

Why be Moral in a Virtual World?

JOHN MCMILLAN

University of Otago

MIKE KING

University of Otago

ABSTRACT

This article considers two related and fundamental issues about morality in a virtual world. The first is whether the anonymity that is a feature of virtual worlds can shed light upon whether people are moral when they can act with impunity. The second issue is whether there are any moral obligations in a virtual world and if so what they might be.

Our reasons for being good are fundamental to understanding what it is that makes us moral or indeed whether any of us truly are moral. Plato grapples with this problem in book two of *The Republic* where Socrates is challenged by his brothers Adeimantus and Glaucon. They argue that people are moral only because of the costs to them of being immoral; the external constraints of morality.

Glaucon asks us to imagine a magical ring that enables its wearers to become invisible and capable of acting anonymously. The ring is in some respects analogous to the possibilities created by online virtual worlds such as Second Life, so the dialogue is our entry point into considering morality within these worlds. These worlds are three dimensional user created environments where people control avatars and live virtual lives. As well as being an important social phenomenon, virtual worlds and what people chose to do in them can shed light on what people will do when they can act without fear of normal sanction.

This paper begins by explaining the traditional challenge to morality posed by Plato, relating this to conduct in virtual worlds. Then the paper will consider the following skeptical objection. A precondition of all moral requirements is the ability to act. There are no moral requirements in virtual worlds because they are virtual and it is impossible to act in a virtual world. Because avatars do not have real bodies and the

persons controlling avatars are not truly embodied, it is impossible for people to truly act in a virtual world. We will show that it is possible to perform some actions and suggest a number of moral requirements that might plausibly be thought to result. Because avatars cannot feel physical pain or pleasure these moral requirements are interestingly different from those of real life. Hume's arguments for why we should be moral apply to virtual worlds and we conclude by considering how this explains why morality exists in these environments.

INTRODUCTION

Our reasons for being good are fundamental to understanding what it is that makes us moral or indeed whether any of us truly are moral. Plato grapples with this problem in book two of *The Republic* where Socrates is challenged by his brothers Adeimantus and Glaucon (Plato 1993). They argue that people are moral only because of the costs to them of being immoral; the external constraints of morality.

Glaucon asks us to imagine a magical ring that enables its wearers to become invisible at will, and capable of acting anonymously. He relates a fiction in which a shepherd named Gyges discovers this ring and uses its powers to seduce the queen of the kingdom, kill her husband, and take control of the kingdom. Glaucon claims that if there were two such 'rings of Gyges' and one was worn by a previously moral person and the other worn by a previously immoral person that the moral person would end up committing immoral actions too. So, the central point of this thought experiment is to claim that people only do the right thing because of the potential rewards of identifiable right action and the potential punishments of wrong action, which we refer to as the sanctions of morality. The internal constraints of morality or moral reasons are weak or nonexistent, causally ineffective.

The ring of Gyges enables its wearers to act without fear of detection and punishment. It doesn't make them omnipotent or omniscient. However, it is in some relevant respects analogous to the possibilities created by online virtual worlds such as Second Life. These worlds are three dimensional partially user-created environments where people who are members of the social network control avatars that 'live' virtual lives. These avatars and the 'lives' they lead need bear no relation at all to

the person controlling them, or their life outside of the virtual world. Avatars can perform a very wide range of actions, interact with others in the virtual world, attend lectures or performances, and engage in many other activities.

Launched in 2003 by Linden Labs, Second Life is one of the earliest and arguably most successful online virtual worlds, with an estimated 1 million regular users in recent years (Levy 2014). Linden Labs has recently announced Project Sansar, which claims to give a higher level of freedom to users to create their own highly detailed virtual content, and to incorporate virtual reality headset technology to create a more realistic experience for users (Linden Labs 2016a). As well as virtual worlds being a very important social phenomenon, the way people use their avatars within these worlds may shed light on what people might do when fear of sanction is diminished.

This paper begins by explaining the traditional challenge to morality as posed by Plato. Then it considers whether the anonymity and avoidance of external sanction possible in virtual worlds is a useful test case for the challenge to Socrates. We argue that although virtual worlds do exhibit the problems of reduced prudential reason to be moral, virtual worlds raise more acutely the question of whether there is any moral reason to act beyond mere prudence. Then the paper will consider the following skeptical objection. As Kant observed, a precondition of all moral requirements is the ability to act. The acts of avatars occurring within virtual worlds are not, and cannot be, acts in the sense intended by Kant. Since it is impossible to act in a morally relevant way in a virtual world, there can be no moral requirements constraining the actions of avatars in virtual worlds. We counter this objection by arguing that it is possible for avatars to act in ways that are relevant for morality, and suggest a number of moral requirements that might plausibly be thought to result. However, since avatars are different from physically embodied people in morally relevant ways these moral requirements are interestingly different from those of real life. While immoral actions such as rape lack some of their physical consequences in a virtual world, their psychological impact and what they express about the attitudes of those who perform them, are good reasons for viewing them as immoral.

Generating an account of the moral requirements of a virtual world is no less difficult than creating an account of those in the actual world. However, in both cases an important place to start is by considering the nature of action. Taxonomizing the actions that are possible in a virtual world is a significant project in its own right and this paper will confine itself to showing how three actions that would be immoral if committed against people in the actual world (murder, rape and slavery) are not the

same kind of action in a virtual world. Then, it will defend two instances of action (promising and asserting) that are similar in actual and virtual worlds. In each case, the relevance of the nature of the act for moral requirements relating to that act will be considered.

We will suggest that, contrary to what we might expect given the nature of virtual worlds, morality can exist and flourish within them. If this is so, then the question ‘why be moral in a virtual world?’ can be subsumed within the more general question, ‘why be moral?’ If there is reason to be moral, then this reason will retain its normative strength within virtual worlds. The final part of the paper revisits David Hume’s discussion of the ‘sensible knave’ and argues that those who act morally in a virtual world experience what Hume calls the ‘invaluable enjoyment of a character’ and those who do not have abandoned this for the sake of a few worthless virtual gewgaws. The Humean observation also explains why it is that morality appears to flourish, albeit in a different form, in virtual worlds.

THE RING OF GYGES

The challenge to Socrates is skepticism about whether those who appear to act on moral reasons are genuinely acting on moral reasons. Socrates is asked to consider how a shepherd would act if he found a ring that enabled him to become invisible. The story goes that upon discovering the power of the ring the shepherd seduced the Queen and murdered the King so that he might take the throne.

It might be objected that the immorality of the shepherd is due to moral weakness on his part and someone moral would have acted otherwise. So as to rule out this possibility, Glaucon introduces the following extension of his thought experiment.

Suppose there were two such rings, then—one worn by our moral person, the other by the immoral person. There is no one, on this view, who is iron-willed enough to maintain his morality and find the strength of purpose to keep his hands off what doesn’t belong to him, when he is able to take whatever he wants from the market-stall without fear of being discovered, to enter houses and sleep with whomever he chooses, to kill and to release from prison anyone he wants, and generally to act like a god among men. His behaviours would be identical to that of the other person: both of them would be heading in the same direction (Plato 1993, pp 47-48).

Glaucon then claims that people never freely choose to act morally; if given the option of acting immorally and doing what was in our interests, without fear of detection, we would act immorally because acting morally is a burden and not prudentially valuable.

One reason that this is such an elegant thought experiment is because if Glaucon's predictions are correct it can account for the fact that many people appear to act morally. The appeal to external sanctions is consistent with the common everyday observation that those who have less to lose are, all other things being equal, more likely to commit a crime. An alcoholic who lives on the street has different reasons for not stealing a bottle of bourbon than the high school teacher who finds herself out of cash and in need of a drink. Glaucon's claim suggests that those who follow the rules of morality do so because they judge the costs of noncompliance to be too great and this seems reasonable given that those costs can differ depending upon the person.

The ring of Gyges enables its wearers to act anonymously and without fear of detection and punishment. It doesn't make its wearers godlike in other ways so its magic is not directly comparable with other corrupting powers in fiction such as the One Ring in *The Lord of the Rings*. Glaucon claims that mere anonymity is sufficient for revealing the true nature of morality and that all apparent moral reasons are reducible to self-interest.

While this is a thought experiment it is not simply an ethical intuition pump. In effect, it is an empirical claim about our psychology and what we would do if placed in a situation where we could act without fear of punishment or criticism. Whether Glaucon is correct depends upon whether it really is the case that this is what we would do, if we could act free from the threat of punishment or criticism.

Wartime atrocities, especially those that have occurred when soldiers believed they were following orders or immune from punishment, show that many apparently normal human beings are capable of appalling actions. While alarming, immoral behavior of that kind doesn't give direct support for Glaucon's prediction because they are only committed by a select group of persons, but also because wartime atrocities occur under a set of unusual conditions. Jonathan Glover describes the process of 'moral distancing' whereby military actors become increasingly alienated from the moral quality of their actions (Glover 2000). This means that those who commit atrocities during wartime may not do so simply because the external sanctions of morality have been removed but also because their responsiveness to moral reasons

has been eroded. It is also possible that those who commit wartime atrocities do so because they don't see what they are doing as wrong, or think that they have been ordered to do so, as was a feature of some of the abuse at Abu Graibh (Brown 2005).

Milgram's obedience studies are better evidence because they show that many people during peacetime would, if ordered to do so by an authority figure, cause grave harm to another person (Milgram 1963). It's unclear whether the experimental subjects that acted immorally did so because they believed that the presence of an authority figure meant they would not be punished. Nonetheless, it seems reasonable to suppose that many of them must have thought this (Gillett and Pigden 1996). Some subjects believed that they gave another person a lethal shock. If the subjects thought that they could be charged with murder it would be irrational, as well as immoral, to give this shock even when being asked to do so by a man in a white coat. Milgram varied the study by removing the 'instructor' from the room so that they gave the research subject instructions via the phone, and thereby could exercise more freedom over the extent to which 'learners' were shocked.

While the Milgram experiment shows that an alarming number of people will act immorally if ordered to, it doesn't show that all persons will act immorally of their own accord, as Glaucon claims. The experiment was designed to test the extent to which an authority figure could influence behavior so isn't an instance where persons can act in whichever manner they chose without fear for the consequences of acting immorally.

Social psychologists interested in the proclivity of men to rape have studied the effect of beliefs about punishment on the likelihood of rape. Some studies have suggested that many men would rape if they believed there was no chance of them being caught (Malamuth 1981). Alarming as these findings are, they are complicated by the possibility that there could be a difference between what men say they would do, and what they would do under these conditions.

While these and other examples suggest that human beings are capable of immoral actions when the usual external sanctions of morality are altered, they are specific to particular actions and are complicated by the context that these people acted in.

Nogami and Takai found that in an experimental game setting, players that were anonymous (and therefore non-accountable and non-identifiable) broke rules to gain monetary reward (Nogami and Takai 2008). In the same game, rule-breaking to gain monetary reward did not occur in players who were only non-identifiable, only non-

accountable, or non-anonymous (both identifiable and accountable). This suggests that anonymity and the removal of external constraint is a critical factor in determining whether people will be immoral.

Virtual worlds enable people to develop new appearances and identities, in effect they can present themselves as a radically different kind of person. While it is easy for an avatar to reveal their actual world identity, the majority take advantage of the chance to be anonymous. This means that virtual worlds can create one of the preconditions of Glaucon's challenge: anonymity is analogous to a ring of invisibility, and what this provides is non-identifiability and non-accountability for actors in virtual worlds. This provides a reason to suspect that moral conduct in virtual worlds (and other settings in which non-identifiability and non-accountability is permitted) may be worse than that in real life. The evidence we have presented here supports this hypothesis.

However, the discussion of wartime atrocities by Glover also suggests strongly that wrongful acts by moral agents may be rationalized through undermining of the moral reasons that count against these acts. In virtual worlds this may be more acute, since the moral qualities of acts in virtual worlds is genuinely uncertain. Not only is the prudential reason provided by sanction weakened, but it is an open question as to whether any other moral value obtains in the virtual world, given its difference from the real world.

VIRTUAL WORLDS: SECOND LIFE

There are a number of online virtual worlds. They differ in their size, number of residents, language and theme. The largest, most global and most relevant for our purposes here is the virtual world owned and supported by the San Francisco based company Linden Labs called Second Life. This virtual world is importantly different from predecessors such as The Sims Online because Linden Labs gave residents the ability to create their own content. In effect this means that the majority of the content of this world is built and owned by its users. There are some important exceptions such as the physical laws of Second Life that were created by Linden Labs, but users have control over the appearance of avatar's, the creation of objects and most elements of their physical environment.

Second Life is used for many purposes. Many educational institutions use it to simulate environments which are difficult in the actual world, for example clini-

cal settings or where students need to learn how to manage hazardous substances. Global businesses use it for meetings because of its ability to provide a simulated conference environment. However, it is also used for purposes that many would consider immoral. Prostitution occurs in Second Life, as do killings, rape, and slavery (Ludlow and Wallace 2007). Needless to say, there are morally relevant differences between the actual and virtual world instantiation of these things but their presence, even within the confines of a computer generated environment is morally debatable at least. Do virtual worlds such as Second Life cause agents to act in ways analogous to the Ring of Gyges and if so does this mean that people are only moral because of external sanctions?

Reliable general data on moral conduct in virtual worlds such as Second Life is not available. However, there is some anecdotal evidence that misbehavior in virtual worlds is more common in anonymous (non-identifiable, non-accountable) participants compared to those that are identifiable (Suler and Phillips 1998). As with anonymity, accountability for the acts of avatars within virtual worlds is variable. In Second Life there is a set of 'Community Standards' that provides guidance on what constitutes objectionable behavior of an avatar, which includes intolerance, harassment, and assault (Linden Labs 2016b). There are sanctions for violations of these standards, such as suspension of the account or expulsion from the Second Life community.

However, it is easy for another account to be created by the controller of the avatar, so the force of even the most extreme sanction may be fairly light. An example of this was described by Julian Dibbell in his article 'A Rape in CyberSpace' (Dibbell 1994). In this case, an online character ('Mr. Bungle') raped other characters in an online world, leading to calls for sanctions and the eventual elimination of Mr. Bungle from the online world (a case of virtual killing) by one of its users. It is alleged that the anonymous person controlling Mr. Bungle later returned to the community with a character named Dr. Jest.

While Dibbell's case showed that individuals at least sometimes *regard* the acts occurring in virtual worlds as morally significant, it is not clear that this view is correct, given the differences between virtual worlds and real life. We turn now to a skeptical objection that can be levelled against the view that moral conduct is possible in virtual worlds.

CAN PEOPLE ACT IN A VIRTUAL WORLD?

While Second Life can offer anonymity, in order for it to be analogous to Glaucon's ring it must be possible for people to act in a morally relevant way. Kant shows that a moral duty necessarily implies the ability to perform the corresponding action (Kant 1998). This is partly because of the connection between obligation and moral responsibility. We can be held to account for moral obligations that we do not fulfill and this implies that we could have in fact acted on that duty. It is also important because, as Kant points out, moral responsibility implies that our will is causally efficacious: unless a moral reason can have an effect in the world it makes no sense to talk of moral action.

Avatars and the world of Second Life are virtual. The world, its objects, avatars and computer-based images are housed on the servers of Linden Labs, but viewed and controlled by actual persons sitting at keyboards in front of computer monitors. All that actual persons can do is control visual representations via a mouse and keyboard and type lines of text that other persons can read, or if they chose, speak with their actual voice. Second Life is a virtual world and not physically realized in the same way as the actual world. By contrast with physical laws that describe the real world, physical laws that govern a virtual world are commands expressed in computer code by programmers or users and resemblances to the real world can vary depending on their purposes.

Given that avatars are controlled by persons and persons can express their will via an avatar, one aspect of the Kantian precondition for moral action can be fulfilled: an avatar can express the will of a person. However, even if we accept that the will can act upon a virtual world it's not clear that its effects upon a virtual world are morally relevant, as is expressed in the following syllogism.

1. Moral action demands that the will be causally efficacious.
2. Even though the will can be expressed in a virtual world, it can only act upon a virtual world.
3. In a virtual world there is no morally relevant causation.
4. Therefore there can be no moral action in a virtual world.

Premise 3 is contentious. Its truth depends upon the kinds of actions and effects that are possible in a virtual world. In the next three sections we will discuss three actions that are usually immoral in the actual world when performed by human beings. In these cases the effects of these actions are radically different in a virtual world, in ways that affect their moral appraisal.

VIRTUAL KILLING

Although there are disputes about the badness of death and the correct account of the morality of killing (McMahan 2002), ending a human life is ordinarily one of the most immoral things that can be done. In a situation where a person's continued existence will produce no value for them and only intolerable suffering then there are good reasons for thinking that consensual ending of that life might be permissible. Likewise in war, there are situations where there are sound moral reasons for thinking that killing might be justifiable in some circumstances.

Of course if an avatar is killed in a virtual world, ordinarily, no person actually dies. It is also pertinent that many user-created environments within Second Life and virtual worlds such as World of Warcraft are combat-based and the possibility of avatars being hurt and killed is essential to this gameplay. These facts might be considered sufficient for denying that killing in virtual worlds resembles killing in the real world in any morally relevant respect. However, the issue is more complicated because there are different kinds of virtual death. They range from a role play death where an avatar might describe their own death with words, through to the permanent deletion of the avatar from that virtual world.

A role play death has few if any future effects upon the avatar or the person controlling it. Given that role play and many combat game deaths occur within the context of a game of sorts and that the avatar is not actually killed (they merely cease to play a role in that episode of game play) it isn't appropriate to describe them as deaths at all. They are 'pretence' deaths and are similar to the pretend killings that occur in children's games. After a few seconds of lying on the floor pretending to be dead, the deceased jumps to their feet and gets on with the next game.

The most common way that an avatar dies in the sense of being deleted is via suicide, i.e. the person controlling the avatar requests that their account be deleted. Then Linden Labs will delete that avatar and they will cease to exist in Second Life. A feature of virtual suicide and virtual pretence deaths is that they are consensual. In

the case of suicide, the death of an avatar is consented to, and may be caused by, the person who has ownership of the avatar. Getting killed in a game may be a set-back to one's interests in continued play within that episode of play, and the player may strive to avoid it, but in consenting to play the game, one is consenting to the possibility of being killed and receiving this set-back. While a player in combat based game or role-play might not want their character to be killed, this is a possibility that was known before the game started. We can view such deaths as falling within what Huizinga calls 'the magic circle', meaning that because they are conduct that falls within the formal or informal rules governing that game-play, a norm has not been violated (Consalvo 2009).

Non-consensual virtual killings are possible too. We mentioned earlier the deletion of Mr. Bungle as a punitive sanction against his conduct, which included instances of virtual rape (which we discuss next). Another prominent case of this is an avatar in The Sims Online (TSO) controlled by Peter Ludlow, professor of philosophy at Toronto. His avatar edited a virtual tabloid called the Alphaville Herald in TSO and it exposed a seedy underbelly to what was supposed to be a G-rated virtual world (Ludlow and Wallace 2007). He published articles detailing how teenagers below the age of consent were providing virtual escorting services and confidence tricksters would coerce other avatars into handing over virtual property, which often had a significant actual world financial value. Ludlow attempted to log into TSO and found that the company who owned this virtual world had killed his avatar and deleted much of his property within the virtual world (*Ibid* pp 5-7).

It is, of course, absurd to think that Ludlow's avatar was in fact harmed in any morally relevant sense. Nonetheless it did harm Ludlow because he lost property, the project that he had developed, and an avenue for play and self-expression. This was a setback to his interests, and therefore a harm. The setback resembles one plausible account of the badness of death, the deprivation account. The deprivation account holds that the badness of death consists in the deprivation for a person of their future existence and the positive value this holds for them (hence consensually ending a life that holds no positive value for the person living it is not bad, and may be permissible). To the extent that Ludlow's future experiences derived from this virtual life held value, it was bad for him that he was deprived of this through the killing of his avatar. Since this act of virtual killing was non-consensual, it seems to be a *prima facie* case of wrongful virtual killing by the owners of TSO.

However, Peter Ludlow is still alive and has moved to Second Life where he

has created a new virtual tabloid and avatar. At least by the lights of the deprivation account of killing, the badness of the wrongful act of virtual killing depends on the degree of deprivation, or setback to interests that it causes. This means that virtual killing in the case of Ludlow's avatar was massively less of a deprivation and therefore less bad for Ludlow than killing in the real world. Despite some resemblances between the two acts, even though the killing of Ludlow's avatar is wrong it is not wrong straightforwardly for the same reasons the killing an actual person usually is. Virtual killings, even those that fall outside of 'the magic circle', do not justify the same moral obligations to not kill in a virtual world.

VIRTUAL RAPE

Rape is unwanted, non-consensual sexual activity. While it *might* be possible to generate some cases where actual world rape is, all things considered, justified these would be rarified cases and rape is an action that is almost always wrong. As might be expected, given the findings of Malamuth and subsequent social psychologists (Malamuth 1981), virtual rape is common in Second Life. It has also occurred in other virtual worlds, such as the case of virtual rape described by Julian Dibbell (Dibbell 1994). Unlike in the actual world, rape appears to have no physical consequences in a virtual world and this might be taken to imply that it is not wrong.

1. Virtual rape has no harmful physical consequences for the person raped (no pain or permanent bodily damage).
2. The only morally relevant features of rape (virtual or actual) are its harmful physical consequences.
3. Virtual rape lacks morally relevant features.

The weakness in this argument is premise two. While there is no doubt that the physical consequences of actual rape can be appalling, the psychological implications of being compelled to perform a sexual act for another person are at least as significant for its wrongness. While persons are not physically realized in a virtual world, the extent to which many identify with their avatars means that we should be more

cautious about the possible psychological effects of virtual rape, especially for those deeply attached to, or identifying with, their avatars.¹

Moreover, we can consider whether real-world rape that caused no harm (physical or psychological) would be wrong. Such a case was imagined by John Gardner and Steven Shute:

It is possible, although unusual, for a rapist to do no harm. A victim may be forever oblivious to the fact that she was raped, if, say, she was drugged or drunk to the point of unconsciousness when the rape was committed, and the rapist wore a condom (Gardner and Shute 2007).

Gardner and Shute defend this possibility in careful detail in their article, and argue that it is the central case of rape, separated as it is from other features that can accompany it, such as harms of a physical or psychological nature - hence they term it the case of 'pure rape'. They argue that pure rape is wrong, and cases where harms are caused along with it aggravate this central wrong. The wrong, they claim, is the sheer (i.e. non-consensual) use of a person. Using Kantian reasoning, they argue that this use is wrong not because it violates the victim's right to control bodily property, but because it denies the personhood of the victim. It does this by treating the victim as a mere source of use-value through her body.

To the extent that one identifies oneself and one's body as extending into the avatar within a virtual world, sheer use of this avatar may amount to pure rape. Moreover, it is likely that the extent to which one has this attachment will increase the likelihood of harmful effects of this virtual act on the person identifying with the avatar, aggravating the central wrong. However, this is entirely contingent upon the psychological relation between a person and their avatar, and so it is not possible to rule that all acts resembling virtual rape are instances of sheer use of a person and therefore akin to rape in the real world.

However, if all interactions in virtual worlds are consensual, this may mean that virtual rape might not be rape at all.

1. Rape is unwanted, non-consensual sexual activity.

1. Wolfendale refers to identification with one's avatar as 'avatar attachment'. For more on this, see Jessica Wolfendale (2006) "My avatar, my self: virtual harm and attachment." *Ethics and Information Technology* 9(2): 111-119.

2. It is always possible to close the Second Life program, turn off the computer, teleport away or simply refuse to enter a virtual world such as Second Life.
3. Actions that model rape in a virtual world must be consensual.
4. Actions that model rape in a virtual world cannot be rape.

This argument stretches consent within virtual worlds beyond plausibility. It is hardly constitutive of life in virtual worlds that one's avatar is subject to virtual rape, so consent to join a virtual world cannot in itself constitute or imply consent to that eventuality. Even if rape is not an eventuality but a foreseeable possibility, consent given this risk is not consent to the act any more than consent to attend any situation in which sexual assault is a possibility is consent to that occurring.² Moreover even if 4 is correct there still could be moral reasons why actions that model rape in a virtual world (pretence rape) should not occur, but, just as was the case with virtual killings they are very different reasons from those that make actual rape wrong.

VIRTUAL SLAVERY

Slavery is even more common in virtual worlds than pretence killing and rape. The slave-based science fiction world 'Gor', as developed in the science fiction novels of John Norman, has been realized in dozens of user-created environments within Second Life. Male and female slaves are captured, bought and sold, and used for whatever purposes their owners see fit. Still, there are some very clear differences between actual and virtual slavery that map the differences between actual and virtual rape. Virtual slavery does not harm physically and may be consensual, which calls into doubt whether virtual slavery can be considered slavery at all.

Even though virtual killing, rape and slavery may lack the effects that make their actual world counterparts so troubling this does not imply that they have no other moral significance. It might be that pretence rape increases the likelihood of an actual world rape by suggesting that because some agree to pretence rape there this is an actual world want for this too. While slave-based worlds such as Gor do have male slaves, they are very patriarchal societies where men own, rule and use slaves who are primarily female. Again, it might be that pretence slavery spills over to the actual

2. We are grateful to an anonymous reviewer for pressing this objection.

world in some way and the attitudes and presumptions about what it is that women want influence actual world behaviour.

Virtual killing, rape and slavery are not causally connected to a kind of world where they are as morally significant as in the actual world, although it would be a mistake to dismiss their wrongfulness as insignificant. On the other hand there are other actions that avatars and the persons controlling them can perform in virtual world that are much closer to their actual world counterparts.

VIRTUAL VERACITY

The persons controlling avatars can use them to communicate with other avatars. Typed text is often shared within the context of role play or a game of another variety. Given that these contexts involve pretence, statements that take the form of propositions shouldn't be taken as literal assertions. We shouldn't attach any great claim to truth of the child playing Monopoly's assertion that another player owes her \$200,000 than we should to a roleplaying avatar's assertion that they are feeling frightened.

However, avatars type text in many other contexts. These might be statements about the actual world, such as where the person controlling the avatar lives or which time zone they are in. But, it is also possible to make many assertions that are about the virtual world, such as the price of virtual goods, the location of a virtual shop.

Whereas killing, rape and slavery are rightly thought of as 'virtual' in a virtual world when it comes to veracity there is nothing virtual or simulated about it. If we type something false with the intention that another believes it to be true it's lying. Deliberately deceiving another in a virtual world is no less real than lying in the actual world. It's a misnomer to speak of virtual veracity.

VIRTUAL PROMISE KEEPING

The possibility of making true and false assertions implies that it is also possible to make true and false promises. These could be promises about repaying a loan of virtual money, Linden dollars, in the case of Second Life. The same reasons why it is wrong to make a lying promise in the actual world would apply to a virtual world. In both worlds the liar exploits the trust of the person deceived for financial gain. The liar disregards the moral status of the person deceived, the possibility that they might

have had plans for the money that has been loaned and uses them as a mere instrument for their purposes.

The permissibility of actions varies greatly in a virtual world with some that should almost never occur in the actual world becoming benign and others retaining the same effects. Given that the effects of actions differ so markedly we should expect the moral requirements of a virtual world to differ too. Does the difference between the actions possible in virtual and possible worlds mean that we cannot use the anonymity of virtual worlds as a test bed for the challenge to Socrates?

VIRTUAL WORLD ACTIONS AND THE RING OF GYGES

The immorality that Glaucon predicts is similar to many of the violent virtual behaviours that occur frequently in Second Life. However because these actions can have radically different moral qualities in a virtual world we can't infer that people performing virtual forms of acts that would be wrong in the real world are acting wrongly, or at least wrongly in the same way, or that they have abandoned the internal constraints of morality.

On other hand, it is clearly possible to setback the interests of people in the real world through virtual acts (examples we have considered include virtual killing and virtual rape), and the extent to which this occurs or people make false assertions and break promises in the absence of external sanctions might provide evidence for or against Glaucon's prediction. Given the anonymity of Second Life, people can make whatever fanciful claim they wish about their abilities or status, in either actual or virtual world. As Ludlow describes there is no shortage of swindlers and con artists who will take every opportunity they get to cheat others out of virtual cash.

It might be objected that while people can hide behind an avatar, the avatar itself has an identity and a reputation. This is true to an extent but it is very easy in a virtual world to reinvent oneself, to reappear with a different identity and appearance. This is the reason we argue that virtual death is less bad than actual death. So, unlike the actual world where damage to reputation can be lifelong, a new life in a virtual world is so easy to create that the external sanctions can be trivial depending upon the investment made in that virtual life.

From even the briefest foray into virtual world it will be obvious that there are those who chose to use their anonymity in deceptive ways. It will also be obvious that there are just as many who see that veracity and promise keeping are just as important

in a virtual world. Of course it is impossible to say whether the proportion of those who act morally is any different in a virtual world, but Glaucon's claim that there would be no difference in behavior between the previously immoral and moral isn't correct.

Hume's sensible knave makes exceptions to general moral rules when it is to his advantage. Hume describes

... the frequent satisfaction of seeing knaves, with all their pretended cunning and abilities, betrayed by their own maxims; and while they purpose to cheat with moderation and secrecy, a tempting incident occurs, nature is frail, and they give into the snare; whence they can never extricate themselves, without a total loss of reputation, and the forfeiture of all future trust and confidence with mankind (Hume 1946, p. 155).

While it is easy for someone to create another identity in a virtual world, if this becomes necessary because of a deception, that particular identity will forfeit trust with other persons. But, the anonymity of a virtual world will mean that this reason is not significant for any but those who have extended business networks or friendships. The second reason that Hume thinks the sensible knave misses is more relevant to why many are moral within a virtual world.

But were they ever so secret and successful, the honest man, if he has any tincture of philosophy, or even common observation and reflection, will discover that they themselves are, in the end, the greatest dupes, and have sacrificed the invaluable enjoyment of a character, with themselves as least, for the acquisition of worthless toys and gewgaws. How little is requisite to supply the necessities of nature? And in a view to pleasure, what comparison between the unbought satisfaction of conversation, society, study, even health and the common beauties of nature, but above all the peaceful reflection on one's own conduct: What comparison, I say, between these and the feverish, empty amusements of luxury and expence? (Hume 1946, p. 156).

Those who cheat and swindle in virtual worlds have given up the pleasures of virtue for virtual gewgaws. Veracity and promise keeping are not virtual, even when they occur within a virtual world. Understanding the possible real-world harms, not

to mention benefits, from some virtual acts, and recognizing these as giving rise to normative reasons is morally significant. Objects in Second Life and other virtual worlds vary significantly in their value. While some of them can acquire some monetary and other value they may be viewed as ‘gawds’ in Hume’s sense: feverish, empty amusements that can arouse temptation and self-interested pursuit that conflicts with virtue.

Virtual actions that are morally different from their physical world counterparts might be achieved via actions such as being deceptive and in such cases virtue has been compromised for the sake of something that is of comparatively little value. So, Hume’s observation has as much or more relevance for the behavior of those who choose to do wrong in a virtual world. His answer to the challenge to Socrates is as convincing a rebuttal to those who think virtual worlds can only foster immoral behavior as it is to the general observation that we only ever appear to do the right thing because of our fear of the external sanctions of morality.

Acknowledgements

We are grateful to Tom Douglas and two anonymous reviewers for the Journal of Practical Ethics for their insightful and helpful comments upon an earlier version of this paper.

REFERENCES

- Brown, M. (2005). “Setting the conditions” for Abu Ghraib: The Prison Nation Abroad.” *American Quarterly* 57(3): 973-997.
- Consalvo, M. (2009). “There is no magic circle.” *Games and Culture* 4(4): 408-417.
- Dibbell, J. (1994). “A rape in cyberspace or how an evil clown, haitian trickster spirit, two wizards and cast of dozens turned a database into a society.” *Annual Survey of American Law* 490.
- Gardner, J. and S. Shute (2007). *The wrongness of rape*. Oxford, Oxford University Press.
- Gillett, G. and C. Pigden (1996). “Milgram, method and morality.” *Journal of Applied Philosophy* 13(3): 233-250.
- Glover, J. (2000). *Humanity: a moral history of the twentieth century*. New Haven, Yale University Press.
- Hume, D. (1946). *An enquiry concerning human understanding and selections from a treatise of human nature*. Chicago, Open Court Publishers.

Kant, I. (1998). *Groundwork of the metaphysics of morals*. Cambridge, Cambridge University.

Linden Labs, (2016a). "Build worlds with us." from <http://www.lindenlab.com/releases/lindenlab-invites-first-virtual-experience-creators-to-project-sansar-testing> [Accessed November 2017].

———(2016b). "Community standards." from <https://secondlife.com/corporate/cs.php> [Accessed November 2017].

Levy, K. (2014). Second Life has devolved into a post-apocalyptic virtual world, and the weirdest thing is how many people still use it. *Business Insider*.

Ludlow, P. and M. Wallace (2007). *The Second Life Herald: the virtual tabloid that witnessed the dawn of the metaverse*. Cambridge, Mass., MIT Press.

Malamuth, N. (1981). "Rape proclivity among males." *Journal of Social Issues* 37(4): 138-157.

McMahan, J. (2002). *The ethics of killing: problems at the margins of life*. New York, Oxford University Press.

Milgram, S. (1963). "Behavioural study of obedience." *The Journal of Abnormal and Social Psychology* 67(4): 371-378.

Nogami, T. and J. Takai (2008). "Effects of anonymity on antisocial behaviour committee by individuals." *Psychological Reports* 102(1): 119-130.

Plato (1993). *The Republic*. London, Pimlico.

Suler, J. and W. Phillips (1998). "The bad boys of cyberspace: deviant behaviour in a multimedia chat community." *Cyberpsychology and Behaviour* 1(3): 275-294.

Wolfendale, J. (2006). "My avatar, my self: virtual harm and attachment." *Ethics and Information Technology* 9(2): 111-119.