

If you like it, does it matter if it's real?¹

1. Pleasure, Reality and the Experience Machine

In his 1974 book *Anarchy, State and Utopia* Robert Nozick introduced one of the most infamous thought experiments in contemporary philosophy: the so-called “experience machine”. Nozick wanted us to imagine that scientists have created a wonderful virtual reality machine capable of providing us with any pleasurable experience we may desire. Then he asked us to think whether we would like to plug in or not. His idea was that, in general, people would feel very little inclination to plug in because, in addition to pleasurable experiences, we also care about living “in contact with reality” (Nozick 1974, 45).

Nozick’s thought experiment was intended to mobilize our intuitions toward a negative claim: that hedonism is false because pleasurable experiences are not the only thing that matter to us. Ever since, the dialectic triggered by the famous thought experiment has mostly been a conflict between two groups. On the one hand, there are those who claim that Nozick’s thought experiment does not work, and/or that, given the correct definition of ‘pleasure’, hedonism may still survive (e.g. Sumner 1996, Silverstein 2000, Sobel 2002, Tännsjö 2007). On the other hand, there are those who believe Nozick’s thought experiment actually works, and that there must be something else—‘reality’, or perhaps ‘authenticity’—that people value more than pleasure (see Crisp 2005 or Brülde 2007 for an overview).

¹ This paper benefited enormously from the comments of my audience at the University of North Carolina, Chapel Hill. Special thanks are also due to Erik Angner, Daniel Gilbert, Adrienne Harris, Eric Mandelbaum, Jesse Prinz, David Ripley, Dylan Sabo, and (very specially!) Joshua Knobe.

I think this dichotomy is misguided. In this paper I present some experimental evidence suggesting, first, that the original experience machine thought experiment is ill-construed insofar as it does not isolate the intuition Nozick wants to isolate. Second, I show that it is possible to provide some contexts that may alter people's preference, such that they end up favoring a virtual life over a real life. I use this evidence to suggest that, for most people, neither pleasure nor reality seem to be indefeasibly fundamental values. Finally, I present an explanatory model, derived from what behavioral economists call *the status quo bias*, in order to account for these results.

2. Experimenting with the Experience Machine

“Would you plug in?” Nozick asked almost rhetorically after introducing the experience machine thought experiment; and, as our mouthpiece, he answered in the negative. It is true that, for a while, philosophers seemed to be of the same opinion. Recently, however, there has been some skepticism as to whether it is the case that people, in fact, share Nozick's intuitions. In a recent paper, for instance, Torbjörn Tännsjö (2007) expresses these doubts persuasively:

“After all, many people choose to use drugs they know are dangerous, such as alcohol, in spite of the fact that they know that it is difficult to give up the habit of using them. So why not opt for a perfect experience machine (that you can opt out from if you like) with no bad side effects—and stay plugged into it?” (94).

I share his skepticism. I do not think it is obvious at all that, when confronted with the possibility of plugging themselves into the experience machine, people will refuse. And even if some do, I think it is also unclear whether they would refuse to do so *precisely* for the reason Nozick alludes to—that is, that they do not want to lose “contact with reality”.

In order to discern whether Nozick's intuitions about people's opinions are on the right track, I conducted the following two studies.

2.1. Study 1

In the first study each participant was randomly assigned to one of two conditions (20 and 21 participants to each condition respectively. Participants were all freshmen from the University of North Carolina at Chapel Hill, with no philosophical training whatsoever. The second condition, which is reviewed below, was given to participants at the same time.) In the first condition (N=20) they were asked to read a vignette with Nozick's original version of the experience machine thought-experiment. They were then required to answer whether or not they would plug in, and to briefly explain their choices. For the first answer, a 1 to 7 scale was provided, with 1 being "No", 7 being "Yes", and 4 "Undecided". Here is what they received:

Original version:

Suppose there was 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 would 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 actually happening. Others can also plug in to have the experiences they want, so there is no need to stay unplugged to serve them. Would you plug in? [Please circle only one option]

○ ○ ○ ○ ○ ○ ○
No Most probably not Maybe not Undecided Maybe yes Most probably yes Yes

Please explain your answer very briefly:

The results were very close to what Nozick would have predicted (Figure 1). With an average of 2.0, participants in the first condition are indeed of the opinion that they would probably not plug into the machine. This, however, does not settle the issue for Nozick quite yet; it is unclear that participants were reluctant to plug in *because* they did not want to lose contact with reality. In fact, when asked to explain their answers, 11 out of 20 had nothing to do with their preference for a real life over a virtual one. Instead, many indicated that they were dubious the computer program could predict everything they wanted, or that they would feel unhappy if they were to plug in by themselves without friends or relatives. Some participants even expressed qualms concerning the continuity of their memories and identities. One answer, in fact, pointed to the following paradoxical implication (I'm paraphrasing): if the experience machine is supposed to provide me with things I find pleasurable, and I find pleasure in surprising experiences, then by pre-programming all my future experiences I will undercut my major source of pleasure, namely the element of surprise. But if the experience machine cannot provide me with these sorts of experiences, then is false that it is able to provide me with everything I want.

Even as a thought experiment, the experience machine seems ill-conceived. If people have so many qualms regarding the proper functioning of the machine, how can we be so sure that the thought experiment effectively isolates the intuition Nozick wants? Some—Sumner (1996, 95) for instance—have pushed this line of argument against the

experience machine. Nozick himself seems to have been aware of this potential difficulty, as the version of the experience machine in his later book, *The Examined Life*, tries to address some of these worries (1989, 104). Still, philosophers are unmoved by such concerns. They keep thinking that, even if we manage to overcome all these alleged difficulties, people will remain reluctant to plug in. As Matthew Silverstein asserts in response to Sumner's worries:

If we exercise a bit of mental dexterity [...] we can make an earnest and largely successful effort to overcome the sorts of doubts raised by Sumner. Even without these doubts, though, most of us continue to share Nozick's intuitions: we remain unwilling to accept a lifetime on the experience machine (2000, 284-285).

Most of us? I'm unconvinced. In order to ascertain whether people would still be reluctant to plug in, even when as many doubts I could think of vis-à-vis the experience machine's mechanism had been resolved, I gave participants in a second condition (N=21) a new vignette with an updated version of the original thought experiment. As in the first condition, they were required to answer, after reading the vignette with the updated version, whether they would plug in or not, and to explain their choice. The same 1-7 scale was used this time as well (so I'm omitting it below). Here is what they received:

Updated version:

Suppose there is 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, making a friend, reading an interesting book, having sex with the lover of your dreams, enjoying your favorite food, or what have you. Your brain will be floating in a tank, secured, and with electrodes attached to it, but you will feel all the experiences as vividly as you feel any experience now. All your loved ones agreed to be plugged into the machine as

well—including friends you have not seen in a while and those that live far away. Your pet will be hooked up too. All your memories of your past will remain just fine, but scientists will erase any information regarding the experience machine, such that the event of your connection won't be for you but a common waking-up-in-the-morning experience—absolutely indistinguishable from any other morning. Of course, since neuroscientists know you like surprises, instead of asking you what are the pleasurable things you'd like to experience in the future, they studied you throughout your life, so they have a pretty accurate picture of the things you have showed preference for and the things you don't like. Based on this information, they have prepared all sorts of pleasurable experience they know you'd like. Would you plug in?

Since the unreality of the experiences in the updated version of the machine remained constant, one would expect to see no difference in the results, if it is actually the case that “contact with reality” is the operant variable. Surprisingly, the average in the updated version scales up to 3.9, barely below the mid point, and significantly above the average for the initial condition ($t(39)=3.51, p=.001$). This suggests that issues having to do with the pre-deterministic character of the computer program, memory continuity, pain, and loneliness, were affecting people's reluctance to plug in. Contra Silverstein, these results suggest that once we take care of difficulties with the machine itself, as well as with the process of getting hooked up to it, the import of reality diminishes. Far from being unwilling to plug in, now the folk are merely undecided².

² Some have suggested that certain dissimilarities between the original and the updated versions may help explaining the difference in the results. Dylan Sabo, for instance, pointed out that while the original version asks participants to plug into the machine for life, the updated version is less explicit about it. Although this factor may have influenced people's decisions, I have my doubts: neither in the first condition nor in the second one participants mentioned any concerns about the duration of the commitment, either as a reason to plug in or not to plug in. Such omission isn't conclusive of course, but it is suggestive. Similarly, Erik Angner worried that participants may have read the second vignette as inviting them to “connect” to a sort of simulated, parallel life—like SecondLife or Facebook—rather than the kind of virtual experience machine Nozick had in mind. Could this have influenced people's responses? Maybe. But then again, such a reading seems also plausible for the first vignette. We philosophers are more familiar with Nozick's own version (and presumably less so with internet communities), and are less prone to read his thought experiment in this alternative way. Participants in this study, however, might. At any rate, whether any of these dissimilarities may be relevant for the explanation of the results, the main point still holds: manipulating variables other than reality swing people's preferences. For reasons that will become clear soon, I think there is one factor that may help account for this phenomenon.

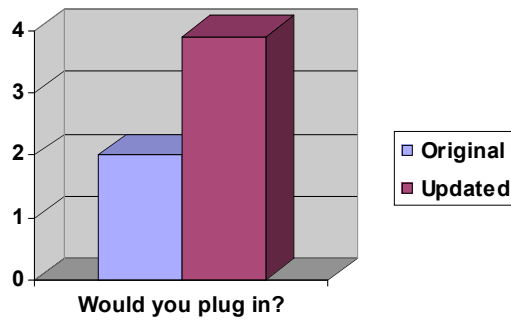


Figure 1: *Original* versus *Updated* versions. Scale runs from 1 to 7, with 4 being the mid point.

Nonetheless, contact with reality alone *may still* be the variable keeping everyone from plugging in unanimously. For, as interesting as the previous results may be, they speak neither in favor of hedonism nor against the importance of reality. So far these results just speak against Nozick’s formulation of the experience machine thought experiment. Of course, one may try to come up with a formulation of the thought experiment which takes care of every potential distraction. And it may be possible that with such a formulation the folk would then judge not to plug in solely from their preference for reality. This task, however, may prove fruitless. For if one fixes the machine to a point in which, *from the inside*, virtual experiences turn out to be indistinguishable from real ones, new skeptical doubts arise. Suppose someone tells you that, unbeknownst to you, you have been living inside an experience machine; that you are a brain-in-a-vat, as it were. Would you feel then that your life is less preferable than a life outside the machine? Would you feel your life has less value to you? Some seem to think so. For instance, Tännsjö reports a conversation with Brülde in which the latter claims that “we are not indifferent to a life in or outside the machine. [That in fact] many

people would find the life outside the machine at least *slightly* preferable to a life in the machine” (2007, 95; but see also Brülde 2007). According to Brülde, this alone would be enough to refute hedonism.

Tännsjö remains unconvinced, and I again share his skepticism. It is not clear at all to me that people would necessarily prefer reality when facing a brain-in-a-vat situation. I do not even think this piece of information would affect, per se, their judgments on their own happiness or well-being. Thus, in order to test people’s intuitions about this claim, a second study was conducted. The results suggest that people’s preferences do not depend initially upon their partiality toward pleasure or reality. Instead, I suggest, their choice is primarily dependent upon their aversion to lose the life they have been experiencing so far. To motivate this alternative explanation, let us move to the second study.

2.2. Study 2:

In this study I presented participants with a backward-looking experience machine thought experiment. I asked them to imagine being already connected to an experience machine, so this time they face the option of disconnecting and going back to reality, or remaining connected and living their virtual lives. There are three variations to this scenario. In the *Neutral vignette* they are presented with the possibility of remaining connected or going back to reality *simpliciter*. In the *Negative vignette*, further information about reality is provided: they are told their real life is different from their virtual one in that, in reality, they are prisoners in a maximum security prison. Finally, in the *Positive vignette*, they also get further information about reality, but this time they are

told that they are multimillionaire artists living in Monaco. Participants were randomly assigned to one of three conditions, with 24 participants in each condition and each participant receiving only one vignette. They were asked to pick one of two options (“Remain connected” or “Go back to reality”) and to explain their answer. As before, participants were undergraduate students from the UNC-CH with no previous exposure to philosophy (and, to the best of my knowledge, none of them were prisoners or millionaire artists). These are the vignettes (titles were omitted; the structure of the question after the vignette is the same in all of them, so I will only include it once):

Neutral vignette:

It is Saturday morning and you are planning to stay in bed for at least another hour when all of the sudden you hear the doorbell. Grudgingly, you step off the bed to go open the door. At the other side there is a tall man, with a black jacket and sunglasses, who introduces himself as Mr. Smith. He claims to have vital information that concerns directly to you. Mildly troubled but still curious, you let him in. “I am afraid I have to some disturbing news to communicate to you” says Mr. Smith. “There has been a terrible mistake. Your brain has been plugged by error into an experience machine created by superduper neurophysiologists. All the experiences you have had so far are nothing but the product of a computer program designed to provide you with pleasurable experiences. All the unpleasantness you may have felt during your life is just an experiential preface conducive toward a greater pleasure (e.g. like when you had to wait in that long line to get tickets for that concert, remember?). Unfortunately, we just realized that we made a mistake. You were not supposed to be connected but someone else. We apologize. That’s why 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.”

What would you choose? [Please circle only one option]

○

Remain connected

○

Go back to reality

Please explain your answer briefly:

The Negative and Positive vignettes read exactly as the Neutral vignette above, except they included the following sentence at the end, i.e. after "...your real life":

Negative vignette:

[...your real life.] By the way, you may want to know that your real life is not at all as your simulated life. In reality you are a prisoner in a maxim security prison in West Virginia.”
What would you choose?

Positive vignette:

[...your real life]. By the way, you may want to know that your real life is not at all as your simulated life. In reality you are a multimillionaire artist living in Monaco.”
What would you choose?

The results were quite surprising (Figure 2). For the Negative scenario, only 13% of the participants said to prefer reality; 87% of them would prefer to remain connected. In the Positive scenario, the distribution evened out: 50% of the participants said they would like to go back to reality, and 50% said they would like to remain connected. Finally, when assessing the Neutral scenario, 54% of the participants said they would like to go back to reality, whereas 46% would prefer to remain connected. (The difference among conditions was statistically significant: $\chi^2(2, N=72)=10.6, p=.005$)

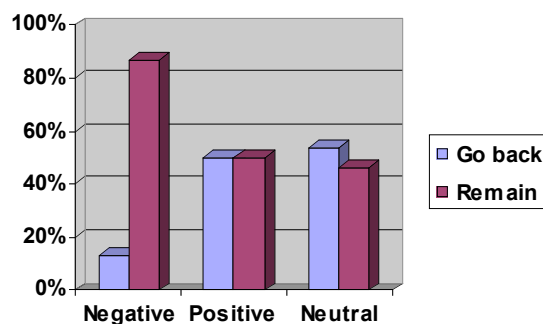


Figure 2: Percentage of responses for *Negative*, *Positive* and *Neutral* vignettes.

Taken at face value these results suggest that the variable “contact with reality” is not operant. If it were the case that people care about reality more than they care about how their life is experienced from the inside, then one would expect that, *regardless* of how reality turns out to be, they would choose it over a simulated life. This prediction was not borne out by the present study, though. Except for the Neutral condition (but see below), not only did participants not seem to put too much value on reality per se, but they also seemed to care about the *quality* of their real life versus the quality of their virtual—albeit familiar—life. But one has to be cautious with this interpretation. It would be a mistake to think that, since the quality of life affected the folk’s decision, their choice was *in effect* dictated by a hedonistic preference. If that were the case, one would expect to see the opposite effect with the Positive scenario than with the Negative scenario, i.e. a strong preference for reality. This prediction, however, was not borne out by the study either. People do not seem as motivated to go for a seemingly better life, as they are to stay away from a seemingly worse one.

Hedonists may try to accommodate this anomalous result saying, for instance, that whereas most people find a life in prison worse than their current life, not everyone finds that being a multimillionaire artist in Monaco is a better one. They may argue, for example, that most people think being an artists is passé, or that having a lot of money is rather stressful, or even that Monaco is a pretty depressing place to live. But—the thought goes—if we were able to find a good example of a kind of life everyone found to be better, we would stumble into the kind of result that most clearly favors hedonism.

Unfortunately, this strategy not only shifts the burden of proof to the hedonist, it also distract us from what I think is a more interesting reading of the result: that, just as with reality, pleasure does not seem to be the tipping point. This may be because—as I will argue in a moment—in addition to what they want, people care for what they have.

Finally, some may contend that the results of the Neutral vignette still show people’s preference for reality. After all, the study showed that most people (54%) preferred to go back to a real life they had no information about versus one they did. Contact with reality *may* still be the variable pushing people’s preferences. I remain skeptical, however. Notice that, unlike the Positive and the Negative scenarios, the Neutral vignette does not imply that the participant’s life is going to change with respect to the life he or she is familiar with. In other words, whereas the Negative and the Positive scenarios invite participants to consider leaving a virtual existence and going back to a life that is very different from the life they have experienced so far, the Neutral case says nothing as to whether the participant will resume his or her current life. What would the results be if I provide participants with a Neutral vignette that only makes explicit a change in their personal “status quo”—i.e. their life as they know it—if they disconnect? In order to answer this question I conducted a follow-up study with 80 participants, all undergraduates from UNC-CH with no previous exposure to philosophy. As before, they were asked to read a vignette, circle one of two options, and explain their choice. This *Second Neutral vignette* was identical to the Neutral vignette above, except for the ending part which I reproduce italicized here:

Second Neutral vignette:

It is Saturday morning and you are planning to stay in bed for at least another hour when all of the sudden you hear the doorbell. Grudgingly, you step off the bed to go open the door. At the other side there is a tall man, with a black jacket and sunglasses, who introduces himself as Mr. Smith. He claims to have vital information that concerns directly to you. Mildly troubled but still curious, you let him in. “I am afraid I have to some disturbing news to communicate to you” says Mr. Smith. “There has been a terrible mistake. Your brain has been plugged by error into an experience machine created by superduper neurophysiologists. All the experiences you have had so far are nothing but the product of a computer program designed to provide you with pleasurable experiences. All the unpleasantness you may have felt during your life is just an experiential preface conducive toward a greater pleasure (e.g. like when you had to wait in that long line to get tickets for that concert, remember?). Unfortunately, we just realized that we made a mistake. You were not supposed to be connected but someone else. We apologize. That’s why 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 *disconnect*. However, you may want to know that your life outside is not at all like the life you have experienced so far.”

What would you choose? [Please circle only one option]

Remain connected

Disconnect³

With the Second Neutral vignette the results shifted strikingly (Figure 3): 59% of participants wanted to remain connected, while only 41% wanted to disconnect. What is more surprising, though, is that only a third of the participants that wanted to disconnect justified their choice by making explicit reference to their preference for reality. Instead, most of them wanted to go back because they were unsatisfied with their current lives, or because they did not want their decisions to be “preprogrammed”, or simply because they were “curious to see what my life outside the machine would be like”. Interestingly, one participant wrote that s/he would like to disconnect to “start all over and try things that

³ David Ripley and Eric Mandelbaum pointed out that framing the question in terms of “Going back to reality”, as I had it before, may affect people’s responses. It is possible that some participants chose that option just because the term “reality” was interpreted as having a positive connotation, regardless of whether they wanted to disconnect or not. Thus, in the Second Neutral vignette the option was changed to the more neutral “Disconnect”.

normally wouldn't have done". Once again, even though it seemed to be important for some, it looks like contact with reality was not the operant variable here.

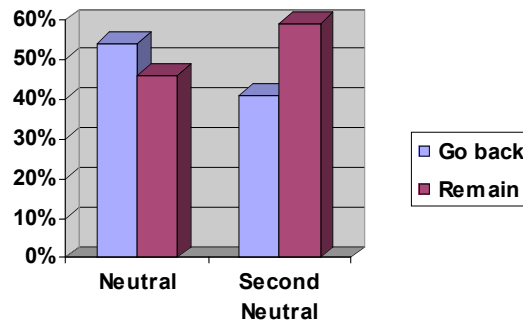


Figure 3: Percentage of responses in *Neutral* versus *Second Neutral* vignettes.

What these results suggests—I contend—is that although people seem to value, at least to some extent, both contact with reality as well as pleasure, it is also true that, given the right circumstances, they are willing to give up either of them. Now, what could these “circumstances” depend on? And—more pressingly—how can we explain the previous results? I think these are important questions, the answers of which may come from a different research area: prospect theory⁴.

3. The Experience Machine and the Status Quo Bias

⁴ The use of prospect theory to account for various intuitions in philosophy and moral psychology is starting to show interesting results. Tamara Horowitz (1998), for instance, famously introduced prospect theory to account for asymmetries in traditional ethical dilemmas (e.g. cases used to exemplify the doctrine of Doing and Allowing). More recently, Bostrom and Ord (2006) used prospect theory to elaborate on whether people would refuse or not to genetically enhance their cognitive abilities. I want to think of the present paper as contributing to this tradition.

Marcia, a philosopher friend, acquired a 1932 edition of Kant's *Critique of Pure Reason* a while ago, at a time in which it cost no more than \$20. Last week, at a party at her house, one of her book-lover acquaintances told her that such a particular edition of Kant's work had significantly appreciated in value, and that any book collector would surely be willing to pay between \$100 and \$120 for a good copy. Marcia knows she could get 3 or maybe even 4 decent newer copies of the same work for that amount of money. However, she is not in the least interested in selling her current copy—even though she would never pay \$100 for the same book, if she didn't already have it.

Marcia's decision to keep her book exemplifies an all too common behavioral pattern known by economists and behavioral psychologists as *the status quo bias* (Samuelson and Zeckhauser 1988): the fact that people tend to prefer the state of affairs they are currently in—their status quo—because, when facing a decision that could alter it, “the disadvantages of leaving it loom larger than the advantages” (Kahneman et al. 1991, 197-8). The status quo bias has been widely tested experimentally. In a famous experiment, Knetsch (1989 [see also Knetsch and Sinden 1984]) gave a questionnaire to two different groups of undergraduates. As a reward, participants in the first group were given a mug bearing the university logotype while participants in the second group received a chocolate bar. Once participants received their remuneration, they also got the option of trading it for the other alternative. Almost 90% of them preferred to keep the present that received initially. People just seem to prefer their current state of affairs because they deem that the pain of giving it up would be greater. This may explain our reluctance to change our current phone plan or our current health insurance plan, even when presented with more attractive alternatives (see Samuelson and Zeckhauer 1988 for

a research on the status quo bias and health insurance plans, and also Gilbert 2007 for the influence of the status quo bias and related phenomena in our affective forecasting).⁵

What does the status quo bias has to do with the results of the previous studies on the experience machine thought experiment? I suggest that what mobilizes people's intuitions against plugging-in in the first study and against disconnecting in the second study is not primarily a reflection on the nature of reality, nor their hedonistic preference for pleasure, but rather a psychological bias toward maintaining their status quo. In this context, I understand "status quo" to refer roughly to the way in which one's own life has been experienced "from the inside". Thus, people's reluctance to plug into the experience machine in the first study, and their reluctance to disconnect in the second study, would be manifestations of the same underlying psychological phenomenon: their aversion to lose their status quo, the life they have been living so far. The explanation for why most people prefer not to plug into the original experience machine (as opposed to the updated version), and why most people prefer not to disconnect after spending their life in it (as in the Negative and Second Neutral vignettes), is not the virtual character of the experience, nor the amount of pleasure they are told they would feel, but rather the simple fact that most people don't want to abandon the life they know, the life they have lived so far, the life they are familiar and comfortable with. After all, as Jason Kawall once remarked, in order to get the thought experiment's point across, an experience machine is not even needed:

⁵ Samuelson and Zeckhauser (1988) suggested that the status quo is a manifestation of a more overarching principle known as *loss aversion*: when facing a decision, agents tend to perceive the disutility of giving up an object as being greater than the utility they see in acquiring it (Kahneman and Tversky, 1984). Loss aversion would also explain the more well-known behavioral pattern dubbed by Richard Thaler (1980) as *the endowment effect*: people tend to increase the value of a good once it becomes part of their endowment. As a result, a person would demand much more to give up a certain owned good than they would be willing to pay to acquire the same good (Kahneman et al. 1991, 194).

We would reject alternative lives in which we would be guaranteed more pleasure, or in which we would maximize overall utility, or in which we would have great accomplishments, *if the alternatives require us to abandon our current lives and commitments*, whether or not an experience machine is involved (Kawall 1999, 383. The emphasis is mine).

The status quo bias may also explain an interesting similarity between the results of the Positive scenario in the second study—where 50% wanted to remain connected, and 50% wanted to go back to a life in Monaco—and the results of the updated version of the experience machine in the first study—where participants were, on average, undecided as to whether to plug in or not. Recall that both vignettes offered participants quite enticing scenarios. Notwithstanding such tantalizing offers, and the potential gains both of them afforded, about half of the participants preferred to remain in their status quo. For them, abandoning one’s own life is taken to be worse than accepting a potentially better yet alien one, regardless of whether the offered life is real or not. As it turns out, some participants in the Positive vignette provided practically the same reasons for not wanting to go back to reality that some participants in the Updated version provided to plug-in to the experience machine: both claimed that, albeit tempting, they did not want to live a life they were not familiar with, and/or that they did not want to lose contact with some of their valuable experiences, quite independently of their unpleasant or virtual nature.

Finally, people’s aversion to relinquish their status quo may also help us explaining why some participants *actually* opted for the change—that is, why some were willing to plug into the machine in the first study, and some willing to disconnect in the second. Behavioral economists tell us that the family of cognitive biases known as *loss*

aversion, of which the status quo bias is but a case (see footnote 3), distort the utility function by making it asymmetric (Figure 4). As a result, when the aversion to abandon the status quo is taken into account, the slope of the value function changes abruptly at the reference point (see below), becoming steeper for losses than for gains. For our purposes, this result has two interesting implications. On the one hand, it tells us that the fact that our decisions are influenced by the status quo bias *does not* imply that people are always unwilling to abandon the situation in which they are or give up a certain good they own. What it *does* imply is that people would demand much more for giving up their own good, or for giving up their current state, than they would offer for the same good (or state) had it not been theirs to begin with. That is why, for instance, you think your beloved car is worth more than what the seller is offering. On the other hand—and more importantly—the phenomenon of loss aversion tells us that people exhibit different attitudes toward risky decisions depending on whether they frame a certain outcome as a potential loss or as a potential gain (Tversky and Kahneman 1991). Risk-prone agents tend to frame outcomes as losses, whereas risk-averse people tend to frame them as gains.

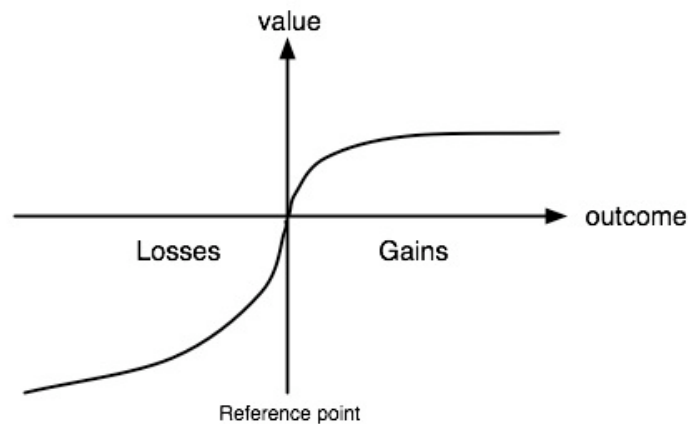


Figure 4: Asymmetric value function in prospect theory (drawing by Marc Oliver Rieger)

Now, if the interpretation of the model I am suggesting is correct, then the following explanation for the results in which participants opted for a change in their status quo is available: in the case of the first study, risk-prone participants framed the outcome of plugging in as a gain, whereas risk-averse participants probably framed it as a loss and thus decided not to plug in. Likewise, in the second study, those who were risk-prone framed the outcome of disconnecting as a gain, whereas those who were risk-averse framed it as a loss and preferred to remain connected. Framing, however, is a two-stage process in which the agent starts off choosing a possible outcome as the neutral or reference point, and then assesses other possible outcomes as losses or gains relative to that reference point (i.e. negative or positive deviations from the reference point [Figure 4]). Which factors could influence the framing process when participants face decisions like those suggested by the studied thought experiments? This, I think, is an interesting question worth investigating.

4. Conclusions and Final Remarks

The results of the reviewed studies suggest that people's reluctance to plug into the experience machine does not depend primarily upon people's preference for a real life over a simulated life, as Nozick thought. Likewise, people's reluctance to disconnect from a life in which they are already connected to an experience machine does not depend primarily upon their preference for a pleasurable life over a life in the real world, as some hedonists may have it. What seems to account for these responses is a psychological bias toward the status quo which motivates people to appraise their current

life, the life they are familiar with, *before* assessing other valuable things like pleasure and contact with reality.

Two final remarks: First, I think that the results of these studies strongly suggest not only that people's intuitions about the experience machine are highly divergent, but also that there are alternative explanations that can account for this divergence. Here I suggested only one. However, I think this must be taken as a cautionary note for philosophers: thought experiments have to be handled with care, because they may fail to mobilize the right sorts of intuitions. And I suspect that the experience machine may not be the only thought experiment not working the way it was supposed to. Second, I also think these results suggest, quite strongly, that moral philosophers must take into account empirical data on human psychology when talking about certain problems, even if that forces them to do what behavioral economists did more than twenty years ago: stop thinking that humans are perfectly rational animals.

Felipe De Brigard
University of North Carolina, Chapel Hill
brigard@email.unc.edu

References

- Bostrom, N. and T. Ord. 2006. The Reversal Test: Eliminating Status Quo Bias in Applied Ethics. *Ethics*. 116: 656-79.
- Brülde, B. 2007. Happiness Theories of the Good Life. *Journal of Happiness Studies*. 8: 15-49.
- Crisp, R. 2005. Well-Being. *Stanford Encyclopedia of Philosophy*.
- Gilbert, D. 2007. *Stumbling on Happiness*. New York: Knopf.
- Horowitz, T. 1998. Philosophical Intuitions and Psychological Theory. *Ethics*. 108: 367-85.

- Kawall, J. 1999. The Experience Machine and Mental State Theories of Well-being. *The Journal of Value Inquiry*. 33: 381-87.
- Kahneman, D. and A. Tversky. 1984. Choices, Values and Frames. *American Psychologist*. 39: 341-50.
- Kahneman, D.; Knetsch, J.L. and Thaler, R.H. 1991. Anomalies: The Endowment Effect, Loss Aversion and Status Quo Bias. *Journal of Economic Perspectives*. 5(1): 193-206.
- Knetsch, J.L. and J.A. Sinden. 1984. Willingness to Pay and Compensation Demand: Experimental Evidence of an Unexpected Disparity in Measures of Value. *Quarterly Journal of Economics*. 99: 507-21.
- Knetsch, J.L. 1989. The Endowment Effect and Evidence of Nonreversible Indifference Curves. *American Economic Review*. 79: 1277-84.
- Nozick, R. 1974. *Anarchy, State and Utopia*. New York: Basic Books.
- Nozick, R. 1989. *The Examined Life*. New York: Simon and Schuster.
- Samuelson, W. and Zeckhauser, R. 1988. Status Quo Bias in Decision Making. *Journal of Risk and Uncertainty*. 1: 7-59.
- Silverstein, M. 2000. In Defense of Happiness: A Response to the Experience Machine. *Social Theory and Practice*. 26(2): 279-300.
- Sobel, D. 2002. Varieties of Hedonism. *Journal of Social Philosophy*. 33(2): 240-56.
- Sumner, L.W. 1996. *Welfare, Happiness, and Ethics*. Oxford: Clarendon Press.
- Tännsjö, T. 2007. Narrow Hedonism. *Journal of Happiness Studies*. 8: 79-98.
- Thaler, R. 1980. Toward a Positive Theory of Consumer Choice. *Journal of Economic Behavior and Organization*. 1: 39-60.
- Tversky, A. and Kahneman, D. 1991. Loss Aversion in Riskless Choice: A Reference-Dependent Model. *Quarterly Journal of Economics*. 106(4): 1039-61.