception.

I want to suggest that the disengagement perspective reduces the game world to devious misrepresentation because it conflates diegetic and non-diegetic forms of representation. It is true that video games incorporate forms of non-diegetic representation. As already noted, a video game may represent a virtual world in its software, and players can certainly interact with that representation. In addition, the graphical assets and sounds of that world are typically approximative representations of objects and sounds found in the real world. All of this, however, does nothing to suggest that the game represents or reinterprets its own story world. It does not; the world is just as real, unreal, or quasi-real, as its characters.

Conflating diegetic and non-diegetic modes of representation is a serious conceptual confusion. It may suggest that the game worlds are representative of some underlying or noumenal reality against which the moral import of the player's actions should really be measured. Hartmann and Vorderer (2010) assume this in stating that

> If users' automatic protections against violations of internal moral standards occasionally fail and guilty or remorseful feelings arise (cf., Haidt, 2001), they can still reframe their wrongdoing and regulate their dissonant state. To fuel this conscious moral rationalization (Tsang, 2002), users can actively recall that they are merely playing a game or that they are fighting for justice. (p.97)

The study references games in which players are situated as actually fighting for good within the virtual world. If an evil alien race or genocidal warlord came to destroy my city, or country, or planet, I would want people to fight back. I would morally praise resistance—violent if necessary—as I imagine most people would. Why should the same response undergo normative reversal in the virtual world? We are never told.

This discussion is not just an inconsequential academic exercise. The misconception that violent video games sell violence *tout court* leads to the false and morally fraught empirical claim that the diegetic contextualization of that violence minimizes players' moral investment. The researchers start from the assumption that violence is bad and that media representations that afford or encourage virtual violence are therefore also bad. The virtual mediation of the violence is taken as the technological equivalent of the psychological processes of moral disengagement because both are seen as forms of *rationalizing misrepresentation*. This false analogy is what enables the researchers to draw on the moral disengagement approach: "These narratives provide an easy way for players to disengage from their own moral sanctions regarding shooting or killing opponents" (Lin 2011, p.695).

On this view, the violent video game stands to actual violence as the unreliable (and morally suspect) narrator of a story stands to the actual events of that story. The equivocation is deeply misleading. Ironically, most violent video games seek primarily to engage, rather than to disengage, players' moral psychology. By typically presenting virtual worlds in which violent opposition is actually just, and not misrepresented in a more acceptable light, violent video games recruit players' moral minds to motivate their involved engagement. As Riddle et al. (2018) note, players typically want to play as the good guy and fight the bad guy(s). In *Halo 3* (Bungie 2007), for instance, they do so by fighting as an elite soldier against alien forces that aim to exterminate the entire human species. Over the course of the campaign, players will fight the aliens through violent but necessary means, aiding numerous embattled compatriots in the process. There is no reason to think that the players are somehow being misled into thinking that this is what they are doing. Within the virtual world of the game, this *is* what they are doing. Players are offered an integrated moral cause, not a superimposed and misleading moral justification.

By contrast, the disengagement perspective considers violent video games as essentially representations of a single, abstract theme: violence, which is a bad thing. This abstraction is then posited as the underlying reality that moral disengagement theory presupposes: "The present findings suggest that FPS games commonly depict violence as morally acceptable behavior" (Hartmann et al. 2014, p.328). The virtual violence is viewed as something logically prior to, and somehow separate from, its situation and purpose. This fallacy leads the researchers to ask how it is possible that players could enjoy violent gameplay. They do not ask why, *under a fuller description of precisely the same virtual actions*, players enjoy fighting injustice or saving the world.

> The question is why virtual violence obviously is enjoyable for many players. (Hartmann and Vorderer 2010, p.97)

> Accordingly, particularly more empathic users tend to feel guilty about harming (seemingly social) video game characters (e.g., Hartmann et al., 2010; Lin, 2011). But why do most users still experience video game violence as fun? (Hartmann et al. 2014, pp.311-312)

> In the context of violent media entertainment […] violence may even function as [a] source or catalyst of positive experiences. The concept of moral management tries to resolve this contradiction. (Klimmt et al. 2006, p.312)

But there is no contradiction. Players typically enjoy fighting for good outcomes, and they typically dislike harming innocent characters or fighting for an unjust cause. That is the story told by the data (Gollwitzer and Melzer 2012; Hartmann et al. 2010; Lin 2010; Tamborini et al. 2013; Weaver and Lewis 2012). This result, however, is neither surprising nor, I submit, particularly worrying. We might even rejoice to learn that players generally prefer to do the right thing. The result only becomes disconcerting when one adds the gratuitous assumption that players are somehow actually doing evil despite their own convictions, or at least that they are violating their "inner moral standards" (Hartmann and Vorderer 2010, p.97), as it is usually put.

What about violent games that seemingly aim to represent the real world in explicitly claiming or presentationally implying historical accuracy? For example, the popular *Battlefield* series of online FPS games can be seen to advertise authenticity through the games' referential titles (e.g., *1942* and *Vietnam*) and correlative scenery and weapons. In these cases, the moral disengagement approach may get some conceptual traction. Insofar as these games are taken to represent historical events and characters, they can be accused of misrepresenting them for the sake of removing immoral, unpleasant, or even just onerous aspects of the real deal. (Of course, existing research has not shown that players approach these games with the callow conception that they tell the full story about what it is like to be in a war). One could run with this argument to suggest that even games that do not represent historical facts still represent and potentially distort facts about what it is like to be in a war or hurt someone, or simply to fire a gun—simulative aspects of gameplay (Fullerton 2008). Nevertheless, such an argument would be silent on the disengagement cues discussed above, which target not the violence itself but its diegetic justification.

According to research in the moral disengagement tradition, then, violence in video games is always and necessarily senseless. The way in which this violence is "communicated" (Hartmann et al. 2014, p.310), that is, narrated and transformed into gameplay, is reduced to misleading appearance—tacked-on excuses and guileful artifices. Indeed, it is viewed as precisely the kind of rationalizing misrepresentation that moral disengagement researchers seek to expose. I hope to have shown that this is an erroneous and deeply fraught way of describing how the typical violent video game works. An inadequate game ontology misleads the researchers into pathologizing players' preference to act morally rather than immorally as moral disengagement.

Moral Disengagement Reinterpreted

To players, the story worlds of video games are directly presentational rather than deviously misrepresentational in nature. A failure to acknowledge this fact leads some researchers to mischaracterize how players conceive of virtual violence. A conception that recognizes the presentational significance of narrative and contextualization in players' experience suggests reinterpretations of many supposed cues to moral disengagement. I will now exemplify how with reference to Hartmann et al.'s (2014) pertinent content analysis of violent video games. Participants in the study coded short segments of 17 popular violent FPS games to explore the prevalence of a range of moral disengagement cues. The researchers, however, do

nothing to show that the cues they identify—including cues to moral justification, euphemistic labeling, diffusion of responsibility, distortion of consequences, dehumanization, and attribution of blame—do in fact morally disengage players and thereby increase enjoyment. The labelling of a disengagement cue, or "indicator," rests on what I have argued to be mistaken assumptions about how players approach the diegetic worlds of video games. Do a significant number of these designations admit of alternative interpretations?

Euphemistic labeling ("language, sounds, or graphical depictions that make harmful conduct respectable if not humorous," Hartmann et al. 2014, p.319): This type of cue references sanitized intradiegetic references to violence, such as when players are tasked with "taking out" or "dispatching" opponents instead of "killing" them. The sanitization is supposed to alleviate players' moral concerns about their violent actions. A problem with this interpretation is that, in many cases, a more plausible interpretation cites game designers' aim to craft a believable fictional setting. Perhaps, in war games such as those in the *Call of Duty* and *Halo* series, the reason for this sanitization has nothing to do with morally disengaging the player and everything to do with attempting a realistic portrayal of how soldiers would talk about war and their roles in it. Does it really make sense to class an instance of such sanitized language as a moral disengagement cue if both game designers and players would perceive its absence to be out of character and immersion-breaking, and if players might *thereby* morally disengage? For example, in *Call of Duty: Modern Warfare 2* (Infinity Ward 2009), in the single-player mission "Contingency," the player's superior, Captain Price, instructs the player to "take out" enemy soldiers at several points. The captain is a professional, conservative soldier. It would have been entirely out of place for him to command the player to "kill," or indeed to "murder," the targets. If for no other reason than its realism, the captain's distancing military register promotes the player's immersion and investment in the story world. I note here that the fact of the captain's wording being *realistic*, that is, representative of perceived real-world martial discourse, is not a counter to my argument that the story that frames this violence is not at all representational. The elements of that story naturally have to cohere with our background assumptions about how humans would interact in such circumstances as are specified by the story in order to make sense at all (Ryan 1980; Walton 1990).

Can something be a moral disengagement cue if it was not included to morally disengage the player, and if it subserves critical functions within the overall structure and narrative setting of the game? In other words, can a disengagement cue be an entirely accidental feature? I have not been able to find a clear definition of the term in the literature, which variously and inconsistently draws on intentional and neutral specifications. An intentional specification labels as a disengagement cue any component of a game *designed* to morally disengage the player. A neutral specification labels as a disengagement cue a finite set of abstracted events and actions assumed to morally disengage the player in any narrative and gameplay context. In cases of individual cue specifications, Hartmann et al.'s (2014) review shifts between intentional and neutral criteria:

> Euphemistic labeling means that an act is given a sanitized label in order to make it seem less severe. [intentional criterion]

Advantageous comparison occurs if one act is compared to an even more heinous one perpetrated by an opponent. [neutral criterion]

Moral justification implies that a violent act is redefined as serving some socially worthy or moral purpose. [ambiguous] (p.312)

Future studies owe it to their readers and critics to front and defend their preferred specification. If the preferred specification is the neutral one, as seems most plausible, then it is in tension with the staging of the research, which often emphasizes game designers' conscious efforts to morally disengage players:

Game designers seem to design violent game play to be enjoyably guilt-free. (Hartmann and Vorderer 2010, p.99)

Most violent video games frame the events that occur during their use in specific ways in order to enable and support players to cope with moral concern. (Klimmt et al. 2008, p.115)

Contemporary FPS games are primarily designed to entertain their users and not to represent reality. Following this argument, it seems reasonable to expect that moral disengagement factors are also frequently embedded in other violent video game genres that seek to entertain their users. (Hartmann et al. 2014, p.327)

In addition, a neutral specification of what constitutes a disengagement cue is lacking theoretical justification, as I have argued.

Diffusion of responsibility ("minimizing personal agency for a violent act by placing responsibility on others who were involved, such as soldiers in a squad," Hartmann et al. 2014, p.320). Labeled as a disengagement cue to diffusion of responsibility is any presence of friendly characters partaking in the player's violent actions. The insensitivity of this cue to diegetic context again invites criticism. Are we really to believe that in all cases the presence of allied fighters in a violent game morally disengages players? Is context—what players perceive themselves to be doing in the virtual world—*never* relevant? In *Left 4 Dead*, the player attempts to survive zombie attacks with the help of human or NPC allies. Arguably, the only thing that morally engages players in this scenario is the presence of these allies. The game prompts the player to help these allies when they are attacked, and the player, as the *de facto* leader of the team, may feel bad for failing to deliver all of his or her teammates safely to the next Safe Zone. This may happen, for example, when a single teammate is swarmed by a horde of zombies near the end of a level, and the player must decide between trying to save the teammate, which might get the rest of the team killed, and leaving the swarmed teammate behind to be eaten alive. Note what the study's coders have had to do in this situation, based on the researchers' contextually blind criteria. The coders have witnessed the band of survivors help each other reach the Safe Zone. Then, because the desperate group fights off ravenous, braindead zombies to stay alive, they have classed what is arguably the only morally motivating aspect of the game—the presence of a number of fellow travelers whom the player may want to aid and protect—as a moral disengagement cue.

Distortion of consequences ("minimizing, ignoring, distorting, and/or disbelieving the effects of immoral actions," Hartmann et al. 2014, p.320). This category labels as

disengagement cues gameplay and presentational elements of violent video games that appear to eschew the true consequences of violent actions, such as bloodshed, suffering, and fear. The researchers' perspective predicts that these displays would only become rarer: Players enjoy guilt-free violence, and game designers should accommodate that preference by minimizing the unpleasant consequences of violence. In fact, the trend has been the reverse. Advances in AI and injury modelling make the virtual enemies respond powerfully, and often quite realistically, to the player's actions; they may flee with fear when outnumbered or walk with a limp when shot in the leg. Virtually all violent video game series have, if anything, seen increases in explicitly violent content over time. This trend includes such massively violent and hugely popular franchises as *Doom*, *Mortal Combat*, and *Grand Theft Auto*. Exceptions to the trend exist, of course, but we need not always assume a direct preference for sanitized violence in order to explain them. For example, the Entertainment Software Rating Board (ESRB) rates games based centrally on their level of explicit violence in order to shield children and adolescents from potentially disturbing material. A consequence of these prohibitive ratings may be that game developers sometimes limit their games' explicit violence in order to be able to target a wider audience.

Especially in older games, there is another plausible explanation for the prevalence of "distortion of consequences," of which players are generally aware: technical limitations. Even modern computers are limited in their capacity to render violent virtual interactions, just as they are limited in their capacity to render other kinds of interaction. In *Left 4 Dead*, for example, the bodies of dead enemies will disappear ("despawn") after some time. This is labeled a disengagement cue because seeing the dead bodies would remind players of their immoral actions by evidencing the gruesome consequences of those actions (Hartmann et al. 2014, p.322). However, the bodies could not *remind* players that they have violated their moral principles because, in all likelihood, players have not violated their moral principles. In fact, having dead bodies stay visible in video games is typically considered preferable because their sudden disappearance makes absolutely no sense and may therefore break immersion. The real explanation for the prompt despawns is that removing the bodies from the player's visual field decreases video-memory load, making the game run more smoothly. In the PC version of *Left 4 Dead*, players may adjust how long the dead bodies stay visible by increasing or decreasing the game's graphics settings, to which the duration is keyed. Higher and therefore more desirable graphics settings equal longer despawn times. Therefore, whether the researchers identified this disengagement cue in *Left 4 Dead* as part of their study (we are not told) would have likely depended on their graphics settings, and hence on their hardware, and only secondarily on the game's creative vision.

All of this is not to say that the consequences of violent virtual actions are never distorted to spare players from the distressing reality of violence and thereby increase enjoyment. Such targeted distortion surely happens in some cases. My points are instead the following. First, the disengagement perspective implies a counter-empirical trend, which shows that, generally speaking, it cannot be right. It simply is not true that players, generally speaking, prefer sanitized violence, though they may well prefer certain types of sanitization. Second, instances of this supposed disengagement cue are frequently more readily explained from outside the disengagement perspective. This is not surprising, since, as I have argued, the

typical player will not feel the least bit guilty about "killing" a mindless and bloodthirsty digital zombie. The context of the violence matters.

Dehumanization ("having enemies that either do not possess human qualities or are stripped of human qualities, therefore seeming to possess fewer moral rights," Hartmann et al. 2014, p.320). I have already discussed how, if a game involves shooting monsters and not humans, it is nonsensical to suppose that it is thereby "redefining," "reconstruing," or indeed "dehumanizing" these monsters. There was nothing objectively or subjectively human there in the first place. In addition, the researchers label as a cue to dehumanization "interchangeable enemies […] making it hard to identify unique personalities or faces" (p.321). For example, the zombies in *Left 4 Dead 2* (Valve Corporation 2009), one of the games coded in the study, are often interchangeable. Their interchangeability is assumed to function as a problematic disengagement cue. The trouble with this designation is that the zombies' interchangeability is narratively necessitated. The zombies are not supposed to have individual personalities. If they did, they would not be zombies. Again, the moral disengagement perspective clashes with the diegetic thrust of the game at a fundamental level. It implies that any violent game that features zombies will *by that fact alone* also feature a cue to the player's moral disengagement. This claim again assumes that violence in video games is somehow completely separable from its diegetic embeddedness. Only then can one answer the question, "What is the player disengaging from?" The answer would have to be something like "the true nature of the violence committed, which the game's presentation distorts." But there is no such deeper and truer level.

The same misconception causes the researchers to conceive of virtual beast and monster opponents as objects of dehumanization that "do not possess human qualities or are stripped of human qualities, therefore seeming to possess fewer moral rights" (Hartmann et al. 2014, p.320). The suggestion truly seems to be that the presence of any conceivable non-human antagonist constitutes a cue to moral disengagement by way of dehumanization. This suggestion only makes sense on the assumption that players perceive these quasi-beasts and quasi-monsters as quasi-human beings at some deeper level. If not, there would be nothing human there to dehumanize. Separate from these beings' perceived existence as humans, however, are their stripped-of-human-qualities representations in gameplay as vicious monsters, which make the thought of harming them palatable, perhaps even attractive. This baffling conceptual farrago is the consequence of abstracting the entire diegetic context of the violence away as a devious spin on a deeper reality. Note again the phrasing: "[Dehumanization] may include instances when enemy groups are *portrayed as* animals or beasts, or in a stereotypical way as out-group members" (Hartmann et al. 2014, p.320; my emphasis). Remarkably, this a widely accepted analysis. For example, Lin (2011) applies it uncritically to *Left 4 Dead*, tellingly labeling it the "dehumanization technique" (p.695). Klimmt et al. (2008) state that "*Dehumanization* of victims is also facilitated by most violent video games. For instance, some games (*Resident Evil*, 1996) introduce monsters and other fantasy creatures that differ in some respect from humans" (p.116).

Attribution of blame ("blaming adversaries or circumstances for the violent action in the game," Hartmann et al. 2014, p.322). This final cue was coded if "voiceovers or text commentaries explicitly blamed the enemies or victims for the violent action (e.g.,

'they deserved what they got')" (p.322). It is a fitting cue with which to end because this article has concerned itself with precisely the question of whether video game characters can deserve what they get, or at least whether violent struggle against them is sometimes morally defensible. The disengagement perspective answers both questions with an unqualified "no." Even when the player is trying to save innocent people from sadistic torturers, any hint from the game that the player's actions are good and noble would count as a morally problematic disengagement cue. I have argued that this implausible view rests on the fallacy of assuming that the diegetic worlds of video games are deviously representational rather than directly presentational.

The only example given of a cue to blame attribution (and one of the only instances in which a disengagement cue is linked to a concrete scenario in an identifiable game) once again illustrates a disregard for diegetic context. The example is from the action-horror video game *F.E.A.R First Encounter Assault Recon* (Monolith Productions 2005). Late in the game, a seemingly friendly character states that "it is the nature of monsters to destroy their makers" (Hartmann et al. 2014, quoted on p.326). The friendly character is here referring to a hostile monster character that the player expects to face later on. According to the researchers, the statement of the friendly character is supposed to justify the player's subsequent aggression against the monster by implying its "destructive" blameworthiness. However, the full quotation and diegetic context—unreported by the researchers—make it abundantly clear that the line is neither meant nor actually functions as a moral disengagement cue. The full quotation is spoken by Harlan, the biological father of the so-called monster:

> She was a very disturbed child. Terrible, debilitating nightmares. Hallucinations. Hysteria. She never had a chance at a normal life. It was Disler who realized that she was telepathic. He speculated that she was extremely sensitive to the negative emotions of people around her. We removed her from that chamber only twice. She was fifteen when the first prototype was born. She was supposed to be comatose, but she started screaming when we induced labor and didn't stop until the liquid filled her lungs. There are plausible scientific explanations for everything that followed, but I think it was really just a question of hatred. It is the way of men to make monsters, and it is the nature of monsters to destroy their makers.

Harlan has subjected the "monster," whose real name, as the player learns, is Alma, to countless invasive procedures and near-total social isolation throughout her childhood and beyond. The girl's own father has used her as a human incubator. He made her what she is. The quotation is not meant to position Alma as the game's blameworthy villain, but to make the player wonder who the real villain is—in other words, to *displace* rather than to attribute blame, and to morally *engage* rather than disengage the player. The researchers mischaracterize this message because of their categorical dismissal of diegetic context as mere devious representation.

Conclusion

A video game's being violent does not imply its being antisocial or immoral (cf. Greitemeyer and Mügge 2014). Some researchers fail to see this because they

consider the story worlds of violent video games to be devious misrepresentations of what any violent game is really about: senseless violence. I have argued that this critical misconception motivates their search for psychologically uninterpretable "disengagement cues" that they have stipulated into existence.

At bottom, this argument is about the phenomenology of gameplay, about how players actually experience and understand the games they play. And no regular player of video games needs to be told that the experience of gameplay varies enormously with the type of game in question. In the single-player modes of the FPS games that are usually studied by moral disengagement researchers, story typically matters: It frames the gameplay experience and fixes the moral import of the immersed player's actions. It is just as clear, however, that there are other possible stances to take toward games, and that these stances are often associated with types of gameplay in which diegetic context is deemphasized (e.g., Aarseth 2014, pp.487-490). For example, players may not care anything about the diegetic setting of the multiplayer mode of *Halo 3*, in which they fight the opposing team's player characters in a series of disconnected skirmishes. They may only care whether their team wins. To habitual players of *Halo 3* and similar games, the suggestion that they ought to feel bad for shooting at the opposing team would likely seem ridiculous. They might reply that you could equally well have asked them why they do not feel bad for invading a country in the board game of *Risk*. The comparison is apt because players may conceive of player-controlled characters in multiplayer skirmishes as mere game pieces—tools used by the player in order to play—and not as autonomous agents (Aarseth 2014; Lazzaro 2004; Perron 2005). The competitive interaction is emphasized enough, and the diegesis deemphasized enough, so that the action does not take on any meaning apart from what competitive accolades players assign to its outcome. Thus, the individual player takes a fundamentally different and amoral stance toward the action than he or she might have taken in other types of games. Could this amoral, distanced mode of engagement somehow carry over into interactions with other virtual agents, or even to interactions with real people? Perhaps, though I see little evidence for supposing so. My point, however, is that the moral disengagement approach can have nothing to say about such a fundamental shift in perspective. It cannot even begin to discuss it because, as a first theoretical premise, it dismisses the player's meaning-making as excuse-making.

If moral disengagement researchers truly want to understand how people are affected by violent video games, they will need to take seriously the phenomenology of gameplay through which any effects of such gameplay, good or bad, are presumably mediated (Tamborini et al. 2013, pp.103-104). In this article, I have emphasized the significance of diegetic context. But even in cases where diegetic context may not matter, such as in a multiplayer skirmish in *Halo 3*, there are frequently reasons to doubt that players feel, and indeed that they ought to feel, that they are violating their own moral standards. I encourage future research in the moral disengagement tradition to form a more adequate conception of its complex research object—the game, the player, and their meaningful interaction (Aarseth 2014)—and to reformulate the approach accordingly.

Games Cited

Bohemia Interactive Studio (2001) *Operation Flashpoint: Cold War Crisis*. Codemasters (PC).

Bungie (2007) *Halo 3*. Microsoft Game Studios (Xbox 360).

Infinity Ward (2009) *Call of Duty: Modern Warfare 2*. Activision (PC).

Monolith Productions (2005) *F.E.A.R. First Encounter Assault Recon*. Vivendi Universal Games (PC).

Valve Corporation (2009) *Left 4 Dead 2*. Valve Corporation (PC).

Valve South (2008) *Left 4 Dead*. Valve (PC/Xbox 360).

References

Aarseth, E. J. (2014) Ontology. In Wolf, M. J. P. and Perron, B. (eds.) *The Routledge companion to video game studies*. New York: Routledge, pp. 484–492.

Bandura, A. (1991) Social cognitive theory of moral thought and action. In Kurtines, W. M., Gewirtz, J., and Lamb, J. L. (eds.) *Handbook of moral behavior and development, Vol. 1: Theory*. Hillsdale: Lawrence Erlbaum Associates, Inc, pp. 45–103.

Bandura, A. (1999) Moral disengagement in the perpetration of inhumanities. *Personality and Social Psychology Review*, Vol. 3 (3), pp.193–209.

Bandura, A. (2002) Selective moral disengagement in the exercise of moral agency. *Journal of Moral Education*, Vol. 31 (2), pp. 101–119.

Bandura, A., Barbaranelli, C., Caprara, G. V., and Pastorelli, C. (1996) Mechanisms of moral disengagement in the exercise of moral agency. *Journal of Personality and Social Psychology*, Vol. 71 (2), pp. 364–374.

Egenfeldt-Nielsen, S., Smith, J. H., and Tosca, S. P. (2016) *Understanding video games: The essential introduction* (third ed.). New York: Routledge.

Fiske, A. P., and Rai, T. S. (2015) *Virtuous violence: Hurting and killing to create, sustain, end, and honor social relationships*. Cambridge: Cambridge University Press.

Fullerton, T. (2008) Documentary games: Putting the player in the path of history. In Whalen, Z. and Taylor, L. N. (eds.) *Playing with the past: History and nostalgia in videogames*. Nashville: Vanderbilt University Press, pp. 215–238.

Gabbiadini, A., Riva, P., Andrighetto, L., Volpato, C., and Bushman, B. J. (2013) Interactive effect of moral disengagement and violent video games on self-control, cheating, and aggression. *Social Psychological and Personality Science*, Vol. 5 (4), pp. 451–458.

Gollwitzer, M., and Melzer, A. (2012) Macbeth and the joystick: Evidence for moral cleansing after playing a violent video game. *Journal of Experimental Social Psychology*, Vol. 48 (6), pp. 1356–1360.

Greitemeyer, T., and McLatchie, N. (2011) Denying humanness to others. *Psychological Science*, Vol. 22 (5), pp. 659–665.

Greitemeyer, T., and Mügge, D. O. (2014) Video games do affect social outcomes. *Personality and Social Psychology Bulletin*, Vol. 40 (5), pp. 578–589.

Grodal, T. (2003) Stories for the eye, ear, and muscles: Video games, media, and embodied experiences. In Wolf, M. J. P. and Perron, B. (eds.) *The video game theory reader*. New York: Routledge, pp. 129–155.

Hartmann, T. (2017) The "moral disengagement in violent videogames" Model. *Game Studies*, Vol. 17 (2). Available at: http://gamestudies.org/1702/articles/ hartmann [Accessed: August 2 2018].

Hartmann, T., Krakowiak, K. M., and Tsay-Vogel, M. (2014) How violent video games communicate violence: A literature review and content analysis of moral disengagement factors. *Communication Monographs*, Vol. 81 (3), pp. 310– 332.

Hartmann, T., Toz, E., and Brandon, M. (2010) Just a game? Unjustified virtual violence produces guilt in empathetic players. *Media Psychology*, Vol. 13 (4), pp. 339–363.

Hartmann, T., and Vorderer, P. (2010) It's okay to shoot a character: Moral disengagement in violent video games. *Journal of Communication*, Vol. 60 (1), pp. 94–119.

Heider, F., and Simmel, M. (1944) An experimental study of apparent behavior. *The American Journal of Psychology*, Vol. 57, pp. 243–259.

Huizinga, J. (1970) *Homo ludens: A study of the play element in culture*. London: Maurice Temple Smith Ltd.

Klimmt, C., Schmid, H., Nosper, A., Hartmann, T., and Vorderer, P. (2006) How players manage moral concerns to make video game violence enjoyable. *Communications*, Vol. 31 (3), pp. 309–328.

Klimmt, C., Schmid, H., Nosper, A., Hartmann, T., and Vorderer, P. (2008) "Moral management": Dealing with moral concerns to maintain enjoyment of violent video games. In Jahn-Sudmann, A. and Stockmann, R. (eds.) *Computer games as a sociocultural phenomenon*. London: Palgrave Macmillan, pp. 108– 118.

Kohlberg, L. (1971) Stages of moral development as a basis for moral education. In Beck, C. M., Crittenden, B. S., and Sullivan, E. V. (eds.) *Moral education: Interdisciplinary approaches*. Toronto: Toronto University Press, pp. 15–98.

Kohlberg, L. (1981) *The philosophy of moral development: Moral stages and the idea of justice* (first ed.). San Francisco: Harper & Row.

Krakowiak, K. M., and Tsay-Vogel, M. (2013) What makes characters' bad behaviors acceptable? The effects of character motivation and outcome on perceptions, character liking, and moral disengagement. *Mass Communication and Society*, Vol. 16 (2), pp. 179–199.

Lazzaro, N. (2004) Why we play games: Four keys to more emotion without story. *Proceedings of the Game Developers Conference (2004)*, San Jose, California. Available at: http://twvideo01.ubm-us.net/o1/vault/gdc04/slides/why_we_play_games.pdf [Accessed August 15 2018].

Lin, S.-F. (2010) Gender differences and the effect of contextual features on game enjoyment and responses. *Cyberpsychology, Behavior, and Social Networking*, Vol. 13 (5), pp. 533–537.

Lin, S.-F. (2011) Effect of opponent type on moral emotions and responses to video game play. *Cyberpsychology, Behavior, and Social Networking*, Vol. 14 (11), pp. 695–698.

Mar, R. A., and Oatley, K. (2008) The function of fiction is the abstraction and simulation of social experience. *Perspectives on psychological science*, Vol. 3 (3), pp. 173–192.

Moore, C. (2015) Moral disengagement. *Current Opinion in Psychology*, Vol. 6, pp. 199–204.

Perron, B. (2005) Coming to play at frightening yourself: Welcome to the world of horror video games. *Aesthetics of Play*. Available at: http://www.aestheticsofplay.org/perron.php [Accessed September 3 2018].

Raney, A. A., and Bryant, J. (2002) Moral judgment and crime drama: An integrated theory of enjoyment. *Journal of Communication*, Vol. 52 (2), pp. 402–415.

Riddle, K., Tay, S. K., and Wu, J. (2018) "It lets me fight the bad guys": An exploration into the factors predicting enjoyment of violent video games. *Communication Studies*, Vol. 70 (1), pp. 36–58.

Ryan, M.-L. (1980) Fiction, non-factuals, and the principle of minimal departure. *Poetics*, Vol. 9 (4), pp. 403–422.

Tamborini, R., Weber, R., Bowman, N. D., Eden, A., and Skalski, P. (2013) "Violence is a many-splintered thing": The importance of realism, justification, and graphicness in understanding perceptions of and preferences for violent films and video games. *Projections*, Vol. 7 (1), pp. 100–118.

Walton, K. L. (1990) *Mimesis as make-believe: On the foundations of the representational arts*. Cambridge: Harvard University Press.

Weaver, A. J., and Lewis, N. (2012) Mirrored morality: An exploration of moral choice in video games. *Cyberpsychology, Behavior, and Social Networking*, Vol. 15 (11), pp. 610–614.