

A Review and Assessment of the Experience Machine Objection to Hedonism

Dan Weijers

Dan.Weijers@vuw.ac.nz

Victoria University of Wellington

Abstract

Prudential hedonism has been beset by many objections, the strength and number of which have led most modern philosophers to believe that it is implausible. One objection in particular, however, is nearly always cited when a philosopher wants to argue that prudential hedonism is implausible—the experience machine objection to hedonism. This paper examines this objection in detail. First, the deductive and abductive versions of the experience machine objection to hedonism are explained. Following this, the contemporary responses to each version of the argument are assessed and the deductive version is argued to be relatively ineffective compared to the abductive version. Then, a taxonomy of the contemporary critical responses to the abductive version is created. Consideration of these responses shows that the abductive version of the objection is fairly powerful, but also that one type of response seems promising against it. This response argues that experience machine thought experiments seem to elicit judgments that are either too biased to be used as evidence for the objection or not obviously in favour of reality. It is argued that only this type of refutation seems likely to convince proponents of the abductive version that the objection is much weaker than they believe it to be. Finally, it is suggested that more evidence is required before anything definitive can be said on the matter.

Key words: experience machine, Nozick, hedonism, well-being, prudential hedonism, experience machine objection to hedonism

1. Introduction

There are many different hedonistic theories, all of which claim that pleasure and pain are the only ultimately important explanatory considerations. This paper concerns prudential hedonism, and internalist accounts of prudential hedonism in particular. Prudential hedonism is the theory that all and only pleasure intrinsically makes people's lives go better for them and all and only pain intrinsically makes their lives go worse for them (Weijers 2011a). Prudential hedonism subsumes several different accounts of what makes our lives go well for us and one relevant way to categorise them is by whether they define pleasure and pain as internal or partly external phenomena. Internalist accounts of prudential hedonism hold that only the internal aspects of pleasure and pain intrinsically affect our well-being. In contrast, externalist accounts of prudential hedonism hold that more than just the internal aspects of pleasure and pain contribute to our well-being. An internalist prudential hedonist would not dispute that external events, such as being promoted, can impact our well-being. However, they would argue that being promoted only affects our well-being instrumentally (and only to the extent that it affects the internal aspects of our pleasure and pain). For example, according to internalist prudential hedonism, being promoted is usually a good thing, not because being promoted is good in and of itself, but because being promoted usually leads to people experiencing more pleasure.

As remarked several times in the literature, so many strong objections have been levelled at prudential hedonism that most modern philosophers believe it to be implausible.¹ Most prominent amongst these objections is the experience machine objection to hedonism, which is nearly always cited when a philosopher wants to argue that prudential hedonism is implausible (Barber

¹ See for example: Bradley (2009, p. 9), Crisp (2006a, pp. 619–620), Feldman (2004, p. 7), Silverstein (2000, p. 279), and Weijers (2011a; forthcoming-a).

2011, p. 257). Indeed, virtually everyone who has written about prudential hedonism since the late 1970s cites the experience machine thought experiment as a (and often the) decisive objection against it (Tiberius 2006, p. 496).²

This paper investigates whether the experience machine objection to hedonism is effective. The approach taken here is to assume that the experience machine objection to hedonism is effective (as most philosophers do) and then to see if the objection can withstand the numerous critical responses to it that have been published in the last 15 years. It should be noted that the majority of the discussion in this paper could be cast slightly differently so that the target theory is mental state theories of well-being, instead of prudential hedonism. If this change were made, the reasoning and results would remain essentially the same.

The investigation begins, in Section 2, by explaining the experience machine objection to hedonism and discussing why it has been so influential. The explanation of the objection includes the discussion of two formal presentations of it; a deductive and an abductive version. In section 3, the contemporary critical responses to the deductive version are considered and it

² Examples of authors who have stated or implied that the experience machine thought experiment is a knock-down refutation of internalist accounts of prudential hedonism (or all internalist mental state accounts of well-being) are easily found. For example: Arneson (1999, pp. 121–123), Attfield (1987, p. 33), Baggini and Fosl (2007, pp. 74–76), Becker (1992, p. 25), Brink (1989, pp. 223–224), Brülde (2007, pp. 26–29, 33), Bok (2010, pp. 24–28), Darwall (1997, pp. 162, 178), Feldman (2002, p. 615), Finnis (1980, p. 33; 1983, pp. 37–42), Griffin (1986, pp. 9–10), Hausman (2010, p. 329), Haybron (2008, p. 21), Hooker (2000, p. 39), Hurka (2011, pp. 68–70), Jollimore (2004, pp. 333–334), Kagan (1998, pp. 34–36; 2009, p. 253), (Kauppinen 2011, n.p.), Kazez (2007, pp. 51–54), Keller (2009, p. 657), Kraut (2007, pp. 124–126), Kymlicka (1990, pp. 13–14), (Nozick 1989, pp. 99–117), Railton (1984, pp. 148–149), Rivera-López (2007, p. 75), Sobel (2002, p. 244), Sumner (1996, p. 98), Thomson (1987, p. 41), Tiberius (2004; p. 311, n. 4), Tiberius & Hall (2010, pp. 214–215), van Wyk (1990, p. 109).

is concluded that the deductive version of the objection is relatively ineffective and should not be used. In section 4, a taxonomy of the contemporary attempts to refute the abductive version is created and the responses are assessed. Consideration of these responses shows that the abductive version of the objection is fairly powerful, but also that one type of critical response seems promising against it. This response argues that experience machine thought experiments seem to elicit judgments that are either too biased to be used as evidence for the objection or not obviously in favour of reality. It is argued that only this type of refutation seems likely to convince proponents of the abductive version that the objection is much weaker than they believe it to be. In section 5, it is suggested that more theoretical and empirical evidence is required before anything definitive can be said about how effective the experience machine objection to hedonism really is. It is also concluded that the evidence, as it stands, gives strong reason to believe that the experience machine objection to hedonism is not nearly as strong as is often assumed.

2. The Experience Machine Objection to Hedonism

In 1957 popular author Ray Bradbury published a short story, *The Happiness Machine*, in his collection *Dandelion Wine*. Bradbury described the happiness machine as a fantastic contraption that can simulate all kinds of wonderful experiences, such as sensing the sights, sounds, and tastes of Paris. Despite the apparent appeal of such a machine, Bradbury's story highlighted the dangers of it, including that the experiences the machine creates only give the illusion of true happiness. Perhaps it was this intriguing story that led to both Jack Smart and Robert Nozick discussing the philosophical implications of such a machine. Indeed, Jack Smart may have been the first philosopher to point out that our disinclination to use such a machine creates problems for prudential hedonism (1973, pp. 19–22). Without a doubt, however, it was

Robert Nozick's (1974, pp. 42–45) vivid description of an “experience machine” that popularised the idea amongst philosophers:

Suppose that there were an experience machine that would give you any experience you desired. Superduper neuropsychologists could stimulate your brain so that you would think and feel you were writing a great novel, or making a friend, or reading an interesting book. All the time, you would be floating in a tank, with electrodes attached to your brain. Should you plug into this machine for life, preprogramming your life's experiences? If you are worried about missing out on desirable experiences, we can suppose that business enterprises have researched thoroughly the lives of many others. You can pick and choose from their large library or smorgasbord of such experiences, selecting your life's experiences for, say, the next two years. After two years have passed, you will have ten minutes or ten hours out of the tank, to select the experiences of your *next* two years. Of course, while in the tank you won't know that you're there; you'll think it's all actually happening. Others can also plug in to have the experiences they want, so there's no need to stay unplugged to serve them. (Ignore problems such as who will service the machines if everyone plugs in.) Would you plug in? *What else can matter to us, other than how our lives feel from the inside?* Nor should you refrain because of the few moments of distress between the moment you've decided and the moment you're plugged. What's a few moments of distress compared to a lifetime of bliss (if that's what you choose), and why feel any distress at all if your decision *is* the best one? (Nozick 1974, pp. 42–3, his italics)

Although Nozick did not originally write the experience machine thought experiment in order to argue for the falsity of prudential hedonism, it was later adopted by others for that purpose, including himself (Nozick 1989, pp. 99-117). Nozick's experience machine thought experiment is equally effective against all theories that claim the internal aspects of our experiences are the only things that intrinsically affect our well-being, but prudential hedonism has attracted the most attention, presumably because it is the best known example of this type of theory (Weijers 2011b).

The vast majority of people who read Nozick's experience machine scenario think that they would choose to remain in reality (Weijers manuscript). Given the extra net pleasure that would be experienced during a life in Nozick's experience machine, when reasonable people choose reality it seems that they must value something other than the internal aspects of pleasure. Furthermore, given that there is no substantial connection with reality in an experience machine life, the strength of the connection with reality is usually assumed to be the only relevant difference between the two options in Nozick's scenario. This assumption has led many people to infer that a strong connection with reality must intrinsically contribute positive to our well-being.

The overwhelming influence of the experience machine objection to hedonism is doubtless caused by the dramatic impact the idea of an experience machine tends to have on our imaginations and memories and its great effectiveness despite the considerable concession it makes to prudential hedonism. On the latter point, consider that an experience machine life is described as a "lifetime of bliss" (Nozick 1974, p. 43), in which you can receive any and all of the best experiences possible. This is indeed a major concession because anti-hedonists would have reason to refute hedonism even if most

people just preferred reality over an experientially identical machine life,³ which is much more likely than preferring reality to a 'blissful' machine life. Consideration of the experience machine objection to hedonism, then, seems to provide a powerful two-pronged attack on the plausibility of prudential hedonism; it seems to show that pleasure is not the only thing of intrinsic value and that a deep connection with reality has intrinsic value. The few hedonists who claim that they wouldn't connect cannot easily avoid this assault on the fundamental tenant of hedonism because their minority view seems insincere or unreasonable to the vast majority of philosophers.

Proponents of the experience machine objection to hedonism very rarely construct it as a formal argument. Indeed, it is often discussed in a page or less before it is acknowledged as a complete refutation of prudential hedonism. Will Kymlicka gives it this treatment:

Now if pleasure were our greatest good, then we would all volunteer to be hooked for life to this machine... But surely very few people would volunteer. Far from being the best life we can lead, it hardly counts as leading a life at all... The hedonistic account of utility is wrong, for the things worth doing and having in life are not all reducible to one mental state like happiness. (Kymlicka 1990, p. 13)

The loose language many philosophers have used to explain the experience machine objection to hedonism has led to two distinct interpretations of it, one deductive and the other abductive in structure.⁴

³ Moral considerations aside.

⁴ As far as I am aware, only Alex Barber (2011), Torbjörn Tännsjö (2007), and Matthew Silverstein (2000) have drawn any attention to the distinction between deductive and abductive versions of the experience machine objection to hedonism (or some similar distinction), which is surprising given the distinction's importance for evaluating the widely-

The deductive version of the experience machine objection to hedonism is formalised below.

DP1. In terms of the internal aspects of our experiences, an experience machine life would be much better than a life in reality. (Stipulated in thought experiment)

DP2. When instructed to ignore their responsibilities to others, the vast majority of reasonable people report preferring reality over a life in an experience machine. (Empirical claim)

DP3. If when instructed to ignore their responsibilities to others, the vast majority of reasonable people report preferring reality over a life in an experience machine, then reality matters intrinsically to the vast majority of reasonable people

DC1. Therefore, reality matters intrinsically to the vast majority of reasonable people. (*Modus ponens* DP2, DP3)

DP4. If something matters intrinsically to the vast majority of reasonable people, then that thing has intrinsic prudential value

DC2. Therefore, reality has intrinsic prudential value. (*Modus ponens* DC1, DP4)

DP5. If internalist prudential hedonism is true, then the internal aspects of pleasure and pain are the only things of intrinsic prudential value (or disvalue) in a life. (Stipulated definition)

DC3. Therefore, internalist prudential hedonism is false. (*Modus tollens*, DC2, DP5)

The abductive version of the experience machine objection to hedonism is formalised below.

cited objection. See also Fred Feldman (2011), Antti Kauppinen (2011), and Dan Weijers (2011b) for alternate constructions.

- AP1. In terms of the internal aspects of our experiences, an experience machine life would be much better than a life in reality. (Stipulated in thought experiment)
- AP2. When instructed to ignore their responsibilities to others, the vast majority of reasonable people report preferring reality over a life in an experience machine. (Empirical claim)
- AP3. The best explanation for AP2 is that reality matters intrinsically to the vast majority of reasonable people
- AP4. Inference to the best explanation: If a hypothesis is the best explanation of an observation, then it is rational to believe that hypothesis is true. (Standard methodological premise)
- AC1. Therefore, it is rational to believe that reality matters intrinsically to the vast majority of reasonable people.
(*Modus ponens* AP3, AP4)
- AP5. The best explanation for reality mattering intrinsically to the vast majority of reasonable people is that reality has intrinsic prudential value
- AP6. Inference to the best explanation. (Standard methodological premise)
- AC2. Therefore, it is rational to believe that reality has intrinsic prudential value. (*Modus ponens* AP5, AP6)
- AP7. If internalist prudential hedonism is true, then the internal aspects of pleasure and pain are the only things of intrinsic prudential value (or disvalue) in a life. (Stipulated definition)
- AC3. Therefore, it is rational to believe that internalist prudential hedonism is false. (*Modus tollens*, AC2, AP7)

A few points about these arguments should be noted. First, many published versions of the experience machine objection simply identify hedonism as

their target, rather than internalist accounts of prudential hedonism (e.g. Kymlicka 1990, p. 13). This is probably due to either ignorance of the different versions of hedonism or simplification, perhaps for educational purposes. These simpler presentations of its target might have influenced the philosophers who argue that some variants of hedonism can avoid the experience machine objection precisely because they are not *internalist* accounts of prudential hedonism.⁵ Since it is not clear that the objection was ever really designed to refute anything other than internalist accounts of prudential hedonism, only the responses that attempt to defend internalist versions of hedonism will be discussed in this paper and any further use of the terms ‘pleasure’ or ‘prudential hedonism’ (or their linguistic derivatives) refer to internalist accounts of pleasure and prudential hedonism.

Second, both arguments state that they are based on ‘reality’ mattering to people, but should be understood as being based on ‘something other than the internal aspects of pleasure and pain’ mattering to people. This later phrase is more likely to be true, better supported by the experience machine thought experiment, and about as good at giving reason to doubt hedonism. However, it is also particularly unwieldy; so to increase readability, ‘reality’ will be used instead.

Third, if something ‘matters intrinsically’, then it seems to have intrinsic value. If reality matters intrinsically to someone, then that person believes that reality has value *qua* reality—that reality has value over and above the value of any other things that might come from it.

Finally, the deductive version of the experience machine objection to hedonism has a much stronger conclusion than the more nuanced abductive version. As would be expected, with this strong conclusion comes a strong evidential burden on the premises of the deductive version. Although, the

⁵ (e.g. Donner 1991; Feldman 2004; Heathwood 2007; Rivera-López 2007; Sumner 1996).

conclusion of the abductive version ('it is rational to believe that internalist prudential hedonism is false') is still strong enough to discredit prudential hedonism, many philosophers have nevertheless discussed the experience machine objection to hedonism as if it were a deductive refutation of hedonism, including Nozick himself (1989, pp. 99–117).

Rivera-López has recently claimed that the experience machine objection to hedonism has "never [been] seriously disputed" (2007, p. 75), even though many academics have published responses which they consider to be refutations of it. The fact that different responses continue to be published to this day probably supports Rivera-López's claim that the objection has never been *seriously* disputed. Furthermore, as we shall see, several of the responses are only effective against the deductive version of the objection and so might not qualify as serious disputations. Responses to the deductive argument are considered first, in Section 3, followed by the responses to the abductive version, in Section 4. This compilation of responses is intended to be exhaustive of all the published paper-length attempts to refute the experience machine objection to hedonism during the last 15 years and indicative of all of the types of critical responses to the experience machine objection to hedonism to date.

3. Responses to the Deductive Version

The responses discussed in this section are limited to those directed against the deductive argument. All of these related responses involve DP4 (repeated below for convenience).

DP4. If something matters intrinsically to the vast majority of reasonable people, then that thing has intrinsic prudential value

Harriet Baber (2008) argues that the experience machine objection to hedonism is unfair because it presupposes some form of preferentism, or

desire-satisfaction account of well-being.⁶ The common thread running through all preferentist, or desire-satisfaction accounts of well-being, is the principle that having (certain kinds of) our preferences satisfied is the only thing that intrinsically improves our well-being. Or, as Baber puts it, “according to [preferentism] what makes a state of affairs good for a person is her desiring it...” (2008, p. 134).

Baber describes the key premise of the experience machine objection to hedonism as follows: “If a reasonable and informed subject, *i*, would choose *S* over *S'*, then *S* would contribute more to *i*'s wellbeing than *S'*.” (Baber 2008, p. 133, her italics). In Baber's interpretation of the key premise, we can see that reasonable and informed peoples' choices (i.e. their preferences) dictate what contributes to well-being. This move should be considered problematic, but not for the reason provided by Baber. The experience machine objection to hedonism does not assume that preferentism is true because, as Nozick points out, our preferences about the experience machine do not directly dictate the values involved:

Notice that I am not saying simply that since we desire connection to actuality the experience machine is defective because it does not give us what we desire... for that would make “getting whatever you desire” the primary standard. Rather, I am saying that the connection to actuality is important whether or not we desire it—that is *why* we desire it—and the experience machine is inadequate because it doesn't give us *that*.” (Nozick 1989, pp. 106–107, his italics)

⁶ Torbjörn Tännsjö (2007, pp. 94–95) offers a similar response to the deductive version of the experience machine objection to hedonism, but he is much more careful to acknowledge that the abductive version of the argument is a better interpretation of it and that the abductive version does not suffer from this kind of response.

Baber's (2008, p. 133) version of the key premise skips an important step in the argument. Baber starts with the equivalent of DP2 ('the vast majority of reasonable people report preferring reality over a life in an experience machine') and skips over DC1 ('reality matters intrinsically to the vast majority of reasonable people') straight to DC2 ('reality has intrinsic prudential value'). The first step is important because it shows how the preference that is revealed by contemplating the experience machine thought experiment is evidence for forming a judgement about what matters to us. That judgment (reality matters intrinsically) then leads to the judgement that it has intrinsic prudential value.

So, while Baber interprets 'choosing' as 'having a preference for' in her version of the key premise, 'choosing' seems better interpreted as 'making a judgment about value'. Of course, that judgment might still be mistaken. But if the vast majority of reasonable and informed people make the same judgment, then we have reason to believe that it is not mistaken. Consider the possibility that Nozick's experience machine scenario is found to elicit extremely biased judgments and that a new bias-free version of the experience machine thought experiment elicits widespread agreement from reasonable and informed people that a life in an experience machine is better than one in reality. This might give us reason to believe that some version of hedonism is true. The method used to reach this hedonism-endorsing conclusion is not preferentism; it is the dominant method in moral philosophy. As Roger Crisp puts it: "Intuitions appropriately reflected upon are unavoidable in ethical theory." (Crisp 2006a, p. 636). After all, could there be any way to endorse hedonism over other theories of well-being without the judgments of reasonable people about whether some X has more intrinsic value than some Y?

Fred Feldman (2011) also criticises the move from 'matters' to 'has value', but on different grounds. Feldman presents a number of different versions of

the experience machine objection to hedonism, but finds none of them particularly convincing. Essentially, Feldman argues that what matters intrinsically to reasonable people is of little consequence to questions of intrinsic value because reasonable people lack certain traits that are required to really know what has intrinsic value (and therefore what should matter intrinsically to us). Most importantly, Feldman insists that the move from 'matters intrinsically to reasonable people' to 'has intrinsic value' only works if the reasonable people are also "axiologically insightful", which is to say that they already understand what has intrinsic value. Therefore, he concludes that his close analogue of DP4 is false.

Feldman (2011) rightly points out that changing the experience machine objection to hedonism so that only what matters intrinsically to people who are both reasonable and axiologically insightful can actually tell us what has intrinsic value would produce a solid (deduction-supporting) bridge between 'matters intrinsically' and 'has intrinsic value'. Feldman is also right that changing the objection in this way would create a new problem for it: now Hedonists could deny his version of DP2 (that reasonable and axiologically insightful people would choose reality over a life in an experience machine) and there would be no suitably axiologically insightful person who could disconfirm this denial!

Feldman's (2011) denial of DP4 reveals a flaw in the deductive version of the experience machine objection to hedonism. His refutation of the experience machine objection to hedonism is only effective against the deductive version, however. We don't have to assume that reasonable people know everything about axiology before we can infer that their agreement gives us a defeasible reason to believe that something has value. Again, how could we know anything about what has value if only the judgments of axiologically insightful people (who already understand what has intrinsic value) were relevant to questions of value? So Feldman's (2011) critique of the

experience machine objection to hedonism gives us reason to believe that the more nuanced abductive version of the objection is much more plausible than the deductive version.

Sharon Hewitt (2009, p. 348) takes a related approach, arguing that: “even if our intuitions [about what matters intrinsically to us] are directly responsive to the existence of real relationships... this should not immediately lead us to conclude that these things are objectively intrinsically valuable.”⁷ Hewitt is arguing that the move from ‘matters intrinsically’ to ‘has intrinsic value’ cannot be deductive because there are reasons other than something having intrinsic value for us that make things matter intrinsically to us, including irrational ones. Or, as Barber (2011, p. 269, his italics) puts it: “It is not enough... [to] *think* that hedonism is mistaken, since we could be mistaken that it is mistaken.” Jason Kawall (1999, p. 385) agrees for two reasons. First, Kawall points out that: “While it is true that we value many things besides our mental lives, it could well be that the other values contribute to our personal well-being only through the effects they have on our mental lives” (1999, p. 385). Kawall also notes that “we can value more than our own well-being” which can lead to people sacrificing their well-being to further some other value, such as “the well-being of their children” or “the search for truth” (1999, p. 385). Hewitt (2009) and Kawall (1999) are correct and, when combined with Feldman’s (2011) critique, their views give us strong reason to believe that DP4 is false.

Analysis of Baber’s (2008), Feldman’s (2011), Hewitt’s (2009) and Kawall’s (1999) responses to the experience machine objection to hedonism show that the move from ‘choosing’ or ‘mattering intrinsically’ to ‘has intrinsic prudential value’ is best understood as an inference to the best explanation,

⁷ This is not Hewitt’s complete strategy to refute the experience machine objection to hedonism. This argument is combined with others that are discussed below.

not as a deduction. Indeed, no careful philosopher should believe that widespread agreement between reasonable people on a matter of value *deductively proves* any particular conclusion. As Silverstein (2000, p. 299, his italics) puts it: “Most philosophers recognize this... and take the experience machine argument to be an *indirect* refutation of hedonism.”

Instead of trying to deductively prove what has value, philosophers should focus on making inferences to the best explanation. What then is the best explanation for something mattering intrinsically to the vast majority of reasonable and informed people? ‘That it has intrinsic value’ is a plausible answer to this question. Furthermore, if competing answers have been sought by reasonable people and none seem more credible, then there is good reason to believe ‘that it has intrinsic value’ is the correct answer. Of course, that belief should be amended if a more credible competing answer surfaces at any stage.

4. Responses to the Abductive Version

Some of the responses to the experience machine objection to hedonism discussed in this section may also apply to the deductive version, but they are included here because the deductive version has already been shown to be relatively ineffective, if not implausible. This section is a taxonomy of all of the responses to the abductive version of the experience machine objection to hedonism.

4.1 Responses to AP5

At least three authors have attempted to defend hedonism against the experience machine objection by denying AP5 (repeated below for convenience). These authors all argue that consideration of the process by which our preferences are formed gives us reason to doubt that reality having

intrinsic prudential value best explains why it matters intrinsically to the vast majority of people.

AP5. The best explanation for reality mattering intrinsically to the vast majority of reasonable people is that reality has intrinsic prudential value

Matthew Silverstein argues against AP5 by appealing to the hedonistic basis of our seemingly anti-hedonistic preferences:

[O]ur experience machine intuitions reflect our desire to remain connected to the real world, to track reality. But... the desire to track reality owes its hold upon us to the role it has played in the creation of happiness. We acquire our powerful attachment to reality after finding again and again that deception almost always ends in suffering. We develop a desire to track reality because, in almost all cases, the connection to reality is conducive to happiness. Our intuitive views about what is prudentially good, the views upon which the experience machine argument relies, owe their existence to happiness. (Silverstein 2000, p. 296)

According to Silverstein, our desires are created by our experience of what has brought us happiness in the past and our intrinsic desires (what matters intrinsically to us/what we desire for its own sake) are created by *repeated* experiences of what has brought us happiness in the past. This account of desire formation is based on the work of Richard Brandt (1979) and an idea of Peter Railton's (1989). Silverstein distances himself from Brandt's view slightly by stressing that his argument does not rely on Brandt's extreme claim that "[happiness-related] conditioning is the *only* fundamental process involved in the acquisition of desires" (Silverstein 2000, p. 293, his italics;

Brandt 1979, p. 100). Instead, Silverstein claims that all desires are created by happiness and that happiness is the main influence on our desires. This allows for some desires to be affected by non-happiness-related factors, although Silverstein doesn't discuss any such factors.

If Silverstein's (2000) account of desire creation is true, then why doesn't everyone simply save time by desiring only happiness? Silverstein quotes both Sidgwick and Mill on what has become known as the paradox of hedonism (and also the paradox of happiness) to answer this question.⁸ The paradox of hedonism is that pursuing happiness directly is likely to bring about much less happiness than pursuing other goods. While not strictly a paradox, there is certainly a lot of *prima facie* tension between the propositions 'happiness is the greatest good' and 'happiness should not be pursued'. The paradox of hedonism is well-supported in the philosophical literature, but it seems to be most true of folk hedonism, which usually entails the greedy pursuit of immediate pleasures at the risk of harm and unhappiness for themselves and others sometime after. Drug-taking (when the drug causes harmful side-effects) is a good example of how taking the direct route to happiness does not pay off in the long run. But even if we wanted to pursue happiness directly, it is not clear how we would do it; there is no genuine 'happiness shop'; so our attempts to attain happiness are always indirect to some extent (including drug-taking).

Perhaps Silverstein's (2000) example of why reality and truth matter intrinsically to us is the best example of the lessons to be learnt from the

⁸ Sidgwick: "[The] fundamental paradox of Hedonism, that the impulse towards pleasure, if too predominant, defeats its own aim." (1907, p. 48). Mill: "I now thought that this end [happiness] was only to be attained by not making it the direct end. Those only are happy (I thought) who have their minds fixed on some object other than their own happiness.... Aiming thus at something else, they find happiness by the way." (1969, pp. 85–86).

paradox of hedonism. Deception has brought pain so consistently, Silverstein argues, that reality and truth have come to matter intrinsically to us; we will experience much more happiness if we pursue truth and reality than if we try to pursue happiness as directly as possible. We might wish to live in a bubble of false beliefs because lots of facts about the world might make us unhappy (e.g. your partner is cheating on you). In the vast majority of cases, however, we will be happier in the long run if we face up to the truth and try to make the best of the situation (e.g. confront your partner and sort out your problems or break up with them and get a new more loyal partner). So, the lessons we should learn from contemplation of the paradox of hedonism are that we shouldn't always prefer what will make us happy in the moment and that certain other goods mattering intrinsically to us is the best way to achieve this.

Silverstein is using his account of intrinsic desire creation to deny AP5 by arguing that the best explanation for reality mattering intrinsically to the vast majority of people is not that reality has intrinsic prudential value. Rather, it is that preferring reality (avoiding deception) nearly always leads to happiness in the long run. Silverstein (2000, p. 297) also argues that all of our intrinsic desires are formed this way and concludes that: "The most plausible explanation is a hedonistic one: the reason all of our desires point towards happiness is that happiness is the only thing that is intrinsically prudentially valuable." So Silverstein has attempted to turn the experience machine objection to hedonism on its head by arguing that the widespread preference for reality over a life in an experience machine (along with all other preferences) actually provides evidence in support of prudential hedonism!

Unfortunately for Silverstein, it is not clear that many philosophers would agree that his 'all desires are created by happiness' view better explains reality mattering intrinsically to the vast majority of people than reality having intrinsic prudential value does. The vast majority of philosophers

would probably accept that the circumstances surrounding our previous experiences of happiness will have affected what currently matters to us, but they would probably not accept that those experiences of happiness are always the *main* cause of what currently matters to us. They could argue that reality matters to us a small amount because of our previous experience of the pain associated with deception, but also that it matters to us a lot because it makes experiences more meaningful (regardless of how much happiness they bring).

This response would leave Silverstein in a difficult position. If strong motivational hedonism is true (i.e. if all of our desires are completely governed by the pursuit of pleasure and avoidance of pain), then Silverstein could ignore this response. However, Silverstein (2000, p. 293, n. 42) does not claim that strong motivational hedonism is true, so he cannot pursue this line of argument. And even if he did, there are good philosophical and introspective reasons to doubt strong motivational hedonism, that appear to require advances in cognitive science to overturn (Silverstein 2000, p. 294, n. 47). This leaves Silverstein the problem of trying to motivate why the preference for reality in particular is likely to be largely or wholly governed by our underlying preference for happiness. Considering the prevalent pluralist beliefs most current philosophers have about prudential value (particularly about real achievements and meaning in life), Silverstein has not done enough to convince philosophers that reality mainly matters intrinsically to us because of our past experiences of reality-related happiness.

Roger Crisp (2006a) and Sharon Hewitt (2009) also argue that AP5 is false because reality mattering intrinsically to the vast majority reasonable people is better explained by evolutionary and psychological explanations. Both Crisp and Hewitt discuss the paradox of hedonism and then use specific examples to argue that many of our judgments about which goods matter intrinsically to us could have developed because they aided our ancestors'

procreative fitness and provided us with pleasure throughout our personal development. Crisp (2006a) mainly discusses how our preference to accomplish real achievements could have developed into a powerful intrinsic desire under the guidance of selfish genes and a pleasure-seeking brain. Hewitt (2009) focuses on how our preference to establish real interpersonal relationships could have developed in a similar way. Since both accomplishing real achievements and establishing real interpersonal relationships require us to live in reality, and on the assumption that evolutionary and psychological mechanisms best explain why accomplishing real achievements and establishing real interpersonal relationships matter to us, Crisp (2006a) and Hewitt (2009) both conclude that evolutionary and psychological mechanisms best explain why reality matters intrinsically to us.

Crisp's (2006a) and Hewitt's (2009) arguments will encounter the same initial problem as Silverstein's—other philosophers are unlikely to be convinced that the evolutionary and psychological mechanisms explain why reality matters intrinsically to us *better* than reality having intrinsic prudential value does (e.g. Fletcher 2007). Again, without the support that the truth of strong motivational hedonism would offer, it seems like proponents of the experience machine objection to hedonism can reply that Crisp's (2006a) and Hewitt's (2009) explanations are plausible, but that connection with reality being valuable for its own sake is *more* plausible and, most importantly, that this is the reason it matters intrinsically to them and why they prefer reality over a life in an experience machine.

So, even when taken together, Silverstein's (2000), Crisp's (2006a), and Hewitt's (2009) denials of AP5 are unlikely to convince many non-hedonists that reality is mainly caused by anything other than it having intrinsic prudential value. This is especially the case because non-Hedonists have both introspective and philosophical evidence (from the experience machine thought experiment, amongst other sources) that connection to reality has

prudential value over and above any pleasure or pain that it might lead to. Indeed, it seems like defenders of prudential hedonism should put this particular argument on the shelf until cognitive science can provide more detailed information on how our judgments and preferences are created.

4.2 Responses to AP3

A slightly more promising group of responses to the experience machine objection to hedonism deny AP3 (repeated below for convenience). These denials all identify features of the experience machine thought experiment that might elicit responses to it that are biased or otherwise corrupted by irrelevant factors. The arguments all deny AP3 on the basis that the best explanation for why the vast majority of reasonable people report preferring reality over a life in an experience machine is that one or more features of the thought experiment (that are irrelevant to the purpose of the thought experiment or to assessing well-being) are what really matters to the vast majority of reasonable people.

AP2. When instructed to ignore their responsibilities to others, the vast majority of reasonable people report preferring reality over a life in an experience machine. (Empirical claim)

AP3. The best explanation for AP2 is that reality matters intrinsically to the vast majority of reasonable people

Most proponents of the experience machine objection to hedonism understand the experience machine thought experiment to isolate a prudential value comparison between reality and how our experiences feel to us on the inside. But the ability of exotic philosophical thought experiments, like the experience machine, to isolate what they intend to is derided by some psychologists and behavioural economists who research how judgements and preferences are formed. Cass Sunstein had this to say:

I believe that some philosophical analysis, based on exotic moral dilemmas, is inadvertently and even comically replicating the early work of Kahneman and Tversky by uncovering situations in which intuitions, normally quite sensible, turn out to misfire. The irony is that where Kahneman and Tversky meant to devise cases that would demonstrate the misfiring, some philosophers develop exotic cases with the thought that the intuitions are likely to be reliable and should form the building blocks for sound moral judgments. An understanding of the operation of heuristics offers reason to doubt the reliability of those intuitions, even when they are very firm (cf. the emphasis on moral learning from real-world situations in Churchland 1996). (Sunstein 2005, p. 541)

Bronsteen, Buccafusco, and Masur (2010, p. 1609) worry that our intuitive reactions to the experience machine thought experiment might be based on more than just an isolated prudential value comparison between reality and how our experiences feel to us on the inside: “When we ask whether someone attached to the machine has greater welfare, we must look behind whatever visceral aversion to the machine we might have and assess (i) whether that aversion relates to welfare and (ii) whether the aversion springs from rejecting the rules of the hypothetical example.” Bronsteen and colleagues’ worry is shared to varying degrees by many philosophers.

Indeed, many potential causes of negative visceral sensations, especially fear, have been identified in the experience machine thought experiment. D. W. Haslett (1990) expressed concern that the experience machine thought experiment is so unrealistic that it demands too much of the readers’ imaginations and that the judgements it elicits should be expected to be unreliable. Wayne Sumner (1996, p. 95) asks: “How do we know that the

technology is foolproof? What happens if there is a power failure?”. Christopher Belshaw (forthcoming) fears that the machine might break down “with the risk of disaster greater than anyone should bear”. Bronsteen, Buccafusco, and Masur note that readers of the experience machine thought experiment “might not be convinced the machine will actually work; or they might fear that while on the machine, they will be vulnerable to harm from those in the real world” (2010, p. 1609). Torbjörn Tännsjö agrees, stating that some people might not choose an experience machine life “because of an (unreasonable) fear that those in charge of the machine... would take advantage of them in some nasty way” (2007, p. 93). Goldsworthy (1992, p. 18) emphasises these points, claiming that a reasonable person might choose not to plug in because of fear of “catastrophic, unimaginably horrible consequences of malfunction or abuse”. Hewitt adds that anyone connecting to an experience machine “must trust those outside of the machine to look out for his interests as well as he could himself if he were living in contact with the external world” (2009, p. 338). Barber (2011, p. 267) makes an analogy with some people’s reluctance to fly: “Some refuse to fly but will drive even if they know this is far more dangerous, and boarding an airplane is nothing compared to volunteering oneself up to a different—and delusional—plane of reality.”

Adam Kolber (1994, pp. 13–14) notes that the fears of machine underperformance or failure are exacerbated by the troubling irrevocability of the experience machine and our general fear of the unfamiliar: “We are hardly comfortable enough with our own world to risk life under totally foreign circumstances.” (1994, p. 13). Belshaw (forthcoming) is also concerned that the thought of spending our whole lives in an experience machine “stacks the odds in favour of the cautious”, deterring many of us from connecting. Both Weijers (forthcoming-a) and Mendola (2006) also emphasise how worries about the machine feed into a general fear of the unknown, with Mendola

(2006, p. 450) claiming that it is the “unfamiliar gadgetry which invokes our fear of the unfamiliar”. These fears lead Crisp to dismiss the question about whether people would or should choose to connect to an experience machine because their choices are likely to be affected by “differing attitudes to risk” (2006a, p. 635). All of these worries about the machine are highly relevant to our well-being, but they should not enter into a direct comparison between the prudential value of reality and how our experiences feel to us on the inside. Therefore, if these worries significantly affect people’s preferences in the experience machine thought experiment, AP3 starts to come under pressure.

However, only a few of the authors mentioned argue that these irrelevant fears might amount to a refutation of AP3. This is because, although these fears plausibly affect people’s preferences in the experience machine thought experiment, they don’t obviously provide a better explanation for them than reality mattering intrinsically to us. Even Silverstein (2000), who attempts to refute the experience machine objection to hedonism, argues that this response to AP3 will not work. Silverstein claims that with “a bit of mental dexterity” (i.e. tweaking the scenario and having faith in the outcomes stipulated in the scenario) we can allay any doubts about our intuitive experience machine-related fears (2000, p. 284). And without these fears, Silverstein believes that “we remain unwilling to accept a lifetime on the experience machine” (2000, p. 285). For this reason, many of the philosophers who argue that these fears might provide a good explanation for the widespread preference for reality over a life in an experience machine also provide a further objection to AP3.

That further objection to AP3 is that people can have preferences for things because they promote several different kinds of value (not just prudential value). Silverstein (2000, p. 290) puts it like this: “One way to lessen the force of the experience machine intuitions is to demonstrate that they are really

about something other than well-being.” This objection argues that people might report preferring reality over a life in an experience machine because they prefer to promote aesthetic, moral, or other non-prudential values, which they could not achieve while connected to an experience machine. If the widespread preference for reality over an experience machine life is best explained by the widespread desire to promote moral values, then AP3 is false and the experience machine thought experiment does a very bad job of isolating a prudential value comparison between reality and how our experiences feel to us on the inside. If AP3 is false for this reason, then the experience machine thought experiment doesn’t tell us much at all about prudential value and, therefore, it can’t provide the basis for a reasonable objection to prudential hedonism.

Silverstein (2000, p. 291) argues that the low aesthetic value of an experience machine life might be heavily impacting people’s preferences in the experience machine thought experiment: “When we entertain Nozick’s thought experiment, we find life on the machine unattractive: it is aesthetically displeasing.” Most philosophers who argue for non-prudential values affecting our preferences in the experience machine thought experiment, however, focus on moral values. Kawall gives the example of a soldier who throws himself on a grenade as evidence that people do sacrifice their own well-being for moral reasons (1999, pp. 385–386). Mendola notes that “things other than our own well-being... matter to us. For instance, our lives have effects on other people which are quite significant, and which we care about” (2006, p. 450). Kolber suggests that some people might not report preferring a life in the experience machine because it’s self-indulgent and self-indulgence is considered immoral (1994, p. 14). Bronsteen, Buccafusco, and Masur note that readers of the experience machine thought experiment “might not want to forego the opportunity to use their lives to improve the lives of others” (2010, p. 1609). Belshaw (forthcoming) agrees that worries

about “free-riding on the rest of society” might deter some readers from connecting and also argues that readers “wanting certain things to happen in the world” (which they could better bring about if they were not in an experience machine) could have the same effect.

Nozick (1974, p. 43) anticipated this objection and attempted to protect the experience machine thought experiment against it by stating: “Others can also plug in to have the experiences they want, so there’s no need to stay unplugged to serve them. (Ignore problems such as who will service the machines if everyone plugs in.)” But Feldman provides two effective examples of why this approach might not work:

A person might still worry about fulfilling his moral obligations even if he were convinced that others would be plugging in. For example, consider someone who solemnly promised his mother that he would rescue her if her Experience Machine should happen to malfunction. Believing that others would be in Experience Machines (or would have the option of plugging in) would not relieve him of his feeling of obligation to keep out of the machine so as to be available for rescue operations, should they be necessary. And even more obviously, suppose he knows that others have been given the opportunity to plug in but have chosen to remain unplugged. His feeling of obligation to them would be unaffected. (Feldman, 2011, n. 17)

Again Silverstein (2000, p. 291) argues that these concerns can be stipulated away, but it is unlikely that they can be *completely* stipulated away. Based on how judgements about thought experiments are actually formed, both Hewitt (2009) and Weijers (forthcoming-a) have argued that simply stipulating that readers needn’t worry about something that they value greatly doesn’t always work. Therefore, there seems to be good reason to think that some

people might prefer reality over an experience machine life because a non-prudential value, probably a morally-related one, matters a lot to them. Having said this, without further evidence about how many people report preferring reality over an experience machine life because a non-prudential value really matters to them, it seems like reality mattering intrinsically to the vast majority of people might still be the best explanation for the widespread preference for reality.

When all of these potential causes of people choosing reality over the experience machine are taken together, however, they begin to pose a potential threat to AP3. The difficulty is in assessing how much effect these causes have individually and cumulatively. For this reason, the philosophers who argue that these alternate causes might provide the best explanation for the widespread preference for reality over a life in an experience machine also try to create experience machine scenarios that isolate and eliminate them to see if AP2 still holds.

4.3 Responses to AP2

Another group of responses to the experience machine objection to hedonism deny AP2 (reproduced below for convenience). The contemporary denials of AP2 are often subtle and indirect, but at least one is incredibly straightforward.

AP2. When instructed to ignore their responsibilities to others, the vast majority of reasonable people report preferring reality over a life in an experience machine. (Empirical claim)

Torbjörn Tännsjö takes a direct approach by claiming that he would not choose reality over an experience machine life (2007, p. 95). He also points out that lots of people take drugs and argues that this is the pharmacological equivalent of connecting to an experience machine (Tännsjö 1998, p. 112).

Combined with his worry about people's fear of abuse while connected to an experience machine, this evidence forms the basis for his claim that "it is far from clear that ... the claim that we would not plug in, is *true*" (Tännsjö 2007, p. 93, his italics). In a similar vein, Belshaw (forthcoming) argues that many people, especially those whose remaining life years look to be somewhere between relatively dull and excruciatingly painful, might connect to an experience machine. However, this kind of approach will do very little to convince anyone who has presented the experience machine thought experiment to students (or any other group of people) that AP2 is true because these people have had first-hand experience of the vast majority of people reporting that they would prefer reality over a life in the experience machine.⁹

A much more promising approach to denying AP2 is to construct an alternate version of the experience machine thought experiment that does *not* elicit a widespread preference for reality over a life in an experience machine, while remaining essentially the same in all relevant ways. Based on the worry that the potential causes of people's preference for reality over an experience machine life, such as irrational fear, are irrelevant to assessing prudential value, several philosophers have developed new experience machine scenarios. These new scenarios attempt to eliminate from consideration all of the factors that are irrelevant to an isolated prudential value comparison between reality and how our experiences feel to us on the inside. If other versions of the experience machine thought experiment can equally or better isolate this prudential value comparison, and people's judgments about them

⁹ The one exception here seems to be Barber (2011, p. 263, n. 7), who claims that his undergraduate philosophy students were fairly evenly split ("52% non-enterers to 48% enterers"), although he implies that his data is not as credible as data from more formal surveys.

significantly diverge, then either AP2 or AP3 might be false. Kolber explains the strategy behind this approach:

[A]ll thought experiments that consider the same issue as the [experience machine thought experiment] must yield the same results... [Otherwise] Nozick's argument has failed. [This is because] it seems unlikely that a reflective, unbiased person could give different answers to two versions of essentially the same question. (Kolber 1994, p. 13)

Kolber (1994, p. 15), De Brigard (2010, pp. 47–49), and Weijers (forthcoming-a)¹⁰ have all created new experience machine scenarios that attempt to reduce interference from irrelevant factors. What is most notable about all of these new scenarios is that they are designed to minimise the impact of unfamiliarity with, and fear of, experience machines by framing being connected to a machine as the status quo. Here is an excerpt from one of De Brigard's reversed scenarios:

"I am afraid I have some disturbing news to communicate to you" says Mr. Smith. "There has been a terrible mistake. Your brain has been plugged by error in to an experience machine created by super duper neurophysiologists. All the unpleasantness you may have felt during your life is just an experiential preface conducive toward a greater pleasure... we'd like to give you a choice: you can either remain connected to this machine (and we'll remove the memories of this conversation taking place) or you can go back to your real life. By the way, you may want to know that your real life is not at all as your

¹⁰ See also Weijers (2011c) for a scenario that addresses the issue of status quo bias using examples based on the movie *Inception*.

simulated life. In reality you are a prisoner in a maximum security prison in West Virginia.” What would you choose? (De Brigard 2010, p. 4)

Kolber’s (1994, p. 15), De Brigard’s (2010, pp. 47–49), and Weijers’s (forthcoming-a) scenarios all appear to isolate a prudential value comparison between reality and how our experiences feel to us on the inside about as well as Nozick’s original scenario. Furthermore, in all of these reversed scenarios, especially De Brigard’s Negative scenario, it seems that most people might prefer a life connected to an experience machine over a life in reality. This is a questionable empirical claim, but if it’s true, then either AP2 is false because the vast majority of people don’t actually prefer reality (at least in all cases) or AP3 is highly questionable because people’s preferences in experience machine cases might be explained better by the status quo (what is most familiar to us) than by reality mattering intrinsically. However, exactly what reasonable people would prefer in these reversed cases is not the only potential problem for Kolber, De Brigard, and Weijers; their scenarios might also have introduced more biases and other confounding factors than were present in Nozick’s original scenario.

Nevertheless, for anyone who endorses the experience machine objection to hedonism and found themselves preferring a life connected to an experience machine over a life in reality in one of these new scenarios, the onus should be on them to point out the confounding feature of the new scenario(s). This apportionment of the benefit of the doubt seems fair because proponents of the experience machine objection to hedonism have already been using a suspect thought experiment (Nozick’s experience machine scenario) as evidence in an argument, thereby making it inconsistent for them to deny the same practice for others. Note that proponents of the experience machine objection to hedonism cannot, in good conscience, claim that making

the experience machine life the status quo is the confounding factor because Nozick's scenario would also take advantage of any bias in favour of the status quo.¹¹

A safer response for proponents of the experience machine objection to hedonism is the direct approach taken by Tännsjö above; to simply deny that they would choose a life connected to an experience machine over a life in reality in the new scenarios. While this approach is safer, it is not necessarily enough to prevent the denial of AP2 because these authors have begun to empirically test their thought experiments and it is no longer clear what the vast majority of people believe about experience machine scenarios.¹² This means that attempts to deny AP2 by creating a new scenario that produces a different result from Nozick's scenario requires a sound empirical result before defenders of prudential hedonism stand a good chance of convincing stubborn proponents of the experience machine objection to hedonism that AP2 or AP3 is false.

Recent work by Weijers (manuscript) is the closest to providing an empirical refutation of AP2 with a test of the following scenario on a group of university students:

A stranger, named Boris, has just found out that he has been regularly switched between a real life and a life of machine-generated experiences (without ever being aware of the switches); 50% of his life has been spent in an Experience Machine and 50% in reality. Nearly all of Boris' most enjoyable experiences occurred while he was in an Experience Machine and nearly all of his least enjoyable experiences occurred while he was in reality. Boris now

¹¹ See Weijers (manuscript) for more discussion of this point.

¹² Having said this, questions have been raised about the validity of some of these specific tests and of testing the experience machine in general (e.g. Smith 2011). But see also Weijers (2012) response to Smith for a defence of empirical tests of experience machine scenarios.

has to decide between living the rest of his life in an Experience Machine or in reality (no more switching).

You have had a go in an Experience Machine before and know that they provide an unpredictable roller-coaster ride of remarkable experiences. When in the machine, it still felt like you made autonomous decisions and occasionally faced tough situations, such as striving for your goals and feeling grief, although you didn't really do these things. Your experiences were also vastly more enjoyable and varied in the machine. You also recall that, while you were in the Experience Machine, you had no idea that you had gotten into a machine or that your experiences were generated by a machine.

Boris' life will be the same length in an Experience Machine as it would in reality. No matter which option Boris chooses, you can be sure of two things. First, Boris' life will be very different from your current life. And second, Boris will have no memory of this choice and he will think that he is in reality.

1) Ignoring how Boris' family, friends, any other dependents, and society in general might be affected, and assuming that Experience Machines always work perfectly, what is the best thing for Boris to do *for himself* in this situation?

Tick only one of these options:

- Boris should choose the Experience Machine life
- Boris should choose the real life

Over half of the respondents to this survey (55%, 42/77) indicated that they thought Boris should choose the experience machine life (Weijers

manuscript). If Weijers' result is valid (if it is not an artefact of some structural bias, perhaps in the scenario or in the sample of respondents), then AP2 is false; when instructed to ignore their responsibilities to others, the vast majority of reasonable people *do not* report preferring reality over a life in an experience machine. But additional tests of Weijers' scenario on other sample groups would be required before AP2 could confidently be rejected. Furthermore, proponents of the experience machine objection to hedonism should analyse Weijers' scenario for new potentially bias-eliciting features and conduct surveys to test any hypotheses to that effect. Finally, other variants of experience machine scenarios that focus more directly on pleasure (instead of experiences) might also get very different results, perhaps resulting in a clear endorsement of AP2 for those scenarios.

So, AP2 should not be rejected outright, but it is certainly on shaky ground because there is currently some uncontested evidence that the vast majority of reasonable people *do not* report preferring reality over a life in an experience machine. Furthermore, Weijers (manuscript) provides some evidence that the status quo bias and certain other confounding factors, such as irrelevant worries about whether the experience machine would work properly, are elicited by Nozick's experience machine scenario. If this is correct, then AP2 might well be true for Nozick's scenario, but AP3 would come into serious doubt; the best explanation for AP2 *might not be* that reality matters intrinsically to the vast majority of reasonable people (because people's preferences might be directed more by confounding factors than by reality mattering intrinsically to them). Again, further testing is required to empirically assess the extent of bias Nozick's and Weijers' scenarios elicit before anything definitive can be said about the fate of AP3 for Nozick's scenario.

5. Conclusion

This paper has introduced the experience machine objection to hedonism, explained why it is widely thought to refute prudential hedonism, and presented deductive and abductive versions of the argument supporting it. These two versions were used to create a taxonomy of responses to the experience machine objection to hedonism that is intended to be exhaustive of all the paper-length attempts to refute the experience machine objection to hedonism from the last 15 years, and indicative of all of the types of critical responses to it.

The discussion of the responses to the deductive version of the experience machine objection to hedonism made it clear that it was relatively ineffective, if not implausible. Most importantly, the deductive version seems to ignore the fairly obvious fact that reality might matter intrinsically to us for reasons other than reality being intrinsically prudentially valuable.

This paper has also shown that the abductive version of the experience machine objection to hedonism might eventually be seen as unlikely to be effective because of the pressure on AP3. That pressure is coming from the numerous possible causes of people's preference for reality over a life in an experience machine. Furthermore, when this pressure on AP3 is combined with the threat to AP2 coming from alternate experience machine scenarios, serious doubts about the experience machine objection to hedonism begin to arise. Most prominent of these threats to AP2 and AP3 is the worry that people's preferences for reality over a life in the experience machine in Nozick's scenario are heavily affected by the framing of the status quo, as demonstrated by the reversed experience machine scenarios of Kolber (1994), De Brigard (2010), and Weijers (forthcoming-a), and especially the empirical evidence from Weijers (manuscript).

However, to be certain of the fate of the abductive version of the experience machine objection to hedonism, the potential effects of biases and other confounding factors on our judgments about experience machine

scenarios should be investigated by considering alternate experience machine scenarios and testing any resulting hypotheses on a diverse sample population. Until this further research has been conducted, however, proponents of the experience machine objection to hedonism might want to use one of the other commonly used objections to refute hedonism.

References

- Arneson, R. (1999). Human Flourishing Versus Desire-Satisfaction, *Social Philosophy & Policy*, 16(1): 113–142.
- Attfield, R. (1987). *A Theory of Value and Obligation*. London: Croom Helm.
- Baber, H. (2008). The Experience Machine Deconstructed, *Philosophy in the Contemporary World*, 15(1): 132–137.
- Baggini, J. & Fosl, P. S. (2007). *The Ethics Toolkit: A Compendium of Ethical Concepts and Methods*. Wiley-Blackwell.
- Barber, A. (2011). Hedonism and the Experience Machine, *Philosophical Papers*, 40(2): 257–278.
- Becker, L. C. (1992). Good Lives: Prolegomena, in E. F. Paul, F. D. Miller, Jr., & J. Paul (eds.), *The Good Life and the Human Good*. Cambridge: Cambridge University Press: 15–37.
- Belshaw, C. (forthcoming). What’s Wrong with the Experience Machine?, *European Journal of Philosophy*. Available from:
<http://onlinelibrary.wiley.com/doi/10.1111/j.1468-0378.2011.00507.x/pdf>
- Bok, S. (2010). *Exploring Happiness: From Aristotle to Brain Science*. New Haven: Yale University Press.
- Bradbury, R. (1957). *Dandelion Wine*. Doubleday.
- Bradley, B. (2009). *Well-Being and Death*. Oxford: Oxford University Press.
- Brandt, R. (1966). The Concept of Welfare, in S. R. Krupp (ed.), *The Structure of Economic Science*. Prentice-Hall: 257–276.

- Brink, D. (1989). *Moral Realism and the Foundations of Ethics*. Cambridge: Cambridge University Press.
- Bronsteen, J., Buccafusco, C., & Masur, J. (2010). Welfare as Happiness, *Georgetown Law Journal*, 98: 1583–1641.
- Brülde, B. (2007). Happiness Theories of the Good Life, *Journal of Happiness Studies*, 8: 15–49.
- Churchland, P. (1996). Towards a Cognitive Neurobiology of the Moral Virtues, *Topoi*, 17: 83–96.
- Crisp, R. (2003). Equality, Priority, and Compassion, *Ethics*, 113(4): 745–763.
- Crisp, R. (2006a). Hedonism Reconsidered, *Philosophy and Phenomenological Research*, LXXIII(3): 619–645.
- Darwall, S. (1997). Self Interest and Self Concern, *Social Philosophy and Policy*, 14: 158–178.
- De Brigard, F. (2010). If You Like it, Does it Matter if it's Real?, *Philosophical Psychology*, 23(1): 43–57.
- Donner, Wendy (1991) *The Liberal Self: John Stuart Mill's Moral and Political Philosophy*. Ithaca, NY: Cornell University Press.
- Feldman, F. (2002). The Good Life: A Defense of Attitudinal Hedonism, *Philosophy and Phenomenological Research*, LXV(3): 604–628.
- Feldman, F. (2004). *Pleasure and the Good Life*. Oxford: Clarendon Press.
- Feldman, F. (2011). What We Learn from the Experience Machine, in R. M. Bader & J. Meadowcroft (eds.), *The Cambridge Companion to Nozick's Anarchy, State, and Utopia*. Cambridge University Press.
- Finnis, J. (1980). *Natural Law and Natural Rights*. Oxford: Clarendon Press.
- Finnis, J. (1983). *Fundamentals of Ethics*. Washington, DC: Georgetown University Press.
- Fletcher, G. (2007). Wrongness, Welfarism and Evolution: Crisp on Reasons and the Good, *Ratio* (new series), 20(3): 341–347.

- Griffin, J. (1986). *Well-Being: Its Meaning, Measurement and Moral Importance*. Oxford: Clarendon Press.
- Haslett, D. W. (1990). What is Utility?, *Economics and Philosophy*, 6(1): 65-94.
- Hausman, D. M. (2010). Hedonism and Welfare Economics, *Economics and Philosophy*, 26: 321–344.
- Haybron, D. M. (2008). Happiness, the Self and Human Flourishing, *Utilitas*, 20(1): 21–49.
- Hewitt, S. (2009). What Do Our Intuitions about the Experience Machine Really Tell Us about Hedonism?, *Philosophical Studies*, 151(3): 331–349.
- Hooker, B. (2000). *Ideal Code, Real World: A Rule-Consequentialist Theory of Morality*. Oxford: Clarendon.
- Hurka, T. (2011). *The Best Things in Life: A Guide to What Really Matters*. New York: Oxford University Press.
- Jollimore, T. (2004). Meaningless Happiness and Meaningful Suffering, *The Southern Journal of Philosophy*, 42(3): 333–347.
- Kagan, S. (1998). *Normative Ethics*. Boulder: Westview Press.
- Kagan, S. (2009). Well-Being as Enjoying the Good, *Philosophical Perspectives: Ethics*, 23: 253–272.
- Kahneman, D. & Tversky, A. (2000). *Choices, Values, and Frames*. Cambridge: Cambridge University Press.
- Kauppinen, A. (2011). How the Experience Machine Works, *Experimental Philosophy* (blog), 9 February 2011. Available from: http://experimentalphilosophy.typepad.com/experimental_philosophy/2011/02/how-the-experience-machine-works.html
- Kawall, J. (1999). The Experience Machine and Mental State Theories of Well-Being, *The Journal of Value Inquiry*, 33: 381–387.
- Keller, S. (2009). Welfare as Success, *Noûs*, 43(4): 656–683.
- Kolber, A. (1994). Mental Statism and the Experience Machine, *Bard Journal of Social Sciences*, 3(3–4): 10–17.

- Kraut, R. (2007). *What is Good and Why: The Ethics of Well-Being*. Cambridge, MA: Harvard University Press.
- Kymlicka, W. (1990). *Contemporary Political Philosophy: An Introduction*. Oxford: Oxford University Press.
- Mendola, J. (2006). Intuitive Hedonism, *Philosophical Studies*, 128(2): 441–477.
- Mill, J. S. (1969). *Autobiography*. J. Stillinger (ed.). Boston: Houghton Mifflin Company.
- Nozick, R. (1974). *Anarchy, State, and Utopia*. Oxford: Blackwell, 1991.
- Nozick, R. (1989). *The Examined Life: Philosophical Meditations*. New York: Simon and Schuster.
- Railton, P. (1984). Alienation, Consequentialism, and the Demands of Morality, *Philosophy & Public Affairs*, 13(2): 134–171.
- Rivera-López, E. (2007). Are Mental State Welfarism and Our Concern for Nonexperiential Goals Incompatible?, *Pacific Philosophical Quarterly*, 88: 74–91.
- Sarch, A. (2011). Internalism about a Person's Good: Don't Believe it, *Philosophical Studies*, 154: 161–184.
- Sidgwick, H. (1907). *The Methods of Ethics* (7th Edition). Macmillan.
- Silverstein, M. (2000). In Defense of Happiness: A Response to the Experience Machine, *Social Theory and Practice*, 26(2): 279–300.
- Smart, J. J. C. (1973). An Outline of a System of Utilitarian Ethics, in J. J. C. Smart & B. Williams (eds.), *Utilitarianism: For and Against*. Cambridge: Cambridge University Press: 3–74.
- Smith, B. (2011). Can We Test the Experience Machine?, *Ethical Perspectives*, 18(1): 29–51.
- Sobel, D. (2002). Varieties of Hedonism, *Journal of Social Philosophy*, 33(2): 240–256.
- Sumner, L. (1996). *Welfare, Happiness, & Ethics*. Oxford: Clarendon press.

- Sunstein, C. R. (2005). Moral Heuristics, *Behavioral and Brain Sciences*, 28: 531–573.
- Tännsjö, T. (2007). Narrow Hedonism, *Journal of Happiness Studies*, 8: 79–98.
- Thomson, G. (1987). *Needs*. London: Routledge and Kegan Paul.
- Tiberius, V. (2004). Cultural Differences and Philosophical Accounts of Well-Being, *The Journal of Happiness Studies*, 5: 293–314.
- Tiberius, V. (2006). Well-Being: Psychological Research for Philosophers, *Philosophy Compass*, 1: 493–505
- Tiberius, V. & Hall, A. (2010): Normative Theory and Psychological Research: Hedonism, Eudaimonism, and Why it Matters, *Journal of Positive Psychology*, 5(3): 212–225.
- Van Wyk, R. N. (1990). *Introduction to Ethics*. New York: St. Martin's Press.
- Weijers, D. (2011a). Hedonism, *Internet Encyclopedia of Philosophy*. Available from: <http://www.iep.utm.edu/hedonism/>
- Weijers, D. (2011b). The Experience Machine Objection to Hedonism, in M. Bruce & S. Barbone (eds.), *Just the Arguments*. Wiley-Blackwell: 229–231.
- Weijers, D. (2011c). Reality Doesn't Really Matter, in D. K. Johnson (ed.), *Inception and Philosophy*. Wiley-Blackwell: 92–107.
- Weijers, D. (2012). We Can Test the Experience Machine: Reply to Smith, *Ethical Perspectives*, 19(2), pp. 261–268.
- Weijers, D. (forthcoming-a). Intuitive Biases in Judgements about Thought Experiments: The Experience Machine Revisited, *Philosophical Writings*.
- Weijers, D. (manuscript). Nozick's Experience Machine is Dead: Long Live the Experience Machine!, under review at *Philosophical Psychology*.