Book Review

Experience Machines: The Philosophy of Virtual Worlds. Edited by Mark Silcox

Reviewed by Christopher Bartel


In Anarchy, State, and Utopia (1974), Robert Nozick offers what has become known as the experience machine thought experiment. Nozick asks the reader to imagine a machine that can provide a virtual reality experience that would be indistinguishable from reality and in which the user could enjoy the most pleasurable life imaginable. Nozick asserted that most people would find the idea of plugging into the experience machine for life to be a horror, or, as he put it, a kind of suicide. Nozick inferred from this intuition that a life spent permanently in the experience machine would be unacceptable, and that something more than mere pleasure is needed to make a life meaningful.

Mark Silcox’s edited volume, Experience Machines: The Philosophy of Virtual Worlds, seeks to reassess Nozick’s thought experiment and the intuitive response that Nozick drew from it. Given the technological advancements in computing since the publication of Nozick’s book, the possibility of an experience machine has become less remote. Indeed, while current video games do not offer anything like the multi-sense-modality experience of Nozick’s hypothetical machine, they do offer an experience that gamers have little trouble becoming deeply immersed in. Through their interactivity, video games provide a sense of agency and, perhaps more importantly, a sense of impact—that is, a sense that the player can have an impact on the world of the game, for better or for worse—and yet, it is a sense of impact that carries no real-world consequence. What often seems so jarring about Nozick’s thought experiment is that the user of the machine enjoys an experience that is free of any real-world consequence. This, I suggest, is where video games intersect with Nozick’s experience machine: video games provide experiences for free. So, a reassessment of the experience machine and our intuitive responses to it is timely.
The first question one might ask is whether Nozick’s thought experiment is one that is worthy of such treatment. Descartes’ evil demon and the trolley problem are two thought experiments influential enough to have warranted numerous books and countless articles devoted to discussion of them. In the pantheon of thought experiments, where does Nozick’s experience machine stand? While many articles have been written about Nozick’s experience machine, Silcox’s edited volume is the first devoted solely to it. What Silcox has demonstrated here with a collection of essays covering a wide and fascinating range of topics is that the experience machine is certainly rich enough to sustain much interest and is central to many pressing issues.

The thirteen contributed essays reexamine the experience machine from numerous perspectives and against a wider range of issues than Nozick likely anticipated. Some of the issues and topics covered are just what one would expect and hope for. Given the similarity between Nozick’s experience machine and Descartes’ evil demon, one should certainly expect some discussion of the metaphysical and epistemological concerns that would arise (Chapters 1, 4, and 10). Nozick offered his thought experiment as an attack on hedonistic theories of meaning and value, and many of the authors in Silcox’s book demonstrate a range of views teasing apart the subtleties of the issues (Chapters 1, 2, 3, and 11). Finally, given the obvious relationship between the virtual reality of Nozick’s experience machine and the virtual worlds of many video games, one would hope for some discussion of Nozick’s machine applied to such a timely contemporary topic (Chapters 7 and 8). In addition, Silcox’s book offers a number of chapters that venture beyond these expected topics: the epistemic status of learning and experiencing from works of fiction (Chapter 5), the nature of friendship within social media (Chapter 6), theological concerns over the problem of evil in the design of an experience machine (Chapter 9), the use of experience machines for palliative care (Chapter 12), and the political status of the citizen living in a virtual world (Chapter 13).

The book is divided into three broad sections. Section 1 (“Virtual Experiences and Human Well-Being”) offers four chapters that each examine what well-being might consist of in an experience machine. In Chapter 1, Peter Ludlow (“Cypher’s Choices: The Variety and Reality of Virtual Experiences”) spells out many different versions of an experience machine. He examines the various levels of agency, personal risk, and possibilities for harm to the user, and argues that, in some configurations of an experience machine, the user’s actions would be genuine in both moral and meaning-giving senses.

In Chapter 2, Daniel Pietrucha (“Intuition and Imaginative Failure”) notes that it is easy to imagine lives that are so wretched and unhappy that it would be irrational for the individual facing such a life to resist plugging into the experience machine; from this point, Pietrucha goes on to argue that, if we imagine an experience machine that could offer similar increases in well-being to any individual, then it may be irrational for any individual to resist plugging in.
Nozick’s thought experiment is often presented as an argument against hedonistic views of value; however, Emiliano Heyns and Johnny Hartz Søraker (“Give Me the Confidence”, Chapter 3) defend a more nuanced conception of hedonism, which they argue avoids Nozick’s objection. Heyns and Søraker defend a version of Fred Feldman’s “intrinsic attitudinal hedonism”, which holds that hedonism should not be understood as valuing merely sensory pleasures, but instead as valuing attitudinal pleasures—that is, an attitude of “taking pleasure in” something. With this account in mind, Heyns and Søraker argue that the problem with taking pleasure in the experiences of a persistent virtual reality machine is simply that the individual may suffer a lack of confidence in the reality of their experiences, and that such lack of confidence would deter the individual from enjoying the appropriate attitudinal pleasure.

Finally, John Cogburn (“Ceci n’est pas une cuve”, Chapter 4) rejects the concern that we might in fact really be living the lives of envatted brains by considering Hilary Putnam’s version of the thought experiment as an inclosure paradox.

Section 2 (“Real-World Experience Machines?”) offers four chapters that each examine problems and concerns that would arise from technologically mediated experiences. In Chapter 5, E. M. Dadlez (“Virtual Reality and ‘Knowing What It’s Like’: The Epistemic Up Side of Experience Machines”) examines the possibilities that experience machines might offer to develop moral understanding through empathy in a way similar to what we currently find in narrative fiction. Her conclusion is more positive than the standard Nozickian reply. While it might be unacceptable to plug into the experience machine for life, plugging in only occasionally might offer significant rewards. Just as we turn to narrative fiction to experience what it is like to inhabit another point of view, experience machines with their greater power of immersion may be an ideal technology for the task.

In Chapter 6, Alexis Elder (“Figuring Out Who Your Real Friends Are”) examines the nature of friendship in an age of social media and argues that, while robot-friends fail to constitute genuine friends, the relationships that one might develop and sustain with others via social media platforms like Facebook can facilitate genuine friendship rather than a mere appearance of friendship.

In Chapter 7, Grant Tavinor (“Welcome to the Achievement Machine: Or, How to Value and Enjoy Pointless Things”) examines the value of virtual experiences. Gamers spend countless hours immersed in the virtual worlds of video games. Is all that time spent wasted on something frivolous? When playing games, players attempt to achieve an end that is in itself admittedly pointless—whether that end is crossing the finish line first, or jumping over the highest bar, or crafting the best suit of armor in Skyrim. Moreover, the achievements that one pursues in games are not only pointless but also difficult. Yet, pursuing difficult goals can provide meaning for the player because the pursuit provides the player with something to strive toward. As Tavinor says, “the striving itself
can more effectively become the locus of value” (110). However, this does not mean that all virtual achievements are equally valuable. What matters is that difficult goals are willfully pursued. If the user of Nozick’s machine is guaranteed to win, then the machine offers no achievement at all. Video games, insofar as they are difficult, can provide meaning pursuits in a way that the mere experience of achievement cannot.

Finally, in Chapter 8, Stefano Gualeni (“Virtual Weltschmerz”) considers a “paradox at the core of Nozick’s thought experiment”: namely, whether the user of the experience machine can remember life unplugged. If one can remember the unplugged world, then likely one would experience a sort of nostalgia for the real world, or perhaps even a sort of “world-weariness”—a feeling that the virtual world is predictable and lacks meaningful depth. No technology—experience machines included—can solve our existential problems. The best that any technology can offer is a tool to explore new possibilities and new ways of being, and yet in the end we will find the same old struggles with meaning and being in the virtual world.

The final section (“Experiential Design: Problems and Prospects”) offers five chapters that examine the subtle differences in the problems that arise from various ways of designing an experience machine. In Chapter 9, Brendan Shea (“The Problem of Evil in Virtual Worlds”) considers an interesting variation on the problem of evil for virtual worlds. A virtual world might offer the user prospects for meaningful activity, Shea suggests, if the world were populated with human-like artificially intelligent beings who could suffer and for whom the user could take an interest in ameliorating their suffering. But this would require the designer to create a world with suffering. It seems morally wrong to create a world with significant amounts of suffering, particularly when the designer has the power to create a world without suffering. So, we are left with a dilemma: to create a virtual world worth engaging with, one must act in ways that are morally wrong.

In Chapter 10, James McBain (“Epistemic Lives and Knowing in Virtual World”) argues that, if we accept an information-based epistemology, then users of the virtual world are in fact not epistemically worse off than the unplugged inhabitants of the real world. Traditional Nozickian worries about the epistemic situation of a plugged-in user of the experience machine, like the traditional Cartesian worries about the individual who has been captivated by the evil demon, proceed from a worry that the contents of one’s beliefs and experiences do not refer to reality. While plugged in to an experience machine, the user is cut off epistemically from the real world; yet the user gains epistemic access to information about a virtual world. That information, McBain argues, is just as epistemically valid as the information that unplugged individuals enjoy in regard to the real world. Issues of realism aside, an envatted individual is not epistemically as badly off as Nozick might lead us to believe.
In Chapter 11, Michael LaBossiere (“Digital Tears Fell from Her Virtual Eyes: Or, the Ethics of Virtual Beings”) considers two arguments in defense of the moral status of virtual beings. First, LaBossiere argues that the virtual beings one would encounter in the experience machine are beings that can serve and emotionally impact the user; therefore, by developing the Kantian account of how moral agents can have indirect moral obligations to non-rational beings like animals (on Kant’s account), agents can have indirect moral obligations to virtual beings too. Second, LaBossiere suggest that virtual beings might even have a moral status of their own. If the virtual beings of Nozick’s experience machine are sophisticated enough to pass the Turing Test, then they may be “human enough” to warrant their own (direct) moral status.

In Chapter 12, Dan Weijers and Russell DiSilvestro (“The Morality of Experience Machines for Palliative and End-of-Life Care”) suggest some reasons why one ought not to resort to the experience machine at the end of life. Weijers and DiSilvestro suggest that a person who has plugged into the experience machine in order to avoid intense, persistent, and incurable pain would (quite reasonably) not wish to unplug again; however, this would mean that such an individual would be cut off from their real-world friends and family. Plugged in for the remainder of one’s life, one would have effectively died already from the perspective of the people one leaves behind. Weijers and DiSilvestro ask whether the avoidance of pain would be worth more to the sufferer than all those relationships, achievements, and goals that seem to hold so much meaning for us in life. Palliative uses of the experience machine may be morally problematic, but Weijers and DiSilvestro suggest that they might be permissible in some severe cases.

Finally, in Chapter 13, Steven Montgomery (“The Experience Machine and the End of the World (As We Know It)”) considers whether the individual who has plugged into the experience machine retains any duties to the state that they might be shirking while envatted. At the heart of this issues are the questions: what sort of citizen-status does the envatted individual retain, and in what ways might the envatted citizen still be dependent on the state? If the individual’s dependence on the state is quite minimal (e.g. one is plugged in to a privately owned and self-sustaining solar-powered machine), then one might be free from nearly all duties to the state. Additionally, one’s duties to the state will depend on the kind of state—one would be shirking one’s duty to vote in a fully participatory democracy.

The essays collected in Silcox’s book present a range of views, both positive and negative—though I often found myself siding with the pro-plugging in crowd, sometimes to my own surprise and dismay! For instance, Pietrucha’s slippery-slope argument (Ch. 2) is surprisingly persuasive (I say “surprising” because I am usually skeptical of slippery-slope arguments). We may intuitively feel that a life so wretched would be better off in the experience machine because no one gains anything by suffering from abject misery and the potential increase in happiness would be substantial. But, then, what
do I gain by suffering from a bad headache? Personally, I often suffer from debilitating headaches, ones that do not respond to painkillers; so, perhaps I should plug in? Gaining relief from my headaches would provide a similarly substantial increase to my happiness. Weijers and DiSilvestro (Ch. 12) argue that palliative uses of experience machines should be reserved only for extreme cases. Yet I now feel quite convinced that the palliative use of an experience machine at the end of one’s life or to avoid persistent and intense pain would be happily tolerable. Those of us who have never suffered intense pain should be very wary of our ability to decide whether some pain is intense enough.

Interestingly, one thread that can be traced between each of the essays is a concern over what sort of experience machine we envision. As Ludlow notes (Ch. 1), many of the issues that arise from our thinking about the experience machine depend strongly on the nature of the machine. When envisioned one way, certain issues and concerns arise; when envisioned another way, some issues disappear while others emerge. For instance, consider the differences between a networked experience machine—one in which the user can interact with other real-life users who are also plugged into the machine—and a non-networked, solipsistic experience machine—one in which the user is the only real-life inhabitant of the virtual world, cut off from contact with any other real-life user. The user of the networked machine may still retain some genuine moral obligations toward other real-life users (i.e. because they meet Kant’s requirement of being rational beings) while it is debatable whether the user of a solipsistic machine owes any moral obligation toward the virtual beings that are encountered within the machine (Cf. Ch. 11).

Moreover, the relationships that one develops with other real-life users in the networked machine have a legitimacy and genuineness that are likely lacking in the relationships that one develops with the computer generated artificially intelligent beings of the solipsistic machine (Cf. Ch. 6). Consider another variation on the experience machine, one in which the envatted user is not fully cut off from the real world, where the boundary between the virtual world and the real world is more porous. Perhaps living my happy life in the experience machine, I can spend some time each morning in the virtual world reading a newspaper that keeps me informed about real-world events. In this sort of machine, a user might be able to keep up to date on political affairs and vote in the same way that citizens living abroad are able to vote in some countries (Cf. Ch. 13); or a user who has plugged into the machine for life for reasons of palliative care (Cf. Ch. 12) might still be able to receive phone calls or Skype with loved ones who are living outside of the machine. Surely a machine as sophisticated as the one envisioned here can handle a phone call. This might not provide the best means of pursuing meaningful relationships with those who matter the most to us, but it would seem to be no worse than the reality that many people face of having to sustain long-distance relationships with others often for great periods of time.
In sum, Silcox’s collection is a fascinating read. While philosophers working in many different sub-disciplines would find much of interest in this book, there are a number of chapters that would be of interest to scholars thinking about video games. In particular, Dadlez’s (Ch. 5) discussion of knowing-what-it-is-like intersects with issues about what one could experientially gain from a video game; Elder’s (Ch. 6) discussion of friendship in virtual worlds is directly applicable to current social games; LaBossiere’s (Ch. 11) discussion of our ethical duties to virtual beings is plainly relevant to debates over the ethics of video game violence; and Tavinor’s (Ch. 7) account of the value of virtual experiences offers a convincing (and reassuring!) account of why we shouldn’t worry about spending too much time playing video games. The essays are all clear, engaging, and surprisingly (and thankfully) lacking in technical jargon. The book would work as an excellent focus of discussion for an undergraduate class. As the technology becomes a greater possibility, Silcox’s book would serve as an excellent starting point for the many discussions that would need to be had.