Welcome to the Machine

J. Adam Carter and Neil McDonnell explore the difficult philosophical questions raised by virtual reality

In his 1974 book *Anarchy, State and Utopia*, American political philosopher Robert Nozick proposed the following thought experiment:

“Suppose there was an experience machine that would give you any experience you desired. Super-duper neuro-psychologists 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 experiences? … Of course, while in the tank you won’t know that you’re there; you’ll think that it’s all actually happening … Would you plug in?”

The thought experiment had obvious theoretical implications. If Bentham’s moral philosophy – hedonistic utilitarianism – were correct, then there’s no good reason not to just plug into the experience machine, let our desires be virtually satisfied, and turn our minds off to the outside (some would call this this the “real”) world. But surely – as a common intuition holds – we shouldn’t turn away entirely – or perhaps at all – from the reality outside the machine. And so there’s something wrong with hedonistic utilitarianism. This, at any rate, was the argument.

It’s been more than forty years since *Anarchy, State and Utopia*, and utilitarians and their critics are still arguing about the (then-purely fantastical) case and what it and similar cases mean in theory. But while they’ve been arguing, one very important thing has changed dramatically – not so much a change in ideas but a change in our machines: experience machines offering various kinds of virtual reality (i.e., computer-generated simulated reality) are now here. If you don’t already have a VR set-up you could easily buy one tomorrow (there are presently over 230 companies developing VR-related products, as opposed to just a handful 15 years ago). And we need to decide now – philosophers and non-philosophers – whether and in what manner to hook ourselves up to them.

“Would you plug in?”

The choice seems to turn on an array of questions that have up to now not been given much consideration: Just how real are the virtual experiences that contemporary VR technologies offer – viz., the kinds of experience we can enjoy through, say, Oculus Rift? Even more fundamentally, how should we view the line between the “virtual” and the “real” – both in terms of the
Thoughts

virtual objects we encounter in such computer-generated simulated realities (i.e., the tables, chairs, and other people we seem to be interacting with), and the (virtual) actions we seem to perform when in VR? Now is probably better than later to try to bring some of these questions under intellectual control.

Experience machines are now here

The recent VR revolution was started – Kickstarted, in fact – by Oculus VR in 2012 when they announced their intentions to develop stereoscopic headsets with high-resolution screens, and low-latency tracking of the user’s movement. The seed-funding campaign for their Rift headset was extraordinarily successful and, within a couple of years, Facebook paid $2billion to buy the company outright. The primary promise of this technology was that one would feel immersed in the content: part of the game, inside the movie, there with your friends. Founder Palmer Luckey’s explicit aim was to replicate the feeling of being in the Matrix.

An important enabler for Oculus was a slightly earlier software revolution in 3D games engines. Such engines had been commercially available since the mid-1990s, but were jealously guarded, and prohibitively expensive to licence. By making their engine free (to some users, at least) Unity stimulated an explosion of content from independent developers and changed gaming forever. Crucially for VR, Unity (and rival, Unreal) calculated realistic 3D light and physics, and rendered the content quickly enough to keep step with changes in perspective from a moving user/camera. By doing so for two viewpoints simultaneously, stereoscopic rendering of a realistic scene became possible.

These innovations made VR possible, but they also paved the way for a sister technology: Augmented Reality (AR). Where VR takes you somewhere else, simulating 100% of the environment, AR leaves you in the real world, but adds virtual elements. AR (sometimes branded MR for Mixed Reality) is set to become integrated into many aspects of our daily lives as it represents data about the world on the world, rather than on screen. Finding your way from A to B, for example, will no longer require the cognitive capacity to mentally translate 2D information on a map, into your native 3D understanding of the world – the route will be shown in 3D, and integrated with the environment before you. In 2015 Microsoft announced Hololens, a prototype MR device that could understand its environment, and render realistic models in it, without the need for a separate PC. The future had arrived.

As of early 2018, there are dozens of VR headsets on the market, but the leaders in terms of quality and experience remain the HTC Vive and the Oculus Rift. New iterations of both are on the way, as is the much anticipated Magic Leap One (AR/MR) headset. Science fiction has become science fact, and the eventual creation of Nozickian Experience Machine now seems inevitable, if still a decade or two away.

Nozick’s thought experiment is supposed to make us choose between the life of pleasure in the machine, and the more authentic life outside it. This presupposes that what goes on in the machine is somehow
Thoughts

less real, and less authentic, than what goes on in the non-virtual world. The choice it is supposed to pose is between the genuine value of ordinary reality coupled with less pleasure, and the ersatz value, with greater pleasure, to be found in the machine.

What if the presupposition is wrong, however, and the virtual is genuinely real, and genuinely valuable? That would quite dramatically shift the calculus of the choice.

Recall the games engines described above. Their purpose is to simulate realistic physics in the virtual environments – balls should bounce like balls really bounce, bricks should shatter windows, unsupported objects should accelerate to the ground at 9.6 m/s/s. In short, the virtual world should have a causal structure very much like the causal structure in the non-virtual world.

Philosophers of a certain stripe – structuralists – take this sort of structure to be all there is to the status of an object. Embracing a form of structuralism, David Chalmers has recently argued that this match in causal structure (isomorphism) between the virtual and the non-virtual worlds means that both are (or soon could be) equally real. What is more, if the virtual is as real as the non-virtual, then it can bear just as much value as the actual world does. Thus, entering the experience machine, and living an authentically valuable life, need not be incompatible.

For some, the above will serve as good reason to abandon that form of structuralism: we know that objects and events in VR are not real, and we know this better than we know any of the reasons in favour of that metaphysical position. That said, there are also considerations around the notion of causation that would seem to tell against this view of the virtual, even by a structuralist’s own lights.

Think back to watching Tom and Jerry visit unconscionable levels of violence upon one another. Now, picture Jerry thumping Tom on the head, then an egg appearing on Tom’s head a moment later. Here, we would gladly say that Jerry caused Tom’s injury. A little reflection should show that, whilst we are happy to say such a thing, it is not strictly speaking true. What we see in a traditional animation is one frame after another, played back at such speed that they appear to form a continuous sequence. The animator draws each frame and, in certain ways, aims to represent a standard causal process when doing so. Hence we get the sequence Jerry hits Tom > Tom gets lump. Note that it is the animator who has chosen this sequence, and it is in her head that the causal connections exist (presumably the frames share a common cause in the script). The point is that there is no genuine causation between Jerry and Tom here, just a good mimic of it. We should beware of such pseudo-causation.

VR is a little different from this, of course, precisely because the games engines (aim to) encode real-world physics such that the virtual objects and events do conform to genuine rules, and are not subject to the whim of the animator. This is part of what makes VR feel real and, the Chalmersian structuralist will claim, is what actually makes it real. Yet, the games engine is just like the animator here. It is what makes the decisions about what to show, and how such things connect with one another. What appears on the screen in VR is not genuine causation, any more than it is in the Tom and Jerry case, because the frames of the VR animation are not causing one another, but are arising from a common cause: from
the games engine. It looks like the apparent causal structure in VR is not genuine, and so, even by structuralist lights, nor is its reality.

Suppose that the objection from causation just outlined against virtual realism holds water, and that the virtual world we inhabit in VR is not as equally as real as the non-virtual. Would reaching such a conclusion “unlock” the rest of the philosophical puzzles VR poses? Not quite.

**What we learn from VR might make us re-think assumptions about the nature of intelligent action and practical knowledge**

Descartes in the Meditations (and more recently, Morpheus in The Matrix) has shown us that if what we are seeing all around us is non-real, then – try as we may – we lack knowledge of what is really around us. At most, we would have knowledge of how things seemed or appeared to us. But to know how things really are, we’d need to first know that our apparent surroundings really are as they seem, rather than some other way entirely.

Granted, we generally do not mistakenly think *that we’re not in VR* when we in fact are (even if we temporarily forget due to immersion!). So that’s good news (to the extent that skepticism is bad!) But the more basic point holds – that if the scenario we seem to find ourselves interacting with is in fact not real, then – at least when in that scenario – we are not in a position to acquire real-world knowledge. This, at least, seems to be a Cartesian-style point embraced by skeptics and non-skeptics alike.

Here, though, a puzzle emerges. If we assume both that (i) the virtual is not real; as well as the Cartesian claim that (ii) we’re not in a position to acquire real-world knowledge (but just, knowledge of how things seem) when what we’re interacting with is not real, it looks like it would follow that (iii) we aren’t in a position to acquire real-world knowledge how to do things on the basis of the interactions we have while hooked up to VR.

This conclusion, though, should sit uncomfortably with us. Pilots, after all, consistently train in virtual reality, and the results have been increasingly positive. John Brooks, the CEO and founder of Mass Virtual (which supplies VR pilot training to the US Navy) says that the more life-like the training environments, the better the results (e.g., accuracy, reaction time, etc.) have tended to be – (the particular set up for the US Navy situates pilots inside a remarkably detailed F-18 fighter jet). Moreover, VR training has been increasingly positively correlated with skill acquisition in (among other things) dentistry, endoscopic sinus surgery and laparoscopic surgery and other medical procedures, as well more generally in mainstream education.

Something has to give. It looks – initially at least – like either know-how can’t be acquired in VR despite all the positive performance results reported; or, that at least some kinds of real-world knowledge actually *can* be acquired (*contra* the Cartesian sceptic) when the environment we are
interacting with is non-real. Which way to go? The first horn seems a bit like denial, the second like epistemological heresy.

The foregoing highlights the contemporary relevance of an old chestnut the famous behaviourist philosopher Gilbert Ryle posed around the middle of the twentieth century: is knowledge-how a matter of knowing facts (i.e., regulatory propositions about how to perform tasks), or is it instead a matter of possessing abilities? More generally, what is it, exactly, that makes human performance intelligent when it is?

How it is that we go about answering such fundamental questions may inform how we should approach puzzles like the one sketched; (after all, perhaps we can explain the transference of know-how from simulation to the real-world better if know-how is fundamentally a matter of non-propositional ability acquisition, as Ryle himself thought)? Though, on the other hand, what we learn from VR might at the end of the day cause us to have to re-think previous assumptions about the nature of intelligent action and practical knowledge.

What, exactly, is the status of the things we are thinking, in response to our virtual surroundings, while hooked up VR? Here’s one simple story that might look initially attractive: when you VR-experience-that-you-are-standing-on-a-mountain, you form the corresponding belief that you are standing on a mountain. Since, in many such cases (at least, unless your VR was somehow hooked up on a mountain!), your belief is false, your mental life in VR often involves a long series false beliefs, one after another.

Such a story – even if a bit depressing – at least fits snugly with one well-known way of thinking about what happens when we dream. The view, held by both Descartes as well as psychiatrist and dream researcher J. Allen Hobson says that when you dream that something is so, you (at the time, anyway) believe that it is so. And – unless what you dream really is happening outside the dream – dreaming often involves false beliefs.

What, exactly, is the status of the things we are thinking, in response to our virtual surroundings, while hooked up VR?

Over the past decade, there has been some notable resistance to the belief model of dreaming. According to epistemologists Ernest Sosa and Jonathan Jenkins Ichikawa dreams are more like imaginings. When you dream, you’re not (for instance) believing that you’re standing on a mountain; you’re just imagining that you are.

Perhaps – if this is on the right track – then a belief model of virtual reality such as the one initially described should, analogously, be jettisoned for an imagination model of virtual reality, according to which the experiences that VR affords you are better understood as a series of imaginings than as a series of (false) beliefs.

Interestingly, whichever way is right may well have downstream moral implications, implications about how we should assess your conduct (and in particular, your mental conduct) when in VR. Consider
again the analogy with dreams: if beliefs, in dreams, really are beliefs (albeit false ones), then presumably the same should go for the other aspects of our mental lives, when in a dream: hopes, desires, fears as well as intentions.

But here things get messy. As Thomas Aquinas had thought, intending to commit an evil act is surely wrong. But merely dreaming you intended to commit such an act cannot be wrong. Aquinas' observation seems to hit upon something that may also apply (all else being equal) to virtual reality. If the distinction between intending to do something and merely dreaming to intend to do something is a morally relevant distinction, then likewise, plausibly, the distinction between intending to do something and virtually intending to do that thing is a morally relevant distinction. If VR-conduct involves forming, while in VR, real beliefs and intentions (as opposed to merely, virtual beliefs and intentions) then these would seem to be candidates for both epistemic and moral appraisal from which they would be otherwise exempt if what we're really doing is merely imagining, while in VR, that we have the beliefs and intentions we seem to have.

This all points to a rather odd conclusion – in order to work out in a satisfactory way how to assess the moral and epistemic relevance of our mental lives while in VR, we first have to figure out – quite literally – what it is that, in VR, we are actually doing in the first place.

Since technology can change how we behave towards each other, or how we pursue the good life, it is perfectly ordinary for new technologies to raise important questions within the domain of practical philosophy. It is far less ordinary, however, for new technology to prompt discussion within theoretical philosophy, and we think it is a mark of just how remarkable these new immersive technologies are, that they have done just that. Note, too, how integrated the questions we have asked here are: questions of knowledge, and mental states in VR seem to hinge in part on the ontological status of virtual entities. Going forwards we think that bringing philosophers from across sub-disciplines together to work on the philosophy of AR and VR is bound to be fruitful, and through our development of VR resources at our lab in Glasgow, and by hosting discipline-wide symposia, we hope to facilitate just that sort of progress.

J. Adam Carter is a lecturer in philosophy at the University of Glasgow, working mainly in epistemology. His book Metaepistemology and Relativism was published in 2016 by Palgrave Macmillan.

Neil McDonnell is the Lord Kelvin Adam Smith Research Fellow in Virtual Reality and Augmented Reality at the University of Glasgow, working mainly in metaphysics and the philosophy of virtual reality.