

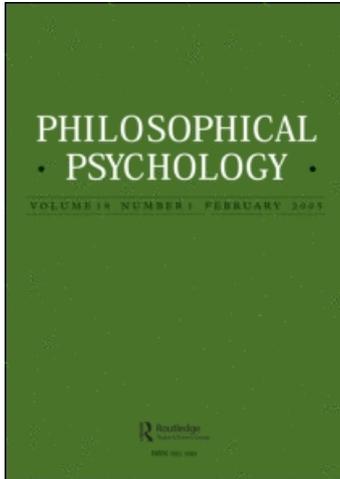
This article was downloaded by: [University of North Carolina-Chapel Hill]

On: 13 November 2010

Access details: Access Details: [subscription number 927346821]

Publisher Routledge

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



## Philosophical Psychology

Publication details, including instructions for authors and subscription information:

<http://www.informaworld.com/smpp/title~content=t713441835>

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

Felipe De Brigard

Online publication date: 08 February 2010

**To cite this Article** De Brigard, Felipe(2010) 'If you like it, does it matter if it's real?', *Philosophical Psychology*, 23: 1, 43 — 57

**To link to this Article:** DOI: 10.1080/09515080903532290

**URL:** <http://dx.doi.org/10.1080/09515080903532290>

## PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: <http://www.informaworld.com/terms-and-conditions-of-access.pdf>

This article may be used for research, teaching and private study purposes. Any substantial or systematic reproduction, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

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

Felipe De Brigard

*Most people's intuitive reaction after considering Nozick's experience machine thought-experiment seems to be just like his: we feel very little inclination to plug in to a virtual reality machine capable of providing us with pleasurable experiences. Many philosophers take this empirical fact as sufficient reason to believe that, more than pleasurable experiences, people care about "living in contact with reality." Such claim, however, assumes that people's reaction to the experience machine thought-experiment is due to the fact that they value reality over virtual experiences—an assumption that has seldom (if ever) been questioned. This paper challenges that very assumption. I report some experimental evidence suggesting that the intuition elicited by the thought-experiment may be explainable by the fact that people are averse to abandon the life they have been experiencing so far, regardless of whether such life is virtual or real. I use then an explanatory model, derived from what behavioral economists and psychologists call the status quo bias, to make sense of these results. Finally, I argue that since this explanation also accounts for people's reaction toward Nozick's thought-experiment, it would be wrong to take such intuition as evidence that people value being in touch with reality.*

*Keywords:* Experience Machine; Experimental Philosophy; Nozick; Pleasure; Reality; Status Quo Bias

## 1. Introduction

In his 1974 book *Anarchy, State and Utopia* Robert Nozick introduced one of the most infamous thought-experiments in contemporary philosophy: the experience machine. He wanted us to imagine a situation in which scientists have created a wonderful virtual reality machine capable of providing us with any pleasurable experience we may desire. Then he asked us to suppose that we were offered the possibility of plugging in to the machine. Nozick seems to suggest that most people's intuitive reaction after considering the thought-experiment would be just like his: they would feel very little inclination to plug in. Importantly, he suggests that the fact that the experience machine thought-experiment elicits in us this particular

---

Felipe De Brigard is a 5th year graduate student in the Department of Philosophy, University of North Carolina. Correspondence to: Felipe De Brigard, University of North Carolina at Chapel Hill—Department of Philosophy, CB# 3125, Caldwell Hall, Chapel Hill NC 27599-3125, United States. Email: brigard@email.unc.edu

intuition gives us reason to believe that, more than pleasurable experiences, people care about “living in contact with reality” (Nozick, 1974, p. 45; see also Nozick, 1989).<sup>1</sup>

Unsurprisingly, Nozick’s experience machine thought-experiment has received a lot of attention in the philosophical literature, and the dialectic triggered by it has mostly been a conflict between two groups. On the one hand, there are those who claim that the intuition elicited by Nozick’s thought-experiment does not necessarily undermine hedonism. They usually offer some account as to why people’s alleged preference for reality ends up supporting, rather than conflicting with, their favored version of hedonism (e.g., Silverstein, 2000; Sobel, 2002; Sumner, 1996; Tännsjö, 2007). On the other hand, there are those who believe that Nozick’s thought-experiment works just fine; indeed, that the intuition it elicits demonstrates not only that hedonism is false (at least in the weak sense of revealing that pleasurable experiences are not the only thing that matter to us), but also that, should they have to choose between pleasure and reality, people would prefer the latter (see Brülde, 2007 or Crisp, 2005 for an overview).

Nonetheless, the assumption that people’s reaction to the experience machine thought-experiment is due to the fact that they value reality has seldom been questioned. This paper wishes to challenge that very claim. More specifically, it attempts to cast doubt upon the effectiveness of the experience machine thought-experiment, not by way of providing an understanding of hedonism that can accommodate the intuition it elicits (whether there is one or not is besides the point), but rather by showing that such intuition does not provide support for the claim that people prefer not plugging in because they value being in touch with reality. To that effect, I present some experimental evidence suggesting that the intuition elicited by the experience machine thought-experiment may be explainable by the fact that people are averse to abandon the life they have been experiencing so far, regardless of whether such life is virtual or real. I present then an explanatory model, derived from what behavioral economists and psychologists call *the status quo bias*, in order to make sense of these results. Finally, I argue that since this explanation also accounts for people’s reaction toward Nozick’s thought-experiment, it would be a mistake to take such intuition as providing us evidence that people value being in touch with reality.

## 2. Experimenting with the Experience Machine

Notwithstanding the considerable amount of philosophical literature it has inspired, Nozick’s original formulation of the thought-experiment was rather short:

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. (Nozick, 1974, p. 44)

"Would you plug in?"—Nozick then asked. As our mouthpiece, he answered in the negative. Many philosophers are of the same opinion. They consider the alleged effectiveness of this thought-experiment a rather convincing proof against the tenability of psychological hedonism. Nonetheless, many philosophers remain doubtful. In a recent paper, for instance, Torbjörn Tännsjö (2007) persuasively expresses some of these doubts; in his opinion, it is far from obvious that people would prefer not to plug into the machine. After all, he says, "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?" (Tännsjö, 2007, p. 94)

If you try to teach Nozick's thought-experiment in an introduction to philosophy class you may hear your students voicing their skepticism along similar lines. In an informal survey I conducted a while ago, I introduced my students to Nozick's original version of the experience machine thought-experiment and then I asked them whether they would plug in or not. Although the majority preferred not to plug into the experience machine, when I asked them why, most of them—60%, to be more precise—gave reasons that had nothing to do with their preference for a real life over a virtual one. Many said they doubted 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: 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 surprises, then it is false that it is able to give me everything I want.

The lack of unanimity in people's answers raises doubts as to whether the thought-experiment actually works as Nozick intended. After all, if people have qualms regarding, say, the proper functioning of the machine, how can we be so sure that the thought-experiment effectively isolates the intuition Nozick wants to isolate—namely, that people's reluctance to plug in is driven by their preference for reality? Some—Sumner (1996, p. 95) for instance—have pushed this line of argument against the experience machine<sup>2</sup>:

How do we know that the technology is foolproof? What happens if there is a power failure? . . . In order to isolate the philosophical point which the experience

machine is meant to illustrate, we have to control for boundary conditions by supposing that all the risks have somehow been neutralized. But this is very difficult to do, since we know that in real life we cannot eliminate all possible malfunctions and screw-ups. For the thought experiment to yield any results at all we must therefore imagine ourselves in a world quite alien to our own—and who knows what we would choose in a world like that? (Sumner, 1996, p. 95)

Still, many philosophers are unmoved by such concerns. They maintain that, even if we manage to eradicate all these misgivings, people will remain reluctant to plug in. In response to these worries, for instance, Matthew Silverstein claims that “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, pp. 284–285).

Perhaps one could try to come up with a formulation of the thought-experiment that takes care of all these potential distractions, and it may be possible that with such a formulation people would then decide not to plug in *solely* based on their preference for reality. But this is not at all obvious, 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 your entire life; 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? If you agree with Silverstein, you may think so. Indeed, Tännsjö reports having had this very conversation with Bengt Brülde, the latter claiming 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, p. 95; but see also Brülde, 2007).

Tännsjö remained unconvinced, and I again share his skepticism. It is not at all clear to me that people would necessarily prefer reality when facing a brain-in-a-vat situation. I am not even sure whether this piece of information would affect, *per se*, their judgments on their own happiness or well-being. In order to shed light on this issue, I decided to study people’s reaction to this last, controversial claim. 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. Participants were undergraduate students from UNC 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 out of 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 you directly. 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; someone else was. 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 maximum 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 1). 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

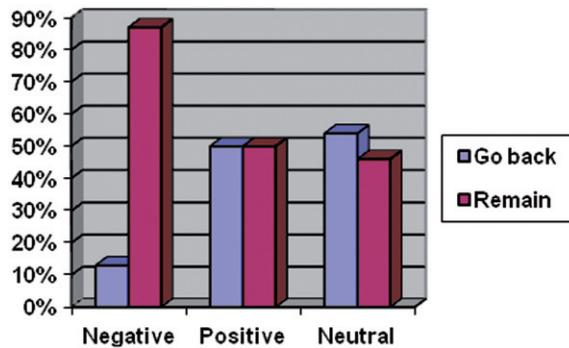


Figure 1 Percentage of responses for *Negative*, *Positive*, and *Neutral* vignettes.

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.<sup>3</sup>

Taken at face value these results show that the variable “contact with reality” is not as critical as the original version of the thought-experiment suggests. 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 all the weight on reality alone, 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, viz. 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.<sup>4</sup>

Still, some may contend that the results of the *Neutral* vignette still show that most people prefer reality. After all, the study showed that the majority of participants (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 main factor 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

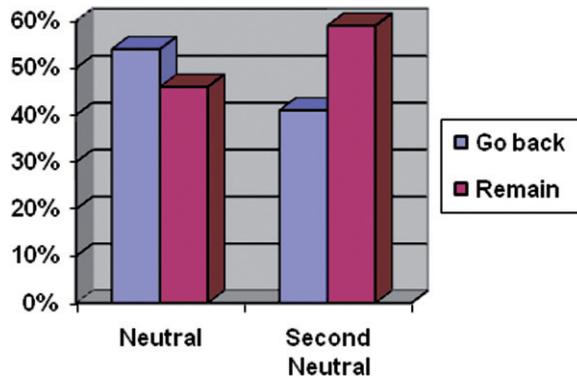


Figure 2 Percentage of responses in *Neutral* versus *Second Neutral* vignettes.

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 with no previous exposure to philosophy. As before, they were asked to read a vignette, circle one of two options, and explain their choice. They all received the same vignette. This *Second Neutral vignette* was identical to the *Neutral* vignette used in the first study, except for the ending part which I reproduce italicized here:

*Second Neutral vignette:*

[...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

Disconnect<sup>5</sup>

With the *Second Neutral* vignette the results shifted strikingly (figure 2): 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 explained 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 I 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 only operant variable here.

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 theory.<sup>6</sup>

### 3. The Experience Machine and the Status Quo Bias

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 three or maybe even four 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 & 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, Knetsch, & Thaler, 1991, 197–198). The status quo bias has been widely tested experimentally. In a famous experiment, Knetsch (1989; see also Knetsch & 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 & Zeckhauser, 1988 for a research study on the status quo bias and health insurance plans; see also Gilbert, 2007 for the influence of the status quo bias and related phenomena in our affective forecasting).<sup>7</sup>

What does the status quo bias have to do with the results of the previous study? I suggest that what mobilizes people’s intuitive reaction against disconnecting is not solely a reflection on the nature of reality, nor their hedonistic preference for pleasure, but also 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” (to use Nozick’s terminology). Thus, in some cases, people’s reluctance to disconnect would be a manifestation of this underlying psychological phenomenon: their aversion to lose their status quo, the life they have

been living so far. Part of the explanation for why most people prefer not to disconnect after spending their life in an experience machine (as in the Negative and Second Neutral vignettes) may not have to do with the virtual character of the experience, nor with the amount of pleasure they are told they would feel, but rather with 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. And if this explanation of my variation on Nozick's thought-experiment is accurate, then it follows that people's reluctance to plug in to Nozick's original version of the experience machine may turn out to be just an effect of the same underlying psychological bias: some people may prefer to remain unplugged, not because they value reality, but because they are averse to losing their status quo. After all, as Jason Kawall once remarked, in order to get the thought-experiment's point across, an experience machine is not even needed:

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* [italics added], whether or not an experience machine is involved. (Kawall, 1999, p. 383)

In addition, prospect theory may also help us explain why some participants *actually* opted for the change—i.e., why some participants were willing to disconnect—while others did not. 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 at the reference point by making it asymmetric (figure 3). As a result, when the aversion to abandon the status quo is taken into account, the slope of the value function changes abruptly at that precise point, becoming steeper for losses than for gains. For our purposes, this phenomenon 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 they are in 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

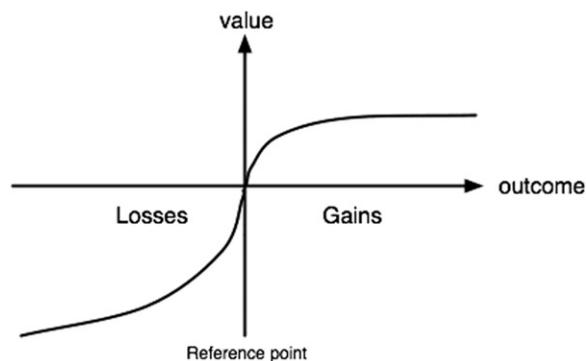


Figure 3 Asymmetric value function in prospect theory (drawing by Marc Oliver Rieger).

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 (i.e., choices in which at least one of the outcomes is uncertain) depending on whether they frame a certain outcome as a potential loss or as a potential gain (Tversky & Kahneman, 1991). If a person frames a potential outcome as a loss, then she will be far more willing to engage in risky behavior in order to avoid that probable loss than she will be willing to take risks to improve her status quo. Similarly, if she frames a potential outcome as a gain relative to her current situation, she will engage in risk-averse behavior in order to secure such gain.

Now, if the implementation of the model I am suggesting is correct, then the following interpretation of the results may become available: if one assumes that participants saw disconnecting as a risk, then those who saw remaining connected as a loss engaged in the risky behavior of disconnecting in order to avoid such a loss. Conversely, those who saw remaining connected as a gain exhibited risk-averse behavior by staying connected and avoiding thus the risk of disconnecting. Another way of making the same point is to say that for risk-averse people the shape of the curve is less concave in the upper right quadrant (i.e., gains) and more convex in the bottom left quadrant (i.e., losses) than it is for people who exhibit a risk-seeking behaviour.<sup>8</sup> As a result, due to a difference in the slope of the curve among participants, the same outcome may elicit risk-averse preferences in one group but risk-seeking ones in the other. This is particularly clear in the Positive scenario, where 50% wanted to remain connected and 50% wanted to go back to a life in Monaco. Notwithstanding such a tantalizing offer, and the potential gain it afforded, half of the participants preferred to remain at their status quo. For them, abandoning their own life is deemed to be worse than accepting a potentially better yet alien one, regardless of whether the newly offered life is real or not. As it turns out, some participants in the Positive vignette provided this sort of reason for not wanting to go back to reality: they 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 virtual nature of those experiences.

Why anyone would frame the alternative of remaining connected to an experience machine as a loss or as a gain relative to his or her status quo, remains an open question. One possibility is that people assign value to a whole set of different things, in addition to the status quo. This set presumably includes values like authenticity and pleasure, among others. For some people, the added value of the elements in this set outweighs the value of the status quo, while for others it does not. That may determine who goes for the risky outcome and who prefers to avoid it. Another possibility is that people's reference point changes as a function of the emotional or cognitive state they find themselves in when facing a particular risky decision, like that of disconnecting from the experience machine. Varying the conditions under which a person assesses her reference point may affect the shape of the curve and,

consequently, her behavioral preferences. After all, framing is a two-stage process in which the agent starts off choosing a possible outcome as the reference point, and then assesses other possible outcomes as losses or gains relative to it. Since there are many factors that could influence the framing process when participants face decisions like those suggested by the studied thought-experiment, there are many ways in which their preferences may fluctuate. These, I think, are all interesting questions worth investigating, but we should get to the philosophical implications of the reported results.

#### 4. Philosophical Implications and Final Remarks

Although the experience machine has received a substantial amount of attention in the recent philosophical literature, most commentators have accepted the assumption that our intuitive reaction to this thought-experiment is due to the fact that we value reality over a virtual life (even if such life is pleasurable). If this claim were true, one would expect that after modifying Nozick's original thought experiment to a backward-looking experience machine—that is, one in which people are asked whether or not they would disconnect had they learned that they were living a virtual life—most of us would be motivated to disconnect and go back to our real lives rather than remaining connected in a virtual one. However, contrary to this prediction, the results of the above study suggest otherwise. Although some people are willing to disconnect from a backward-looking experience machine—presumably because they do, in fact, value reality—most people prefer to remain connected. Interestingly, it seems as though people's reluctance to disconnect from a life in which they are already connected to an experience machine does not depend solely upon their preference for a pleasurable life over a life in the real world either, as some hedonists may have it. This invites us to move beyond the pleasure/reality dichotomy that has dominated the literature on the experience machine. For 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, even *before* assessing other valuable things like pleasure and contact with reality. Moreover, the reported results also suggest that if people are inclined not to disconnect from the backward-looking experience machine as a result of a psychological bias toward remaining in the status quo, then it is likely that (at least some) people may be choosing not to plug in to the original experience machine thought-experiment as a result of the same psychological bias. And if this is the case, then we have reason to believe that people may be choosing not to plug in to the experience machine because of their aversion to relinquish their status quo, and not—as Nozick and many others think—because they prefer to be “in contact with reality.”

I think this possibility should make us question the way in which the intuition elicited by the experience machine thought-experiment has been traditionally put to use in the philosophical literature. In general, philosophers have taken Nozick's suggestion at face value. That is to say, they usually assume that people's intuitive

reaction against plugging in to the experience machine give us reason to believe that they do so because they find value in being in touch with reality. However, the results of the present study suggest a second alternative: they may be reluctant to plug in simply because they value their status quo, the life they have been living so far “from the inside,” regardless of the fact that such life may be completely virtual. But if this second account also explains why people may be having such intuitive reaction, then it is unclear whether the experience machine thought-experiment gives us sufficient reason to believe that people truly value being in touch with reality.

Someone may contend that we need not be worried about this possibility on the grounds that not all intuitions are created equal, and that we only have to take into account *some* of the intuitive reactions elicited by the thought-experiment. More precisely, one could say that the intuitions swayed by the status quo bias are *irrational*, and that only intuitions elicited by the considered judgments of rational people should be taken as good reasons not to plug in.<sup>9</sup> Presumably these people may be choosing not to plug in solely from their preference to being in touch with reality.

I am suspicious of this claim on two grounds. On the one hand, it isn't clear that by the mere fact of being swayed by the status quo bias an intuition ought to be considered tout court irrational. If I was offered the possibility of exchanging my newly born but defective baby for another perfectly normal one, I would be inclined to say that it would be irrational—let alone, immoral—for me to accept the deal, even if I have no reason for keeping the baby other than the fact that she is mine. Or consider people that prefer to remain in loveless marriages because changing their status quo looms larger than staying together. Are they being irrational? Not necessarily. When it comes to morality, many things we value we just don't value as commodities. When these things feature in our decisions, certain choices that may seem as irrational from an economic point of view—that is, as diverging from the behavior of ideal rational consumers—may seem perfectly rational from a moral, or even a psychological, perspective. My status quo, i.e., my own life as I have been living it “from the inside,” seems to be precisely a case in point.

On the other hand, I am not sure that Nozick's thought-experiment holds as much water even if we only accept the evidence provided by the considered intuitive judgments of highly rational people—like well trained philosophers—while excluding the knee-jerk intuitive reactions of the folk. If one takes the point of Nozick's thought-experiment to be a psychological one—that is, as attempting to show the fact that people, in general, value reality (even over virtual pleasure)—then discounting psychological counterevidence that does not conform to the theory because it comes from a certain population, seems unduly arbitrary. Perhaps a good way of avoiding such arbitrariness is by offering an argument as to why the knee-jerk intuitive reactions of the folk shouldn't be accounted for when philosophers talk about the intuitions elicited by the experience machine thought-experiment. I am afraid, however, that even if such an argument was plausible, we still have to face the problem of explaining the divergent intuitions the experience machine elicits from philosophers themselves. Consequently, I believe that those who accept Nozick's explanation of the alleged intuitions need to take into account that people's

preferences may not at all be motivated by the fact that they care about reality—they may just be the product of our human tendency toward the status quo. However, in so doing, they will also have to seriously entertain the possibility that this very same evidence may diminish the credibility of the experience machine as a reliable thought-experiment.

Let me conclude, then, with two brief 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. 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 may benefit from taking into account empirical data on human psychology when talking about certain problems, like people's alleged preference for reality in addition to pleasure.<sup>10</sup>

### Acknowledgments

This paper benefited enormously from the comments of my audiences at the *Work in Progress* series at UNC-Chapel Hill, the *Workshop on Subjective Measures of Well-Being and the Science of Happiness* at the University of Alabama, Birmingham, and *The 34th Annual Meeting of the Society for Philosophy & Psychology* in Philadelphia. Special thanks are also due to Erik Angner, Daniel Dennett, Daniel Gilbert, Adrienne Harris, Bryce Huebner, Edouard Machery, Eric Mandelbaum, Jesse Prinz, David Ripley, Dylan Sabo, Elanor Taylor, Valerie Tiberius, and (very specially!) Joshua Knobe.

### Notes

- [1] Nozick's exact meaning of "living in contact with reality" is unclear. He seems to offer three possible alternatives: (1) living in contact with reality may mean *doing* certain things, as opposed to merely having the experience of virtually doing those things; (2) it may mean *being* a certain way, insofar as (allegedly) there is nothing one can *be* in the virtual world of the experience machine; or (3) it may mean being in *actual* contact with reality, as opposed to being in contact with a simulated or virtual reality (1974, p. 43). At the foundation of these possible alternative accounts rests a simple contrast between virtual and non-virtual experiences. For the purposes of this paper, I understand valuing or caring about "living in contact with reality" in terms of this contrast. That is, the sort of experience—or contact—with reality that people supposedly care about differs essentially from the experience—or contact—a person could have in the experience machine because the latter is virtual while the former is not.
- [2] 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, p. 104).

- [3] The difference among conditions was statistically significant:  $\chi^2(2, N=72)=10.6, p=.005$ .
- [4] 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 could 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. This is certainly possible. However, as I mentioned at the beginning, the point I am trying to make is independent of whether there is or not such an example. My purpose is to move beyond the pleasure/reality dichotomy by suggesting that what may be mobilizing our reaction to the experience machine thought-experiment is the fact that, in addition to what they want, people care for what they have.
- [5] Eric Mandelbaum and David Ripley 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.”
- [6] 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 accept or refuse to genetically enhance their cognitive abilities. I want to think of the present paper as contributing to this tradition.
- [7] 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 & 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, p. 194).
- [8] Mutatis mutandis, one could explain people’s diverse intuitions toward Nozick’s original thought-experiment along similar lines. If one sees connecting to an experience machine as a risk, one can either frame the outcome of not connecting as a loss and thus opt for plugging in, or else frame it as a gain and opt for remaining unplugged. Many thanks to Adrienne Harris and Edouard Machery for pressing me to be more explicit on this point.
- [9] I want to thank Valerie Tiberius for the comments that gave rise to this argumentative line.
- [10] After I finished the final draft of this paper, I read Adam Kolber’s excellent piece “Mental Statism and the Experience Machine” (1994). Those readers interested in the line of argument I developed here will find Kolber’s paper illuminating and inspiring.

## References

- Bostrom, N., & Ord., T. (2006). The reversal test: Eliminating status quo bias in applied ethics. *Ethics*, 116, 656–679.
- 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*. Retrieved Spring, 2008, from <http://plato.stanford.edu/entries/well-being/>.
- Gilbert, D. (2007). *Stumbling on happiness*. New York: Knopf.

- Horowitz, T. (1998). Philosophical intuitions and psychological theory. *Ethics*, 108, 367–385.
- Kahneman, D., Knetsch, J. L., & Thaler, R. H. (1991). Anomalies: The endowment effect, loss aversion and status quo bias. *Journal of Economic Perspectives*, 5, 193–206.
- Kahneman, D., & Tversky, A. (1984). Choices, values and frames. *American Psychologist*, 39, 341–350.
- Kawall, J. (1999). The experience machine and mental state theories of well-being. *The Journal of Value Inquiry*, 33, 381–387.
- Knetsch, J. L. (1989). The endowment effect and evidence of nonreversible indifference curves. *American Economic Review*, 79, 1277–1284.
- Knetsch, J. L., & Sinden, J. A. (1984). Willingness to pay and compensation demand: Experimental evidence of an unexpected disparity in measures of value. *Quarterly Journal of Economics*, 99, 507–521.
- Kolber, A. J. (1994). Mental statism and the experience machine. *Bard Journal of Social Sciences*, 3, 10–17.
- Nozick, R. (1974). *Anarchy, state and utopia*. New York: Basic Books.
- Nozick, R. (1989). *The examined life*. New York: Simon & Schuster.
- Samuelson, W., & 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, 279–300.
- Sobel, D. (2002). Varieties of hedonism. *Journal of Social Philosophy*, 33, 240–256.
- 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., & Kahneman, D. (1991). Loss aversion in riskless choice: A reference-dependent model. *Quarterly Journal of Economics*, 106, 1039–1061.