Philos. Technol. (2014) 27:561–579
DOI 10.1007/s13347-013-0137-4

RESEARCH ARTICLE

Against Brain-in-a-Vatism: On the Value of Virtual Reality

Jon Cogburn · Mark Silcox

Received: 11 April 2013 / Accepted: 7 October 2013 / Published online: 16 November 2013

Abstract The term "virtual reality" was first coined by Antonin Artaud to describe a value-adding characteristic of certain types of theatrical performances. The expression has more recently come to refer to a broad range of incipient digital technologies that many current philosophers regard as a serious threat to human autonomy and well-being. Their concerns, which are formulated most succinctly in "brain in a vat"-type thought experiments and in Robert Nozick's famous "experience machine" argument, reflect a fundamental misunderstanding of the way that such technologies would probably have to work. They also considerably underestimate the positive contributions that virtual reality (VR) technologies could make to the growth of human knowledge. Here, we examine and critique Nozick's claim that no reasonable person would want to plug into his hypothetical experience machine in light of a broadly enactivist understanding of how future VR technologies might be expected to function. We then sketch out a tentative theory of the phenomenon of truth in fiction, in order to characterize some of the distinct epistemic opportunities that VR technologies promise to provide.

Keywords Virtual reality · Brain in a vat · Nozick · Artaud · Enactivism · Truth in fiction

Antonin Artaud first coined the expression "virtual reality" in the context of a subtle and fascinating analogy between theatrical performances and alchemical manuals. "There is a mysterious identity of essence," Artaud wrote,

> between the principle of the theater and that of alchemy. For like alchemy, the theater… is developed from a certain number of fundamentals which are the same for all the arts and which aim on the spiritual and imaginary level at an

J. Cogburn
Department of Philosophy and Religious Studies, College of Humanities & Social Sciences, Louisiana State University, 102 Coates Hall, Baton Rouge, LA 70803, USA

M. Silcox (✉)
Department of Humanities and Philosophy, University of Central Oklahoma, 100 N. University Drive, Edmond, OK 73034, USA
e-mail: msilcox@uco.edu

> efficacity analogous to the process which in the physical world actually turns all matter into gold (Artaud 1958, p. 48).

In the performance of a play, he suggests, "the characters, objects, images, and in a general way all that constitutes the *virtual reality* of the theater" are directed toward bringing about a reaction in the "spirit" of the properly attuned audience member that is analogous to the transmutation of lead into gold (Artaud 1958, p. 49).

At the risk of draining some of the color from his language, Artaud's view might be paraphrased as follows: the central aim of both theater and alchemy is to reveal the existence of a latent, inner tendency of objects to transform into something more valuable than they first appear. However, since the personal computer revolution of the mid-1980s, the term virtual reality (hereafter VR) has come to mean something quite different from what Artaud originally intended.

In his 1998 book, *Virtual Realism*, Michael Heim takes considerable pains to distinguish what he takes to be canonical uses of the term from "loosely associated meanings and spin-offs." Understood in its "strong sense," the expression virtual reality refers to the following:

> an immersive, interactive system based upon computable information… Immersion comes from devices that isolate the senses sufficiently to make a person feel transported to another place. Interaction come from the computer's lightning ability to change the scene's point of view as fast as the human organism can alter its physical position and perspective… Constantly updated information supports the immersion and interactivity, and to rapidly update the information, computers are essential (Heim 1998, pp. 6–7).

Modern computers are already able to provide enormously complex responses to user input and thus seem to foreshadow future technologies that will eventually outpace the human sensorium itself. Such machines might one day not only be able to present us with plausible representations of the world around us but, additionally, to more or less instantaneously modify those representations based upon how we respond to them.

Heim's and others' pleas for precision in usage have done little over the past few decades to stop journalists, salesmen, and aspiring futurists from characterizing everything from 3D movies to hair replacement products as engines of VR. But ironically, along with this increased laxity of usage, alarmism has grown. Many observers of contemporary culture fear that the phenomenon of VR represents a real threat to human well-being. These concerns all to some extent have their roots in an epistemic worry raised by the prospect of VR technology.

The original Cartesian skeptical hypothesis involved the possible existence of a supernatural Evil Genius manipulating our perceptions so that all, or perhaps most, of our beliefs are false. But this, Tony Brueckner observes, allowed naturalistic philosophers to dismiss such concerns too easily:

> Materialists who hold that the mind is a complex physical system deny that it is possible for there to be an Evil Genius world, since, on their view, your mind could not possibly exist in a matterless world. Accordingly, a modern skeptic will have us consider an updated skeptical hypothesis that is consistent with materialism. Consider the hypothesis that you are a disembodied brain floating

> in a vat of nutrient fluids. This brain is connected to a supercomputer whose program produces electrical impulses that stimulate the brain in just the way that normal brains are stimulated as a result of perceiving external objects… (Bruekner 2012, p. 2)

In this manner, contemporary technological innovations and science fictional speculation about VR render live a hypothesis that might otherwise be regarded as only realizable in a very faraway possible world and hence perhaps only of concern to philosophers. Furthermore, as Charlie Gere has observed, the image of the brain in the vat has acquired its own irreducible cultural resonance:

> it is also a persuasive representation of how our social relations are mediated through technologies of communication, from telegraphy and telephony through to email and the world wide web. In turn it is in the context of such mediation that our contemporary understanding of the brain has been developed. (Gere 2004, p. 365)

Brain-in-a-vatism is an understanding of VR that encourages two related forms of philosophical skepticism, one epistemic and the other axiological. The first type is the traditional sort of worry about our knowledge of the external world. The second type is the concern that no genuine value could attach either to a life lived in VR or to any of the objects that populated a virtual environment. We will begin our exploration of these two sources of concern by looking at Robert Nozick's infamous "experience machine" argument against hedonism, which in many ways represents a paradigmatic symptom of the post-Artaudian moral reversal. We think the type of *axiological* brain-in-a-vatism that is presupposed by Nozick's thought experiment gets the metaphysics of VR disastrously wrong. But the concerns we will bring to bear on Nozick's argument are not sufficient on their own to undermine *epistemic* brain-in-a-vatism. We will therefore proceed to follow up our discussion of the experience machine by describing some positive reasons why the development of greater and greater VR should be expected to radically increase, rather than decrease, the sum of human knowledge. This will necessitate some sustained philosophical reflection upon a few of the surprising ways that works of fiction often teach us truths.

Our ultimate aim in pursuing these two distinct lines of enquiry will be to demonstrate some ways in which the contemporary attitude of alarmism about VR is vastly overblown. We think that there remains much wisdom in Artaud's optimistic attitude toward the prospect of VR, even though his work significantly predated the widespread use of digital technologies.

1 VR in Popular Culture and Philosophy

The anti-Artaudian alarmist tendency is nowhere more pronounced than in recent popular science fiction (e.g. films such as the *Matrix* series, *Existenz*, *The Thirteenth Floor*, and *Avalon* and novels like Jeff Noon's *Vurt* and Ernest Cline's *Ready Player One*). All of these works depict heroic protagonists who must choose between direct contact with "reality" and the allure of a virtual environment. Although the later Matrix films contain many ethical ambiguities, it is absolutely clear in the first movie

that Neo's selection of the red over the blue pill is meant to be viewed as a heroic act. There does seem to be an informal consensus among both science fiction authors and fans of the genre that any serious commitment to living in a virtual environment comes at a very steep cost to one's basic humanity.

Philosophers who have written about the subject generally share this alarmist attitude (McMahon 2002; Griswold 2002). In *Surviving the Age of Virtual Reality*, Thomas Langan acknowledges that realistic works of literary fiction are similar to VR in that they can summon sensory images in a reader so that she feels like an inhabitant of a "verbally aroused imaginative space." It would be mere philistinism, he thinks, to view this psychological effect as a reason to never read fiction at all. But he argues that such a "space" is never summoned by mere language with the "realism of a delusion or hallucination." By contrast, he argues that genuine "virtual reality effects" are always to some extent experienced *involuntarily* and therefore should be viewed as "symptoms of mental illness" (Langan 2000, p. 125).

In spite of his professed nihilism, Jean Baudrillard achieves an even more panicky tone in *Simulacra and Simulation* by associating a wide variety of cultural artifacts that approximate the effects of VR with what he refers to as the "disappearance" of the self. "[N]ow the media," he says, "are no longer a stage where something is played, they are a strip, a track, a perforated map of which we are no longer even spectators: receivers" (Baudrillard 1994, p. 160) "What are we to do," he goes on to ask,

> when… nothing really ever takes place, since everything is already calculated, audited, and realized in advance (the simulacrum preceding the real, information preceding the event, etc.)? (Baudrillard 2000, p. 37)

Baudrillard seems to regard the increased use of VR technologies as being fundamentally inevitable. We are irredeemably destined to become "fascinated by all forms of disappearance, of our disappearance. Melancholic and fascinated, such is our general situation an era of involuntary transparency" (Baudrillard 1994, p. 160).

Other thinkers offer criticisms of VR as part of broader agendas in cultural criticism. Perhaps the most pervasive such agenda is the deep suspicion of "mediated" (see Mander 1978, p. 24) or "synthetic" (see Virilio 1994, p. 59) perceptual experience expressed by authors such as Neil Postman, Jerry Mander, and Paul Virilio. This sort of suspicion is usually manifested in connection with some particular medium of representation, e.g., photography (for Virilio) or television (for Mander and Postman), but the tone of their invective is often reminiscent of Plato's dismissal of the entire realm of physical sensation as something that "both is and is not" (Plato 1992, p. 155). Some interpreters of the later thought of Martin Heidegger have also seen evidence that he would have regarded VR and its widespread use as symptoms of a malign and hubristic "productionist metaphysic" that treats nature in general as a something "constructed" (Coyne 1994, p. 68).

All of these thinkers concentrate more or less exclusively upon the illusory or deceptive possibilities of VR technologies. They have almost nothing to say about the capacity of these technologies to *illuminate* aspects of reality that are normally concealed from ordinary perception. Their epistemic and axiological perspectives on VR are therefore diametrically opposed to Artaud's. Our aim in this paper will be to argue that many of their worries about VR are unfounded.

2 Axiological Brain-in-a-Vatism: Nozick's Experience Machine

For epistemologists, responding to the skeptical challenge is supposed to require showing how we are not brains in vats. If you know that you are not a brain in the vat, then you can claim to know that skepticism is false. For Nozick, if you can show that being in an experience machine is disfavorable, then you can show that hedonism—the view that pleasure is the only intrinsic good—is irrational. Instead of a supernatural Evil Genius, we are to consider:

> 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 (Nozick 1974, p. 42).

Although we think that there are plenty of good responses that the hedonist can make to Nozick's argument (see Silverstein 2000), we want to present a somewhat more internal criticism of the conclusions that he tries to draw from his thought experiment. Nozick's three more specific objections to the decision to live in the experience machine are all problematic in interesting ways.

2.1 Doing

Nozick's first claim is that "we want to *do* certain things, and not just have the experience of doing them" (Nozick 1974, p. 42). The person in the experience machine is not really writing a novel, just having the same experiences of someone who is writing a novel. Because of this (Nozick thinks), there is something ridiculous about the experiencer taking pride in having written a novel, since it was actually a product of the machine. And clearly one would rather be justified in one's pride.

What Nozick is describing might be called "passive virtual reality." People wired into the machine are receiving experiences in the same way that a moviegoer views a movie, except that the external sensations are far more realistic, and via direct stimulation of the brain, the experience machine can also present absolutely realistic inner experiences, such as the decision to pick up a pen and begin writing a novel. Attempts to describe this kind of passive VR often appear in science fiction that predates the rise of video games and the Internet. One of the most famous examples is Philip K. Dick's 1966 story "We Can Remember it For You Wholesale," upon which the movie *Total Recall* is loosely based. In both story and movie, the company "Rekall" gives people cheap vacations by implanting false memories (Dick 1966). Similarly, in Cory Doctorow's novel *Down and Out in the Magic Kingdom*, psychologists and cyberneticists redesign the Hall of Presidents ride at Disney World as a "cerebral flash-baking" device that instantaneously provides one with the sensory gestalt of what it would have been like to be Abraham Lincoln (Doctorow 2003).

Ever since the rise of video games, the norm in literary and film depictions has been a different type of VR, one that allows for interaction between player and virtual environment. One might think that this is just an interesting cultural fact about the evolution our concept of VR. But it is in fact much more philosophically significant.

Recall that the Matrix we see in the eponymous film is actually described as being the *second* such machine. Our robot overlords first tried to provide us with a paradise, but it did not take. The idea in the movie is that we just could not handle being happy. But perhaps, programming a paradise involved too much passive VR. Instead of living in happy immersive movies, people in the second matrix were allowed to make all the kinds of decisions that humans of the real 1990s did. This is not quite the whole story though. When Neo's body is unplugged from *The Matrix*, he must initially receive days of acupuncture because he has never used his muscles before. Let us, then, call this "semi-passive VR": active in that the player's brain is autonomously making choices in the virtual realm, but passive in that the player's body is doing nothing. We think that this is, in fact, how many people currently believe VR would actually work. But the depiction is radically implausible, at least if we think of VR as the result of incremental improvements of technologies already available to us that engage directly with our perceptual faculties.

As Alva Noë has demonstrated in his writings on the topic of "enactive cognition" (Noë 2006), human perception is profoundly tied to our ability to manipulate things in the environment. At the most basic level, in binocular perception, we determine how far away something is in part by the angle of the focal point, determined by the position of our eyeballs. The angle is greatest when we are cross-eyed, and as things get further away, the angle decreases as we get less cross-eyed. But how then could virtual reality truly capture depth perception, if, like Neo, we had never used our eye muscles?[1]

In a canonical experiment from 1964, Ivo Kohler had people wear lenses that reversed the light taken in by their eyes in a left/right direction. The null hypothesis of such lenses is that while wearing them you would feel your right arm extending and see your left arm extending. But this is not what happened. The lenses threw test subjects into a funhouse nightmare:

> During visual fixations, every movement of my head gives rise to the most unexpected and peculiar transformations of objects in the visual never before seen. At times, parts of figures run together, the spaces between disappearing from view; at other times they run apart, as if intent on deceiving the observer. Countless times I was fooled by these extreme distortions and taken by surprise when a wall, for instance, suddenly appeared to slant down to the road, when a truck I was following with my eyes started to bend, when the road began to arch like a wave, when houses and trees seemed to topple down, and so forth. I felt as if I were living in a topsy-turvy world of houses crashing down on you, of heaving roads, and of jellylike people (Kohler 1964).

Noë argues that the reason for this strange result is that bodily movement was untethered from the light waves hitting the eyes.

This hypothesis is greatly strengthened by what happened the longer test subjects wore the glasses. Initially, they traversed the landscape like drunken toddlers randomly trying to grab things. But they gradually reached a second stage, during which they did experience

[1] Note also that children with congenital strabismus, the inability to exert appropriate muscular control over an eye, do not see a moving landscape through the wandering eye. Instead, they develop amblyopia, which is standardly understood as the brain's being unable to process the information provided by the eye.

mirror vision: letters and numbers looking reflected, seeing the left hand extended even while feeling the right hand extend. Then, even weirder, as test subjects got better and better at navigating their mirrored environment, gestalt shifts happened such that (even though they were still wearing the glasses) when they extended their right hands, they saw their right hands extending. These two stages overlapped somewhat. Subjects who did not read very much would still see letters as reversed even while nothing else in the environment seemed reversed any longer.

Noë uses such experiments to argue that perception is nothing over and above knowledge of potential movement (the enactivist theory of perception). But we need not subscribe to enactivism to see the challenge to passive forms of VR. Test subjects were only able to perceive correctly by physically moving around in an environment. This strongly suggests that there is nothing the evil supercomputer could do to the recumbent infant Neo to get him to have the visual experiences depicted in the Matrix.

The defender of passive VR has a ready response here. Recall that, in the brain-in-a-vat thought experiment, perception is supposed to be taking place as a result of neurons firing. Since the evil supercomputer could cause the same brain waves to occur that normally take place accompanying eye movement, it could supposedly convince a brain in the vat that it was perceiving objects near and far. We know that there are "maps" of the body in the brain such that sensation in various parts of the body corresponds to neural excitation in the corresponding part of the brain. And when we move our muscles in certain ways, it feels like something, exciting these parts of the brain. So why could not a passive VR machine just excite those areas of the brain that correspond to the movement of certain muscles, along with whatever parts of the brain that fire when we feel we are doing something autonomously? Surely, this would fool the non-writer into thinking she has decided to start writing a novel!

But recent research on "muscle memory" complicates the story considerably. Consider the hypothetical guitar virtuoso Steve. Phenomenologically, it feels as if his hand somehow remembers how to play certain scales and chords. If he has not played for a while, when he initially picks up the guitar, he might sound like a novice. But if Steve and the novice play for an hour or so, by the end of the hour, Steve will be "warmed up," playing vastly better, while the novice will only be incrementally better than when they both started. The defender of passive VR is committed to the view that there is a neural explanation for all of this, and that the Matrix could therefore replicate the experience by just stimulating the correct parts of Steve's and the novice's brain. For philosophers of a certain reductionist persuasion, this kind of move would seem entirely plausible.

But it is not. Recent research on muscles (Andersen and Aagaard 2000) explains the difference between Steve and the novice in terms of protein levels in the muscles themselves. What happens is that when a muscle is immediately able to move in a certain way with a certain force, the level of the protein MHC IIA is much higher in that muscle. At the beginning of practice, the novice and Steve have similar levels of MHC IIA in the muscles used in playing an E chord. However, previously, when Steve could play the E chord much better, he had much higher levels of MHC IIA than a novice. As he got out of practice and the MHC IIA protein levels declined, they were replaced by MHC IIX proteins. MHC IIX proteins prime the relevant muscles to ramp up levels of MHC IIA. Since Steve had higher MHC IIX proteins, he was able

to warm up and regain his mastery of the E chord in the hour while the novice only got incrementally better during the same amount of time.

Since the mechanism for muscle memory is MHC IIX protein levels, there is very little reason to think that the ability to play guitar is fully recoverable from neural activity. If Steve's brain could somehow be put into the body of the novice, it would take him vastly longer to learn how to play as well as he used to. But then, there is no reason to think that mere stimulation of neurons could actually simulate the experience of playing the guitar.

One of the very best recent science fiction novels, Ernest Cline's *Ready Player One*, decisively moves away from semi-passive models of VR. At the point in the novel where the narrator has amassed enough wealth to use a state of the art VR rig, he describes his haptic suit as follows:

> The outside of the suit was covered with an elaborate exoskeleton, a network of artificial tendons and joints that could both sense and inhibit my movements. Built into the inside of the suit was a weblike network of miniature actuators that made contact with my skin every few centimeters. These could be activated in small or large groups for the purpose of tactile simulation—to make my skin feel things that weren't really there (Cline 2011, p. 192).

This is followed by a description of his retinal display, surround-sound speakers, and smell-o-vision. In addition, walking is simulated by an "omni-dimensional treadmill," which is later upgraded to a stationary hamster ball in which the player resides.

Clearly, Cline is at the front end of novelists picking up on the way video games have been moving away from the semi-passive paradigm. The user interfaces employed by the Nintendo Wii, recent versions of *Guitar Hero*, newer industrial training sims, and military simulations all force the player to engage many of the same muscles as are engaged in corresponding real-world tasks. And anyone who has ever knocked over a glass in the real world while aiming a virtual rifle knows that active VR is more immersive.

Even if enactivism does not tell the whole story about perception, its central insights into embodiment must be regarded as a constraint upon speculation about the future of VR, at least until the emergence of some radically new research paradigm into the human/machine interface.[2] This being the case, if the experience machines were capable of giving us *any* experience rather than (assuming this is possible) just pleasant, delusional beliefs about the past, then it would have to allow us to "do" things in the actual world as well as the virtual world.

[2] A referee for this journal has called our attention to the way that optogenetics might eventually bring about such a transformation. Researchers in this field have already had some success at helping lab mice to better navigate mazes (see Zackaib 2013) and at modifying the behavior of nonhuman primates (see Berdyyeva and Reynolds 2009, p. 159) by introducing microbial rhodopsins directly into brain tissue and then activating the neurons there with pulses of laser light. We remain skeptical about whether future iterations of these technologies could be used to generate the types of virtual environments that would qualify as "interactive" in Heim's sense of the term, but to the extent that they could, what we have called "passive" VR would not end up being quite such a counter-intuitive prospect.

2.2 Being

Nozick's second reason for why we should not plug in is that by doing so we would give up the ability "to *be* a certain way, to be a certain sort of person" (Nozick 1974, p. 42). This is very straightforwardly problematic in the case of passive VR. If VR were just like a vastly more immersive and invasive movie, then presumably the same movie could be played for different people, all of whom might wrongly think that their own virtues of character are being manifested in the movie. So, a coward in real life could be convinced that she was brave and vice versa. Moreover, since the experience of manifesting the same virtues is open to anyone who plugs into the VR movie, there really is no sense in which the plugged-in person can claim that those virtues are her own.

The same point seems to hold for the players of some video games. The 100th level Paladin strong enough to fell dragons in an *Elder Scrolls* game might really be a philosophy professor sitting at his desk at night, who during the day is completely beaten down by the vast number of assessment reports demanded (but never read) by thoroughly un-dragon-like administrators. But other video games are somewhere between semi-passive and active, as we have characterized these notions with respect to VR. And for a number of skills, the more active (in our sense) the game, the less the Nozickian disconnect between the characteristics of real and virtual persons. The latest version of *Rock Band* now has an interface that is the same as an actual guitar: the Fender Rock Band 3 Squier works as a game controller, but it can also be plugged into regular amps and played.

This suggests that one measure of how "active" a particular VR technology is might be the extent to which actual and virtual personhood become indistinguishable. We can say that VR is active to the extent that exercising some virtue in the virtual realm (such as getting a decent score on Black Sabbath's "Iron Man" in *Rock Band*) has a similar result for the player in the real world (i.e., actually being able to play the song).

But in Nozick's thought experiment, the person is choosing whether to spend her *whole life* in the experience machine! Does not this make it irrelevant whether the person in the real world could manifest the same virtues developed in the machine? If her choice is to never again leave the machine, then all of her activities there would be in some sense fictive because nothing she's doing in the experience machine makes any difference to anything real. Even if we grant that the machine might be an instance of active VR, she is not really achieving anything, since nothing she does will affect the real world. To the extent that we follow the existentialist's advice to define ourselves in terms of what we do—not in terms of what we try, but rather what we actually achieve—surely we should conclude with Nozick that life in the experience machine is the life of a morally desiccated being.

The proper response to this worry is two-pronged. First, we should note that many virtual realms are not solipsistic in anything like the way that Nozick suggests. In *Second Life* and *World of Warcraft*, players play in groups and people rightfully judge one another by the behavior of their avatars. Some players have gotten married in the real world after forming virtual relationships in *Second Life*. So in active, non-solipsistic VR, the player is not only "doing something," but is also "being" a certain way at least insofar as her/his behavior has ethically significant effects upon the lives of others.

What about the sorts of realms that each player traverses by him/herself? Anyone who has played such deeply solipsistic video games as *Myst* or *Journey* will probably

be able to envisage the possibility of deeply life like and absorbing VR environments within which the experience of human contact is extremely fleeting and rudimentary. Nozick is probably right that something is irretrievably lost by anyone but the most bizarrely stalwart introvert who plugs into such a world never to return to reality. But in 1950, Alan Turing notoriously argued that if a machine ever *could* converse well enough to fool a human into thinking that it is human, then that machine should be counted as intelligent (Turing 1950). What about VR environments run on computers that could perform this feat? Might they allow for the manifestation of virtue in a way that was robust enough to undermine Nozick's point about "being a certain sort of person" inside of VR?

We should begin by emphasizing that no computer thus far designed has ever come even remotely close to passing the Turing test, so the hypothesis just proposed about as-yet-undeveloped, non-"solipsistic" VR environments is more radically speculative and provisional than any other possible extension of present-day technologies we have envisaged here so far. But John Searle has attempted to criticize Turing's suggestion in a more radical way with his Chinese room thought experiment, which involves an English speaker looking up answers to written queries in Chinese and handing back answers (Searle 1980). Searle's thought was that even if the system could pass the Turing test, it is clear that nothing about it really understands Chinese. One way to look at Searle's thought experiment is to see it as simply posing the following question: if the ability to converse like a human were insufficient for possessing a mind, what else would be required?

Several other philosophers (Margaret Boden, Daniel Dennett, and Hans Moravec, amongst others) have argued that all that is missing from the original version of the test is a robotic body that can move through its environment in such a way that it is rational to attribute beliefs and desires to the robot.

> The Robot reply is responsive to the problem of knowing the meaning of the Chinese word for hamburger—Searle's example of something the room operator would not know. It seems reasonable to hold that we know what a hamburger is because we have seen one, and perhaps even made one, or tasted one… Given this is how one might come to know what hamburgers are, the Robot Reply suggests that we put a digital computer in a robot body, with sensors, such as video cameras and microphones, and add effectors, such as wheels to move around with, and arms with which to manipulate things in the world (Cole 2012).

The prospect of non-passive VR suggests the possibility of an interesting variation on the robot response to Searle's Chinese room argument. If non-passive VR's virtual humans must respond intelligently to whatever human players throw at them, this would obviously include passing the Turing test but also much more.[3] Their facial expressions, motor skills, and the ways they adapted to the players' own presence (including when the

[3] Let us be clear: this is only meant to apply to (actual and virtual) agents who use language and is only meant to be a necessary condition for intelligence even then! Neither of us thinks the test is remotely plausible as a necessary or sufficient test for intelligence: not necessary because nonhumans such as crows and humans such as non-linguistic deaf adults and aphaisics are clearly intelligent and not sufficient for reasons to do with what we have already described as the "enactive" nature of much human thought and perception.

player interferes with their goal-directed activities) would have to be indistinguishable from the analogous human behaviors.

If all this is right, then the Nozickian claim that one cannot really *be* a certain way in a virtual environment fails to hold even for solipsistic VR. This is because the human's actions in the virtual realm would affect intelligent creatures living within that realm. If these virtual creatures were genuinely thinking beings, we can surely infer that virtues such as bravery and honesty manifested toward them by a human in the experience machine would be just as real as virtues manifested toward other players in non-solipsistic VR or toward other human beings in the real world.

2.3 Reality

Nozick takes it to be self-evident that a "man-made reality, a world no deeper or more important than that which people can construct" is ultimately unsatisfying. His reasoning here is a bit opaque. Of the experience machine he writes that

> There is no *actual* contact with any deeper reality, though the experience of it can be simulated. Many people desire to leave themselves open to such contact and to a plumbing of deeper significance (Nozick 1974, p. 43).

He identifies fans of the psychoactive drug LSD and advocates of what he calls "traditional religious views" as defenders of this thesis about the importance of maintaining contact with "a deeper reality."

There is something right about Nozick's inchoate intuition. Nozick's denigration of a world "no deeper or more important than that which people can construct" makes undeniable sense with respect to passive VR, where people would live merely in the scaffolding, scaffolding they could not walk behind or investigate. In non-passive VR, the player can autonomously explore and experiment on the virtual world, determining whether the buildings she sees are merely a movie set or can be investigated like real buildings.

Nonetheless, one might argue that the player in a non-passive VR environment is still just investigating scaffolding, in the sense that her experiences are the result of a finite set of possible choices determined completely by the computer code. Everything done in the virtual realm is just moving down a search tree already determined by the programmers. In contrast, one might think the real world offers genuine novelty and surprises. Of course, a strict determinist might balk at this last claim, but the criticism then can be reformulated in a manner consistent with determinism. The determinist critic of VR could argue that even if a Laplacean demon could predict the distribution of elementary particles and behavior of fields throughout the universe at any point in the future, because of the nature of epistemically emergent properties, this is not something any creature like us could hope to do. But then, the critic would argue that the algorithmic nature of virtual realms entails that we ourselves can perfectly predict their evolution over time.

The determinist who relies upon this observation to argue that VR lacks the sort of "novelty" that can be found in real-world experiences presumes that VR is something like a highly detailed video game, insofar as both are at root deterministic algorithms. But this is actually as wrong as wrong can be, as a cursory examination of how video games are designed will show. At the lowest level of description, a run-through of a

video game is nothing over and above an algorithm moving numbers around among Von Neumann-style registers. But, it is simply fallacious to infer that video games themselves are nothing over and above the algorithms necessary for their execution.

First, note that the same video game can often run on different machines. Once everything is compiled down to the executable code of the machine language for the machine running the game, the algorithm will be unique to that machine. And the effect is even more pronounced when the same game is being run in different operating systems. The point is not only that video games are "multiply realizable" in the sense that musical works are. Rather, when we look at both how video games are designed and their interactive nature, it becomes clear that games themselves (as opposed to the compiled executable code of a game instance) are not algorithmic.

Consider how video games are designed. Most of the work is not low-level coding, but rather uses high-level proprietary engines and graphics programs such as Maya or 3D Studio Max to determine how virtual objects interact with one another. Even at the semantically freighted level of description (e.g., "my tank attacks your infantry"), for games such as those of the *Civilization* series, the search tree is enormously complicated. If we consider that a brute force algorithm for chess that can look ten moves in the future has more nodes in its tree than there are elementary particles in the universe, we see how combinatorially unmanageable *Civilization* is. In such a game, there are minimally hundreds of discrete moves that a player can make during any given turn, moves that come nowhere near exhausting the full set of possible moves in a game. Because of this, with any sufficiently complicated computer game, the search tree gets away from the designers. Part of the lusory joy of playing such games is beating the algorithm, finding strategies that the designers and playtesters did not foresee. And there is narrative joy too because doing so (in this case) ends up crafting a world history not foreseen by the designer. This is not a merely incidental property of one game, but rather the *raison d'être* of a whole host of games, including tabletop role playing games such as *Dungeons and Dragons* and card games such as *Magic: The Gathering*. So in this sense, at least, contemporary video games already offer genuine novelty.

A much stronger case can be made here if one attends to some basic results from computability theory. Consider the following game. A computer picks a sentence of classical first order logic and poses the problem to the players. If a player finds the claim to be provable, she is to provide a proof. If she thinks it is not provable, she enters "no." So, there are four possibilities, with the scoring as follows:

1. A player turns in a correct proof, and the computer gives her a point
2. A player turns in an incorrect proof, and the computer subtracts a point
3. A player says there is no proof when there is not and gets no points
4. A player says there is no proof when there is and gets no points

Such a game would be trivial to program and has the property that that even though the computer's realization of the game is algorithmically implementable, the game itself is not. The task set before the players involves determining whether arbitrary sentences of first-order logic are theorems. But there is provably no algorithm for this. The game itself can be instantiated by the human–computer interface, and at one level of description, everything the computer is doing is algorithmic (there is, after all an algorithm for checking whether purported proofs are correct or not even as there is no

algorithm for discovering such proofs). But the game itself has non-algorithmic properties; consider, for example, the property of "being a winning strategy."[4]

We have thus refuted each of Nozick's three claims about why a brain in the particular type of vat he envisages would have less valuable experiences than an ordinary person like you and me. We have focused our attention in this last section upon the notion that VR cuts off the user's access to "deeper reality" specifically because it removes the possibility of a certain type of *novelty* in one's experiences. There may, of course, be other senses in which the experience of VR is disvaluable on account of its sheer disconnection with the "real" world. But this would be significantly harder to demonstrate if the types of VR we have been describing were in fact no more of an impediment to *knowledge* about the actual world than eyeglasses or cochlear implants. Our next major task, therefore, will be to argue that VR needs not function as an impediment to knowledge any more than it does to overall human well-being.

3 Epistemic Brain-in-a-Vatism: Truth in Fiction

In the previous section, we focused on undermining axiological brain-in-a-vatism, the metaphysical supposition underlying a certain kind of ethically censorious alarmism about VR. But this at best only accomplishes half the job, for at root brain-in-a-vatism trades on the seemingly obvious epistemic point that the brains in the hypothetical vat are systematically *deluded*. And since, all else being equal, it is not a good thing to be deluded, axiological brain-in-a-vatism will always retain at least some of its appeal unless the epistemic challenge is faced head on.

Before proceeding, we must distance ourselves from one train of thought. Hilary Putnam famously attempted to use externalist theories of knowledge to undermine a type of skepticism that the brain in a vat thought experiment makes attractive. Putnam's arguments have been examined in detail and remain controversial.[5] Nothing we are going to say presupposes their soundness. Without trading on content externalism, we want to argue against a deeper presupposition behind epistemic brain-in-a-vatism, viz. that someone who confused fictional and actual truth to the point of wrongly believing that a work of fiction was a work of history or biography would be *systematically* deluded as a result. Our proposal is that, if the relevant work of fiction were good enough in certain respects, then such a deluded person might end up with a better understanding of the actual world than could otherwise be achieved. In "On Truth and Lies in a Nonmoral Sense," Nietzsche observed that "all the material within and with which the man of truth, the scientist, and the philosopher later work and build, if not derived from never-never land, is a least not derived from the essence of things" (Nietzsche 2012). Our more modest claim is that *some* types of

[4] This argument raises many broader issues about computationalism that would take us much too far afield. However, we have put forward the argument in detail and considered many of its surprising ramifications in three other publications: Silcox and Cogburn (2006), Cogburn and Silcox (2008), and Cogburn and Silcox (2011). See also Cogburn and Silcox (2005). For further discussion of Nozick's "experience machine" argument in a slightly different context, see Silcox and Cogburn (2009).

[5] For a useful survey of the debate, see Bruekner (2012).

knowledge might only be attained by accepting as veridical the illusions made possible by active VR.

This is because part of the job of fiction is to teach truths. A great deal of academic discourse about fiction has elided this central point.[6] Our remarks on how this happens will thus be somewhat provisional, but we will try to indicate a promising direction for further exploration. We will describe a sense in which a fictional work can be true and make some suggestions about the mechanisms that make this possible. We will describe how we think that the possibility of VR should influence the relevant aesthetic debates, bringing up new issues that an account of truth in fiction must face.

3.1 Truth in Fiction

Since the mid-twentieth century, analytic philosophers have largely focused on three concerns about the relationship between truth and fiction: (1) how to develop a modal semantics to represent the possible worlds described by fictional discourse, (2) how to understand our emotional reactions to fiction, and (3) the extent to which the moral properties of fictional narratives are relevant to their aesthetic properties. It is more than a little strange that none of these debates centrally involve the idea that a fictional work can *actually* (as opposed to fictionally) be true or false. Thus, we must begin by suggesting our own lacunary definition of truth-in-fiction:

Fictional work X is true/false with respect to discourse Y if, and only if:

> X succeeds as a work of art in its genre for person A only if
> A's imaginative complicity with respect to X, all else being equal, leads A to have non-trivial true/false Y beliefs.

To see how this works, consider Ayn Rand's *Atlas Shrugged*. The book is false with respect to economics because it can only overcome a reader's imaginative resistance if she accepts that all of a modern, industrialized society's large-scale infrastructure could be privately financed. People who already accept this will not be jarred out of their aesthetic experience by Rand's dishonesty. People who like the novel and have no previous opinion about the feasibility of private financing of large-scale infrastructure are much more likely to believe the economic falsehood. But anyone familiar with the history of rural electrification in any country in the world knows that the proposition is hogwash[7] and will balk at such passages as the following:

> Nathaniel Taggart had been a penniless adventurer who had come from somewhere in New England and built a railroad across a continent, in the days of the first steel

[6] One important exception to this generalization is Gaut (2007). See especially chapters 7 and 8.

[7] Unfortunately the hogwash has affected Wikipedia. Rand's Nathaniel Taggart is based on James J. Hill, creator of the Great Northern Railway. Both Hill's Wikipedia page and the page for the Great Northern Railway (echoing misleading claims in the objectivist blogosphere) make much of the fact that the Great Northern was privately financed and didn't receive land grants. But this is extraordinarily misleading, since the Great Northern was initially created by changing the name of the Saint Paul and Pacific Railway, which Hill purchased in a fire-sale. But the Saint Paul and Pacific was formed initially from the Minnesota and Pacific Railway, which (and it is to Wikipedia's credit that they continue to admit this) was a public railroad formed out of massive land grants and a five million dollar bond in taxpayer money. Without "loans, bonds, subsidies, land grants" and "legislative favors" the Great Northern would never have existed.

> rails… He never sought any loans, bonds, subsidies, land grants or legislative favors from the government (Rand 1999, p. 59).

Some have argued that *Atlas Shrugged* is also morally, psychologically, and historically false. The substance of these charges is exactly what one would expect, given our definition of the truth or falsity of a fictional work with respect to a given discourse: people who find themselves imaginatively complicit with the work are more likely to, partially as a result, have false beliefs about morality, psychology, and history.[8]

Likewise, people imaginatively complicit in *The Scarlet Letter* are more likely to arrive at the deep ethical and psychological insight (or so it seems to us) that "unacknowledged guilt leads to perdition, whereas expiated guilt leads to salvation" (Sirridge 1975, p. 455). Because of this, we can say that Hawthorne's work is psychologically and morally true.

Before we discuss the more general consequences of our definition, we should be clear about what we have *not* provided. While we have defined the extent to which fictional works are actually true with respect to some given subject matter, we have not tried to characterize the truth or falsity of propositions supposedly contained in a work of fiction. The simplest account of this would be that the true/false propositions that imaginatively complicit readers are likely to believe are just those propositions contained in the text. This is probably fine as far as it goes, but the notion of "containment" that we have just appealed to here will have to be left irremediably vague. It is not, in other words, that any work contains a finite set of actual truths that are extracted by readers. Rather, a work of fiction is a kind of engine that allows good readers in different milieus to come to believe some subset of the important truths that it may be said to contain.

This is no concession to relativism. It is a fact that unacknowledged guilt leads to perdition, and expiated guilt leads to salvation. But radically different readers might learn different facts as the result of their different, but equally defensible readings of *The Scarlet Letter*. There seems to us to be no reason why the same sort of claim could not be made about VR environments. To see why this is so, it will be worthwhile to reflect upon the epistemic usefulness of works of fiction in a little more detail.

3.2 Fiction as Thought Experiment

In her canonical paper "Truth from Fiction?," Mary Sirridge points out that appeals to semantic relationships such as such as entailment, presupposition, and ascent to a meta-language have not been of any help in elucidating how a good reader infers actual truths from fictional texts. But she does not think this means that fictions cannot teach us about the actual world:

> works of fiction are by no means alone in not being able to serve as direct evidence about the actual world. Cooked-up counterexamples may defeat proposed criteria

[8] Their plausibility might depend upon the somewhat delicate question of whether Rand's work is better classified as realistic or fantasy fiction, since a crucial difference between these two genres might (or might not: the authors disagree) be the degree of flexibility in imagination that each demands from its readers.

> meant to apply to kinds of things in the actual world. Thought experiments are often used to clarify hypotheses and to do them in. Counterfactual analysis is often used to support the corresponding positive claim. No one supposes that the "facts" adduced in these cases are genuine—in fact, they are usually so chosen that we can assume that certain things are unproblematically true, as we could not in actual life… We seem to have no satisfactory explanation of how these methods work epistemologically (Sirridge 1975, pp. 470–1).

In defense of the cognitive status of thought experiments, Sirridge notes that they "form one of the mainstays of philosophical thinking." Eva Dadlez follows Sirridge here, persuasively arguing that if one takes thought experiments in ethics to have any positive epistemic weight, then one cannot gainsay the positive epistemic weight of fiction (Dadlez 1997).

Naturalistically inclined philosophers might just conclude that this is so much the worse for philosophical thought experiments. But this is overwhelmingly implausible, given their manifest usefulness in many of the "hard" sciences. Discussions of weird imaginary entities such as Newton's bucket, Schrödinger's cat, and Einstein's ladder clearly do have positive epistemic value and do provide counterfactual information about how the actual world is disposed to behave.[9]

Wherever they are utilized, thought experiments have two moments: a *setup*, which in physics often involves physical impossibilities, and an *evolution*. There may be a unique evolution, if the setup is idealized enough, and if the discourse in question has the resources to treat the idealized setup deterministically. For example, consider Simon Stevin's seventeenth century thought experiment, the setup of which involved a world consisting solely of a frictionless, scalene triangle with one side facing downward. Objects placed on either side will slide down in accordance with the mechanical laws of our world. We also assume that there is no perpetual motion in this possible world. The system *evolves* when a large chain is placed around the triangle. Stevin reasoned that if the forces on the two planes constituting the upright sides of the triangle did not equal one another, the chain would rotate perpetually as the plane suffering the stronger force pulled the chain down (remember, there is no friction in the world in question). Here, once the chain is added, the evolution is entirely deterministic: nothing whatsoever happens. From this, it may be inferred that equal weights act with force inversely proportional to the lengths of planes of equal height.

For a fiction to be true in some regard, it first must be the case that the actual world could be the way the setup is in that regard. A science fiction novel with impossible physics could still have a psychologically true setup. Second, the evolution of the system from setup to end state must be something that could plausibly happen in the relevant respects in the actual world. If the characters of a science fiction novel do not develop in psychologically plausible ways, the novel will be psychologically false.

In works of fiction, the evolution is usually, if not always, just one of many plausible ones. Nonetheless, for a fiction to succeed, it must be the case that, were the actual world setup in the manner of the fictional setup, it could plausibly evolve in the manner described by the writer as the narrative progresses. In the *Poetics*, Aristotle develops this

[9] For detailed arguments to this effect, see Brown (2010) and Sorenson (2008).

point into a normative constraint upon good storytelling when he remarks that "[w]ith respect to the requirements of art, a probable impossibility is to be preferred to a thing improbable and yet possible" (Aristotle 2012).

One might misread the foregoing as implying that one must *already know* the truths of psychology, economics, history, etc. in order to detect truth in fiction. If that were the case, we would not actually *learn* anything from fiction. But psychology, economics, history, etc. all themselves consist partially of fictional thought experiments! It also just is a fact that most humans possess a nose for truths, even when they are stated in the languages of highly technical disciplines. Equipped with little other than basic common sense, contemporary readers can detect economic falsehood in Ayn Rand's works even without having studied the history of rural electrification. The general level of implausibility on a wide variety of other topics that pervades her work provides much (defeasible) evidence that she gets economics wrong too.

Keeping all of these considerations in mind, it is hard not to see video games and VR as improvements upon more conventional fictional media, epistemically speaking. While a novel typically just gives the reader one evolution from the setup, a good video game allows the player to test the plausibility of a huge number of possible evolutions from a single given setup by replaying the game while subtly altering one's inputs. One could, of course, try to do this for regular fictions set in the contemporary world by (say) seeing what happens when you treat your wife the way a John Updike character does, or eating nothing but pumpkins for a week to see if doing so has the effects described in J. M. Coetzee's *The Life and Times of Michael K*. But in the real world, such actions are often irreversible and hence impossible to repeat. Video games allow one to perform them repeatedly *in the fictional world*. Of course, if the virtual entities within such a world have moral status, then there will also have to be severe limits on such trials, for the same reason that there are severe moral constraints governing experimentation on human subjects.

We have provided no account of how one abstracts true propositions from these aspects of a fiction, and there are admittedly many different ways one could go here. The most successful VR-type technologies that have been invented so far are at any rate far more directed toward producing knowledge-how than knowledge-that. A good flight simulator teaches one how to fly the relevant kinds of planes that are simulated. To the extent that mastering the simulator gives one more true propositional beliefs, this is a side effect of how one's knowing-how constrains one's knowing-that.

A physicist we know researches the fluid behavior of stars, a task that consumes mammoth computational resources. When asked about what exactly is being crunched by the computers, he usually says things like, "I'm trying to figure out what would happen if you could knock your hand against a star. If all of those computers are employing the relativity equations correctly then the star would actually jiggle like gelatin."

Another friend of ours ran a "VR video arcade" during the early 1990s. He hooked up immersive headsets and improved controllers onto high-performance computers running games like *Doom* and *Descent* and called it "VR." His most loyal customer was an elderly woman who was confined to a wheelchair. She referred to playing *Doom* as "having my legs back." Clearly, controlling a superhuman avatar running around and shooting at enemies is radically different from actually running around with your own legs. But the similarity was strong enough for our friend's customer

that she singlehandedly kept his business afloat for a while even after the technology he used had become affordable to middle-class households.

We have described how VR technologies are crucial in helping players develop new know-how appropriate to dealing with the actual world and, consequently, learn truths about the actual world. But there is also a phenomenological aspect to this. Flying a DC-7 feels a certain way and good flight simulators try to recapture this. Perhaps in the case of flight simulations, the feeling of what something is like is analyzable without remainder into the activity of exercising one's know-how. To hold that this is in general true would be to undertake significant philosophical commitment to something like Noë's enactivism. But pre-philosophical commonsense at least tells us that VR does not just get the player to move in certain ways, but also stimulates the physical sensory systems in reaction to how the player is moving.

Part of the promise of VR is the delivery of knowledge about what certain otherwise unavailable types of experiences feel like. At the most optimistic edge of this, a player could be (for a time) a bat, an octopus, a dog, or a person knocking his hand against a star. Unfortunately, there is no clear and uncontroversial philosophical account of the relation between knowing how and so-called phenomenal or qualitative knowledge. But it is not wholly irrational to hope that further development of VR technologies will shed light on this very issue.

4 Concluding Thoughts

We hope we have dealt two significant blows to brain-in-a-vatism. VR should not be a mere bugaboo of alarmist philosophical and fictional dystopias. The reasons that have traditionally been given for the idea that it will rob us of some irreplaceably valuable aspect of our contact with the real world, or that it will radically undermine our epistemic capacities, are founded upon little more than prejudice and confusion. On the contrary, when fully-fledged VR technologies are developed, they may well seem to us to serve as a kind of alchemy, just as Artaud originally speculated. There are excellent reasons for believing that such technologies will not only enrich our hedonic and moral lives, but provide us with surprising new propositional and phenomenal insights into both ourselves and the world at large.

References

Andersen, J. L., & Aagaard, P. (2000). Myosin heavy chain IIX overshoot in human skeletal muscle. *Muscle and Nerve, 23*(7), 1095–1104.

Artaud, A. (1958). *The theater and its double*. New York: Grove. Trans. Mary Caroline Richards.

Aristotle. (2012). *Poetics*. Trans S.H. Butcher. Project Gutenberg. http://www.gutenberg.org/files/1974/1974-h/1974-h.htm. Accessed 20 Oct 2012.

Baudrillard, J. (1994). *Simulacra and simulation*. Ann Arbor: The University of Michigan Press. Trans. Sheila Faria Glaser.

Baudrillard, J. (2000). *The vital illusion*. New York: Columbia University Press.

Berdyyeva, T. K., & Reynolds, J. H. (2009). The dawning of primate optogenetics. *Neuron, 62*, 159–160.

Brown, J. R. (2010). *The laboratory of the mind: thought experiments in the natural sciences* (2nd ed.). London: Routledge.

Bruekner, T. (2012). Skepticism and content externalism. Stanford encyclopedia of philosophy. http://plato.stanford.edu/entries/skepticism-content-externalism. Accessed 13 Sep 2012.
Cline, E. (2011). *Ready player one*. New York: Random House.
Cogburn, J., & Silcox, M. (2005). Computing machinery and emergence: the aesthetics and metaphysics of video games. *Minds and Machines, 15*(1), 73–89.
Cogburn, J., & Silcox, M. (2008). *Philosophy through video games*. New York: Taylor & Francis.
Cogburn, J., & Silcox, M. (2011). Computability theory and ontological emergence. *American Philosophical Quarterly, 48*(1), 63–74.
Cole, D. (2012). The Chinese room argument. The Stanford encyclopedia of philosophy. http://plato.stanford.edu/archives/win2009/entries/chinese-room/. Accessed 23 Sep 2012.
Coyne, R. (1994). Heidegger and virtual reality: the implications of Heidegger's thinking for computer representations. *Leonardo, 27*(1), 65–73.
Dadlez, E. (1997). *What's Hecuba to Him?: Fictional events and actual emotions*. University Park: The Pennsylvania State University Press.
Dick, P.K. (1966). We can remember it for you wholesale. *The Magazine of Fantasy and Science Fiction,* April, 4–23.
Doctorow, C. (2003). *Down and out in the Magic Kingdom*. New York: Tor Books.
Gaut, B. (2007). *Art, emotion, and ethics*. New York: Oxford University Press.
Gere, C. (2004). Brains in vats, giant brains, and world brains: the brain as metaphor in digital culture. *Studies in the History and Philosophy of Science Part C: Studies in the History of Biological and Biomedical Science, 35*(2), 351–366.
Griswold, C. (2002). Happiness and Cypher's choice: is ignorance bliss? In W. Irvin (Ed.), *The Matrix and philosophy: welcome to the desert of the real* (pp. 126–137). Peru: Open Court.
Heim, M. (1998). *Virtual realism*. New York: Oxford University Press.
Kohler, I. (1964). The formation and transformation of the perceptual world. *Psychological Issues, 3*(4), 1–173.
Langan, T. (2000). *Surviving the age of virtual reality.* Columbia: University of Missouri Press.
Mander, J. (1978). *Four arguments against television*. New York: HarperCollins.
McMahon, J. (2002). Popping a bitter pill: existential authenticity in *The Matrix* and *Nausea*. In W. Irvin (Ed.), *The Matrix and philosophy: welcome to the desert of the real* (pp. 166–177). Peru: Open Court.
Nietzsche, F. (2012). On truth and lies in a nonmoral sense. http://faculty.uml.edu/enelson/truth&lies.htm. Accessed 20 Oct 2012.
Noë, A. (2006). *Action in perception*. Cambridge: The M.I.T. Press.
Nozick, R. (1974). *Anarchy, state, and utopia*. New York: Basic Books.
Plato. (1992). *Republic*. Indianapolis: Hackett. Trans. G.M.A. Grube.
Rand, A. (1999). *Atlas shrugged*. London: Penguin.
Searle, J. (1980). Minds, brains and programs. *Behavioral and Brain Sciences, 3*(3), 417–457.
Silcox, M., & Cogburn, J. (2006). Computability theory and literary competence. *British Journal of Aesthetics, 46*(4), 369–386.
Silcox, M., & Cogburn, J. (2009). Azeroth vs. The experience machine. In L. Cuddy, & J. Nordlinger, (Eds.), *World of warcraft and philosophy* (pp.55–66). Peru: Open Court.
Silverstein, M. (2000). In defense of happiness: a response to the experience machine. *Social Theory and Practice, 26*(2), 279–300.
Sirridge, M. (1975). Truth from fiction? *Philosophy and Phenomenological Research, 35*(4), 453–471.
Sorenson, R. A. (2008). *Thought experiments*. London: Oxford University Press.
Turing, A. (1950). Computing machinery and intelligence. *Mind, 59*(236), 433–460.
Virilio, P. (1994). *The vision machine*. Bloomington: Indiana University Press. Trans. Julie Rose.
Zackaib, G.D. (2013). Let there be light for new and better mind control (in mice). http://www.alzforum.org/new/detail.asp?id=3462. Accessed 15 June 2013.