# INO Masterôs Thesis

## The Ethics of Avatars in Virtual Worlds

**Jaqueline Lynch**

June 9, 2011

**LNT Advisor:**
Dr Eric Hadley-Ives

**Thesis Advisor:**
Dr Keith Miller

**Dean's Representative:**
Dr Roxanne Marie Kurtz

**Committee Member:**
Dr Peter Boltuc

Print Form

**UNIVERSITY OF ILLINOIS AT SPRINGFIELD**
**Office of Records and Registration**

# Master's Closure Approval Form

Please complete, print (see "Print Form" button above), sign, and return this form to the Office of Records and Registration, University Hall, One University Plaza, MS UHB 1076, Springfield, IL 62703-5407. Forms are processed on a weekly basis. Please be sure to carefully complete all information prior to printing the document.

Master's degree candidates are required to complete a closure exercise demonstrating mastery of some area within their major field of study. The exact nature and format of these exercises are determined by individual programs, but all of them must have an identifiable academic focus and must include a written component.

Last Name: LYNCH  First Name: JAQUELINE  UIN:

Degree (MA, etc.): M.A.  Major: Liberal and Integrative Studies

LNT (Liberal and Integrative Studies) students must state the title of their degree so that it can be posted to their official academic transcript. The title must be no longer than 26 characters.

LNT: VIRTUAL WORLD PHILOSOPHY

The following are the types of closure exercises currently utilized by program at the University of Illinois at Springfield. Please check the appropriate option listed below.

☐ Capstone Course ☐ Comprehensive Examination ☐ Creative/Artistic Work Product ☐ Project ☑ Thesis

Project or Thesis Title (if applicable)

The Ethics of Avatars in Virtual Worlds

The following signatures signify that the student listed above has completed the required closure exercise in order for him/her to earn a Master's degree at the University of Illinois at Springfield. The number of signatures may vary depending on the program and type of closure exercise completed. In order for the Office of Records and Registration to consider the closure requirement met, the signature of the department chair is required.

| Appropriate Approval Signatures | Role/Title (Committee Chair, etc.) | Date |
|---|---|---|
| Keith Miller signed on line 5 | Academic Advisor / Thesis Sponsor | |
| Eric Hadley Iver | LNT Advisor | June 9th 2011 |
| Peter Boltuc electronic submission EHI | 2nd Academic Advisor | Date " |
| Roxanne Kurtz electronic submission EHI | Dean's Representative | Date " |
| Keith Miller | Thesis Advisor | 9 June 2011 |
| | | Date |
| | | Date |
| Eric Hadley Iver | | June 9, 2011 |

Department Chair (Required) — Date

THIS FORM SHOULD BE SENT TO THE OFFICE OF RECORDS AND REGISTRATION. TO BE ELIGIBLE FOR GRADUATION IN A PARTICULAR SEMESTER, ALL GRADUATION PAPERWORK (INCLUDING THIS FORM) MUST BE RECEIVED BY THE OFFICE OF RECORDS AND REGISTRATION NO LATER THAN THREE WEEKS AFTER THE LAST DAY OF CLASSES OF THE SEMESTER.

For Office Use Only
Processed By:
Date:

Original: Permanent File

**Records (217) 206-6709** **Registration (217) 206-6174** **Degree Audits (217) 206-8216** **Graduation (217) 206-7730**

5-06-11

## Contents

**Abstract**

With the emergence of virtual worlds such as Second Life and World of Warcraft, there has been a significant convergence between the real and virtual worlds. From an ethical perspective this has given rise to a number of questions about the status of avatars in these worlds and the obligations of avatars and avatar controllers in both the virtual and real worlds. This thesis attempts to answer three questions related to this topic:

1. *In what way do avatars exist?*
2. *Are avatars moral agents capable of causing harm?*
3. *Are avatars moral patients capable of being harmed?*

I will argue that avatars exist as artifacts but that this does not make them ontologically inferior to natural substances. I will then argue that it is the drivers of the avatars who are the moral agents and moral patients and the avatars themselves are an extension of the driver that (the avatar) is used to enact the will of the driver. Finally I will argue that real moral harm can result from interactions within virtual worlds such as Second Life (SL) and World of Warcraft (WOW).

**Introduction**

For the purposes of this thesis we will define an avatar to be the visual embodiment of a person in cyberspace made up of computer code and data. A full metaphysical and technical analysis of what an avatar isis beyond the scope of this thesis. However, a key question to be answered is whether an avatar itself has the right to be treated

respectfully or whether it is just an artificial entity with no ethical significance whatsoever. It is clear that an avatar exists in some manner as it takes up memory in a computer somewhere. Their existence as artifacts is supported by Lynne Rudder Baker [12 and 37]. In both of these articles, Baker responds to five criteria used by David Wiggins [38] to determine whether something is a natural substance. I will go into Bakerôs arguments in some detail to show that artifacts can qualify as natural substances or at least cannot be totally ruled out as such. In the second article on the ñOntological status of artifactsò[37], Baker provides the four conditions she sees as being necessary for something to be viewed as an artifact. I will go into the details in both of these articles in more detail in Part I. The intent of this thesis is to look deeper into whether existence for an avatar goes beyond that of a mere artifact and what the ethical significance of that existence might be. In order to do this we need to look at the various ways in which avatars could claim existence.

Ontology looks at what categories of being are fundamental and what it means to say that something "is" or "exists". It is the study of what there is (or existence) and, in the broader view of ontology, it is also the study of problems around entities that do exist. In order for us to have a consistent picture of the world around us we need to make ontological commitments which involve affirming the existence of the things and people around us. In general, we have no issue making such commitments to physical or tangible entities such as people, trees, cars and so on, but other more abstract entities such as properties, numbers, imaginary beings and avatars create ontological challenges for us. In order to make such commitments regarding avatars we need to

determine which theories best apply to avatars.

Questions about the existence and nature of minds, bodies, god, space, time, causality, unity, identity, and the world are all metaphysical issues. Appendix B contains a summary of some of the philosophers and how their theories might apply to the existence of avatars.

This thesis will look at avatars in virtual worlds, the role of avatars in virtual worlds, the ontology of avatars, and questions about the nature of avatars. I will argue that Avatars do not exist in the same way that people exist, and nor do they exist in the same way as tangible entities such as tables or trees exist. Instead they should be viewed *as morally significant* artifacts and, as such, should be included in ontological schemes, but should not be regarded as being inferior to natural objects. I will argue that the avatar should be seen as an extension of the driver and that the avatar has an increased ontological status beyond that of a mere artifact because of that.

Initially I will review what is meant both by a virtual world and an avatar. After exploring various views on existence, I will argue that avatars are artifacts and, while they do not exist in the same way that persons exist, they do exist in a morally significant way. Lastly I will look at issues around virtual harm, specifically moral harm in virtual worlds: Can avatars do harm to others - either other avatars or to humans? Is it possible to do harm to avatars? I will argue that the answer to both of these questions is that the harm is really performed by and to the drivers of the avatars rather than the avatars

themselves and that the moral obligations are twofold ï the driver has obligations as a moral agent and others have obligations to the driver as a moral patient.

## PART I

In order to investigate the status of avatars letôs first look at virtual worlds and what I mean by Avatars. I will also look at the economies behind virtual worlds so that we have a better understanding of virtual worlds and avatars before I move on to the discussion around existence and moral harm.

### Description of Virtual Worlds and Avatars

In this thesis I will focus primarily on two different virtual worlds - World of Warcraft (WOW) and Second Life (SL). WOW and SL are sets of programs that run on computers, and create a 3D grid structure to provide ñonline worldsò where people can interact with each other via avatars. WOW and SL are like virtual stages that provide people with a virtual life where they can escape the real world and can have experiences removed from physical risks and consequences. In these worlds, a person expresses identity through the avatar and the clothes it wears, its body language, its career and hobbies, and how it communicates with others online. This is how the avatar communicates who or what it represents ï its form of identity.

It is important to differentiate between worlds like WOW or SL and other environments. SL is what is referred to as a metaverse and online games like WOW or Eve Online are 3D gaming worlds. Lastly there are video games like Grand Theft Auto (GTA) which are

not covered in this thesis. Metaverses and online gaming worlds differ from video games in several important ways:

1. They provide a seamless continuous world that continues whether the avatar is logged in or not.
2. They have no predefined objectives and there is no concept of winning or losing
3. They focus on developing relationships and characters over a period of time
4. The users generate and (at least in SL) retain ownership of almost all of the content
5. They are very complex environments and it takes a great deal of time to become absorbed into the complex social groups therein.
6. They exhibit a growing economy where users buy and sell in-world content and real world goods ï this can occur within the game or you can buy some of the content on sites like EBay and transfer it into the game later.
7. They are dynamic with ever changing objects, places and users.

These differences are important when trying to determine the moral agency of those involved ï using games like GTA or Starcraft as an example, you have the option to restart the game or reload from a saved point if you make a mistake or get zapped in the game. So, if you make a mistake or get killed (your avatar or character is terminated is some manner) you can just restart the game and make different choices to avoid the negative outcome. However, WOW, EVE Online and SL are different because of their persistent nature ï there are no do-overs and change is permanent.

Throughout this thesis I will mention persistence several times. So I will start here by defining what I mean by persistence. Persistence, when applied to personal identity, refers to what it takes for the same person to exist at different times or how we attribute identity over time. A full analysis of persistence is beyond the scope of this thesis, however the key question is around what is necessary and sufficient for a past or future being to be you. When I refer to persistence regarding identity for an avatar I mean that the avatar is the exact same avatar quantitatively (not necessarily qualitatively) as it was at a different time. For virtual objects persistence is used to mean that the object is permanent - when you log back in it will still be there exactly where you left it. To lessen confusion I will use continuity when referring to virtual worlds which are continously growing and changing worlds.

Both Azeroth (WOW) and Second Life are continuously running digital environments where the players have to deal with the long-term effects of their mistakes. Relationships become complex over time and people invest a significant amount of emotion, time and money into their characters and into obtaining and keeping their virtual property.

The citizens within these virtual worlds are called avatars. Technically, the avatar is a mixture of lines of code and an animated 3D image or object that is used to provide a digital expression of the person driving the avatar. It is the clearly intended creation of the driver who has to design not just the avatarôs looks and gender, but who also has to determine their personality and behavior. The avatar is not autonomous and requires a

human driver who makes subjective judgments and determines all actions taken ï rarely are they AI (artificial intelligence) driven. In both WOW and SL there is a TOS (terms of service) and a EULA (end user license agreement) that spell out the general rules and expectations of the games owners. Additionally, depending on the groups the driver has their avatar join, there will also be societal expectations on behavior. In SL, the avatar can change its appearance and persona, but it cannot change its birth date (date created) or its name. Those continue throughout the life of the avatar.

There are three kinds of avatars ï those controlled by an individual driver (the focus of this thesis), those controlled by multiple drivers (not included in the thesis) and those controlled by AI or known as non-player characters (NPCs) (also not included in the thesis). We differentiate NPCs from player characters, especially when it comes to evaluating the impact of actions. An NPC is an avatar that is not associated with a human driver and, as such, is simply a program that reacts automatically with no human intervention, based on a set of preprogrammed rules ï there is no rational being involved in the decision making process. In WOW the NPC is indirectly controlled by the Blizzard programmer and is clearly labeled as an NPC, so anyone seeing the avatar knows there is not a human driver behind it. In this thesis, the focus is on the moral behavior of the player character and those are the avatars that I will focus on.

In both WOW and SL the worlds are heavily tilted toward social and business interactions between drivers through their avatars. Within SL there is a whole system based on status ranging from newbies (lowest) to the Lindens (owners). WOW has its

own cultural caste system based around experience, professions, race, and a number of other factors. Drivers, through their avatars, can also buy, own and sell virtual land, property and goods and the funds earned therein can be transferred to and from the real world. However, it is really the driver performing the actions and the virtual objects are õownedò by the digital account. Basically the avatar is a creation of the self we want to represent online and it takes actions on behalf and under the control of its human driver or drivers. This means that the avatar is what is termed a mind or intention dependent object.

**Virtual World Economies**

Virtual worlds, social or gaming, are commercial enterprises. They rely on lots of users who are willing to pay to play and who are willing to purchase in-world virtual objects to continue to fund the economy of the virtual world. Through their avatars, those users hold jobs, run businesses, buy and sell land and virtual objects, and exchange currency, all within the virtual environment. These worlds can be used to generate significant amounts of real world wealth. In 2006, Anshe Chung became the first person within SL to accumulate over $1 million (real world, U.S. dollars) in assets, the bulk of which were in virtual real estate holdings as well as stock in SL virtual corporations and shopping centers. In the real world she ran a spin-off corporation that generated content and immersive 3D environments for corporations. She has been featured in Business Week magazine which stressed the fact that she was able to attain this level of assets in less than 30 months from an initial investment of $9.95.

According to the Linden Labs website, in 2008, SL had over 2 million unique users with a peak of 77,000 users logged in at one time and over 397 million hours were spent online in SL. Residents leased over 1724 million square meters of land from Linden Labs and over $350US million was spent on selling virtual goods, land and services with over $100US million exchanged via the Linden Labôs exchange (this is real to virtual money and vice versa).

According to Warcraftrealms [30], there are over 5 million active WOW players who are level 10 or higher and who have been active in the last 30 days. About 1.5 million players reside in North America and 1.2 million in Europe. It is estimated that WOW is producing over $800 million per year in revenue which is more than some small countries.

Until 2008 there were unregulated banks and financial institutions in SL. After SL banned all gambling institutions there was a run on one of the banks to the extent that they were unable to repay $750,000 to those who invested in them. From 2008 the only banks and financial institutes allowed to set up a presence within SL are those who have a real-world presence and are licensed and regulated by real governments.

The effects of virtual worlds have been significant from a financial perspective alone. According to appdata.com, over 92 million users log into Cityville every month and over 49 million into Farmville. Mark Pincus, the owner of Zynga and creator of those two games, has made a fortune greater than $1 billion from selling virtual items in Farmville

alone. Users pay *real money* to buy *virtual goods* to decorate their virtual worlds. I confess to having used real money in Farmville ï it is a social experience for me as I use the chat interface with family during the game and we share and trade virtual goods between us.

According to Gigi Wang, chair/emeritus of the MIT/Stanford Venture lab, it is estimated that the US virtual goods market will exceed $2.1 billion this year.. According to Kzer [26], US revenues from Virtual Worlds was around $2 billion in 2010 and is expected to rise to $4 billion in 2011 and $6 billion in 2012. Virtual goods have become a booming industry whether it be the trade of virtual homes, tables and chairs or the design and trade of clothing, hair, or skins for Avatars. These virtual goods are now being integrated into social networks and virtual worlds and have become a significant part of our economy. The value of these virtual environments should not be underestimated ï the week of May 20, 2011, Zynga partnered with Lady Gaga to use their games like Farmville to help promote her most recent album. They did this by integrating quests into the game that provided the ability to get free clips from her new album. What artist would not want the potential of having upwards of 49 million people being exposed to their music?

Of course, this is causing some to view these social worlds as business constructs. In some real world countries, transactions that take place within virtual worlds are recognized as taxable entities and the users are expected to pay sales tax on them in the real world, even though they are virtual objects paid for with virtual money. The

argument is that real money was used to purchase the virtual money and that money flows between both the real and virtual worlds.

In particular, SL has seen millions of dollars poured into it by companies like IBM, Cisco, American Apparel and Honda, as well as by a significant number of learning and research institutions. SL has also been used by hospitals and towns to simulate emergency responses to disaster situations. The activities within these worlds can be based on social interactions, but there is a significant amount of commerce that takes place and that commerce has a direct impact on real world economies, one of the reasons that many of these companies are spending time there.

Now that we have a better understanding of the virtual worlds and avatars I will move on to trying to determine their ontological status, namely what kind of existence they have. In order to do this I will first look at the various views out there of artifacts and will then go into more detail relating these specifically to avatars.

**Discussion on the ontology of Avatars**

*"If you will practice being fictional for a while, you will understand that fictional characters are sometimes more real than people with bodies and heartbeats."*

- *Richard Bach*

According to Barry Smith [8] ñOntology as a branch of philosophy is the science of what is, of the kinds and structures of objects, properties, events, processes and relations in

every area of reality.ò More specifically, ontology concerns not only ñwhat isò and being in general, but also looks at what categories of being are fundamental and asks whether, and in what sense, the items in those categories can be said to "be". Basically it asks the question, ñWhat does it mean to exist and in what manner does an object exist?ò When looking at the ontological status of avatars I review whether an avatar is merely a personal projection of the player, entitled to be treated respectfully, and what kind of existence an avatar has. Lynne Rudder Baker [11 & 12 & 37] has already argued that avatars exist as artifacts. She asks whether artifacts are less real than natural objects because they are mind dependent and then provides arguments to show that they are in no way metaphysically deficient because humans created them. I will cover her views in more detail later in this thesis, but for now I will just say that I concur with her that avatars exist as artifacts and will augment that view to show that avatars deserve to be treated as ontologically significant.

The first step is to identify what it means to be an artifact. In the ñOntological status of Artifactsò [37 Page 7], Baker proposes a list of four conditions that she thinks can be used to determine whether something is an artifact. Those four conditions are:

“

*(A1) x has one or more makers, producers, or authors. Designers and executors of design (perhaps the same people) are authors.*

*(A2) x’s primary kind (its essence, its proper function) is determined in part by the intentions of its authors.*

*(A3) x’s existence depends on the intentions of its authors and the execution of those*

*intentions.*

*(A4) x is constituted by an aggregate that the authors have arranged or selected12 to serve the proper function entailed by the artifact's primary kind.*

*"*

If we analyze an avatar using these criteria then it is clearly an artifact as is shown below:

(A1) An avatar has a maker or author ï it is the product of human creation.

(A2) An avatar is totally dependent for its proper function on the intentions of its human creator. By proper function we meant that it does what it was designed to do.

(A3) An avatar only exists if the human creator intends to create it and the actions that it takes are based on the intentions and will of that creator.

(A4) An avatar is an aggregate of computer code, computers, design elements, etc that the creator combines together in order to create the avatar.

So according to Baker [37 page 7], avatars would qualify as artifacts.

So now we need to identify what kind of artifact an avatar actually is, since it is something that is intangible. Letôs first look at works of art. In his article on "the Ontology of the Work of Artò [9], Roman Ingarden argues that in order to determine the ontological status of a work of art we need to look at the work of art, its relation to concrete entities (such as the score, canvas, etc), the creative acts of the artists and the conscious states of the viewers. Although Ingardenôs views are more relative to characters in novels or musical scores, they can also be applied to virtual entities like

avatars. An avatar could be viewed as a work of art ï it is created at a set time by an artist and its performance is controlled by the artist. The artist also has significant time and labor invested into the avatar. It does not have autonomic functions nor does it have its own consciousness. The driver, through their avatar, relates with other avatars in SL or WOW, and, through them, with the people who are driving them. The actions and behavior of the avatars have a direct impact on the other avatars around them.

According to Ingarden, music is a purely intentional object that comes into being at creation time and its ontic foundation is in the score. The avatar can be likened to a piece of music (where the code implementing the avatar is analogous to the score) and can be viewed as a purely intentional artifact or object that comes into being when it is created, and its ontic foundation could be viewed as being in the logs and blueprint for the avatar. This is an interesting viewpoint as it means that an avatar should be viewed at least as having the same status as a piece of art or music. It can be seen to exist in some manner, even though it is primarily an online entity.

However, unlike a piece of music, avatars have a form of persistence where changes that are made in gameplay continue on with the avatar and have long-term consequences for the avatar and the virtual environment around it. A piece of music exists in the performance or in the score (when not being played) and, with the exception of improvised music, the musical piece does not change and grow over time although interpretations by artists may show differences. Improvised music is different again as it has no permanence ï it is played once and is then done.

Avatars are different from works of art. Anything can be deemed to be a work of art regardless of what it depicts. However, an avatar has a specific make-up or set of rules around it. Additionally, the virtual world differs in that it continues on regardless of whether any specific avatar is currently online. This continuity affects the future for the avatar and those who interact with it. While this can be seen as similar to a musical instrument in that the instrument is built and a performer plays it (replace instrument with avatar and performer with driver), there are some fundamental differences. The same applies to thinking of the avatar as a puppet with the driver as a puppet master. It can be argued that playing an instrument or putting on a puppet show changes the instrument or puppet such that future actions can be constrained, however, where this is normal wear and tear this would be seen as the kind of change that does not change the ability of the puppet or instrument to be able to continue to perform the same show over and over again until such time as it is so worn down that it cannot be used any more.

The actions taken by an avatar on behalf of a driver are permanent actions that will impact future decisions for that avatar and that also impact the inventory and ongoing relationships associated with that avatar. There is no reset back to before those actions were taken ï  any actions taken are permanent and the exact performance is not repeatable since the effects were permanent. As an example, in SL I may have my avatar purchase a piece of land ï  there is no way to go back to prior to that purchase. So until I sell the land using the avatar I own it. In WOW I may give someone my sword

ï there is no way to get it back unless they decide to return it. Basically the actions that I had the avatar take are now permanent and there is no way to undo them or to go back to how things were prior to the actions being taken. In the case of a puppet or a musical instrument any performance can be repeated as there are no persistent effects from a performance, beyond the wear and tear mentioned above. Just consider puppet shows such as Punch and Judy ï the show is repeated many times and is almost identical each time. In the case of a virtual world it is more like living in the real world where each day and each show is different and they build on each other. The one exception would be if the performance destroyed or damaged the instrument or puppet which is similar to the avatar dying. Otherwise, it is possible to go back and replay the same scene or the same piece of music repeatedly ï with virtual worlds and avatars this is not possible as the virtual worlds continue on regardless and changes cannot be undone, so the players have to deal with the long-term effects of their actions. This is also true of any entity that a person uses to interact with the world to express their thoughts and feelings ï an example here might be ventriloquistsô dummies. Such dummies differ from puppets and are actually much more like avatars except they have a physical presence in the real world. Like avatars they tend to be an extension of the driver (in this case the ventriloquist) and act on the intentions of the driver. Like avatars, these dummies have an enhanced significance because of the attachment that their drivers (ventriloquists) have to them.

Another key difference is around the way that people use avatars. Puppets and musical instruments and even books are forms of entertainment. While virtual games are a form

of entertainment the versions like WOW that are immersive 3D worlds do not have an end goal and are more like continuously developing virtual worlds. According to Mark Silcox and Jon Cogburn (13) a movie, book or piece of music is all already there, whereas a virtual world is constantly changing and developing. While artworks are open to interpretation, they do not change in the way that virtual worlds do. SL and WOW and similar worlds are open to creative intervention and manipulation that allows for discovery and new forms of self-expression and the player characters change over time as a function of each others behavior. This is an important differentiator as the player characters cooperate to complete quests and often have permanent social group memberships. Additionally, their actions affect how other characters develop in the game. Although avatars can be seen as a form of entertainment, for many people they are experientially real and they are actually living their lives online through the avatars. With puppets people are exercising their imagination and exploring possibilities or they are replaying stories and shows that have been around for a long time. But with virtual worlds and avatars, people are actually living their lives there ï for many people these worlds are not games or a form of entertainment. So when bad things happen to avatars they have the potential to negatively impact the drivers behind those avatars.

Another key difference is that dummies and puppets are not usually an important means of interacting with the real world whereas avatars are primarily used for that kind of interaction. However, attitude and context are critical. It should be noted that avatars can be like dummies or puppets if there is no emotional connection to them and

dummies and puppets can be more like avatars when there is an emotional attachment.

Avatars are said to ñdieò when the driver stops using them, if someone steals them, if the virtual world owner (i.e. Linden labs or Blizzard) kicks them out, or if the person driving them dies. When an avatar ñdiesò, that virtual death is connected with a loss of power and status in the virtual community and it removes the avatar from the social community that itôs driver had been using it to participate in. Sometimes when that happens there are online funerals and periods of mourning. Those who had been ñfriendsò with that driver-avatar combination experience a sense of loss. I will talk more about this later as part of the discussion on avatar attachment. Additionally, when an avatarôs driver dies in real life, something has to be done with the land and businesses they own. Often the drivers digital account and associated avatar is left to other people in the driverôs will so that the avatar can continue to exist online. Lastly, there are also real legal and financial implications. Appendix A lists a few recent legal cases ï reviewing these in depth is beyond the scope of this thesis, but it should be kept in mind throughout that actions taken in the virtual world are causing actions in the real world as well.

With the convergence between the real world and virtual worlds like SL and WOW, the driver, through the avatar, can own virtual land, work, make virtual and real money, create and sell products and buy goods, all within the virtual world. In reality it is the driver who owns these objects as they are associated with the driverôs user account and

the person driving the avatar retains all of the intellectual property rights for those creations and can convert their game dollars to real dollars at any point in time. So there is a thriving economy that is linked to real world money. There are people driving avatars in these worlds, who earn all of their money for real life there and spend more time with their friends in the virtual world than they do with people in the real world.

The question of whether virtual worlds are real is related to the question of whether other non-physical entities are real. We can ask whether non-physical entities like religion and music, or organizations like stock markets or companies are real. In all of these cases we answer that they are real because they affect the real world and people take them seriously even though they are social and cultural constructs. In ñShrinking differenceò[12], Baker quotes Jaegwon Kim, saying that if we apply Alexanderôs Dictum ï ñto be real is to have effectsòï then it can be argued that artifacts are real. Since avatars are artifacts and they can be used to affect the economy as well as a personôs social community then they are also real. In fact, I would argue that avatars have causal powers between the economy and the real world, powers invoked by their drivers. If avatars are real and having effects then clearly they exist in some form. Given the number of users and the amount of money involved (see virtual world economies Page 9), then it is clear that the real world is affected.

As mentioned earlier, for some people a vast majority of their life is lived within virtual worlds and some of those people have become millionaires in real life because of their businesses there. So, clearly life in these worlds can have effects on real life ï another

argument for avatars to have ontological status. So the question becomes whether the avatar is ontologically inferior to other natural objects and whether they are even ontologically robust. I will now go on to look at some views on existence and reality and then determine how they apply to avatars.

### *Is reality just an illusion?*

In the 1999 movie *the Matrix*, Morpheus says the following:

> *"What is real? How do you define, 'real'? If you're talking about what you can feel, what you can smell, what you can taste and see, then real is simply electrical signals interpreted by your brain."*

This is an interesting lead into a discussion on reality. After all, if reality is simply electrical signals interpreted by the brain, then anything we can perceive can be seen as being real, whether it is a physical entity or a virtual one. All the data we have about the physical world is interpreted based on signals received by the brain and how we perceive them. This can be different depending on the chemical or other makeup of the person perceiving the event. Justin Holme [20] makes the point that we effectively see a simulated version of the real world, and that what we experience is the result of gathering and interpreting data from the real world rather than the world itself. Holme invokes Kant, who said that existence should be seen as a concept corresponding (or not) to the world. Kantôs view was that saying that something exists is to say the concept of the thing is exemplified in the world. Kant, in Critique of Pure Reason, differentiated between the noumenal and phenomenal worlds, pointing out that our

experiences are dependent on our perceptions and that what we experience is the phenomenal world. The noumenal or mind-independent world, as defined by Kant, is beyond our perception of it. He thought that we could never have direct experiential access to the noumenal world, as its fundamental nature is hidden from us behind the veil of our perceptions. Instead he thought that the world as we see it is really the world of appearances or what he called the phenomenal world, which is a mind-dependent view of the world.

Holme points out that the simulation of the world that we see is mind-dependent, not mind-independent. If Holme is correct and we only experience simulations of the world, then virtual worlds like SL and WOW are in many ways not that different from what we think of as the real world, and the objects in them could also be deemed to be real. I think Holme is on the right path here in thinking that mind-dependence does not rule out the ability for something to be seen as real. This does not mean that all mind-dependent objects are real, it merely does not rule out the possibility that some are real.

Mark Silcox and Jon Cogburn [13] go a step further, making a nice comparison to Nozickôs experience machine [24 and Appendix B], pointing out that Nozick said that most people would shun the experience machine because ñ*you don't get to be any particular way*ò By this they are referring to the fact that, in the experience machine, the person does not control the experiences and has little say over what happens. However, avatars in role playing games (RPGs) are not like the experience machine. The driver can set their avatar up to be any way they want it to be within the restrictions

of the virtual world, and to have many experiences that are not available in real life. The characters can do things their real players cannot physically do such as flying or walking on ceilings, albeit virtual ceilings. This is different to being a tool such as a hammer or a screwdriver. While tools allow the person to perform actions that would be difficult without them, they do not provide the same kind of social interaction with others that a driver can experience when interacting through their avatar online. If I was to try to interact with others through a hammer or a screwdriver I highly doubt that it would be a positive interaction, since the purpose of the hammer is to hit things for example. The avatar has a persona associates with it that is created by the driver and, in most cases, the driver will ensure the avatar behaves in ways consistent with that persona.

This means the avatar has a reality or identity separate to the driver ï the avatar cannot be ñyouò because the basic traits that dictate what it can do in the game are incompatible with the ones that characterize you in the real world. This is a very strong argument for the avatar to have its own form of identity or at least to be a significant extension of the player.

Dario Compagno and Patrick Coppock [14] go even further, pointing out that fictional game objects, including avatars, are immaterial or intangible objects, but they are ñexperientiallyò real for the players who relate to them or use them in the game. Although they refer to them as fictional I think it is better to call them intangible objects as it is clear they exist, which fictional implies is not the case. Players are able to purchase or earn objects, and to buy and sell these virtual objects in the real world as

well as in the game. The continuous nature of the virtual world, and actions taken in it, mean there are long-term consequences for actions and that those consequences feel real to the players and affect their ongoing relationships with other players.

Compagno & Coppock compare WOW players to the people chained in Platoôs cave who only experience the shadows. In book VII of the Republic, Plato uses the ñAllegory of the Caveò to describe our situation regarding knowledge. He describes an environment where prisoners are chained to a wall and they only see the digital shadows cast by real people. It is not until they are liberated and they go outside that they see the real objects in the world. Compagno and Coppock point out that avatars in virtual worlds can be like those shadows, where the drivers forget that there are actually real people behind the avatars. They go on to make an argument that avatars and game objects should be viewed as intangible cultural artifacts and that their ontological status should be based on that. I think this is a reasonable argument and it supports the argument that avatars are artifacts and that they exist.

So what kind of existence does an avatar have and why do I think it has some form of persistence? Firstly, the avatarôs virtual body is linked to the avatar permanently so there is a form of continuity there. The driver can change the avatarôs gender or hair or clothing, but certain things cannot be changed. Avatars have names that stay with them throughout their lifetime as well as a fixed ñbirth dateò in SL. This is a form of continuity. It is also rare for an avatar to exhibit a change in their online personality or behavior ï it is more likely that the driver will create an ñaltò who exhibits the new personality

continually. Avatars (and alts) tend to be consistent and to live continuous lives with set social groups and communities.

Continuity is also found in the logs of SL and WOW, and in the inventory associated with the avatar. The avatarôs inventory functions as a permanent indirect record of activity as it records new and used items and provides for continuity when the avatar is not logged in. The logs also provide a record. When not logged in, the avatar does not cease to exist; it still exists in the form of lines of code that are not currently being executed, and the virtual world around it continues on. In fact, since virtual worlds exist on software and hardware that is setup by the designers, the virtual world continues on regardless of which players are currently connected. During that time property and land can still be bought and sold even though the avatar is not currently present.

Now that I have discussed some definitions around avatars and virtual worlds, this is a good point to look at how we determine existence. Earlier I reviewed an article by Lynne Rudder Baker that I used to determine that avatars are in fact artifacts. In the next section I will look at her arguments as to why existence should not be limited to natural substances and why artifacts should be seen as having a robust ontological status.

In the ñOntological status of personsò [11], Baker points out that the existence of a person requires a first person experience, and that a person is a primary kind with human animal being the bodyôs primary kind. This definition clearly rules out avatars as existing in the same way that human persons exist. I have already argued that an avatar

is a mind-dependent artifact and does not have a first person perspective, so their ontological significance must derive from something other than that. This might lead some to say that this mind-dependence means that avatars do not exist in a robust way, especially since any existence they do have is only in virtual worlds. However, if we look at Baker further [12 and 37] it is clear that avatars really do exist.

In the article on the ñshrinking difference between artifacts and natural objectsò [12], Baker argues against the requirement that, in order to have ontological significance something must be a natural substance. She looks at the five standard conditions used by David Wiggins [38] for determining ontological genuine substances and then looks at how they apply to natural objects and artifacts. She points out that there are natural objects that fail some of these tests and artifacts that could arguably pass them. Artifacts are defined as objects that are intentionally created to serve a given purpose whereas natural objects come into being without human intervention. Human beings themselves are deemed natural objects, as are their children (regardless of the fact that intervention is required to create children). Artifacts are therefore intention dependent as they depend both ontologically and causally on humans for their existence. In this article Baker [12] is specifically talking about the ontological status of tangible objects like trees and tables, however we should be able to analyze avatars using them.

Baker [12 Page 3] lists Wigginsõ criteria as follows:

ñ

1. *Fs are genuine substances only if Fs have an internal principle of activity.*

2. *Fs are genuine substances only if there are laws that apply to Fs as such, or there could be a science of Fs.*
3. *Fs are genuine substances only if whether something is an F is not determined merely by an entity's satisfying a description*
4. *Fs are genuine substances only if Fs have an underlying intrinsic essence.*
5. *Fs are genuine substances only if the identity and persistence of Fs is independent of any intentional activity.*

ñ

If we analyze avatars according to those criteria then we see some interesting results:

1. *Fs are genuine substances only if Fs have an internal principle of activity.*

   Aristotle (Physics 2.1) said that nature is an inner principle of change and being at rest. This means that when an entity moves or changes, or is at rest according to its nature, then it is possible to refer back to its nature to explain the change. Such changes come from within the substance and do not have external causes although there may be preconditions that we need to account for. Aristotle thought this condition could be used to distinguish objects from nature from those with other causes. He considered nature itself as an internal principle of change and that natural substances were essentially material substances. So the principle could be seen as the actual laws that determine the properties of the substance. For a natural substance these do not change over time. So what Wiggins seems to have been saying is that for something to be a natural

substance then its potential states and changes need to be intrinsic to it by its very nature.

Baker says that gold is a natural object that has no internal principle of change and compares it to a heat seeking missile which is an artifact that Baker says has an internal principle of activity. Baker interprets this as meaning that this condition cannot be used to distinguish artifacts from natural objects. Peter Kroes and Pieter Vermaas [10 page 29] point out that ñ*it is not clear how Baker understands the internal principle of activity of a heat-seeking missile*ò and they argue that gold would be seen as having an internal principle of activity, based on ñ*the psychochemical laws that determine its properties*ò They point out that the missile may have a principle of activity as a physical object that has a certain makeup and adheres to certain laws but that does not necessarily mean that it has an internal principle of activity as an artifact, namely as a heat seeking missile as it gets its principle of activity for that role from its maker, not from its nature.

I concur with Kroes and Vermaas that it is not clear how Baker understands the internal principle of activity, but I am not sure that matters here as an avatar does not have an intrinsic internal principle of activity. All changes come from outside and are not intrinsic to the nature of the avatar. All activity is controlled by the person or persons driving the avatar and is limited by the rules programmed into

the virtual environment. That person sees the avatar as an extension of themselves and they certainly (as persons) have an internal principle of activity.

However, if we view this in the manner argued for by Kroes & Vermaas [10 page 29] then *"the proper function of a technical artifact can indeed be taken as a principle of activity of the artifact qua artifact because the function of an artifact is not reducible to the physical properties of the artifact"*. They argue that artifacts are ontologically different from natural objects but that they are not inferior. They concur with Baker that we can take the proper function of an artifact to be its internal principle of activity and thus it passes test 1 above. The key difference from a natural object is that the principle of activity is internal but it is not intrinsic to the artifact, as it is dependent on the intentions of the creator. The proper function of an avatar is to do whatever it is the driver intended it to do when they created the avatar. Since an avatar is controlled by the driver and the rules built into the game, it is clear that the avatar will always act according to its proper function. In this view, as long as the avatar is fulfilling its proper function, then it has an internal principle of activity and avatars pass this test.

2. *Fs are genuine substances only if there are laws that apply to Fs as such, or there could be a science of Fs.*

   This refers to the fact that natural substances have laws that determine their makeup and these substances can be studied using scientific methods to determine what kind of substances they are. As examples, water has a specific

makeup and carbon has a specific atomic weight. There are sciences built around these substances such that they can be analyzed and studies.

Baker rightly points out that many artifacts are also subject to laws and that engineering and other similar schools have set up fields of study for many artifacts. In the case of avatars, there are laws controlling the avatar ï there is a TOS (terms of service), there are real laws and there are physical laws around designing and controlling and driving the avatars. In particular, the physical laws that put constraints on the design, control and driving of the avatars significantly affect what the avatar can and cannot be made to do. The computer and graphics cards put additional controls on the environment and the actions of the avatars and the virtual worlds themselves have a set of boundary rules. There is a whole population dedicated to the building and designing of avatars, skins and buildings so it could be argued that a science of avatars could be possible and, therefore, avatars pass this test.

Additionally, Kroes and Vermaas [10] point out that there are certainly laws that apply to artifacts, however they can change over time whereas the laws around natural objects do not change. They use these points to support their argument that natural objects and artifacts exhibit some differences but that artifacts can still pass the tests for being real.

3. *Fs are genuine substances only if whether something is an F is not determined merely by an entity's satisfying a description.*

   Here, Baker seems to be pointing to how we refer to natural kinds. Continuing the example of gold she points out [37 page 13] that the distinction between using a description versus an indexical reference is *"Only a matter of the state of our knowledge and of our perceptual systems".* Her point is that sometimes we will use a description to refer to gold and sometime we will use its atomic weight. Something like fools gold will match the description for gold but it is not the same atomic weight and it is not gold and never will be.

   Avatars are much more than descriptions. A webbot could be mistaken for an avatar if the only thing that mattered was that it matched the description of an avatar. However, the webbot is not the same thing as it is missing the basic nature of an avatar. Avatars are designed to exhibit personality traits that the person driving them wants them to portray and they are used to build relationships over long periods of time with other driver/avatar combinations. This goes beyond just putting a smile on their face, as the driver designs a persona and story around the avatar which involves interactions with other driver/avatars. The story around an avatar can change over time depending on the actions the driver decides they should take, and based on interactions with other avatars and their drivers. This is different to literary characters where the story is fixed in the novel. The story for an avatar is not fixed until the actual actions take place ï it can change right up to that point. Drivers, through their avatars, can also own

property and conduct business, something a description cannot do. The driver can also have their avatars belong to kinds or groups within SL/WOW and, in fact, much of the status within virtual worlds is gained from the groups and communities the driver associates their avatar with. Avatars also have a form of persistence and, on command of their drivers, have the ability to make lasting changes, something that is beyond the ability of descriptions. Avatars pass this test too.

4. *Fs are genuine substances only if Fs have an underlying intrinsic essence.*

An intrinsic essence is that which makes something what it is. However, Baker does not think an intrinsic essence should be a requirement. She points out that not all natural objects have intrinsic essences ï using the example of mountains she points out that a mountain is not a mountain because of an intrinsic essence ï rather it is a mountain because of its relationship to some other entity (i.e. flat land). But a mountain is still a genuine substance.

If we accept Bakerôs argument (and I do) that something is an X, not in virtue of what it is made of but in virtue of the relational properties that it has, then the avatar can be seen as having an underlying essence, just not one that is intrinsic. An avatar has a specific look and feel to it and, as determined by the driver, has ongoing permanent relationships with the virtual entities (land, property and other avatars) with which it interacts, so there are some specific relational properties that would imply that it passes this test and has essence. An avatar always has

the same birthdate ï this cannot be changed. It also has the same owner so an argument could be made that these are actually an intrinsic essence since they are always part of that avatar.

5. *Fs are genuine substances only if the identity and persistence of Fs is independent of any intentional activity.*

This is the argument about mind-dependence. The principle here is that something can only be a genuine substance if it is not dependent on human intervention to be created. This raises some interesting conundrums. As an example, water is a natural substance. However, I can create water in the lab so does that mean we have water that is a natural substance (i.e. comes from rain) and water that is not (I created it in a lab)? This is inconsistent as a requirement. Additionally, this could also lead us to argue about the ontological status of children who are born because their parents were trying to have children (intentional), or who were test tube babies, or who were accidents (not intentional). Is the child who was born unintentionally superior ontologically? This is clearly a somewhat ridiculous distinction but I am using it to illustrate the problems with the mind-independent requirement for being recognized as having ontological status.

The identity of the avatar is dependent on the driver i.e. intention or mind-dependent. The driver creates the avatar and the avatar takes actions based on

the intentions and commands of the driver. So the avatar, just like a puppet or music or art, could not be seen to be independent of intentional activity ï in fact it is dependent on it. Baker invokes Alexanderôs dictum ï ñto be real is to have effectsòï and uses it to argue that if the criterion for being real is to have causal effects then artifacts (mind dependent objects) are as real as natural objects. Because of this impact and the inconsistencies identified above, I think it is becoming more difficult to require non-intentionality as a criterion for being ontologically significant. Since the avatar can and does have an effect on both the virtual and real worlds, I would argue that avatars are real and are not ontologically inferior to natural objects.

In three of the five cases (2-4) above it is clear that avatars pass the test for being a substance and therefore existence. It is arguable that the avatar could pass the first test depending on how activity is interpreted. I take the view that if the avatar is performing its proper function then it has an internal principle of activity and, therefore, passes that test. A key point that Baker was trying to make was that some natural substances fail some of these tests and some artifacts pass some of them so perhaps this requirement is not the best differentiator. I concur with Baker that, while the avatar is clearly a mind-dependent object, the fact that the avatar and its behavior do have an impact on the real and virtual worlds can be used to argue that it is as real as natural objects. This impact leads me to agree with Baker that avatars need to be given status as real entities in the form of artifacts. In this view avatars should be seen as existing in a robust sense

beyond just being artifacts in virtual worlds as they have the ability to impact things and people beyond their narrow environment.

Bakerôs work was written primarily to address artifacts in the physical world but Amie Thomasson [15] takes Bakerôs work and applies it to imaginary objects, pointing out that they too are the product of thought and are intention or mind-dependent entities. She differentiates them from physical artifacts because physical artifacts need physical acts as well as thought ï as an example a chair needs to be put together as well as imagined. She also argues that any account of artifacts needs to be able to account for these abstract artifacts including novels, laws, songs and corporations. Based on her article she would categorize an avatar as an abstract object.

I disagree with Thomasson and would argue that she missed a valuable point when it comes to avatars and virtual objects. Just as one has to physically build a chair as well as designing it, there is a considerable amount of work that goes into designing and building an avatar and its online persona. The avatars, and any associated virtual tables, etc. have to be designed and then hand-built using online tools. I think that this puts them in the same category as physical artifacts rather than abstract artifacts.

David Leech Anderson [18] takes this a step further when he points out that virtual objects should actually be seen as having more ontological significance than physical objects. He argues that a table, once created, is no longer dependent on the mind that created it, however a table in SL or WOW is still dependent on the mind that created it

and it will cease to exist if the player stops thinking about it. I do not agree with that view, given that I can log out and log back in several days later and the virtual table is still there, I would argue that the virtual table and the physical table have the same level of significance. In fact, if every mind was to disappear a physical chair would still exist, but so would the virtual table, as long as the virtual world was still around. The table in SL will only disappear if the player destroys it, sells it, kills their avatar, shuts down their SL account or if the virtual world disappears for some reason. The same is true in WOW ï my inventory will still be there when I log back on. That is not all that different from a physical table that we would have to destroy or take apart to make it disappear.

Leech Anderson [18] also makes the point that while players prize their virtual objects for various reasons, they also place commercial value on them and are able to buy or sell them for real money, both inside and outside of the game. He says [18 page 16]: *"if you can buy them and sell them for hundreds of millions of dollars then they are real"*. I think Leech Anderson has a valid point here and it concurs with Bakerôs earlier invocation of Alexandersôs dictum, that to be real is to have effects.

Lastly, Leech Anderson [18] alludes to Descartes, saying that one can experience the tree in a virtual environment even though there is no physical tree. The same is true when looking at artistic representations of trees (e.g. paintings and sculptures) ï they are not an actual tree, but can still be used to experience ñtreenessò. Although virtual trees are not physical they can be viewed as ideal as they depend on human intentions. He argues that virtual objects definitely exist and are not fictional like those in a novel ï

in a virtual world real betrayal can take place and virtual objects can be left for others to use or can be bought and sold. Using the virtual sword as an example he points out that if I can commit an act of betrayal with a virtual sword then it is real and it exists. So again we have support for a robust ontological status for virtual objects, including avatars.

Finally, Roxanne Kurtz [17] also supports Bakerôs contention that being intention or mind dependent does not make an artifact inferior to a natural object. Since the players are part of nature she points out that this means that their intentional activities are also part of nature, so the issue of being mind dependent should be deemed a non-issue and artifacts that we create should not be seen as ontologically deficient.

**Comments**

So what does all this mean? It is clear, after analyzing the articles mentioned above, that virtual objects and fictional entities like avatars are artifacts rather than persons. They do not have first person experiences or self-awareness and are mind-dependent entities. They are similar to fictional entities like ñSherlock Holmesò in that they are the creations of an author, but they differ from fictional entities like ñSherlock Holmesò or ñSanta Clausò as they can (controlled by their drivers) have direct financial and other impacts on the real world as well as the virtual world ï this convergence of the two is a major differentiator. By direct impacts I mean that the driver, through the avatar, participates in those impacts ï real financial and other transactions take place under the direction of the driver. In the case of ñSanta Clausò, there are financial impacts in the

form of Christmas shopping, movies, etc but the fictional entity ñSanta Clausò does not participate in the actions in the same manner as the drivers of avatars do. Fictional characters in books (Sherlock Holmes) and movies would be similar to ñSanta Clausò ï they impact economies in the form of people purchasing books, comics, movies, etc but they do not actually conduct those transactions. The transactions are *about* them not *by* them whereas the avatar combined with the driver, actually performs, or at least participates in, the acts involved in the transactions.

We also drew parallels to puppets and musical instruments, pointing out that the key difference here is that there is a continuity present in virtual world performances that is not present in puppet shows or in performing on a musical instrument. While puppets and musical instruments may persist as physical objects I am referring to the permanent nature of actions taken during performances. Actions taken in a virtual world have a permanent effect on the avatar and those around it whereas the same puppet show or musical performance can be replayed many times.

As artifacts, avatars are subject to the rules of the virtual world they inhabit as well as dependent on the intentions of their creator (driver or player). They exhibit persistence and can have a real impact (when driven by the driver) on the real world as well as the virtual world. The drivers, through their avatars, are also subject to social pressures from others in their guilds or neighborhoods. As will be seen in the next section the driver/avatar combination also has the ability to cause harm (moral agents) and to be the victim of harm (moral patients). All of this leads to the conclusion that avatars and

virtual objects should be treated as ontologically significant artifacts that are real and that exist. The combination of the driver and avatar should be expected to behave as moral agents who have a moral obligation to abide by the rules of the societies within which they exist. They should also be viewed as moral patients to whom others have obligations, since the actions of others can have a permanent effect on the avatar, the driver and others around them.

This is particularly true if we look at what David Chalmers and Andy Clark [7] say about the extended self. Here they argue for what they call ñactive externalismò which refers to when a human is linked to an external entity in a two way interaction. If this interaction creates a coupled system such that the removal of the external component causes a drop in the behavioral competence of the system then they claim this should be seen as a cognitive system.

Using the example of a notebook, Chalmers and Clark point out that the notebook, in the case of Otto, plays the role that would usually be played by a biological memory. I liken this to my dependence on my blackberry, which has become part of my brain to the point that I no longer have to keep track of appointments or contacts. Instead, my blackberry is a constant in my life that makes information available to me easily and accurately. Without it I would not be as efficient or as productive as I am with it, which Chalmers and Clark would say means it is an extended part of my brain and my cognition. The one advantage of the blackberry over my brain is that it can be backed up, upgraded and restored, which is pretty useful.

So, if someone were to interfere with my blackberry it would have the same significance to me as them interfering with my person, as it would make it difficult for me to function normally. The point Chalmers and Clark are trying to make is that some tools take over what used to be functions of the brain and that we should no longer see our mind as ending at the skull/skin boundary. This clearly impacts how we should think of moral relations.

If we view the avatar as a form of self-expression then the hours of work and effort that were put into its creation and actions become extensions of what I am. The avatar is really the combination of the 3D image, my labor and time, and my intentions ï basically extending myself into the virtual world. It is not something I purchased but rather a form of self-expression that I have significant equity in and, as such, is viewed as an extension of me.

If we agree that the person is not held to a skull/skin boundary and that entities we depend upon to fulfill our lives, which would include avatars, are a part of our identity then the driver is definitely the moral agent behind the activities of the avatar and is accountable for the actions taken by the avatar. I will look at this in more depth in the next section on virtual harm.

To summarize, avatars exist as artifacts but should be seen as having ontological significance due to the continuous nature of the worlds within which they exist, as well

as the impact they have on the real and virtual environments (and those who live in them) with which they interact through the drivers. Since they are driven by a human being they should be seen as an extension of that person. The work of Chalmers and Clark supports this view and I concur that the avatar is tied to the driver in such a way that the two constitute an organic whole.

Now that we have determined that avatars exist and have a robust ontological status we can go on and look at what this means in terms of moral relations between moral agents.

## PART II

### Discussion of virtual harm as it relates to Avatars

It is clear from the previous discussions that actions taken in virtual worlds like SL and WOW have the ability to impact people in a multitude of ways, ranging from economic to psychological to social. As with any kind of interaction there will be disagreements and there are always those who seek to take advantage of others. In the real world there are laws and social customs that protect people from others criminal activities. These laws and customs are in place for many reasons, but typically protect basic human rights, ownership of property and intellectual property/copyright.

Because of the social nature of worlds like SL and WOW, it is also necessary to have similar rules in those worlds. The equivalent to laws typically exist in the form of terms of service (TOS) that are spelled out in end user license agreements (EULAs) that the users agree to prior to participating in the virtual world or game. Additionally there are social rules depending on the environment or region (SL) or guild (WOW) that one chooses to belong to. There have even been a number of cases now where behavior in the virtual world has ended up being legislated in the real world courts (Appendix A).

However, before we can determine whether avatars can harm or be harmed I first have to explain what I mean by harm, specifically virtual harm. When I talk about harm here, I am referring to moral harm that comes out of actions within virtual worlds. In some cases that moral harm also causes psychological harm and, in extreme cases, there

have been physical harms as well (Appendix A). A few examples of actions that might be seen as moral harms would include theft of property, ñkillingò avatars in WOW in non-PVP areas, exposing players to pornographic actions without their agreement and interfering with a playerôs ability to interact and control their avatar. In WOW and SL such actions would be deemed as morally unacceptable. However, morality refers to a code of conduct that applies to everyone who can understand it and their behavior is expected to be governed by it. Nearly all the standard accounts of moral accountability therefore require that the moral agents are rational beings. This raises questions in virtual worlds as we will see below when we explore who the moral agents and patients are.

In order for moral harm to occur there has to be a moral agent (who enacts the harm) and a moral patient (who is the lucky recipient). Moral agency refers to the ability to act and it needs an intentional mental state to occur that leads to the performance of an action. In order to be held accountable for an action moral agency requires that the moral agent has the ability to freely choose what actions to take and that it knows the difference between right and wrong. As mentioned above, the standard accounts of moral accountability require that the moral agents are rational beings. If we look at babies we do not hold them accountable for their actions because they are too young to reason and do not tend to know (yet) right from wrong. Psychopaths tend to lack a deep knowledge of right from wrong, so they too are not held accountable in the same way that rational adults are. Avatars are in a similar position as they are not rational beings and they do not know right from wrong. An avatar can no more be moral agents than

sociopaths or psychopaths or babies can. However, their drivers are moral agents.

There are several ethical theories around moral harm ï I am going to quickly mention the two that I think most apply when looking at virtual worlds. The first is utilitarianism. In utilitarianism all participants are equal and an action is deemed to be moral if it maximizes aggregate happiness. Utilitarianism focuses on consequences, not motives and those consequences involve maximizing happiness, however one chooses to define it. As an example, stealing money from a rich person to give to a poor person may seem like a good thing to do. However, allowing such actions would mean that property rights are not protected and this would have serious economic implications that could make a majority of people very unhappy. So in utilitarianism actions like theft would be seen as a moral harm. For worlds like SL and WOW, this is important as these worlds have the potential to make people very happy or very unhappy. Griefers are other players who harass other residents or who intentionally do things in a game or virtual world to deliberately cause annoyance to other players. If we look at the role played by griefers who intentionally try to make other players unhappy, then this is clearly a moral harm when analyzed from a utilitarian perspective, as griefers do not try to maximize pleasure for all, only for themselves which is an egoist viewpoint.

The second ethical theory that can be applied is the deontological theory espoused by Kant that focuses on motives and duty. In deontology actions are good or bad based on the good will of the agent ï consequences do not matter here only the motives.
As an example, if we look at the first formulation of Kantôs categorical imperative

ñAlways act in accordance with a maxim such that it can be made a universal lawò

This is basically saying that in order for an action to be moral, then we have to be able to turn that action into something that can be made universal.

Applying this rule to virtual worlds is important as the drivers need to remember that they are interacting with real people using real social interactions. I will talk more about Kantôs views a little further down.

Moral harm can be seen as occurring whenever a human being is impacted by the actions taken in such a way to cause psychological or physical harm. From a Kantian perspective there is no way to universalize such actions and from a utilitarian point of view causing distress to others would not maximize aggregate happiness. It should also be noted that moral harm can occur when someone simply observes an immoral action and that moral harms can range from theft, betrayal, assault, rape to murder to name just a few. It should be pointed out that rape and murder in a virtual world are not the same as in the physical world, however they can be as experientially real to the drivers and they can have a significant psychological impact on the drivers. Most harms that can be acted out in the real world seem to also occur in a virtual form that impacts those experiencing the virtual environment.

When we talk about moral harms and punishments, we normally refer to harms to flesh and blood people and we hold people accountable for those harms. If a machine injures someone we do not punish the machine or hold it accountable, we look to the driver, builder or designer instead. When looking at harms in virtual worlds it is important to

remember that the moral agents and moral patients are the drivers behind the avatars, and these are flesh and blood people, not automatons. In short the avatars are neither moral agents nor moral patients ï the drivers fill that role.

In their article on virtual harm, Keith Miller, Chuck Huff and Deborah Johnson [21] discuss various types of harm and they categorize and rank them. They say that when looking at a harm we first need to determine if it is physical or virtual, then we need to see if it is direct or indirect, and finally whether it is intentional, negligent or accidental. As an example if a driver uses their avatar to abuse another avatar in a way that upsets the latterôs driver then this would be deemed indirect physical harm as the harm ended up being to a person (physical) even though it was done through an avatar (indirect). They also go on to point out that virtual harm is real moral harm because it affects human beings. There are consequences for the drivers when they have their avatars perform functions that impact other driver/avatars.

In real life people find meaning and value in their ability to affect the world around them. They also value things they create themselves, whether it is a work of art or a flower garden, more than things they just go out and buy. This is because of the labor and time they voluntarily invest in these entities. This investment leads to what Jessica Wolfendale [22] calls avatar attachment. In evaluating harms in virtual worlds I tend to agree with Wolfendale [22] that virtual harms are also real harms, due in some part to this avatar attachment. People who participate in virtual worlds do so over a period of time, carefully building up relationships and assets like homes and other virtual objects.

In WOW those assets affect the status of the player within the game as well as enabling them to perform new activities and quests based on that status. It can take hundreds of hours to attain certain levels. In SL players often make their real life living from selling virtual objects or renting out virtual property. They also build up social relationships over long periods of time that lead to social constructs with a great deal of trust. Anything that impacts these can be seen as moral harm. However, as we will see below this is more than just property damage to the avatar or associated virtual objects.

In her discussion on avatar attachment, Wolfendale [22 page 112] says*:" Avatar attachment is expressive of identity and self-conception and should therefore be accorded the moral significance we give to real life attachments that play a similar role".* By this she is referring to the fact that one cannot dismiss avatar attachment without also dismissing other forms of attachment such as our attachment to people, possessions, communities and cultural objects. She points out that people tend to value things that have sentimental meaning over those that they just invest money in. This is one reason avatar attachment is different to toys and other memorabilia. While it is true that people can become very attached to toys and memorabilia they do not tend to be used to interact with the world in the same way that people use avatars. Memorabilia is about things from the past and toys tend to be complete already. Unlike toys, attachment to an avatar is to a personally created and chosen object that the driver can control, act through and use to interact with others and that has a developing story. People use their avatars to reflect their online identity, and they spend considerable time building the story and social structure around those avatars, building extensive

webs of relationships. It is the combination of the 3D presence, identity and communications that leads to avatar attachment being so common. Avatar attachment should be viewed as a significant aspect of identity expression that is morally important. Actions that affect that attachment will cause moral harm to the driver.

Wolfendale also adds that *"avatar attachment takes place in the context of a community with shared expectations"* ï these expectations lead to moral expectations for acceptable behavior for the driver/avatars. Anything that disrupts these relationships can cause psychological damage to the driver which is a real harm and a moral one as well. It should be noted that these relationships often extend into virtual marriages and in some cases into real world relationships as well.

Additionally, since a considerable amount of time and effort is spent in growing virtual assets the real person becomes very attached to those assets. In worlds like WOW and Eve Online those assets make it possible to access content and to build social communities that are not available without those assets. So theft of those virtual assets reduces the utility for the driver and also reduces their pleasure in the game they were playing. This is one reason that drivers get very upset if someone decides to remove or destroy their assets without their permission. Our conception of self is tied up (rightly or wrongly) in our social context along with our assets and beliefs. This applies in both the real and virtual worlds. In the real world flesh and blood people are both moral agents and moral patients. It turns out to be the same with virtual worlds as the driver for the avatar is a person and they are the one that determines the actions, although the avatar

is the entity that represents them online. The driver is also the one who benefits or suffers from the actions of both their own avatar and those of other peopleôs avatars.

Virtual theft is the theft or unauthorized removal of virtual items from another avatarôs inventory. These items have value and utility and gathering them took a great deal of time and skill and effort. These assets have value in both the virtual and real worlds as, in many cases, they can be bought or sold in either or both worlds. (It is possible to legally purchase goods in-world as well as externally on sites like eBay.) In his article on ñSin in Cyber-edenò, Ashley John Craft [23] points out that the persistence of these assets is one of the reasons people pay the monthly fees to participate in these worlds. Because of this Craft argues (and I concur) that *"virtual theft is morally wrong as it harms the victim by depriving them of the utility derived from the object and wastes the time and money spent acquiring it"*.

Avatars have a special role ï they are not just imaginary friends but rather they are extensions of our own identity. As drivers we design their look and feel, make decisions about which virtual islands they will inhabit and which guilds or social communities they will join. They represent who we want to be in the virtual world and that identity tends to be consistent. Drivers rarely change the behavior of their avatars ï they are more likely to create a new avatar. For many people the avatar is tightly linked to their sense of self and others relate closely to itï this can be seen clearly in SL where they hold funerals for avatars when their drivers pass away or an avatar ñdiesò, and where real life people often leave their avatars to others in their wills so that the avatar can continue to exist,

even if a new person is driving the avatar.

Non player characters (NPCs) are viewed differently from player characters as they are simply programs that react automatically with no human intervention, based on a set of preprogrammed rules. Since there is no human driver associated with an NPC then there is no driver to suffer psychological harm and this makes a difference to how they are viewed as there is no moral patient. For NPCs the programmer could be seen as the moral agent. We differentiate NPCs from player characters because the NPC avatar is not associated with a human driver who is making the moral decisions ï it is simply a program that reacts automatically based on a set of pre-programmed rules.

Conversely, a player character is one that is driven by a human driver (similar to a puppet-master) who is making the decisions. In our case we are looking at player characters when I refer to avatars and we hold the drivers accountable as moral agents. We also view the drivers as the moral patients who are affected by the actions of others.

So we have determined that the drivers are the moral agents and the moral patients. This makes it much easier to determine whether an action creates a moral harm. A moral harm is something that impacts the flesh and blood controller, either physically or psychologically. It also includes behaviors that impact the driverôs autonomy or their utility and pleasure in the game or world.

Miller et al [21] talk about this when they discuss the virtual rape that took place on

Lambda Moo in 1992. In this case a driver named Bungle used a voodoo doll program to take over the characters of two other players. Using the voodoo doll, Bungle caused those two characters (Legba and Starsinger) to perform increasingly violent and sexual acts in the virtual world, against the wishes of their drivers. Even though it took place online it involved removing the autonomy of the drivers of the avatars when Bungle took control of their avatars, and it also exposed the other drivers on Lambda Moo to a violent and pornographic experience without their permission. Those who experienced the situation were outraged and experienced feelings of violation that affected them for some time. While a rape in a virtual world is not the same as rape in the real world it still has a significant psychological impact on those involved, and this was the case with Lambda Moo. Having their autonomy taken away and then being forced to view something that was abhorrent to them caused significant distress to the drivers of the avatars. Yes they could have logged off, but by the time they realized what was happening the damage was already done. So there was both moral harm and psychological harm involved in this situation for the drivers.

In gaming worlds or themed islands, the morality of an action can change a little, depending on context. As an example, games like WOW and Everquest all permit violence in specific instances and are based around guilds or groups of avatars going to war with each other. Islands in WOW and most 3D gaming worlds are normally classified as follows:

PVP ï Player versus player where it is expected that player characters will go to battle with each other and try to ñkillòthe avatars.

PVE ï Player versus environment where it is expected that player characters will not try to ñkillò other avatars.

RP ï role playing where battles are forbidden.

There is an expectation that players will honor the rules of the island and environment they are in and there is also an expectation within guilds that those who are members will not turn on each other. Since the only way to move up after certain levels requires cooperation, then guilds take a utilitarian viewpoint and expect that those who join will work in the spirit of cooperation to further the interests of the whole guild.

Within SL there are islands that have themes around slavery or subservience. Those who sign up for those areas are aware of that and agree to those terms. Since the expectations are set going in then it is fair to argue that actions such as killing that take place in this kind of area are not murder as long as they are done according to the socially expected rules of the game. In this case killing and murder refer to virtual killing which involves disabling the other personôs avatar in some way. From a moral standpoint actions such as ganking (killing lower level players who have no chance of defending themselves against higher level players) and griefing (intentionally aggravating other players just to spoil the game for them) are deemed socially unacceptable and immoral. This even applies in the PVP worlds where attacks are to be expected. Actions such as killing a team mate or friend or betraying your guild is never ethical and causes harm to the guild and psychological harm to those who were betrayed in such a manner. Since SL is modeled more on the real world than WOW,

actions like betrayal have a major impact on the drivers. But context definitely matters. If you belong to a virtual world where killing any avatar you want is permitted and it is in the EULA or clear in the culture that you joined, then it is difficult to argue that it is immoral in that world.

In games like WOW there are risks that the players accept when they play. It is not seen as immoral to attack another player controlled avatar when in a PVP environment, nor is it seen as immoral to kill another's avatar in a consensual fight. Morality becomes an issue when the rules are broken or when psychological harm is done to the person controlling the avatar outside the conventions of the virtual world. Because of the social nature of these worlds and the long term relationships that get built up as players progress in the world these conventions are important to players.

As an example, according to Craft [23], in 2005 there was outrage on Eve Online when members of an organization (Guiding Hand) infiltrated a rival organization, spent over a year building up trust and relationships, then killed the leader and stole all their in game assets which were valued at $16,500.00. In Eve Online drivers have their avatars join guilds to increase their collective power. They spend time using chat and other interfaces to get to know each other and build a trust basis that allows them to interact with other guild members. The actions of Guiding Hand shattered the trust in the organization that was held by the players, and it also was a breach in the EULA. Since users agree to the TOS in the EULA prior to entering the environment they have a moral obligation to abide by that agreement. But more than that, it was also a significant

breach of the social rules. While Eve Online is a world designed around betrayal and underhanded behavior, that behavior is expected to be directed at those outside the guild to which the driverôs avatar belongs.

In Eve Online, users gain significant benefits from interacting with others. These interactions take place within the world but are sometimes supplemented using real world techniques such as instant messaging, voice chat, email and even real world meetings. This involves a great deal of trust as people are exposing their real world identities. By destroying the trust that had been built over time in the guild, significant psychological damage was done to the real life drivers of the affected avatars and it affected their ability to continue in the online world. It is often forgotten that the users of these worlds exist simultaneously in the real and virtual worlds and that player avatars are backed by a real human being. This virtual betrayal was seen by the participants to be as devastating as it would have been if such a betrayal had taken place in the real world. As with real world betrayal the psychological impact was such that it was a long time before those affected felt able to trust others again and it affected their ability to trust both in the game world and in the real world.

Schulzke [27] looks at violent video games and analyzes their morality from several perspectives. He says that from a Kantian perspective our moral actions are judged by how we treat others and our intentions. Kantôs account requires that the agents are rational beings, but we have already determined that the moral agents and moral patients are the drivers of the avatars (who are rational beings) so if we analyze virtual

world behavior according to Kant we see the following:

First formulation of the categorical imperative

ñAlways act in accordance with a maxim such that it can be made a universal lawò

By this we mean that in order for an action to be moral then we have to be able to turn that action into something that can be made universal. As an example if I think that stealing someoneôs shoes should be moral then I have to be willing to concede that it is moral for anyone to steal anyone elseôs shoes, including mine. Since that would make me unhappy, especially if it was winter, clearly I canôt make it a universal law and will have to agree that it is immoral to steal shoes.

Again, if we use killing avatars as an example allowing one such act would make killing avatars generally acceptable, otherwise we would have to agree that it is immoral. Schulzke says that killing of avatars could be universalized by placing limits on it similar to: ñ[i]t is acceptable to kill the avatars of non-player characters or willing combatantsò He argues that the players intend to kill the avatar not the real person, and since there is no intent to kill the real person then there is no real harm done.

While I agree with Schulzke when he says murder in virtual worlds could be universalized by placing limits on it such as only killing NPCs, I do not agree with him that since players intend to kill the avatar and not the real person then there is no real harm done if they do kill an avatar. In this case I think Schulzke undervalues the impact of avatar attachment. While killing NPCs or players in a known war zone is one thing, killing avatars outside of places where it is expected or part of the clearly explained

rules, is immoral and causes psychological harm to the player, whether it is intended or not. It is against the rules that the driver agrees to when joining the game and is also against the social expectations of the other drivers impacted. As discussed earlier, this view is supported by the work of Wolfendale [22] in her discussions around avatar attachment, by Craft [23] in his analysis of the betrayal on Eve Online, and by Miller et al [21] in the discussion on the impact of the Lambda Moo virtual rape. In all three cases they found that harms that were enacted in virtual worlds had real effects on the drivers of the avatars and that these effects could be profound. In his article on online gaming, Dowling [31] sums this up nicely when he says: ñReal people are doing real things to real people when they do them in gameé The only difference is that when it's done online, there's more space between you and the person you're speaking with.ò

So we now have an idea of what we mean by moral harms and who the moral agents and moral patients actually are. But is damage to an avatar really more than just property damage? Earlier we looked at the views of Chalmers and Clark on the extended mind. Chalmers and Clark [7] say that a person is not only a mind but also their entire body and that what counts as us may extend beyond the body. In the case of a virtual world this would mean that damage to an avatar can disrupt the playerôs environment and lead to harm to the player. Since the avatar is a device for directly expressing and communicating the player/driverôs will, then it is clear that the driver should be seen as the moral agent for the actions of the avatar and that any harm against the avatar is potentially harm against the player. Add this to Wolfendaleôs comments regarding avatar attachment and it is fairly clear that any action by an avatar

(directed by a driver) that impacts another avatar (and hence its driver) in a negative way should be viewed as a moral harm.

This view that the avatar is an extension of the driver is not just limited to Chalmers and Clark. In his article on the law and the emotive avatar, Gibbons [29] points out that the best way to look at an avatar from a legal perspective is to ñconsider the avatar as an extension of the individual or an agent of the individual in virtual spaces and then to shift the focus of analysis away from the avatar and back to the individual because it is the potential negative effects that virtual behavior may have on real-world individuals that the law seeks to regulateò. This fits in with the view taken by Chalmers et al [7] where the avatar can be seen as an extension of the individual driver as the avatar is a device for directly expressing and communicating the player/driverôs will and makes it clear that the driver should be seen as the moral agent for the actions of the avatar.

**PART III**

**Conclusions**

In the introduction I said we would answer the following three questions:

1. In what way do avatars exist? Are they more than just puppets?
2. Are avatars moral agents capable of causing harm?
3. Are avatars moral patients capable of being harmed?

I have looked at multiple theories and argued that avatars exist as an extension of their drivers. It is this view that avatars are an extension of the driver combined with the concerns around avatar attachment that leads me to believe that avatars and virtual objects should be treated as ontologically significant artifacts, and their drivers, who are the moral agents, should impel them to act in such a way as to abide by the rules of the virtual societies and communities within which they exist. I have also shown that the drivers (through their avatars) are also moral patients and that others have moral obligations to treat them as such. Finally, I have argued that virtual harm is possible and can have effects in the real world as well as the virtual world where it is originally enacted. This harm can be performed by the driver (through their avatar) as a moral agent or the driver (through their avatar) can be the victim or moral patient.

I looked at the views of Ingarden, Wolfendale and Baker, and also analyzed avatars in terms of what it meant if they died as well as according to Alexanderôs Dictum. In all cases the argument could be made that avatars exist as artifacts, and that they should be treated with respect. It was also pointed out that avatars differ from puppets and instruments because of the persistent nature of the worlds within which they exist and the persistent nature of actions taken therein.

There is no disagreement that avatars are mind-dependent objects that are controlled by their creator. However, this is no reason to see them as being ontologically uninteresting or as having no status. I would argue that avatars should be included in

any reasonable ontology as respectable inhabitants that are not inferior to natural objects or other artifacts. They are really an extension of the driver and are part of the driverôs identity in that they are created to represent who the driver wants to be online. However, they can impact the real world in significant ways. To deny their significance and to give them the same status as an artifact like a table would be to seriously underestimate the impact avatars have on those who use or experience them and on the attachment drivers have to their avatars. Avatars and virtual objects should be viewed as artifacts that exhibit the properties of persistence, but that are causally mind-dependent.

Finally, I also determined that the drivers of avatars can do harm (moral agents) to the drivers of other avatars (moral patients) and can also be harmed themselves (moral patients), and that those harms can have a significant psychological impact on the driver behind the avatar. In virtual worlds drivers/players use avatars to represent them and to claim personal space and build assets. Those avatars are capable of producing harms only in so far as their driver has them perform those acts. While certain harms, such as rape, do not include the physical impact that they include in the real world, they still cause significant distress to the moral patient and this needs to be considered when looking at harm in virtual worlds. Since avatars are extensions of the person driving them, the moral agent is the driver and the driver should be held morally accountable for any harm perpetrated by the avatar.

In all of the cases discussed we were looking at single human driver cases for avatars. There are cases where avatars are shared between drivers, even though this is against the EULA, however, these are beyond the scope of this thesis and are a subject for future work. Additionally, the moral status of NPCs and AI driven avatars are also subjects for future research.

To summarize, single human driven avatars exist in a robust way, and their drivers can be both moral agents and moral patients. The drivers of other avatars need to remember that they have obligations to these avatars and the drivers behind them.

## REFERENCES


1. Sweeney, Leo SJ, *"Authentic Metaphysics in an age of unreality 2nd Edn"*, 1993, Peter Lang Publishing, ISBN 0-8204-2278-9
2. Metaphysics defined at freedictionary.com http://www.thefreedictionary.com/metaphysics
3. Hylomorphism defined at Stanford Encyclopedia of Philosophy http://plato.stanford.edu/entries/aristotle-psychology/#2
4. Frege, Gottfried, "Die Grundlagen der Arithmetik", 1987, Reclam Ditzingen, ISBN 978-3150084250
5. Sartre, Jean-Paul, "Existentialism is Humanism", 2007, Yale University Press, Translation of 1966 French edition of 1945 speech in Paris.
6. Quine, William, "From a logical point of view", 1953, Harvard University Press, Essay entitled "On what there is"
7. Chalmers, David and Clark, Andy, "The Extended Mind" http://consc.net/papers/extended.html
8. Smith, Barry, "Ontology", *"Blackwell Guide to the Philosophy of Computing and Information"*, 2004, Blackwell Publishing - Floridi, Luciano (editor)
9. Ingarden, Roman, "The Ontology of the Work of Art", Translated by Raymond Meyer with John T. Goldthwait. Athens, Ohio: Ohio University Press, 1989.
10. Kroes, Peter & Vermaas, Pieter, "Interesting differences between artifacts and natural objects", *American Philosophical Association Newsletter on Philosophy and Computers*, Fall 2008, Volume 8, Number 1
11. Baker, Lynne Rudder, "The Ontological Status of Persons," *Philosophy and Phenomenological Research* 65 (2002): 370-388.
12. Baker, Lynne Rudder, "The shrinking difference between artifacts and natural objects", *American Philosophical Association Newsletter on Philosophy and Computers*, Spring 2008, Volume 07, Number 2
13. Silcox, Mark and Cogburn, Jon, "Chapter 6 - Azeroth versus the experience Machine", *World of Warcraft and Philosophy*, edited by Luke Cuddy and John Nordlinger, 2009, Open Court Books
14. Compagno, Dario and Coppock, Patrick, "Chapter 8 - World of Warcraft as a collectors paradise", *World of Warcraft and Philosophy*, edited by Luke Cuddy and John Nordlinger, 2009, Open Court Books
15. Thomasson, Arnie L, "Artifacts and Mind-Independence: Comments on Lynne Rudder Baker's The shrinking difference between artifacts and natural objects", *American Philosophical Association Newsletter on Philosophy and Computers*, Fall 2008, Volume 8, Number 1
16. Preston, Beth, "The shrinkage Factor: Comment on Lynne Rudder Baker's The shrinking difference between artifacts and natural objects", *American Philosophical Association Newsletter on Philosophy and Computers*, Fall 2008, Volume 8, Number 1
17. Kurtz, Roxanne, "Ontologically Tough Artifacts and Non-Spooky Intentions – Comments on Baker", *American Philosophical Association Newsletter on Philosophy and Computers*, Spring 2009, Volume 8, Number 2

18. Leech Anderson, David, ñA Semantics for Virtual Environments and the Ontological status of virtual objectsò, *American Philosophical Association Newsletter on Philosophy and Computers*, Fall 2009, Volume 9, Number 1
19. Baker, Lynne Rudder, ñShrinking difference ï Response to repliesò *American Philosophical Association Newsletter on Philosophy and Computers*, Spring 2010, Volume 09, Number 2
20. Holme, Justin, ñThe Simulated Selfò, *Philosophy Now*, Issue 81, 2010
21. Miller, Keith; Huff, Chuck;, Johnson, Deborah, ñVirtual Harms and Virtual responsibility: A rape in Cyberspaceò, *IEEE Technology and Society Magazine*, Summer 2003
22. Wolfendale, Jessica, ñMy avatar, my self: Virtual harm and attachmentò, *Ethics and Information Technology*, 2007, 9:111-119
23. Craft, Ashley John, ñSin in cyber-eden: understanding the metaphysicsand morals of virtual worldsò, *Ethics and Information Technology*, 2007, 9:205-217
24. Nozick, Robert, *Anarchy, State and Utopia*, 1974. New York Basic Books, ISBN 978-0465002702
25. Cuonzo, Margaret, ñVirtual Homes and Sherlock Holmes ï on the existence of virtual and other abstract entitiesò, *American Philosophical Association Newsletter on Philosophy and Computers*, Spring 2010, Volume 9, Number 2
26. Kzer Worldwide http://www.slideshare.net/nicmitham/virtual-worlds-2010-2098472
27. Schulzke, Marcus, ñDefending the morality of violent video gamesò, *Ethics and Information Technology*, 2010, 12:127-138
28. Johansson, Marcus, ñWhy unreal punishments in response to unreal crimes might actually be a really good thingò, *Ethics and Information Technology*, 2009, 11:71-79
29. Gibbons, Llewellyn Joseph, ñLaw and the emotive avatarò, *Vanderbilt Journal of Ent. And Tech. Law*, 2009, Vol 11:4:899
30. Warcraft Realm Stats http://www.warcraftrealms.com/quickstats.php
31. Dowling, Steven, ñOnline Gamingò, http://journals.sfu.ca/eludamos/index.php/eludamos/article/viewArticle/vol4no2-11/190
32. Malleck,Shaun, ñTruth and Lies: Pragmatism and Paradoxò, Bryn Mawr Collegeò http://www.udel.edu/GPPC/malleck2007.htm
33. Kurtz, Roxanne, Lecture notes from PHI482, "Metaphysics of Things", Week 2: Ontological Commitment
34. Yablo, Stephen, ñA paradox of existenceò, MIT, 1998 http://www.mit.edu/~yablo/apex.html
35. Discussion on Kant and Ontology at Stanford Encyclopedia of Philosophy http://plato.stanford.edu/entries/kant-metaphysics/
36. Discussion on Frege at Stanford Encyclopedia of Philosophy http://plato.stanford.edu/entries/frege/
37. Baker, Lynne Rudder, ñThe Ontology of Artifacts,ò *Philosophical Explorations 7* (2004): 99-112.
38. Wiggins, David, *Sameness and Substance*, 1980, Harvard University Press

39. Margolis, Eric & Laurence Stephen (editors), *Creations of the Mind: Theories of artifacts and their representation*, 2007, Oxford University Press

**ADDITIONAL READING**

1. Blanchette, Oliva, Philosophy of Being ï a reconstructive essay in metaphysics, 2003, Catholic University of America Press, ISBN 0-8132-1096-8
2. Boon, Stuart and Sinclair, Christine, ñA World I donôt inhabit: Disquiet and Identity in Second Life and Facebookò, Proceedings of the 6th International Conference on Networked Learning, ISBN No: 978-1-86220-206-1
3. Jones, Donald E, ñI, Avatar: Constructions of Self and Place in Second Life and the Technological Imaginationò, http://gnovis.georgetown.edu
4. McIlvenny, Paul, ñAvatars R Usò, http://www.immi.se/intercultural/nr1/mcilvenny.htm
5. Amatrudo, Anthony, ñThe illusory nature of ontological status and its implications for legal, moral and social Organizationò, International Journal of Law in Context, 4,1 pp. 63ï 77 (2008) Cambridge University Press
6. Boellstorff, Tom, ñComing of age in Second Lifeò, 2008, Princeton University Press
7. Turkle, Sherry, ñLife on the Screen ï Identity on the age of the Internetò, 1995, Simon and Schuster
8. Turkle, Sherry, ñThe Second Self ï Computers and the human spirit ï 20th anniversary6 editionò, 1995, Simon and Schuster
9. Stanovsky, Derek, ñVirtual Realityò, ñBlackwell Guide to the Philosophy of Computing and Informationò, 2004, Blackwell Publishing, Floridi, Luciano (editor)
10. Suler, John, ñThe Psychology of Avatars and Graphical Space in Multimedia Chat Communitiesò, http://www-usr.rider.edu/~suler/psycyber/psyav.html
11. Suler, John, ñIdentity Management in Cyberspaceò, 1996 http://www-usr.rider.edu/~suler/psycyber/identitymanage.html Fleissner, Peter, ñMulti-User Dungeonsò, http://members.chello.at/gre/fleissner/documents/muds.html
12. Floridi, Luciano, ñUnderstanding Information Ethicsò, American Philosophical Association Newsletter on Philosophy and Computers, Fall 2007, Volume 07, Number 1
13. Luck, Morgan, ñThe gamerôs dilemma: an analysis of the arguments for the mioral distinction between virtual murder and virtual paedophiliaò, Ethics and Information Technology, 2009, 11:31-36
14. Furber, Musa, ñEthics and virtual worlds ï Second life as a case studyò, Tabah Analytic Brief, June 2009, No. 9

**APPENDIX A**
**Virtual Worlds and the Law**

Below is a list of examples where virtual world incidents have resulted in real world court cases or actions:

1. October 2008 ï a 43 year old Japanese woman is arrested and charged with illegal access to a computer and with manipulating electronic data. When she got divorced without warning by her online husband in an online world (Maple Story) she hacked the game and destroyed his avatar after taking over all his assets. The penalty for this in Japan is up to 5 years in real world jail or a fine of up to $5000US.

   http://www.cbsnews.com/stories/2008/10/23/tech/main4540763.shtml
2. 2005 ï A Chinese man was arrested on suspicion of carrying out a virtual mugging in the online game Lineage II and exchanging the virtual goods for real money.

   http://www.techspot.com/news/18422-lineage-ii-gamers-mugged-in-virtual-crime-spree.html
3. 2007 ï an avatar in Second Life raped another Avatar

   http://www.nsvrc.org/news/news-field/2306
4. 2007 ï a Dutch teenager was charged with the virtual theft of furniture from rooms in a virtual hotel.

   http://www.bbc.co.uk/news/10207486
5. 2008 PC World ï a 16 year old boy is charged with beating another teenager and

forcing him to log in to Runescape and hand over all the virtual money and virtual goods he possesses. Ignoring the physical assault for now, this is the first time a judge in the Netherlands has ruled that the theft of virtual property is illegal and should be treated like stealing real world objects, saying that ñgoods don't have to be material for the law to consider them stolenò

http://www.pcworld.com/businesscenter/article/152673/netherlands_teen_sentenced_for_stealing_virtual_goods.html

6. 2008 ï Lori Drew was indicted and convicted on three misdemeanor accounts of accessing computers without authorization. She had created a MySpace identity and then used it to bully a classmate of her daughterôs, which ultimately led the classmate to commit suicide.

   http://www.nytimes.com/2008/11/27/us/27myspace.html

In his article on the law and the emotive avatar, Gibbons [29] points out that the best way to look at an avatar from a legal perspective is to ñconsider the avatar as an extension of the individual or an agent of the individual in virtual spaces and then to shift the focus of analysis away from the avatar and back to the individual because it is the potential negative effects that virtual behavior may have on real-world individuals that the law seeks to regulateò

This leads to the ability to take the laws of the real world and apply them to virtual worlds. He points out that the EULAs and social norms can be seen as private laws and those private laws can be used to support existing real world laws such as the Computer Fraud and Abuse Act in the US. The key is removing the veil of the avatar

and focusing on the real person driving the avatar. In the Drew case above I am surprised she was not charged with involuntary manslaughter as her acts led to the death of the classmate, even though that was not the intention.

There are a number of questions that need to be answered from a legal perspective that are somewhat unique to virtual worlds. As an example which countryôs laws apply to actions between two avatars when they are in two different countries? Do real world laws overrule virtual world EULAs? If you are married in real life and have an affair in a virtual world is it cheating and grounds for divorce? Is it moral? And what happens if you get married in the virtual world ï did you just commit bigamy? And then we have the disposition of assets when you get a virtual divorce or you die. Do those assets go to the online wife, real life wife or someone else? These are all interesting questions to be addressed in the future but that bear thinking about.

## APPENDIX B

### Various theories of existence

*“Don't believe what your eyes are telling you. All they show is limitation. Look with your understanding, find out what you already know, and you'll see the way to fly.”*

- *Richard Bach*

Metaphysics is the branch of philosophy concerned with explaining the nature of the world. It is the study of being or reality. It answers questions such as: What is real? Is it natural or supernatural? Sweeney [1] further defines metaphysics as the acquired intellectual ability by which one knows existents (entities that exist) precisely as actually existing. Other definitions include those at the Free Dictionary [2] such as ñthe branch of philosophy that examines the nature of reality, including the relationship between mind and matter, substance and attribute, fact and value.ò

Below is a short list of philosophers and a brief summary of some of their thoughts.

### Ancient Views

### Plato (428 – 348 BC)

*‘Reality also involves stability’*

Plato thought that the universe was full of change, flux and fluency. Reality, for Plato, existed in a separate realm from the real world, a realm he called the forms. Effectively Plato split up existence into two realms, the material realm which was constantly

changing and was perceived through the senses, and the transcendent realm of the forms which was permanent and unchanging and was perceived through the mind.

For Plato, the concept of a dog (or form) was real, as it was stable and unchanging. Individual dogs were not truly real because they were changeable and material. So Plato saw the visible or material world as not being genuinely real since reality requires stability (immutability), self-identity, intelligibility, permanence, truth and eternity. The forms were transcendent, which means they do not exist in space and time. A material object was seen to exist in a different manner to, and independently of, the forms it might manifest. All material objects were copies of the forms and the forms were the causes for everything that existed in the world. The forms also contributed to our understanding of those material objects.

So, if we agree with Plato, material objects are images of collections of forms, and are not real. They have no true worth as they are changeable and they only exist in the sense that they have structure or form. Sweeney [1] suggests that this can be interpreted as saying that a soul, by becoming part of a human individual, actually destroys its true reality. Plato raises some interesting points ï while he cannot account for individuals, his theory does work for the existence of universals such as numbers, trigonometry and geometry.

In discussions, Dr Miller[1] made the point that the forms could almost be viewed as a template for the instances. Using the example of a computer program he suggests that the forms are similar to abstract data types or object definitions that take up no CPU or memory resources on the computer. The actual instances are the entities that use the resources.

If we analyze avatars using this metaphor then the definitions and code for the avatars would be the forms and the avatars themselves, since they use resources, would be the actual instances or images of the forms. The driver would be the originator of the change that caused the instantiation of the avatar.

**Aristotle (384 – 322 BC)**

*'Reality involves both change and stability'*

To Aristotle, the study of metaphysics meant the study of things beyond nature and the sensible world. For human beings this included the soul and discussions of what it meant to exist or to be. Where Heraclitus thought reality involved change and Plato thought it involved stability, Aristotle merged the theories of Heraclitus and Plato into a combined theory. This was called his hylomorphic theory and he argued that both changes and permanence were factors in reality. Hylomorphism is a compound word consisting of [3]: ñHylomorphism (Greek - hylo-, ómatterô + -morphism óformô). In philosophy this is the view that every natural body consists of a potential form (primary matter) and an actual form (substantial form). It is a philosophical concept that highlights

[1] Dr Keith Miller, Professor Computer Science and Philosophy, University of Illinois Springfield

the significance of matter in the composition of being, regarding matter to be as essential to a being as its form. In simple terms, hylomorphism is the view that a substance is defined by a combination of the matter from which it is made and the form which that matter takes.ò

For Aristotle this meant that human beings have a material and a formal reality. He pointed out that for Heraclitus actual existents remain primarily stable while changing and, for Plato change is real, because it is part of the nature of individual things. He differed sharply from Plato in his belief that stability is within actual things, not in a separate realm of forms or essences. Thus, an individual is a combination of stability and stabilized changes.

According to Aristotle there were two kinds of change and stability

a) Those on the substantial level
b) Those on the accidental level

Substantial changes were those that changed the actual substance itself. Changes on the accidental level did not change the actual substance ï they referred to changes such as color or height, etc.

Additionally, he argued for a unity or state of cohesion where the whole was always distinguished from its parts, but where there were two kinds of unity possible. The first kind was the kind of unity found in a house, machine or team. This was called accidental or extrinsic and was basically a unity where the parts worked together to form

the whole, but where those parts did not substantially change the final unity. The second kind was the kind of unity found in living beings and was called intrinsic or essential. It refers to the fact that some parts were actually essential to the existence of a living being i.e. a human without some kind of heart is not going to be alive.

Aristotle described the relationship between substance and accidents as being one of potency to act. In this case, act was a perfection and potency was the capacity to receive a perfection. A perfection was defined simply as that which an item ought to possess in order to be what it is.

Thus, Aristotle thought a material thing was made up of two real and diverse components

a) Substantial form

That by which a thing is of a specific nature

b) Prime matter

That by which a thing receives a substantial form, and yet which is capable of receiving other substantial forms.

So, if we look at an individual person: they are an intrinsic unity where the substantial form is that which makes them human and their prime matter is that by which they have received their substantial form and are now this person. Looking at this description we can see that prime matter must therefore be pure potency or a recipient and a capacity, and substantial form must be an act.

Aristotle's hylomorphic theory was designed to answer questions about change, and it can be difficult to apply to human beings. As an example, a lump of bronze has the potential to become a statue but it does not come into being as a statue until it receives its substantial form. The human body is a different entity as it does not lie around waiting to receive its substantial form. Aristotle, as with Plato, seemed to think that human existence required a body and a soul and that the body came into existence once it received a soul. It is not so clear where the body came from (what caused it to come into existence).

Applying the theory to avatars is somewhat challenging since they are clearly not human or alive in the way we normally think of with people. We think of prime matter as potency, something that receives a substantial form to become the final specific person. For the avatar, the closest parallel would be to think of the code (and computers etc. necessary to run it) as prime matter and the driver who acts to run the code as the substantial form or act. In this way, we could argue that the instance of the avatar is the unity of the code and the driver.

**Avicenna (980 - 1037)**

*'Essence precedes existence'*

Avicenna was inspired by Aristotle's distinction between form and matter, and was adamant that essence was quite different from existence. To Avicenna essences could be present in things (where they were part of the reality of things) or in intellect (where

they were conceived of). Essences were viewed as being possibilities that entered the realm of the actual only if existence, which was seen as accidental, actually occurred. He seemed to view existence as a property that was quite different from essence, which was quite different from how Aristotle viewed it. As an example, a dog has a nature or kind that makes it a dog (dogness) and this is its essence. However, it also has an existence ï an actual dog to which the ñdognessô adheres. Avicenna saw existence as an ontological part of any existent.

In this view there would be a kind or nature that makes something an avatar that would be seen as the essence of what it means to be an avatar. And existence would involve an actual avatar being instantiated.

**Rene Descartes (1596 – 1650)**

*'Existence is a perfection'*

Rene Descartes took a mechanist view of nature, and saw the world as a series of mechanical interactions that could be explained, rather than the mysterious world described by earlier philosophers. He applied that view to all things in nature including animals and humans. However, he was also well aware that humans had some additional capacities over animals ï they can consider the past, the present and the future; they can resist their desires or decide to act on them and they are capable of understanding intangible entities such as mathematics and the minds of others. His explanation for this was that humans possess both a body and a soul. Descartes separated the subject and its world into two categories, the first being *res extensa* (the

physical) and the second being *res cogitans* (the mental). He argued for a cartesian separation of the mind and body.

Descartes was quite clear that there is no real distinction between a thing's existence and its essence. He argued that essences or natures only exist when they are realized in particulars. In this world the perceiver has no real way to distinguish between real and dreamed or imagined experience. He went on to argue for the existence of a benevolent God as such a God would not allow him to be so deceived all the time.

Finally, he also argued that the mere fact that humans can think and be mistaken is indubitable proof that they actually exist ï hence ñI think therefore I amò

In this case the essence or nature of the avatar would only exist when the specific avatar is instantiated by the driver.

**Immanuel Kant (1724 – 1804)**

*'Being is a logical, not a real predicate'*

The ontological argument tries to prove Godôs existence through abstract reasoning alone, which ends up in an *a priori* argument with no empirical evidence supporting it. Kant argued against the version of the ontological argument that can be summarized as follows [35]:

ñ

1. God, the *ens realissimum*, is the concept of a being that contains all reality/predicates.
2. Existence is a reality/predicate.
3. Therefore God exists.

ñ

Kant contended that the argument treats existence as a property that entities can possess or lack, and he pointed out that this leads to inconsistency. To say that God exists is to attribute the property of existence to God. That is to say that there is something in the world that has the name God and that it satisfies the description ñs existentò. The problem is with statements like ñGod is not existentò as this can only be true if there is something in the world that has the name God and that satisfies the description ñs not existentò. Kant denied that existence is a predicate and instead said that existence should be seen as a concept corresponding (or not) to the world. So to say that something exists is to say that the concept of the thing is exemplified in the world.

Kant (Critique of pure reason ï 1781) also differentiated between the noumenal (physical) world and the phenomenal (appearances) world, pointing out that we do not actually experience the noumenal world; our experiences are dependent on our perceptions ï thus the world we experience is the phenomenal world. He thought the noumenal world was real but that its nature was hidden from us and that we could never have certain knowledge as to whether it really existed or what features it had.

.

Using Kant, saying that an avatar exists would be the same as saying that the concept of an avatar is exemplified in the world. Since our experiences are based on the phenomenal world of appearances and the noumenal world is hidden from us, we can never really know if avatars exist or not.

**Contemporary Views**

**Existentialism**

Existentialism is the view that understanding all the truths in natural science is not enough to understand what a human being is. It does not deny the importance or validity of the sciences ï it just claims that human beings cannot be understood completely that way. The existentialist emphasizes the primacy of individual existence over any essences ï one of Sartreôs key phrases was ñexistence precedes essenceò and, in this view, that which is essential to a human being is determined by what they make of themselves, not a kind or nature or essence in the terms that earlier philosophers used. For the existentialist my existence as a human being entails my ability to make of myself what I want and the responsibility to employ that freedom properly.

An avatar has no ability to determine its future ï it is totally dependent on the driver and the intentions of the driver. To that end the existentialist would most likely argue that the avatar does not exist since it cannot act independently and only acts on the desires of

the driver. It is the combination of the driver and the avatar that is necessary for existence.

**Friedrich Frege (1848 – 1925)**

*'Affirmation of existence is in fact nothing but denial of the number zero'*

Frege [36] applied his logic of quantification to existence where he suggested that existence was not a concept that objects fall under. He specifically distinguished between names of objects, names of first-order concepts and names of second-order concepts. Frege argued that existence is a second-level concept rather than a first-level one. For Frege, a first-level concept and predicate need to attribute something to the subject such as ñFrege is wiseò. Stating ñFrege isò does not do that, so he claims it is a second-level concept. He argued that ñaffirmation of existence is in fact nothing but the denial of the number noughtò [4]. So to say something like ñMartians existò is to say there are more than zero Martians. This makes existence a property of a concept and thus it is a second-level concept.

Frege uses his logic to show that the word ñisò has four different meanings as follows:

1. Existence
    a. Frege is ï this is a statement that Frege exists.
2. Identity
    a. Cicero is Tully ï this is a statement that Cicero = Tully, that they are numerically identical to each other.
3. Predication

a. Frege is wise ï this is a statement about a quality held by Frege.

4. Inclusion or generic implication
    a. Man is an animal ï this is a statement about classifying Man into a group of some kind.

In all four cases ñisòhas a different meaning and these meanings can be ambiguous.

Frege said that by moving to a conceptual notion of existence that is dependent on descriptions rather than named objects, it was possible to avoid Platoôs Beard (predication about non-existent objects).

In this case, a statement like ñavatars existòis the same as saying that there are more than zero avatars.

**Bertrand Russell (1872 – 1970)**

*'Universals do not exist but rather subsist or have being'*

Russell took a very simple view of meaning with respect to objects and predicates. Names get their meaning from the objects they stand for and predicates get their meaning from the universals they stand for. Russellôs theory of singular descriptions allows us to use names for supposed entities without presupposing that the actual entities exist. Names refer to a specific entity but descriptions are more general. Effectively the named object gets replaced with a description (e.g., Pegasus might be replaced with a winged horse). It allows for discussion of an entity at a more abstract level such that the non-existence of an object can be discussed without getting tangled

up in Platoôs beard, where mere discussion of the object would require acceptance of its existence.

So to summarize Frege and Russell, they would maintain that a statement like ñthe avatar adespota existsò is not about the avatar itself, but is about the various properties of the avatar. So the avatarôs existence is seen as a second-order property of those first-order properties that happen to be instantiated.

**Jean Paul Sartre (1905 – 1980)**

*“Human beings are not what we are”*

*‘Existence precedes essence’* [5]. This is actually a statement about the nature of human freedom rather than the nature of existence. Sartre thought that a person defined themselves by a) acting and b) being responsible for his or her actions. So for Sartre people define their own essence by how they live a life and the choices they make. Therefore existence must precede essence, as essence is developed over time but one has to exist in order to develop it.

Avatars do not control their own actions so Sartre would most likely say that an avatar exists when it is instantiated by the driver and code, but that its essence is dependent on the driver over time, since the driver is the one determining how the avatar will behave.

**Friedrich Nietzsche (1844-1900)**

*“Existence really is an imperfect tense that never becomes a present”*

*“The irrationality of a thing is no argument against its existence, rather a condition of it.”*

Nietzsche, along with Kierkegaard, was one of the first to focus on the human experience rather than the truths of mathematics and science as a way of understanding human nature. Nietzsche founded his philosophy on a conception of human existence where people defined the nature of their own existence by the choices they made. Individuals invented their own values and created their own goals and thus determined their own existence. In this view the existence of the avatar would be as an extension of the driver, since the driver is the determining factor for behavior.

**William Quine (1908 – 2000)**

*‘To be is to be the value of a bound variable’*

Quine [6] raised two important points when evaluating how we deal with abstract objects and their potential existence. In his first thesis, he responded to statements regarding the ability to discuss abstract objects and the statements made by other philosophers by arguing that the mere fact that we can discuss them means they must exist which is a contradiction. Quine makes two points about discussing the existence of objects like Santa and Pegasus and other abstract objects such as the universals.

Quineôs first point is that saying an object does not exist does not imply that the thing actually exists and so there is no contradiction. His argument is based on using Russellôs Theory of Descriptions to analyze ontological assertions, showing that this

preserves the semantics of the assertions without implying that the object under discussion exists.

His second point relates to non-platonistic properties which are properties such as colors that describe something about an object. His point here is that saying something like ñthe house is redò does not imply that the property (in this case redness) exists as an independent entity.

Quine argued for such properties as descriptions but not as individual entities. A red house is an individual entity, but red itself is not ï redness would be a non-Platonistic property of the house in question. Later he went on to show that the statement ñsome dogs are whiteò does not entail that we commit to doghood or whiteness. The argument was somewhat convoluted but rested on the bounding variable ñsomeò. The statement was interpreted as saying: in order for the statement ñsome dogs are whiteò to be true there must be some white dogs, but that is not a commitment to whiteness or doghood as the statement could be false.

Quine illustrates these points in ñOn what there isò [6] using Pegasus and a fictional philosopher ñMcXò as an example. In ñOn what there isò [6], McXs argument seems to hinge on an argument referred to as Platoôs Beard. According to McX, the mere fact that we are discussing Pegasus means that Pegasus exists, otherwise we are talking nonsense. Given that denial of the existence of Pegasus would then be irrational, then Pegasus must exist. While not arguing for the actual physical existence of a winged horse by that name, McX argues at least for the existence of Pegasus as an idea.

Quine argued that McX had confused the named object Pegasus with the meaning of the word Pegasus, and pointed out that there is a vast difference between meaning and naming. He noted that while McX does not confuse the existence of the Parthenon with the idea-of-the-Parthenon, he does so with Pegasus, as when one refutes the existence of Pegasus one is not referring to the idea-of-Pegasus, but to the physical animal.

Quine also rejected the coupling of meaning and naming even for objects recognized to exist. He cited the classic Hesperus/Phosphorus, Evening Star/Morning Star example, which, although they refer to the same object, have different meanings. With meaning not requiring naming, we can say that Pegasus is not without tripping into meaninglessness.

Quine showed that you can rephrase the word Pegasus as a description and then analyze whether Pegasus is or is not by using Russellôs theory of singular descriptions. The premise Quine was rejecting is that something must exist in order to even discuss it and his point is that the distinction between descriptions and particulars allows us to deny the existence of Pegasus without asserting his existence.

Clearly it is important to make a careful distinction between what we should believe there is (what our ontological commitments should be) and what actually exists. Ideally our ontological commitments to objects would match up with objects that actually exist, but first we have to figure out what actually exists.

Quineôs answer to how one should figure out what actually exists was that we should look at our very best theories of how the world is, and see what would have to exist if the claims those theories make are true. Once we understand the ontological commitments of those theories, we should make them our own ontological commitments because we don't have a better answer about what exists.

Quineôs theory of ontological commitment was that we commit ourselves only when we say something is. We do not commit ourselves when we say something is not. When we look at intangibles or abstract objects we can review them according to the Quine-Putnam indispensability argument for mathematical reason. This argues that we should believe that numbers exist and the argument goes like this:

1. We should have ontological commitment to entities indispensible to our best scientific theories
2. Mathematical entities are indispensible to our best scientific theories
3. Conclusion: We should be ontologically committed to mathematical entities

This is an *a posteriori* argument.

So to summarize Quine ï If we think X is necessary for our theories then X must exist. Thus, in terms of abstract objects a Quinean would argue that numbers exist. And in the case of avatars we can argue that they exist as they are necessary to our theories about virtual worlds.

**Robert Nozick (1938 – 2002)**

*'Pleasure is not the only answer'*

To answer Jeremy Benthamôs hedonistic view that pleasure was the only good, Nozick attempted to show that pleasure is not the only thing people care about. Nozick suggested a thought experiment called the experience machine [24], in which he imagined people making a choice between reality or a more pleasurable simulated reality. According to Nozick, most people would choose to experience reality rather than the virtual world, even though the virtual world would provide more pleasure to them. His argument went something like this:

P1. The most important thing is pleasure.

P2. If an action A provides more pleasure than an action B then we have reason to take action A.

P3. The experience machine provides more pleasure than the real world since we can program into it that which is pleasurable to us.

C: If pleasure is the most important thing to us then we would choose the experience machine

Nozick argued that if all that mattered to us was pleasure then we would want to plug into the experience machine. However, Nozick said that people would not choose the experience machine for several reasons, most of which were based on our desires to actually do things not to just experience them in simulations. Therefore, there are things that matter to us besides pleasure.

The experience machine was one of the themes in the 1999 movie, the Matrix. When Neo is given the choice between remaining in the Matrix (the simulated world) or experiencing the less pleasurable world he chooses the real world, rather than living in the computer generated dream world that he is currently in, saying:

*'Because I don't like the idea that I'm not in control of my life.'*

Conversely, after experiencing the real world, Cypher chooses the other direction:

*ðYou know, I know this steak doesn't exist. I know that when I put it in my mouth, the Matrix is telling my brain that it is juicy and delicious. After nine years, you know what I realize?*

*[Takes a bite of steak]*

*Ignorance is bliss.'*

A virtual world is similar in some ways to the matrix ï the computers and code are necessary for the instantiation of the avatars and the avatars actions are controlled by an external entity (the driver). Some drivers approach virtual worlds in the way that Neo does, wanting control of their life, something they have in a virtual world. Others are more like Cypher, wanting a world to escape into.

**Stephen Yablo (Current)**

*'Just pretend Platonic objects exist'*

Yablo [34] says that we can understand the world by simply pretending that Platonic objects exist. His main argument is that pretend Platonic objects are just as useful in

understanding the world as real Platonic objects would be. Thus, there is no reason to posit their actual existence.

In his Paradox of Existence, Yablo [34] points out that you have to choose between Quineism and rationalism - Quine says that the only way to establish existence of those abstracts is via the *a posteriori* indispensability argument. Rationalism, however, says the existence of these abstracts follows from reason or what Yablo calls *a priori* bridge principles. A bridge principle is one that establishes a relation between empirical evidence and the independent world so that we can deduce a truth about the world. An example would be something like: Pete watched television quietly if there was a watching of television by Pete and it was done quietlyò Basically you are saying that x is true if x happened and it happened in the manner described.

Yablo [34] states that: ñAbstract objects are *a priori deducible* from (i) *a priori* premises and/or (ii) obvious, uncontroversial empirical facts.ò This leads to the paradox as one is *a posteriori* and is seen as what Yablo refers to as ñreceived opinionò, whereas the rationalist view is *a priori* and Yablo points out that it is a consequence of received opinion. This leads to a form of infinite regression which is described by Malleck [32] when Malleck says: *"Yablo's paradox creates an infinitely descending chain of Liar-like statements which are not self-referential or directly circular."*

To summarize Yablo ï weird objects like numbers or sets are make believe and that we can understand the world by simply pretending that Platonic objects exist. So, in terms

of avatars, Yablo would most likely say they that we can treat them like Platonic objects and just pretend they exist.

**David Chalmers (1966 - Current)**

In his article on the extended mind, David Chalmers and Andy Clark [7] argue for what they call active externalism. This is where the external features play a role in determining outcomes. In order to qualify the features have to provide what they call reliable coupling ï i.e. they have to be there when you need them He uses the example of a calculator or slide rule ï if they are always there and I use them all the time then they are part of my cognition. He provides examples of features he thinks are constituted in part by the environment ï one such example is beliefs. Beliefs are critical to our decision making ï they can be based on things in our memory or they can be based on things we have written down in notebooks or stored on a blackberry or similar. He is basically arguing that external things that affect decisions we make or that we depend upon are really part of our mind even if they are external.

As an example, in the article [7], he argues that Ottoôs notebook is a central part of his identity and thus is part of Ottoôs ñextended selfò. He points out that this has philosophical implications such as the moral implications of interfering with someoneôs environment. If someone was to remove or damage Ottoôs notebook then they would be doing harm to Otto as it would affect his ability to function.

The concept that the mind can extend beyond the body is interesting. People used to have to remember all kinds of things, but today they have computers and smartphones to do that for them. Many could not function without their Blackberry or iPhone ï not only is it their email and web browser and alarm clock, but it is also an extension of their memory. Every name and number that is important to them is in there, as is every place they have to be and when. They no longer store that in their brain. The good news is that, unlike their mind, people can back up their Blackberry or iPhone data and restore it.

Clark [7] provides a different perspective on identity that applies to virtual objects. An avatar can be seen as an extension of the person, a form of extended self that is able to perform functions that the player or driver cannot. This leads to some interesting moral implications with respect to virtual harm and accountability.

**Summary**

After reviewing the views of these philosophers it is apparent that it is possible to apply the views of some of the ancient philosophers to the question of the existence and status of avatars. For Plato, the data types, code and object definitions could be seen as templates for the instances and thus could be viewed as the forms. The avatars themselves would be the actual instances as they are the entities that use resources and that are modeled on the forms. For Aristotle, the data types, code and computers could be seen as prime matter as they provide the potential for the avatar. The driver could be seen as the act or substantial form. For Kant we have no way of knowing if the

avatar exists as we only experience the phenomenal world of appearances and have no way of knowing what really exists in the noumenal world. So stating that an avatar exists merely means that the concept of an avatar is exemplified in the world.

Moving on to more modern philosophers, the existentialists could argue that avatars do not exist as they do not determine their own behavior. For the existentialist existence entails having the ability to make of oneself what one wants and an avatar is dependent on the driver for that so, from an existentialist perspective, the driver would exist but the avatar would not. Nietzsche thought that people defined the nature of their existence by the choices they made so he would most likely say that the avatarôs existence depends on the driver as the driver determines behavior, hence the avatar is an extension of the driver. Frege and Russell would say that a statement like ñthe avatar adespota existsòis actually about the various properties of the avatar and that its existence is a second order property. Quine thought that if we think something is necessary for our theories then X must exist, so in the case of abstract objects like avatars he would probably argue that they exist as they are necessary to our understanding of virtual worlds. And Yablo could argue that, as with the Platonic forms, you can just pretend they exist so that you can then use them to understand the world.

Baker dscusses five ways to determine the existence of artifacts and avatars pass four of them clearly and the fifth can be explained. Bakerôs arguments are the foundation for our argument that avatars are indeed artifacts and that, while they exist in virtual worlds, they still exist in a robust sense. They are experientially real and they have persistent

effects in the virtual world and sometimes in the real world. Finally, Chalmers et al could argue that the mind-dependent avatar can be seen as an extension of the driver; an entity that takes action on behalf of the driver, performing actions that the driver is unable to do themselves.