Life Inside The Experience Machine

Benjamin Aaron Sales

University of Colorado

Department of Philosophy

April 1st, 2013

Advisor: Dr. Alastair Norcross, Philosophy

Committee Members: Dr. Wes Morriston, Philosophy

Dr. Michael Zimmerman, Philosophy

Dr. Jennifer Wolak, Political Science

Contents

Introduction

In this paper, I will analyze and critically evaluate the thought experiment of the experience machine first proposed by Robert Nozick in his book *Anarchy, State, and Utopia*. The central question that I aim to answer is whether or not what is "real" is intrinsically preferable to what is "simulated" or fake. I will argue that if you were given the option of using the experience machine as presented in Nozick's book to create a new and better permanent life for yourself, you would be irrational to refuse it. My argument will be developed primarily within the context of a utilitarian consequentialist framework. I will begin by elaborating upon Nozick's argument in which he intended to demonstrate why we would not, and should not choose to enter the experience machine. In addition to consulting the wide array of inquiries that have been carried out on Nozick's thought experiment by contemporary scholars in support of my position, I will also be formulating a new thought experiment of my own. I will then reply to numerous objections that will be raised against my position. My thought experiment will assume that the "reality" we are currently experiencing is in fact one produced by an experience machine. I hope this experiment will replace Nozick's conclusions with my own. Finally, in addition to reaching conclusions based upon my proposed thought experiment, I will complicate the inquiry by introducing new considerations, such as age and social status, which an individual must take into account before first deciding whether or not to enter the experience machine. I will also examine real-world applications of the experience machine, the possibility of a communal experience machine, and how it may or may not hinder the conclusions previously reached.

In *Anarchy, State, and Utopia,* Nozick presents the thought experiment of the experience machine as a criticism of utilitarianism. In this experiment, we are to picture a world in which a

new technological achievement called the experience machine could change the way in which we live. If we so choose, we would be able to connect to this machine and live any sort of life that we could imagine for ourselves. (Nozick 1974, p. 644) For example, I could program a life for myself in which I am a highly successful businessman, a hero, a great author, or virtually anything that I want. Once I enter the experience machine, I lose all memory of entering the machine and instead believe the simulated reality to be the real one. I need not worry about the life I am leaving behind or about my duties to others, as my friends and family are able to enter the machine as well. Nozick argues that if it were the case that pleasure is the greatest good and the only thing that matters to us, then living in the simulated reality would clearly be a person's best plan of action. However, since he is criticizing utilitarianism, he gives us a few reasons why he believes we would choose not to enter the experience machine. He hopes to prove once and for all that there are things that matter to us other than pleasurable experiences.

The first reason that Nozick gives us is that we actually want to do things in "reality", not just have the experience of doing them. It is only because we actually want to do certain things that we also want the feeling or experience of doing them. Second, he claims that we want to be a certain way. Nobody would be happy with the idea of just floating in a tank for the rest of her life. Such a person can express no real characteristics such as courage, kindness, or love. In this sense, entering the machine is a sort of suicide. Lastly, he claims that we wouldn't enter the experience machine because doing so would limit us to a man-made reality. In his view, a world without any real contact with our surroundings, other people, or anything of deeper significance would be repulsive to us. (Nozick 1974, p.645) Upon reflection, Nozick concludes, all people should find these arguments so compelling that they will find hedonism to be implausible. This is not the intuition to which I will be appealing. The not-so-friendly character Cypher, from the

film, *The Matrix,* appeals to a different intuition when he says, "You know, I know that this steak doesn't exist. I know when I put it in my mouth, the Matrix is telling my brain that it is juicy and delicious. After nine years, do you know what I've realized? Ignorance is bliss." The theory to which Cypher is appealing here is known as the mental state account of well-being which purports that "a person's well-being is improved or diminished only by events which enter into her mental or psychological life. If an event does not have an impact upon an agent's mental life then it can have no impact upon her well-being." (Kawall 1999, p.382) Which of these moral intuitions is correct, and how do we know? To begin with, we must understand what a moral belief is, and how such a belief can be justified.

A moral belief is a belief that "something is morally right or wrong, good or bad, virtuous or vicious, and so on for other moral predicates." (Sinnot-Armstrong 2008, p.1) A moral intuition is merely an example of a moral belief that is strong and immediate, and is not brought about by any inferential process or stemming from any other belief. (Sinnot-Armstrong 2008, p.1) So how can a moral intuition be justified? Well, the two general approaches are that a belief can either be justified inferentially or non-inferentially, the former being what I hope to achieve by the end of this discussion. I also hope to show that the process Nozick follows that produced the moral beliefs he espouses is not reliable, which Nozick ought to know, and as such, he is not justified in holding those beliefs. The master argument following this line of reasoning is outlined By Walter Sinnot-Armstrong in *Framing Moral Intuitions,* and is as follows:

1. If our moral intuitions are formed in circumstances where they are unreliable, and if we ought to know this, then our moral intuitions are not justified without inferential confirmation.
2. If moral intuitions are subject to framing effects, then they are not reliable in those circumstances.
3. Moral intuitions are subject to framing effects in many circumstances.
4. We ought to know this (that is, (3)).

5. Therefore, our moral intuitions in those circumstances are not justified without inferential confirmation.

(Sinnot-Armstrong 2008, p.5)

    Much of what I attempt to show in this paper will illuminate the real extent to which framing effects alter the way we think about the experience machine, specifically *context* framing effects. A person's belief is subject to a context-based framing effect when the reason that the belief is held is based on the context outside of the actual content of the belief. My thought experiment will completely alter the context of considering the experience machine and will show that the traditional context of the experience machine as presented by Nozick is a framing effect that adversely alters the way people react to the idea of it. With this in mind, I will now explore some of the core problems and biases that naturally result from the way Nozick presents his experiment.

Status Quo Bias and Experimental Philosophy

    Before we begin our larger inquiry, it is important to first understand whether or not the question can be reasonably answered. Although you, the reader, have only a basic understanding of the experience machine, ask yourself whether you would choose to enter. If you share the inclination of the majority of people who are asked this question, then your answer is probably "no." But why is this the case? Is it because you agree wholeheartedly with the reasons that Nozick presented in his essay, or is it the cause of something else? One possibility is that people tend to fall victim to what is called the status quo bias. Maybe the reason that people don't want to abandon their current lives has nothing to do with an innate desire to be in touch with reality, but rather that people have a psychological bias in favor of the life they currently know. (De Brigard 2010, p.43) (Weijers 2011, p.101) Adam Kolber tells us to

"Imagine an investment banker with no relatives, working for twenty-five years with little or no job satisfaction. Her only pleasure in life is to come home after a twelve-hour work day and read passages from Zen Buddhist philosophers. In fact, she's come to believe that her life would be much better if she used her considerable wealth to move to Asia and study Zen Buddhism. Though she could have reason to believe that such a life would be better (given whatever conception of the good she has), she does not necessarily feel comfortable with such a drastic life change." (Kolber 1994-1995, p.13-14)

While the investment banker undeniably feels fear at the thought of such a drastic life-style change, the change and fear brought on by the prospect of the experience machine is even greater. We likely all have instances in our lives where we have realized we can make a decision to make ourselves happier, but the fear of breaking from the norm and abandoning the status quo is too strong to overcome.

In his paper, *If You Like It, Does it Matter if it's Real?* Felipe De Brigard tests this concept by proposing a reverse experience machine scenario in which we assume we are already in the experience machine, and then asks people to choose whether or not they would return to the real world if given the option. (De Brigard 2010, p. 46-50) The purpose of this is to appeal to the intuition that when the experience machine option is presented in this manner, people would choose to stick with their current lives regardless of knowledge that it isn't actually real. This would demonstrate that the status quo actually matters more to people than being in touch with reality. De Brigard reports that students randomly received one of three questionnaires: one told them that the life they would return to is one in a maximum security prison; another told them they will be a successful millionaire, and the last one told them that their life is very different from the one they are currently living. In all three scenarios, the results were skewed such that people preferred to remain connected to the virtual reality, albeit much more for the negative scenario than the others.

But perhaps the status quo bias, while plausible, isn't the whole story. Maybe the results of De Brigard's study are inaccurate because the students who were surveyed did not answer the prompt correctly. Not that the students lied or believed that they were giving false answers, but maybe they were simply incapable of answering the question because they weren't answering from the position of a "confronted agent". Basil Smith argues this point, making the claim that a major flaw in testing these sorts of philosophical thought experiments is that it asks people to make hypothetical life-altering decisions from the position of someone actually facing the dilemma, with all of the emotional conflict with which such decisions would come. Since students who are surveyed on these questions are not facing the dilemma, they tend to answer quickly, with little deliberation, and certainly without the emotional distress that would accompany a decision of whether or not to enter the experience machine. For these reasons, the students aren't in a position to give a credible answer with regard to De Brigard's or Nozick's thought experiment. (Smith 2011, p.30) In this respect, many "what would you do?" philosophical thought experiments are flawed, including the experiment of whether or not to enter the experience machine.

This is an argument against a certain kind of philosophical practice known as experimental philosophy. Experimental philosophy is similar to the more traditional mode of philosophical practice known as conceptual analysis, but with a key difference. Where conceptual analysis consists of "learning important philosophical lessons by looking carefully at ordinary people's intuitions about cases" (Knobe 2008, p.4), experimental philosophy adds the fact that actual experiments are conducted, such that the conceptual analyst might write, "In this case, one would surely say…" while the experimental philosopher would write, "In this case, 79% of subjects said…" (Knobe 2008, p.4) Much of the academic community is split on whether

or not experimental philosophy actually belongs in the realm of philosophy or whether it is better left out of it entirely. The goal of the experimental philosopher is to show, after conducting a series of experimental investigations, that traditional assumptions that have been made about the way people think about classical issues, and the underlying psychological processes at work, are not what we had believed them to be. (Knobe 2008, p.3) The most common objection to this sort of inquiry is that it actually misses the whole point of philosophy. Instead of simply trying to understand how people think by looking at their intuitions, we need to subject those intuitions to criticism and rational argument to determine when intuitions may be mistaken. (Knobe 2008, p.10) Kauppinen argues that experimental philosophers are mistaken in their assumptions about the relation between people's concepts and their linguistic behavior. (Kauppinen 2007, p.1) Of course, a fair reply to this objection is to assure the skeptic that the traditional method of moral philosophy is not being thrown out in any way, as experimental philosophy simply adds to the arsenal of ways in which the philosopher can pursue his goals. (Knobe 2008, p.10)

But is the inability to respond to such thought experiments in the form of a "confronted agent" necessarily a fatal flaw in experimental philosophy? To the contrary, it appears that Smith might be missing the point completely. As Dan Weijers points out in, *We Can Test the Experience Machine*, adopting the role of a confronted agent requires one to encompass the feelings of "confusion, incredulity, fear, and uncertainty" that come with making that sort of decision. (Weijers 2012, p.4) If someone makes his decision about whether or not to enter the experience machine based on these emotions, we may certainly get the correct answer to what he **would** do, but we may not get the correct answer to what he **ought** to do. Indeed, when posing these sorts of moral thought experiments, the ultimate goal is to discover what the morally rational decision is, or what the most objectively correct answer is, not what the emotionally

charged individual is most likely to do. If this is the case, then it is actually a good thing that we don't answer philosophical thought experiments as confronted agents, since doing so may compromise our ability to approach such experiments with the rational intuitions to which they are meant to appeal.

Defense of Utilitarianism

Before we go any further and begin a larger discussion of the implications of the experience machine and more focused reasons for entering it, it is important to understand the broader underlying theory that values happiness as the greatest good. The utilitarian doctrine has been highly criticized throughout most of its history and famously dismissed as "a doctrine worthy of swine."(Mill Ch.2, p.4) However, it has become more important in recent times as academic and government measures of general welfare have begun to switch from capability and income-based economic welfare measures to the tracking of subjective mental states and happiness as indicators of general welfare. (Angner 2010, p.1) According to Mill, the principle of utility says that actions are right as they tend to promote happiness and wrong as they tend to promote unhappiness. (Mill Ch.2, p.3) The ultimate goal in life is pleasure and happiness, the opposite of which is pain and unhappiness. So, pleasure is the only thing that is intrinsically valuable, or valuable in and of itself, and nothing else. When critics make the claim that utilitarianism is a doctrine worthy of swine, they are objecting to the apparently demeaning proposition that the purpose of everything done in life is to increase pleasure. They contest that life for humans must have some greater purpose or some higher goal than something as animalistic as mere pleasure. Their criticism, according to Mill, says "To suppose that life has (as they express it) no higher end than pleasure- no better and nobler object of desire and pursuit- they designate as utterly mean and groveling; as a doctrine worthy only of swine." (Mill Ch.2 p.

3) In the definition of pleasure or happiness as defined by the critics, we are only referring to those pleasures that are limited to the physical, ignoring the intellectual pleasures which Mill would later articulate as being far more important. This argument, however, has been criticized heavily over time and Roger Crisp uses a wonderful example to show why:

"Haydn and the Oyster. You are a soul in heaven waiting to be allocated a life on Earth. It is late Friday afternoon, and you watch anxiously as the supply of available lives dwindles. When your turn comes, the angel in charge offers you a choice between two lives, that of the composer Joseph Haydn and that of an oyster. Besides composing some wonderful music and influencing the evolution of the symphony, Haydn will meet with success and honour in his own lifetime, be cheerful and popular, travel and gain much enjoyment from field sports. The oyster's life is far less exciting. Though this is rather a sophisticated oyster, its life will consist only of mild sensual pleasure, rather like that experienced by humans when floating very drunk in a warm bath. When you request the life of Haydn, the angel sighs, 'I'll never get rid of this oyster life. It's been hanging around for ages. Look, I'll offer you a special deal. Haydn will die at the age of seventy-seven. But I'll make the oyster life as long as you like'." (Crisp 2006, p.630-631)

The intuition to which this example is appealing is that if we take the reductionist approach to hedonism with the claim that pleasure is the only thing that matters, then at some point we must concede that the oyster's life is more desirable than Haydn's, because an infinite amount of lower pleasure experienced by the oyster will eventually outweigh that experienced by Haydn. This does not sit very well with us, but perhaps we are falling victim to a status quo bias once again. The mere fact that the two options we are given are split between a human being and a minimalistic animal causes a sort of framing effect that naturally results in a bias towards the life of Haydn. Perhaps we aren't thinking properly about maximizing pleasure when presented with this prompt because although we are briefly told that the oyster has more pleasure in the long run, we are still talking about an oyster, an extremely low form of life that a human being would never in his right mind want to take over a human life. The reason for this is that we, of course, have never really experienced what it is like being an oyster living an infinitely long life. But presumably, thinking about it from our perspectives as human beings, it appears that it

would be incredibly boring and unengaging, or in other words, not pleasurable at all. As such, we are not capable of getting outside of our own heads enough to see the oyster life as something we would find acceptable, because as far as our human-based limited perspective permits us to perceive, it is not. I am not personally going to come out and argue that I would take the oyster life, or that a true hedonist would certainly pick the oyster life after doing the hedonic calculus, but I am willing to admit that the reason is most certainly because I am biased in a way that is unavoidable. Perhaps, if the prompt were written in a way that eliminates the bias that is undeniably sparked within us as we read it, we would choose the utilitarian answer.

Now there is another common objection to the experience machine to which I would like to respond. When formally written out, it can be analyzed as an anti-utilitarian argument that conforms to Nozick's approach. Upon discussing the experience machine with numerous colleagues of mine and asking whether or not they would choose to enter the machine, I would commonly get a response that went something like, "Well I wouldn't personally choose to enter the machine but I can understand why some people would." This response confused me because the way I am describing the experience machine and the way I make my argument is meant to show that *everyone* would maximize their overall happiness by entering the machine, not just some people. This follows Nozick's line of reasoning in the following way;

1. If utilitarianism is true, then **you** would plug into the experience machine when given the option as presented by Nozick.
2. But in fact it's not the case that you would plug into the experience machine when given the option as presented by Nozick.
3. Therefore, utilitarianism is false.

(Feldman 2010, p.10)

Now we already know that there is a big distinction to be made between what one would do and what one ought to do, and that is where I believe Nozick and those who agree with him have been mistaken. Just because you wouldn't choose to enter the machine doesn't mean

utilitarianism is false because we know that entering the machine would result in more overall pleasure. There could be many other reasons for this and the status quo problem is one such argument. The argument from my colleagues leans in this direction. Their claim seems to indicate that someone would choose to enter the machine only if their current lives have reached some sort of threshold for unhappiness that makes the transition appealing, while if you are very happy, or even relatively happy in your current life then you should not enter the machine even with the guarantee of more happiness, because you would prefer not to leave the status quo.

There is also an argument to be made for choosing **not** to enter the experience machine that is consistent with utilitarianism. One might argue, from the utilitarian perspective, that when we look at the whole of humanity over time we see cultural, technological, and overall human progress taking place that has resulted in increasing overall happiness. Technological and societal progress has made it possible for human beings to experience new, different, and greater kinds of pleasures that have not always been available to human beings. So, one might argue, if society as a whole decides to plug into the experience machine, we necessarily bring humanity's progress to a standstill, which limits us only to the kinds of pleasures currently known to us, and alienates us and future generations from experiencing the different and potentially greater forms of pleasure that would be available in the future if we continue to progress naturally in the "real" world. This is certainly an interesting idea and one that some people might find attractive, and if true, really throws off the goal of Nozick's experiment. Since the whole purpose of the experience machine was to show that utilitarianism was false because people would choose not to enter the machine, the credibility of the experiment really begins to degrade if utilitarianism actually turns out to be a good reason not to enter the machine. I, however, do not adhere to this belief and will continue to argue that everyone ought to enter the experience machine. While

societal progress is undeniably important and we don't know what sorts of potential sources of pleasure and happiness are ahead in the future, I don't believe they can add up to the amount of happiness that can result from a world designed specifically for the purpose of maximizing your personal happiness. But this will be expanded on in a later section.

There is also a response that can be made to Nozick's claim that we really want to do things in reality and be a certain way, not just have pleasurable experiences. This kind of argument can be written as the following:

(1) We would not abandon our commitments and values, even if such behavior were to lead to better mental states for us. This is reflected clearly in our reactions to the experience machine.
(2) Hence, it is clear that we value more than our own mental states.
(3) Therefore, there is more to our well-being than mere mental states.
-(Kawall 1999, p.384)

Much of what I argue in this paper is that these intuitions are untrue, and my thought experiment is meant to show this. But for the sake of argument here, let us assume that this is true. Why do we value really doing things, being a certain way, and being in touch in with reality? Why would we prefer to actually write poetry instead of just have the experience of writing poetry? Is it because these things extend from some existential form of the good that exists outside of human beings? Maybe, but there is probably a better explanation. Coming from your perspective as the person outside the experience machine, you value actually doing things and not just having the experience of them because you **believe** that actually doing things in reality will make you happier. The intuition that life in the experience machine is not one we would want to live can stem from all sorts of reasons, but when we brush everything aside and boil down to the bare essentials, the one real reason is that we **believe** we will be happier remaining in reality, and happiness stemming from real accomplishments matters more than happiness resulting from simulated ones. For example, we would prefer to actually write a good novel instead of just having the experience of writing a good novel because, living in reality, we naturally value our

accomplishments in reality over a simulation, even if the conditions surrounding the writing of said novel are exactly the same. As such, we believe that writing the exact same novel in the experience machine would be less of an accomplishment than if it were done in reality, and if we believe this, then the resulting psychological state of a person writing a novel in the experience machine is "less happy" than the person writing a novel in reality. (Jolimore 2010, p.334) Of course, this line of reasoning only applies to the person who is currently in reality and is pondering the experience machine, as the person already in the experience machine would be completely unaware of the fact that he is in the experience machine in the first place. As such, the reasons we prefer reality over the experience machine are actually because of how we believe our happiness will be affected. (Silverstein 2000, p. 296) If this is still unconvincing then consider the reverse. If simulated happiness and accomplishments do not matter compared to real ones, then simulated pain and suffering would not matter compared to real pain and suffering either. Consider the following:

"A person is hooked up to an experience machine by a sadist and is made to believe that she is being physically tortured. Given that the pain of the simulation feels just as real, and just as distressing as pain brought about by means of actual physical torture, it is not clear that her situation is significantly less bad than that of an "actual" torture victim, and it would be highly implausible to say that her experience is not intrinsically disvaluable at all." (Jollimore 2010 p. 334)

Using the negative perspective to analyze the positive, it is less clear that "real" happiness and accomplishments should necessarily matter more than simulated ones.

Dan Weijers, in *Reality Doesn't Really Matter,* asks us to consider the character Cobb's scenario in the film *Inception.* For those of us who have seen the film, no-matter how many times we have re-watched it to try and answer all the questions, we are still left at the end of the film wondering whether or not that spinning top is going to fall. We desperately want to know whether or not Cobb ended up in reality or in limbo. According to Weijers, we are all completely

missing the point! For all the questions left unanswered, the one thing that we can say with absolute confidence is that Cobb ends up happy. Whether he ends up happy at home with his kids in reality or in limbo doesn't matter because in the end, he got what he wanted. (Weijers 2011, p.92-93) Like me, Weijers takes the argument a step further and argues that Cobb will actually be happier and be better off if he really is in Limbo instead of reality, but I will explore that argument later.

The Epicureans were also rigorous defenders of the doctrine of utilitarianism. They respond to criticisms by pointing out that the accuser is making a fundamental error in assuming that physical pleasures are the limits of what human beings are capable of experiencing as well as being the only type of pleasure that is worth pursuing. In fact, the Epicureans go even further to say that even if it were the case that human beings were only capable of experiencing physical pleasures and no others, then pleasure would still be the ultimate aim in life, and the extent to which we can lead good lives would just have to be as good as swine. The only reason that such a notion appears degrading is because as human beings we are, in fact, capable of much higher pleasures, making it the case that the pleasures that would satisfy a mere beast couldn't possibly satisfy us as well. (Mill Ch.2, p.5)

Another defender of utilitarianism, Bentham, agrees with each of the other utilitarian writers that humans are different from animals because of their ability to experience different sorts of pleasures that include, but are not limited to, physical pleasures. According to Bentham, following the ethical theory of utilitarianism, the best act is the one that produces the greatest amount of pleasure for the greatest number of people. However, in the case of the experience machine, we can dismiss this notion since there would be only one person in any given simulated

reality. The doctrine can then be modified to say that the best act is the one that produces the greatest amount of pleasure for the one individual in the experience machine, and nobody else.

In his paper *Hedonism Reconsidered*, Roger Crisp defines hedonism in a way that attempts to avoid the sorts of controversial words like "pleasure" and "pain" that make it much easier to criticize and dismiss the theory. He defines hedonism as, "the view that what is good for any individual is the enjoyable experience in her life, what is bad is the suffering in that life, and the life best for an individual is that with the greatest balance of enjoyment over suffering." (Crisp 2006, P.622) Sticking with the term enjoyment, Crisp applies this definition to the Stuart Rachels account of the unity of enjoyment and internalist account of well-being: (Crisp 2006, p.627)

1. Enjoyments are just those experiences that are intrinsically good due to how they feel.
2. Enjoyments are just those experiences that are good for the people who have them due to how they feel.
3. Enjoyments are just those experiences that one ought to like merely as a feeling; liking is an appropriate response to enjoyments alone considered merely as a feeling.

The internalist account being argued here leads to some disturbing conclusions when certain examples are analyzed through the hedonic position. Crisp asks us to consider a person who "writes a great novel, is kind, intelligent, witty, and loving, and makes significant scientific discoveries." (Crisp 2006, p.635-636) This person is meant to embody all of the qualities that Nozick claims are reasons for wanting to remain in reality, and we are to assume that this person makes all of her life decisions autonomously. Now, what would happen if the scientists who created the experience machine were to copy this person's life experiences from beginning to end, and insert them into another individual who is plugged in to the experience machine from birth? Keeping in mind that the hedonic claim, as a form of mental state theory, is that what matters to well-being is experience alone, it must be concluded that the two individuals have

exactly the same level of well-being. (Crisp 2006, p.36) On my account of the best

implementation of the experience machine, I find it to be far more important to let individuals

program their own experiences rather than be implanted with the experiences of others, in order

to achieve the greatest possible level of happiness. In the next section, I will expand more on the

possibilities of dealing with individuals who might be plugged into the experience machine from

birth.

Age and Moral Considerations

Perhaps the thought experiment of the experience machine calls for questions more

complicated than just "should I enter it?" Another question that should be answered is "who is

allowed to enter it?" For the sake of argument, let us assume a world in which the invention of

the experience machine is widely accepted and large portions of the population have decided to

make the switch to permanent life connected to the machine. In such a world, is everyone

allowed to connect? If the purpose of the machine is to allow everyone to program their ultimate

utopia, does that mean that we allow people to program lives in which they are permitted to

commit heinous crimes? If perfect freedom is allowed in this respect, then it certainly isn't far-

fetched to assume that some people will want the ability to experience things in the machine that

would be considered morally unacceptable in reality. This is an interesting question because the

response, coming from a utilitarian standpoint, may seem unappealing to many people.

From the utilitarian perspective, when given two options A and B, the best choice is the

option that results in the greatest overall happiness. Since we normally agree that the happiness a

murderer receives from murdering someone is less than the suffering caused to the victim and

the victim's loved ones, murder is wrong from the utilitarian standpoint. However, in the case of

the experience machine, the happiness a murderer receives from murdering someone is

unchallenged because no suffering is actually incurred by the other party. Given that the victim is essentially a virtual automaton and cannot experience suffering, murder becomes the best option for maximizing utility. The punch line here is that in a world in which everyone has plugged into an experience machine, what we end up with is a perfect hedonistic society. Everyone does exactly what they want, only what they want, and get everything they ask for, in whatever shape or form that may take. Consequently, there is no need for anyone to develop the social skills and virtues necessary to function in the real world such as teamwork, patience, empathy, sense of fairness, tolerance, open-mindedness, etc. In a world programmed specifically for you, the only qualities necessary are selfish desire and gratification.

This leads to another issue concerning the age of those entering the machine. If the experience machine were to become a large part of modern society, at what age should one be allowed to enter? Before answering, let us take a moment to set up this hypothetical world. More and more, people are switching to permanent life inside the machine, eventually to the point where almost every single human being on the planet is living in their own simulated utopia. The first issue to solve is how to get human beings to continue their existence. There are a number of possible solutions to this problem but for the sake of argument we can assume that populations are sustained by the machines keeping us alive, similar to the film, "The Matrix" without the malevolent intent to use human beings as batteries. Or, if we prefer something slightly less unnerving, we can suppose that people are randomly chosen to disconnect temporarily from the machine for mating purposes and are then reconnected thereafter, allowing the machines to take care of everything that comes after. (Perhaps that is still too unnerving) Regardless, the point is that the human race is continuing to reproduce even after connecting to the experience machine and, as a result, we must address the issue of children.

At what age do we allow children to enter the machine? More specifically, at what point can a human being be said to be mature enough to program a good life for himself, one with which he can live for years to come? Certainly we should think that children ought to have some sort of minimum educational experience before entering the machine. The reason for this is that education serves as a tool for expanding the horizons of the human experience and opens all sorts of possibilities for how one may better program the best possible reality. But how would this educational process be implemented? One possibility is that we design a mandatory version of the experience machine in which all children are plugged. This acts as a guiding first educational and life experience to have before unplugging and later designing their personal reality. This, however, poses a new problem as well as a solution to an old problem. The good part about this is that by placing children in a controlled environment most akin to proper development, (whatever that may be) we can better influence the kind of reality they will eventually want to create for themselves, avoiding the adverse scenario of people designing worlds in which they will perform morally and socially undesirable acts. On the other hand, by placing children in such a strictly controlled environment, we inherently restrict freedom and individuality. By completely controlling what children are exposed to in the mandatory experience machine, it is fair to conclude that when they finally get out and are ready to program their own utopia, it won't be a design made completely out of free will and subjective experience, but some design heavily manipulated by the environment that children are exposed to in the developmental simulation. In this respect, we may find whole generations of children ultimately designing realities for themselves that are more or less exactly the same across the board because they all experienced the exact same upbringing, resulting in the complete loss of individuality. While an interesting thought for science fiction, this scenario is probably more of a straw man argument

for why we shouldn't put children in the experience machine. In actuality, it is reasonable to assume that a development reality could be designed in such a way to raise children with proper values while still encouraging individuality and creativity. This way, people don't all end up exactly the same and still exercise their free will to create perfectly unique realities for themselves.

On the other hand, it is possible to avoid the entire unsettling scenario (and perchance replacing it with another) by supposing that there are no children. Perhaps when we enter the experience machine we completely leave our bodies behind and fully upload our consciousness into the machine, immortalizing ourselves in the process. In this sense, human beings will no longer reproduce, but the upload and immortalization process ensures that human beings, at least in terms of their consciousness, do not go extinct. Variations of this scenario have been theorized by many singularity theorists such as Ray Kurzweil, but I won't be taking it any further here.

The Shared Experience Machine and The Matrix

Another possibility to entertain is one in which the experience machine is not actually an isolated experience, but rather, is a shared one in which the other people you interact with are real, thinking people instead of simply programmed automatons. This scenario comes closer to a version of the experience machine that is already being experienced today, which is known as the Massively Multiplayer Online Role Playing Game, or MMORPG. I will expand more on this example later. In his essay, *Never the Twain Shall Meet: Reflections on the Very First Matrix*, Richard Hanley discusses the implications of having a communal experience machine versus a solitary one. He begins with these two cleverly chosen quotations which serve to shed light on what the discussion will entail: '"Did you know that the first matrix was designed to be a perfect human world, where none suffered, where everyone would be happy? It was a disaster.' – Agent

Smith, to Morpheus, in *The Matrix*." The quote subsequent to this states, "'Hell is—other people.' –Garcin, in *Sartre's No Exit*." Following with a discussion of the problem of evil and the nature of free will, Hanley attempts to show that the idea of a perfect utopia and ultimate happiness is incompatible with the communal experience machine and, analogously, with the concept of the Christian heaven.

Hanley relies on a fairly standard conception of heaven, conceding that there is no worldwide consensus on what heaven truly entails. But for the purpose of this thought experiment we are to assume these four commonly held beliefs about the Christian heaven.

1. It's possible for a human being to be in heaven. More precisely, if all goes well, your self as you know it will survive bodily death and go to heaven.
2. Human beings in heaven will experience happiness, but no unhappiness.
3. Human beings in heaven possess free will.
4. Human beings in heaven interact with other human beings in heaven.

(Hanley p. 125)

This sets the stage for the problem that Hanley has with the Christian concept of heaven. To begin with, people generally believe that we will meet up with all our loved ones in heaven, which means that we will remember the lives we shared with loved ones on Earth. If certain loved ones, or other individuals who inspired us in life do not make it into heaven, we would experience some sort of sadness as a result of their absence, conflicting with premise 2. We could get around this problem by supposing either that God suppresses knowledge that certain people are not in heaven, or by supposing that our desires and things we care about change such that we no longer take issue with the fact that certain people aren't in heaven, but both of these conflict with premise 3. We need premise 3 because free will is generally considered to be one of the greatest goods, so much that God found it better to create a world where people have free will but experience evil versus a world without free will where there is no evil. Now we run into the following incompatible statements.

1. Evil is explained by the presence of free will.
2. Human beings in heaven possess free will.
3. Heaven is a place without evil

So, in order for heaven to be in accordance with the premise that human beings in heaven will experience happiness, but no unhappiness, we need to either alter the wills of human beings so that no evil is incurred, or, put human beings in a solitary heaven where they don't really interact with other human beings, only virtual ones. However, in the latter scenario we must forfeit another commonly held belief that God is no deceiver, since we would necessarily have to believe that the humans we interact with are indeed real in order to experience the greatest happiness. In the former scenario, if God is no deceiver then we would have to be aware of the fact that our wills have been altered so that no suffering is incurred, but we would not care about it. (Hanley p.126-127)

The conclusion to draw from this thought experiment is that perhaps we are foolish to believe in the possibility of an existence where "human beings interact with each other and everyone has libertarian free will and nobody suffers." (Hanley p.129) If this notion is so contradictory that God can't even actualize it, then it certainly is the case that human beings could not design a communal experience machine that shares all of these values. Therefore, a solitary experience machine in which we believe we are interacting with real human beings, is the most likely application of the experience machine to achieve our goal of ultimate happiness.

Longevity And The Real World

Another aspect of Nozick's version of the experience machine that calls for attention is the measure of time spent within the machine. In order to partake in the experience machine, Nozick hints to a once-only lifetime commitment to the machine with small increments of awakening to decide what to program for the next extended period of time. This time factor has a

large impact on what a confronted agent would decide to do versus a student deliberating in a philosophy class. Arguably, one of the most difficult things about the experience machine for those posed with the question of whether or not to enter, is the fear of leaving everything you know in one instant and not getting it back for an extended period of time, or maybe never. (Kolber 194-1995, p.13) In *Hedonism and the Experience Machine*, Alex Barber refers to this as "the enterers problem." (Barber 2011 p.10) But why should this have to be the case? If the initial decision to enter the machine is the largest hurdle to overcome for those deciding whether or not to enter, why not remove all cause for concern and introduce a smaller trial period? In this scenario, all of the common concerns with the experience machine, including worries that it won't feel like genuine reality, won't truly make me happy, etc can all be put to rest without having to make a giant life-altering leap to do so.

As an analogy, this scenario is actually quite similar to looking for an apartment to rent in Boulder. Apartment complexes all over Boulder will want you to sign a one-year contract for renting an apartment, but because they are all currently occupied, they don't let you look at the place first! They give you plenty of descriptions of how great the amenities are, the square footage, and how happy you will be there, but until you make the commitment by signing a one-year contract, you'll never know what the place is actually like! Of course, I want to believe that the apartment will be as great as I am told, but there is just no way I feel comfortable making such a huge commitment without getting a preview first. "Humans are naturally risk averse, especially in situations that, as here, involve large and unfamiliar changes." (Barber 2011, p.267) With the experience machine, just as in the apartment scenario, if people could just get a chance to take a look around, all of their fears could be put to rest.

I would also like to respond to one of Nozick's claims that is meant to appeal to the fearful and emotional aspect of entering the experience machine. One of his claims is that we want to be a certain way and so we would not be happy with the idea of floating in a tank for the rest of our lives. On the face of it, I find this emotional appeal compelling, but when explored a little more it doesn't seem so bad. Framing the experience machine in this way would make just about anyone wary of it. But, as Barber points out, "we all already own and enjoy using something like an experience machine on a daily basis: our beds." (Barber 2011, p.267) We spend approximately a third of our entire lives sleeping in a bed, having dreams akin to the experience machine albeit with less conscious control. But nobody finds this idea strange or repulsive. The only difference with the experience machine is that we will be asleep for a larger fraction of our lives than a third, and instead of dreams we will be in a world of our own creation designed specifically to maximize our happiness. If we take away all the talk of "floating in tanks" and "brains in vats," and approach the scenario from the analogy of sleeping in our beds, the concept immediately feels less sinister. Once again, Nozick seems to be stacking the cards in favor of the fearful.

Now let us look at the prospect of a preview option from Nozick's perspective. Nozick appears to be making the claim that because we value actually doing things, being a certain way, and being in touch with reality, that we wouldn't choose to connect to the machine for any period of time at all, including a short preview over an afternoon. (Belshaw p.2) Not only is this an extreme position to take, but it is also incorrect. I can say this not only because I am confident that the reader of this paper would feel more comfortable about connecting to the machine with a preview option, but also because there is already solid evidence of this in real life! Having written about the experience machine in 1974, Nozick had no idea that his thought experiment

could actually be put to the test in the not-so-distant future. In modern times, the rapid advancement and explosion of online gaming is a giant step towards achieving equivalence to the experience machine. Massively Multiplayer Online Role Playing Games (MMORPGs) constitute a relatively recent phenomenon, which is already changing the way many people live their lives. In such simulations, people are encouraged to create and develop their ideal avatars in stunning detail; choose their lifestyle, class, profession, etc; and generate their own adventurous experiences in a fantasy world filled with so much content and detail that you could spend years exploring and developing your simulated life. Unlike a traditional video game where the user mindlessly performs the tasks set before him by the developer, MMORPGs require user participation in making personal, conscious decisions about how every simulated experience will be played out. Stunningly, studies have shown so far that 45% of all those who play MMORPGs play for more than 20 hours per week and 50% play for more than 10 hours continuously. (Hsu 2009, p.990) To give those statistics some numerical context, a popular MMORPG called "World of Warcraft" has currently over 10 million subscribing players, and that is only one of dozens of these games that people play all around the world! Of course these simulations of avatars are merely characters on a computer screen, but what will such simulations look like in ten, twenty, or thirty years from now? "As virtual technology makes the flight from reality ever more familiar and unthreatening, we can expect the percentage of yea-sayers to increase." (Barber 2011, p.10) In the documentary *Second Skin,* Edward Castronova puts MMORPGs in an interesting perspective when he says "In my mind a competition is afoot between synthetic spaces and this which we call the real space. We have one world which seems to be more fair and more fun, for lots of people, than our real world. If that's true, that means we could have hundreds of millions of people spending lots of time in there, which would have a dramatic

impact on life out here." As time goes on, these simulation games may begin to look more and more like Nozick's experience machine, and there isn't a doubt in my mind that people will be plugging in.

New Thought Experiment

Let's consider the same thought experiment as Nozick provides but with one key difference, which will constitute my thought experiment. Let us assume I am already in an experience machine and that the reader of this paper doesn't actually exist. (The reader of this paper, if he actually exists, is to assume that Ben Sales doesn't really exist.) This is actually perfectly compatible with Nozick's original thought experiment because as far as I know, I have already entered the experience machine, programmed this life for myself, and of course forgot about it the instant I entered the machine. Hence, I believe that this life is, in fact, my real one. We will also assume, for the purpose of the experiment that entering the experience machine means that everyone I know enters one as well, such that everyone is living in his own personal simulated world. By entering the machine, we are not leaving behind loved ones who might mourn the lack of our presence in their world. With this in mind, we can begin to explore the implications of such an experiment and how it shows that Nozick's reasons for not entering the experience machine are less than convincing.

Alright, so now that we have it in mind that we are already in the experience machine, a few glaring questions really need to be answered. But before getting to them, just consider what this actually means thus far for the sake of the experiment. You are in the experience machine right now. Take a minute to look around you. Does anything seem to be fake, lacking, or incomplete? Do any of your memories of this life feel less than absolutely genuine? Do your relationships with friends and family feel like merely an imitation of what real relationships

would be like? For the sake of my argument, the answer should be "no". It appears that what initially makes Nozick's argument so appealing is that for some reason we assume that whatever virtual reality we program for ourselves will be in some way inferior to "real" life. In some way or another it couldn't possibly be as real as what we are currently experiencing. This actually makes perfect sense as a natural response to the prompt because these days any sort of "simulation" is still inferior to real life. It is for this reason that assuming you are in the experience machine right now is so powerful and disorienting. If it were true, it would be very hard to claim that simulated reality is possibly inferior to actual reality since throughout our lives until now we have operated under the assumption that this is as "real" as it gets. If the only reason you wouldn't step into the experience machine is that you would rather live a "real" life, but you have just discovered that the entirety of your memories have arisen within a simulated reality, it would be difficult for you to even make the claim that "real" life is somehow better or superior in any way, as you would have no experience of this kind of life to support such a claim.

Ok, now it's time to answer one of the glaring questions. If this reality is the experience machine, why do I experience so many awful things? I don't mean to sound cynical or pessimistic but if I really could program any kind of life imaginable for myself then why would I make one where I am a poor philosophy undergrad trying desperately to get through my senior year at CU? There are a few ways in which to answer this. One possibility is that the original reality I came from is tremendously worse than this one, and I only wanted to program a life for myself that corrects for problems that occurred in the former reality. Another possibility is that my former reality is so vastly different from this one, or that the whole of my collective experiences were so limited, that this reality is the best I could come up with for myself. Due to my new collective experiences in this reality, I can conceive of something even greater. This

would actually be quite unsurprising, since human beings are notoriously poor at figuring out what makes them truly happy. To give another great example from "The Matrix", Agent Smith said to Morpheus about the first version of the matrix, "Some believe that we lacked the programming language to describe your perfect world, but I believe that, as a species, human beings define their reality through misery and suffering. The perfect world was a dream that your primitive cerebrum kept trying to wake up from." Undoubtedly, we would all find something to complain about in a reality programmed by ourselves, and I will discuss the solutions to this problem in a moment. What this prepares us for, however, is another interesting possibility that I was forced into the experience machine reality by someone with malicious intent. I did not, in fact, design my own ultimate paradise, but instead someone else programmed this reality and stuck me in it. As we shall see, this becomes a most important and chilling prospect.

Now that we have assumed we are already in the experience machine, came to realize that this is really as "real" as it gets, and pondered the possibilities for why this reality is the way it is, it's time to answer another big question. If someone walked up to you right now with a *second* experience machine and gave you another chance to program a better life for yourself, would you do it? After heavy consideration, the conclusion that I have come to is that you must say "yes." Clearly, the programmed realities are more real and vivid than anything you could have imagined, given that until a few moments ago, you believed this programmed reality to be real life. Nothing about your experiences, relationships, or accomplishments in this simulation show any signs of being fake, incomplete, or less than genuine. So, the only difference between this reality and the next one you program for yourself is that in the next one, you intend to be happier. If this really is the case, then I can see no reason why you should choose not to adopt a new life. In fact, this is the perfect solution to what you would do if you discovered that after

entering your perfect paradise, something is still wrong or lacking. Perhaps you could not foresee everything that would make you happy on your first attempt at designing your ideal reality, so now you are given a second chance making an even better one. Assuming that we allow for this possibility in each reality that we create, this process may be repeated as many times as is necessary.

Now let us look at it from a different perspective. Before, I said that an interesting possibility for why this reality is not a paradise is because someone with malicious intent designed it and placed me in it. This being said, does it change the way I initially approach the question of whether or not I should enter the second experience machine? I would think it does. On top of the fact that I know that the virtual reality is for all intents and purposes completely real, I would also believe that this reality acts as a sort of prison restricting my freedom. After all, I didn't choose this reality. Some malicious entity put me here and now I am stuck for the rest of my life unless I choose to enter the second experience machine and have a new life of my own choosing. It appears that in this scenario, I have every reason to take the opportunity to use the second experience machine, because once again I am assured that nothing of the "real" element will be lacking; I will be happier there than I am here, and I will achieve a greater freedom there than I do here because I was placed in this undesirable reality against my will. We can think of this reality as a mental prison where the virtual programmer is keeping me captive against my will and the only way to escape to freedom is to enter the next experience machine.

Finally, let us return to real life, outside the experience machine. With everything we have just discussed in mind, would it be rational to program a better life for yourself and step into the experience machine for the **first** time? I say, "yes." If it was rational to step into the experience machine when you assumed "real life" to be a simulation, and the only thing that has

changed between then and now is your knowledge of the "real," why should it make a difference? If I switch off between telling you one minute that your life is real and the next telling you that it is a simulation, has anything really changed? I think not. Either way, your memories are the same, your surroundings are the same and just as real and vivid, and your relationships with friends, family, and significant others feel just as authentic. Once again, the only difference is that you will be happier in the experience machine. Even in my case as an over-zealous undergrad, I might have programmed my life to result in success and riches which will feel far more satisfying if I throw in a few early years fraught with stress and troubles. (I can dream right?) Not convinced? Remember, unlike Nozick, you need not take my word for it and believe everything you're told about the experience machine up front. First-timers are free to take a trial run of the experience machine for as little or long as you desire to alleviate the fears of making a longer commitment.

But let us take it one step further. In the scenario where someone with malicious intent forced me into the experience machine, I concluded that I have every reason to enter the second experience machine described above because I also have a feeling of imprisonment, that the reality I was forced into is a sort of mental prison keeping me captive. Let us apply this thought process to real life. In real life, just as in the scenario described above, I was placed into a reality which I did not choose, which is obviously not designed specifically for my well-being. There may not be some evil computer programmer holding me captive in the real world, but the mere notion of one sounds strikingly similar to a monotheistic God-like figure. Now we start getting into tricky territory because if being forced into an experience machine reality can be interpreted as a sort of mental prison, perhaps real life can be too. Following this analogy, it would follow that in real life, God placed me in this reality against my will, which was not designed for my

well-being, and I am a prisoner here as a result. Now if we remove our preconceived notions about God and view him as Gnostics, then, like the evil programmer scenario, I should plug into my own experience machine to escape my prison! The inclusion of the God analogy in this thought experiment is sure to be quite controversial, but if you take the utilitarian position, and entering the experience machine truly does result in greater overall happiness, then doing so is the proper course of action. Regardless of how far you choose to take the analogy, I will remain steadfast in my conviction that Nozick's argument for not entering the experience machine fails to provide compelling reasons why.

Conclusion

Overall, I have argued that the reasons Nozick provides for not entering the experience machine are unconvincing. Using my own thought experiment in which I assume I am already in the experience machine, I have shown how any fear that life within the experience machine would be lacking or unreal in any way, is irrational. The belief that I wouldn't be able to connect with our surroundings, have real accomplishments, and have genuine relationships is misguided as well. I have shown that given the option of using the experience machine in the way described by my experiment, it would be irrational to refuse it under the assumption that we are already in an experience machine. Analogously, it is rational to enter it for the first time as well. I have gone through a few possibilities for why a simulated reality might be less than a perfect paradise to give the experiment more credibility and in particular, I discussed the solutions for how we could correct for such error. For future work, it would be interesting to conduct a survey experiment asking people whether or not they would enter the machine, and analyze the results across age groups. I hypothesize that people from a younger age group will look more favorably

at the prospect of experience machine life because of being more closely acquainted with advanced simulation technology such as modern video games.

In truth, we may never achieve the technology to completely implement the experience machine as I have described here. If this is the case, then perhaps the lesson to learn from all this, is that we need to do a better job of making this reality a better one to live in. Millions of people are already finding it more appealing to spend their time in a simulated environment than in the real world. But in a world without a true experience machine, simulation can only be a temporary solution to our problems, with reality always waiting on the other side. The main argument of this paper attempted only to show the failures of Nozick's argument to be convincing, however, the extent to which such a thought experiment can alter our beliefs about morality and reality is tremendous, and should not be taken lightly.

Works Cited

Angner, Erik. "Subjective Well-Being." *Journal of Socio-Economics* 33.3 (2010): 361-68

Barber, Alex. "Hedonism and the Experience Machine" Philosophical Papers, 40(2), (2011): 257-278.

Belshaw, Christopher. "What's Wrong With The Experience Machine?" *European Journal of Philosophy* (2012): 1-18

Crisp, Roger. "Hedonism Reconsidered." *Philosophy and Phenomenological Research*LXXIII.3 (2006): 619-45.

De Brigard, Felipe "If You Like it, does it matter if it's real?", Philosophical Psychology, 23: 1, (2011) 43-57

Feldman, Fred. "What We Learn From The Experience Machine." *The Cambridge Companion* (2010): 1-31.

Hsu, Shang Hwa, Ming-Hui Wen, and Muh-Cherng Wu. "Exploring User Experiences as Predictors of MMORPG Addiction." *Elsevier: Computers & Education* 53 (2009): 990-99. Web

Inception. Dir. Christopher Nolan. Perf. Leonardo DiCaprio. Warner Bros., 2010. Film.

Jollimore, Troy. "Meaningless Happiness and Meaningful Suffering." *The Southern Journal of Philosophy* 42.3 (2010): 333-47

Kauppinen, Antti. "The Rise and Fall of Experimental Philosophy." Philosophical Explorations 10 (2007): 95-118.

Kawall, Jason. "The Experience Machine and Mental State Theories of Well-Being." *The Journal of Value Inquiry* 33 (1999): 381-87

Knobe, Joshua & Nichols, Shaun. "An experimental philosophy manifesto." In *Experimental Philosophy*, *ed.* J. Knobe and S. Nichols (Oxford University Press). (2008): 3-14.

Kolber, Adam. "Mental Statism and the Experience Machine." *Bard Journal of Social Sciences* 3.3-4 (1994-1995): 10-17.

Mill, John Stuart, and George Sher. *Utilitarianism*. Indianapolis: Hackett Pub., 2001. Print.

Nozick, Robert. *Anarchy, State, and Utopia*. New York: Basic, 1974.

Second Skin. Dir. Juan Carlos Escoriaza. Perf. Dan Bustard. South By Southwest, 2008. Film.

Silverstein, Matthew. "In Defense of Happiness: A Response to the Experience Machine. Social Theory and Practice, 26 (2000) 279-300.

Sinnott-Armstrong, Walter. "Framing Moral Intuitions." *Moral Psychology* 2 (2008): 1-27

Smith, Basil. "Can We Test The Experience Machine?" *Ethical Perspectives* 18.1 (2011): 29-51

The Matrix. Dir. Andy Wachowski and Larry Wachowski. Perf. Keanu Reeves. 1999. Film.

Weijers, Dan."Reality Doesn't Really Matter", Inception and Philosophy, D. Kyle Johnson (ed.), Wiley-Blackwell, (2011): 92-107.

Weijers, Dan "We Can Test the Experience Machine", Ethical Perspectives, 19(2), (2012): 261-268.