

## WHAT'S WRONG WITH THE EXPERIENCE MACHINE?

Christopher Belshaw

### *Abstract:*

Nozick's thought experiment is less effective than is often believed. Certainly, there could be reasons to enter the machine. Possibly, life there might be among the best of all those available. Yet we need to distinguish between two versions. On the first, I retain my beliefs, memories, dispositions, some knowledge. On the second, all these too are determined by the scientists. Nozick alludes to both versions. But only on the first will machine life have appeal.

Does reality matter? Or truth? Or knowledge? Is it important that there is a world beyond us, that we encounter it in apparently familiar ways, and that it is, more or less, as we take it to be? Nozick's well known thought-experiment involving the experience machine suggests that these things do matter, and are important. You can, after programming it to provide any series of experiences that you choose, enter this machine for some lengthy period. Once in the machine, you'll believe that merely seeming things are real. Would you enter? Think about it, the claim goes, and you'll see that you prefer to remain where you are. You wouldn't trade reality for a life in the machine.

The argument is believed to threaten at least some forms of hedonism, or the view that we are – and rightly – concerned with pleasurable experiences, with elements beyond this, notably here connections with reality itself, having value only insofar as they further such experiences. And it is generally well received.<sup>1</sup> Though extended discussions are surprisingly few, most commentators have looked favourably on Nozick's account, insisting, variously, that it 'refutes the mental state account of well-being', that his case 'strikes against any hedonistic theory', dealing it 'a near-fatal blow', that 'plugging in would be supremely undesirable', and that we learn from it that 'a happiness that is based on illusion is not what we desire'.<sup>2</sup>

I am less impressed. Claims made for this thought experiment are, I argue, seriously overrated. And both Nozick and his commentators run together a number of issues that need to be kept separate, with the result that it can be difficult to decide, not only what is achieved, but even what is attempted. I try to clarify some of this. And then, starting with Section III, I argue for the following. First, on a standard reading, it just isn't true that no one will enter the machine. Life without might be worse than life within. If so, you might make the best of it, and accept the offer. It may well be true that for most of us – most people reading this – the cost of entering the machine is clearly too great. So we should stay out. But there's no blanket argument against machine life. Second, in Section IV, and more tentatively, life in the machine may be not only of some value, but among the best of available lives. None of us should care about reality for its own sake, or want, for its own sake, that our beliefs about it be true. Nor, independently of experience, should we care about reality, or truth, either for our own sake, or for the sake of others. On neither count, then, should we want to be outside rather than inside the machine. Third, as well as the standard reading, there is available a revised reading of just what the machine offers. Given this reading – something I turn to in Section V – what the machine offers is so meagre that almost no one could be tempted. And no hedonist, or

supporter of the mental state view, will need, in light of rejecting this meagre offer, to rethink their position.

In sum, on the standard reading Nozick's claim that we won't enter the machine is just false. On the revised reading his claim is more or less true, but trivially so. Or so I'll argue.

## I

Nozick discusses the experience machine in *Anarchy, State, and Utopia*, mentions it briefly in *Philosophical Explanations* and then returns to it in *The Examined Life*.<sup>3</sup> The first is both the fullest and the most familiar, and I'll focus on that:

Suppose there were an experience machine that would give you any experience you desired. Superduper neuropsychologists could stimulate your brain so that you would think and feel you were writing a great novel, or making a friend, or reading an interesting book. All the time you would be floating in a tank, with electrodes attached to your brain. Should you plug into this machine for life, preprogramming your life's experiences? If you are worried about missing out on desirable experiences we can suppose that business enterprises have researched thoroughly the lives of many others. You can pick and choose from their large library or smorgasbord of such experiences, selecting your life's experiences for, say, the next two years. After two years have passed, you will have ten minutes or ten hours out of the tank, to select the experiences of your *next* two years. Of course, while in the tank you won't know you're there; you'll think it's all actually happening..... Would you plug in? *What else can matter to us, other than how our lives feel from the inside?.....*

We learn that something matters to us in addition to experience by imagining an experience machine and then realizing that we would not use it. (Nozick 1974: 42-44)

This is sketchy, with many important details underdescribed. Some will need to be looked at more closely later on, but others can be considered now. And, at least for now, such consideration assumes a standard reading of the example.

First, there's a point about time. Nozick seems to be offering a once-only deal.<sup>4</sup> You can, apart from brief reprogramming intervals, spend the rest of your life in the machine, or you can stay put. But this stacks the odds in favour of the cautious – in general lifetime commitments are unsurprisingly harder to make. And such a constraint is surely unnecessary. A lifetime commitment might not be a big commitment, depending how much life you have left. And if there is serious disvalue to the machine life, this is unlikely to be countered by spending one's last few days in the real world. So I'll take it that the focus ought to be on the unacceptability of machine life for any substantial period. Maybe it's alright for a wet afternoon, but not for six months, or longer.

Second, there are possible confusions about experience itself. Someone might think that the machine provides only certain comforting sensations – the kind of feeling you might get when things go well. But it isn't just a fuzzy glow machine. Nor, importantly, is it merely a pleasure machine.<sup>5</sup> It offers us particular experiences, in full detail, and not just their generalised after-effects. And these could be highly

complex, challenging, and even to some degree unpleasant, just as you choose. A related but contrary worry is that experiences are in part constituted by their objects, and that you can, for example, only experience climbing Everest if you actually work your way up that particular mountain.<sup>6</sup> But Nozick's idea is that you might at least seem to climb Everest and thus have at least the experience *as of* the real thing. So programme the machine, step inside, and it seems to you that you are actually there, struggling to the top of mountain, then feeling good, admiring the view. And then, as in reality, it seems that you're aware of the cold, the dangers, the need to get down. It offers neither reality nor merely its broad effects, but a detailed seeming reality along with its seeming consequences.

Third, there are important points to make about the last sentence of the quoted passage above. Nozick insists that we wouldn't get into the machine. There are questions about both 'we' and 'wouldn't'. Does he mean no one would get in? That cannot be right. There's no offer so unattractive that no one would take it up. Or does he mean that no thinking, reflective person – neither he nor any of his readers, perhaps – would accept the offer? Perhaps he means this. But, as I shall explain, this also cannot be right. Or is his claim only that anyone who does reject the machine thereby learns something about what matters to them, what they take to be important? He could make this claim while having no view about the numbers involved. But it is, or so I'll suggest, wildly implausible.

Suppose he is claiming that most people, or all thinking people, would stay out. How can he know this? It is, after all, a thought experiment, with no actual survey involved. Well, he might infer it from something else he knows; for example, that most people, or all who think, have concerns that are incompatible with our entering the machine. Should it again be asked, how can Nozick know of these concerns? The situation is different. Like the rest of us, he can claim some knowledge of widespread human characteristics, or human nature. He knows, as do we all, something of what matters to us, and what we care about. So perhaps it is not unreasonable to attribute to him the following claim: those of us acknowledging such concerns will not get into the machine. And now he might go further. If these concerns are ones we ought to have – if what matters to us really does matter – then, even if fools and madmen might succumb, no one should get into the machine. So both a minimal and a more substantial normative claim are implied by best accounting for Nozick's insistence that no one would plug in.

Consider now Nozick's contention that in declining the machine we learn that more than experience matters to us. This is in some respects modest. There is no explicit claim about either how widespread or how mandatory this further mattering is. But what he says is pretty much wrong. Most of us already think that things other than experience matter. This is a familiar idea, with a long history. It's not something discovered only by those reflecting on the machine. And, indeed, the reverse – our having some such view about what matters surely best explains why we won't get in. Perhaps, though, the idea is this. Those tempted to think that only experience matters – those tempted by hedonism or mental state theories – might be disabused of their view by reflecting on the machine, and discovering they wouldn't enter. And certainly some of those commentators quoted at the outset appear to have some such idea in mind. But even this can't be altogether right. For otherwise, given the widespread familiarity with Nozick, hedonism would be dead in the water. Again a more modest account is to be preferred. Perhaps merely some of those tempted by hedonism have been disabused, via reflection on the machine.<sup>7</sup> Yet even here there's risk of a

backfire: thinking about the machine might, equally, lead to hedonism in those previously not so inclined.

## II

More matters than experience. This is susceptible of various readings. We care about more than our own experience, or we should care about more than this. Or, we care about more than experience generally, or we should care about more than this. That many or most of us do care about more than our own experience, how our lives feel from the inside, and that all of us should so care are, I think, relatively uncontroversial claims. We do care, and should care, about, for example, other people. That most of us do care about more than experience generally is, again, relatively uncontroversial. Most people, as a matter of fact, think that mental state theories are false. That we should care about more than experience, that such theories really are false, is a more controversial claim. And in later sections I suggest that we might reject it.

We shouldn't get into the machine. I've argued that we ought to read Nozick as claiming this, rather than simply that we won't get in. So I focus on this, and ask, beginning with the following section, whether it is true. And I argue that there are people who ought to get in the machine. For them, in their circumstances, machine life is the best available. But there are others who, in different circumstances, ought not to get in the machine.

How are these claims connected? For Nozick and his supporters there's intimacy here. Because more than experience matters, and that is all the machine offers, so we shouldn't get in. And because we, or most of us, won't get in the machine then clearly we believe that more than experience matters. But I'll question some of the links. Some people ought to get in the machine. But nothing of importance can be inferred from this about what matters. Most of us, perhaps, should stay out. Again, not much follows about what matters. Suppose things beyond experience do matter. Perhaps we should get in, perhaps we should stay out. Suppose, in contrast, that experience is all that matters, and that some version of hedonism is true. Then there may be no better life than one in the machine. But a reminder: all this assumes the standard reading.

## III

Why stay out? There are some bad reasons, some that are fairly clearly superable. Certain of the bad reasons have been warned against above. We think it might be wishy-washy or dreamlike, and so unsatisfying, within. This is to misunderstand what is offered by the machine. We misunderstand, too, if we think simple pleasures are all that is offered. Another bad reason is that anyone getting in the machine would be free-riding on the rest of society, failing to pay their way.<sup>8</sup> But no one is suggesting that such machines are reserved for the few, and maintained at taxpayers' expense. Perhaps the machine might break down, with the risk of disaster greater than anyone should bear. This would be a reason to stay put. Yet we should remember this is only a fiction. And we can write the story so that it is, and is known to be, reliable.

Set such reasons aside. Still, we'll avoid the machine, according to Nozick, for the far better reason that it offers us only experiences, and more than experience matters to us. We want, he says, to do things, to be a certain way, to contact deeper

reality (Nozick 1974: 43). Life in the machine precludes our getting what we want, and so we should give it wide berth. I'll consider these claims as to costs, but spend only a little time on the last, and defer the second until later. And I'll add a fourth – we want certain things to happen in the world. This, also, is something that matters to us.

Do these concerns imply I should give up on machine life? Take my case. I have various other-regarding desires. I want Venice, and the albatross, to be saved. I want my kids to do well at school, and then in their careers beyond. Evidently what I want here is considerably more than that my experiences should have a certain shape. Yet it doesn't follow from this that I shouldn't enter the machine. Some of the things that I want will happen anyway, wherever I am. Others won't happen anyway, whatever I do. It's looking increasingly likely that Venice will survive, but tough, I suspect, for the albatross. Elsewhere I play a more substantial role. How my children fare will to a considerable degree depend on what I do to help. But that their success matters doesn't straightforwardly give me a decisive reason to avoid the machine. I could, before entering, set up an account, and caring relatives, to see them through. Suppose the machine scientists need volunteers, and will pay. Perhaps, by my entering the machine, their chances of success will be improved. Similarly for Venice. I can take what they pay and donate it to the fund.

And there are things I want for myself. I want my life to be a certain way, and to do things, and to know things. I'm after happiness, or pleasure, or, if I am considerably less fortunate than most, an end to the agony. Even if I doubt whether genuine happiness is available in the machine, pleasure presumably is, and certainly I can avoid there the constant pain. Some of the things I want to do will never happen. I'm not going to write a best-selling novel. Others, thinking about the experience machine, or openings in chess, can happen, apparently, wherever I am. Still others can occur only outside the machine. If I am to climb Everest, catch some late night jazz, play another game of squash, then I'll need to give the machine a wide berth. What about knowledge? Not only do I want my kids to succeed, I want to know that they do. Could I know this, if in a machine? I could ask the scientists to produce in me, and after some suitable interval, the belief that they succeed. But this isn't knowledge. I'll believe in their success even if they fail. I might ask that they give me this belief only if their success is genuine. And then, perhaps, I will know of their situation – knowledge needn't always be produced by ordinary means. But this is now to violate at least the spirit of the thought experiment. Nozick clearly intends to offer a stark choice between our continued and multifarious engagement with the real world on the one hand, and a string of mere appearances, on the other. It is inappropriate to introduce procedures for a mediated contact with reality, as if seeing it in mirrors, or touching it with gloves. So while some knowledge – of the past, of one's own mind, of unchanging facts – might survive the machine, knowledge of contemporary events will be among its casualties.

What about a deeper reality? Nozick thinks we'd be limited to just the man-made in the machine, but that is neither obviously true nor obviously important. Man-made reality includes Beethoven and Titian and Proust, and they are far from shallow. And it borrows from nature: watching a film of a sunset is something like watching a sunset. Communing with God, if that is possible and deep, can presumably occur wherever you are.<sup>9</sup>

So, should I go in, or stay out? It will depend on circumstances. Even supposing that life in the machine is of no intrinsic, or direct value, it still might be of instrumental value. I want my children, or Venice, to do well. And scientists will pay me to enter the machine. I have reason to enter. Or, my life out of the machine is one of endless pain. Even if there's nothing in itself good about machine life, it can be better than a bad life outside. If some will, not irrationally, choose death, either to maximise benefits for others, or to minimise their own harms, then, it seems, there must be circumstances in which we might, with reason, enter the machine. This in part presupposes, of course, that machine life isn't worse than death, and isn't worse also, than a life of agony. But it's hard to believe either of those suggestions could be true.

It's hard to believe, too, that time spent in the machine could be, in itself or directly, at best of neutral value. If it's fun to think about the experience machine, or problems in chess, I could surely have that fun inside the machine. If I'd enjoy the view from the top of Everest, I'd similarly enjoy the seeming view inside the machine. So not only might I enter the machine self-sacrificially, to promote the pleasure of others, or to have done with my pain, I might enter also in pursuit of my own pleasures. And, as pleasure is at least one component in the good life, we shouldn't believe that anyone interested in living this life, in enhancing their own wellbeing, will always avoid the machine. If things are fairly humdrum on the outside, then even if there is some cost involved, you might overall be better off in.

So consider again the examples given above. I can't climb Everest in the machine. But even if I think it second rate, seeming to climb it, given the extreme unlikeliness of my ever attempting the real thing, may be my best bet. Getting to a concert is a lot easier, and often far preferable to a CD. But the machine offers a simulation of much more than this, and given the costs of the real thing – travel, queues, ticket prices – a virtual attendance might reasonably be preferred. It may be different with the game, where actually beating someone may give me much more of what I want than merely seeming to beat them – both my physical engagement, and the physical and psychological reactions of my partner, seem to be important here. Yet even if doing something is often better, and sometimes much better, than merely seeming to do it, there may well be reasons, over a range of cases, to opt for second best.

You want things in the world, and the lives of others, to be a certain way. Whatever the intrinsic value of life therein, your desires here may be best satisfied by entering the machine. And you want a life worth living. If machine life can be of positive value then it can be rational to choose for yourself such a life, and so to enter the machine. It will depend on what is available outside. Your entering the machine will reveal something about the options available to you, but nothing of significance about what you believe to be the constituents of the good life. In particular, it won't show you to be any kind of hedonist, or supporter of the mental state view.

Can Nozick simply concede all this? He could, I think, readily agree that some irrational people might enter the machine. Could he agree also that certain people, perhaps those in one or other way unfortunate, might well have reason to enter? Of course, I think he should concede. And if he does, he still has a substantive point remaining. But I don't find any evidence for such a concession. And the more fine-graining needed for the in or out question, the more the thought experiment loses its heuristic value.

#### IV

What does the mental state theorist believe? Certainly not that the only thing I value is either my own experience, or experience in general. Perhaps he thinks that the only thing of value is experience, but it's possible he could be a moderate pluralist here, allowing that some things – art, nature – are intrinsically valuable, as well as, like many things, of value in allowing for or helping generate good experiences. What he certainly does believe, however, is that all that contributes directly to the value to me of my own life – all that determines whether and to what degree my life is going well – is my experience, or how things seem or feel to me 'from the inside'.<sup>10</sup> And so what doesn't contribute, and doesn't in this way matter, are the sources of experience, whether it is generated by scientists, directly intervening in the brain, or by the everyday furniture of the world, through sense perception. But if so then not only might a machine life be better than the actual alternative on offer, it might as well be the best life possible.

Few of us think this. We care about our experiences, but we care about other things besides. And what often we want is a conformity between how things are out there in the world, and how they appear to us on the inside. Even if it is of some value, we don't think that a life in which we merely seem to do things, merely seem to know what's happening, can possibly be as good as a corresponding life in which the doing and knowing are real.<sup>11</sup>

Yet little follows. That we value certain things, and think them implicated in the good life shows neither that their value, nor their implications, are genuine. Think carefully about what the machine offers, and we might be persuaded that it isn't after all, second rate. There are a number of points. First, while a life that advertises itself as one of mere experience, one in which it's clear we do nothing but observe the passing show, will strike us as far inferior to one of engagement, this isn't the choice here. Machine life will appear, as much as does this life, as one in which we do things, rather than have merely the experience of doing them. Second, and related, as we want really to do things so in the machine we don't get what we want. But one of the bad things about not getting what we want is, typically, that this involves feelings of frustration, disappointment, regret. And there will be no such feelings in the machine. Third, and again related, being able to act, and having true beliefs about the world, are standardly of considerable instrumental value. They help generate further good experiences. But, of course, the machine hypothesis doesn't appear to fall short in this respect. Good experiences will continue, and will appear to be appropriately linked to previous good experiences which will, in turn, appear to be knowledge and action involving. Fourth, concerns about anxiety prior to entering the machine can, rather than assessed, simply be set aside.<sup>12</sup> Just adopt a third person perspective. And suppose that you put someone else, without their knowledge, into a machine. They believe, seamlessly, that they are doing things, and knowing things. And there is no distress. Perhaps it's hard to see how what you do can be bad for this person, or that it should harm them.<sup>13</sup>

Understanding what, if anything, is wrong with a machine existence might be helped by contrasting pairs of lives, seemingly the same, but one in, the other out. Consider first a bad life. Bill is in prison, being tortured, with, as he knows, more torture to follow. Tomorrow, as he knows, he will die. Ben is in the machine, having experiences as of prison, and torture. He too, as he believes, will die tomorrow. Bill's

beliefs about his condition are true, Ben's are mostly false. But it's hard to believe that Bill's is the better life. It's hard to believe that imagining you're being tortured is worse than really being tortured. Perhaps there's a mistake here. Bill's life is in one way good – he has true beliefs, but in another way bad – he is being tortured. Ben's life is bad in the first way, but good, or at least better than Bill's in the second. It's better not to be tortured. So we might still think that Bill's is overall the worse life. But is it better not to be tortured? Normally it is. Torture is painful, and can bring about either death or long term damage. But here both are in pain, and both will die. Perhaps Ben only seems to be in pain? Even so, when there aren't differences in terms of consequences, then seeming pain is at least as bad as real pain.

Consider now a good life. Carol mixes work and pleasure – teaching in an inner city school, and persisting with the ukelele – in a highly satisfactory manner. Alice, in a machine, seems to do these things. Again, one has mostly true beliefs while the others are mostly false. But why should that matter, and make it the case that Carol's is the better life? You might think it's better, and clearly so, because of the good she does in her work, and the benefit she brings to her pupils. But in order to focus on the critical issue here we need to suppose there aren't differences of this kind. As in an earlier example, the scientists ensure that no one loses from Alice's being in the machine. Avoid this confusion, and still it's tempting to think Carol has the better life. Unlike with the bad life, it isn't strongly counterintuitive to suppose that really teaching, and really playing is preferable to the illusion of the same. Even so, it remains hard to explain or account for this temptation.<sup>14</sup>

What this issue might seem to come down to concerns the intrinsic value of one's beliefs being true. A familiar idea, that grasping truth is good for us, that ordinarily it helps our lives go better, is not in question here. Knowing what people think of us, what to eat, where we are, is typically of straightforward instrumental value. Get these things right and we'll be likely to improve our future lives, and so to have – at least among other things – further good experiences. But now supposing the quality of experience is unaffected, then this accessing of truth lacks such instrumental value. And so it has to be tempting, for those hostile to the experience machine, to suppose it is in some way intrinsically valuable, just good in itself, that our beliefs are true, and that we should know what at least some nearby bits of reality are like. The best life, then, will have a shape, a character available to us within the machine, but then will have also the further feature, available to us only with the collaboration of matters without, of most of its component beliefs being true.

Appeal to intrinsic value isn't, however, quite what is needed here. For the idea that it might somehow be good in itself, or for the universe, that, other things equal, our beliefs should be true doesn't sit well enough with this current concern for our living the good life. As the mental state theorist might, I've suggested, allow that there is intrinsic value elsewhere in the universe, so he might allow, too, that there is such value nearer to hand, in our having true rather than false beliefs. But what Nozick's supporters need is the related, but nevertheless distinct, claim that such grasping of the truth is of personal value, and is good for us, even if not in straightforward ways instrumentally good for us, in making our lives go better. Now this idea of something being neither intrinsically nor instrumentally valuable, but rather of personal value, is not unfamiliar.<sup>15</sup> But most of the examples of such value that might be offered – having friends, accessing and appreciating works of arts, pursuing some larger scale projects – significantly affect the experiences we have, and help fashion and deepen our inner lives. As is clear, having true rather than false



beliefs can do no analogous work here as, *ex hypothesi*, our experiences are the same either way. The idea that, in spite of this, having such beliefs is somehow good for us remains mysterious, as of course is the view that having false beliefs harms us.<sup>16</sup> And so it remains unclear why a machine life shouldn't be among the best of those available.

Consider a further point. You might be inside such a machine even now. There is no way of telling. But if you oppose the mental state view you'll believe that however good your life appears it would be better if you were leading the corresponding life outside the machine, such that most of your beliefs are true. Presumably there is at present just no question for you that the world is real. But now suppose scientists persuade you that you're in fact a machine inhabitant. This comes, of course, as bad news. And they offer you the choice of two futures. In the first you remain in the machine while in the second you enter the real world. Either way, you forget completely your encounter with the scientists, and once again take things as real. Either way too, your experience carries on, more or less, as now. Given this choice, you may well opt for reality. There's no reason to choose against and, as you thereby gain true beliefs, there just may be a reason, in terms of your own well-being, to choose for. But suppose there is a cost. If the best life is one outside the machine, then it's presumably worth paying something to live it. So choose reality and, although your experience will be exactly the same until that point, you die five years earlier. It isn't clear to me how it could be rational to pay that price to bring about this result. Suppose, though, it's just that the price is too high. Reality is better, but not that much better. So should we give up a year, or month, or a day? It isn't clear to me how even a much lower price is one we should pay. Yet if it's a mistake to get in the machine, it must be worth some sacrifice to remain out.

As I said at the outset, I make these points only tentatively. I continue to suspect that there is something amiss in the mental state theorist's position, and to believe that ordinary knowledge of, and engagement with, the world contributes significantly to our lives being worthwhile. But this belief may lack any firm foundation. And certainly, reflecting on the experience machine doesn't show it to be true.

## V

The conclusions so far, that worthwhile lives certainly, and best lives possibly might be led in the machine, both assume what I referred to at the outset as the standard reading of what the machine provides. There is as well a revised reading that gives a quite different result. I need now to explain and contrast these readings, to show how they involve different notions of experience, and then to point out which elements in Nozick's account support each of the pair.

The standard reading assumes a narrow account of experience. The scientists simulate ordinary sensory input, making it seem to you as if you're climbing a mountain, beating a friend, impressing a girl. This is all they do. Your memories, dispositions, and most of your beliefs remain intact. And thus your reactions to these simulated events remain as they would were the events really to occur. It might, for example, seem to you surprising that you should be climbing, winning, impressing. Or, if you are generally fearful, you might be afraid on 'climbing' now.

Understanding experience, and what the scientists do, in this way is consonant with ordinary language. If someone says they're looking for new experiences they would normally be taken as seeking out different sorts of encounters with the external world, and thus different sensations, or sensory input, rather than different beliefs, or memories, or dispositions. (If, in contrast, someone says they want to experience the world in different ways they may well be looking for different responses to the same, or similar input). And so understanding experience fits well with the target views. For surely the mental state theorist's claim that all that matters is how things seem, or feel to us 'from the inside', denies merely that it matters how they really are, on the outside, and so denies merely that it matters how sensory input is produced.

Contrast this with a revised reading, using an extended notion of experience. Here the scientists have a wider brief, and are responsible for engineering the whole of your mental content. For under this reading, to believe something, to recognise or respond to someone, to remember some event is, in each case, to have some experience. As in the machine all your experiences are artificially produced, so you no longer bring your ordinary beliefs, memories, etc. to bear on simulated sensory input. And if, earlier, we preferred to speak about experience *as of* rather than *of* the mountain, we might now think that scientists give you experiences *as of* believing that it's cold, or that you're in danger, rather than ensuring that you really believe such things.

Given this revised reading, with its broader construal of experience, will anyone accept the scientists' offer? Only a few. Even if, as before, it might be the best way to achieve some further end, no one looking to live the good life will accept. To agree to, even to initiate, this entire restructuring of one's mental content is in effect to agree to become a sort of puppet or automaton, dancing to, and also yourself singing, another's tune. Someone might urge a stronger objection, that given such a restructuring, and the consequent break in psychological continuity, then you will simply cease to exist with, perhaps, another person or proto-person coming to exist in your place. But even if you survive, as on body or animal views, still this machine existence can't appear to be of value to you. Once the details are made plain your rejection, setting self-sacrifice cases aside, is inevitable.

It may be thought the differences here are exaggerated. The restructuring need not be wholly novel. You could want to preserve much of the mental content – memories, dispositions, personality – you already have. But the critical issue here is the manner of access. And if, say, a certain past event is brought to mind, not by you, in the ordinary manner, but by scientists' direct intervention, then there is at least a question as to whether you are genuinely remembering the event at all. Similarly for changes. It may seem there's nothing problematic about a direct, internal and quasi-mechanical modification of beliefs, desires, dispositions. It may be quicker and more effective than the therapist. But it is one thing for scientists to intervene in mental functioning, and, for example, to make you less afraid of spiders. Subsequent reactions here are, seemingly, yours. It's another for them to engineer each particular response, as each encounter with a spider, or seeming spider, occurs. These reactions seem no longer to be yours.

Although many of the details here could be further explored there is, I hope, already enough to show that the differences between the two readings are genuine and substantial. And while we might well enough understand why someone should give instructions to others to fake some sensory input, so producing experiences as of an

encounter with the external world, it isn't at all easy to see why they might prefer only simulations of beliefs, memories, desires and so on.

## VI

Most discussion of the experience machine presupposes the standard reading, and so allows for, even if it by no means always involves, the sorts of objections made in Sections III and IV above. But is this fair? When Nozick refers to what we will think and feel when in the machine, there is no suggestion – even though they are not linked with ordinary objects in ordinary ways – that anything other than standard thoughts, and genuine feelings are involved. When he says the scientists might offer us the experiences drawn from the lives of others, the natural interpretation fastens on things others might have done – shot lions, dated Marilyn Monroe – rather than ways they might have been. But elsewhere things are different. Recall his listing of the machine's limitations. I said I'd come back to this. And now:

A second reason for not plugging in is that we want to *be* a certain way, to be a certain sort of person. Someone floating in a tank is an indeterminate blob. There is no answer to the question of what a person is like who has long been in the tank. Is he courageous, kind, intelligent, witty? Loving? It's not merely that it's difficult to tell; there's no way he is. Plugging into the machine is a kind of suicide. (Nozick 1974: 43)

It's impossible to make sense of this objection on the standard reading. Someone in the machine thinks, feels, has experiences. They know some things and believe more. They achieve, if Hanfling is right, and as I've suggested, a kind of happiness, even if that happiness is based on an illusion, and even if it is therefore second rate. All this seems a good way from suicide, and from blobbiness. And it seems, too, that this machine inhabitant is still a person of a certain sort, still has a character, and can still be described, at least in principle, in ordinary ways.<sup>17 18</sup>

Although the blob objection makes no sense on the standard reading, it falls into shape when that reading is revised. Someone who has no ordinary thoughts, beliefs, dispositions, but merely a series of simulations, engineered by scientists, is no longer a particular kind of person, and is, perhaps, no longer a person at all. Someone choosing this kind of existence has opted for a kind of suicide. And this is so even if the simulations are not merely random, but have a pattern and coherence, and remains so even if that pattern was originally selected by the machine dweller.

Consider also Nozick's curious illustrations: '....you would think and feel you were writing a great novel, or making a friend, or reading an interesting book'. The first and last of these are odd. For unless we focus on incidental details – pen, paper, the printed page – and insist on their being essential, then these literary activities seem perfectly able to occur in the machine. The scientists feed you simulations of page after page of *The Scarlet Letter*. On leaving you can talk intelligently about characterisation and plot. You've read it. Perhaps evidence of having written a novel would require prodigious memory, but we can well imagine someone reciting a handful of poems, and claiming, plausibly, to have composed them while inside. If such things are to be impossible then it may be that Nozick is again assuming the revised reading. Perhaps the idea is that scientists feed you with the momentary and discrete experience as of believing you're at a certain point in the novel (whether reading or writing) with only the illusion of the relevant earlier experiences having

occurred.<sup>19</sup> When you exit you find, as we almost always find with dreaming, that you just cannot bring the whole to mind.

Still, this doesn't yet clinch it for the revision. The passage I quoted earlier continues thus:

We can continue to imagine a sequence of machines each designed to fill lacks suggested for the earlier machines. For example, since the experience machine doesn't meet our desire to *be* a certain way, imagine a transformation machine which transforms us into whatever sort of person we'd like to be (compatible with our staying us). Surely one would not use the transformation machine to become as one would wish and thereupon plug into the experience machine. So something matters in addition to ones experiences *and* what one is like. (Nozick 1974: 44)

Yet on the revised reading this machine is surely redundant or useless, as either the experience machine can itself be used to bring about internal changes, refiguring our beliefs, desires, dispositions etc., or its deployment will inevitably undo work done elsewhere, and leave us merely as blobs. If the transformation machine is to have any point, then, it seems the standard reading, with its narrower interpretation of experience, needs to be maintained.<sup>20</sup>

It isn't, I think, in the end clear just how Nozick wants the experience machine to be understood. Suppose, though, that he does have this revised reading in mind. He says that no one will enter the machine. Given this reading then, self-sacrifice cases apart, I agree. Why hasn't he thus won the argument after all? Well, because there's no longer an argument to be had. When mental state theorists insist that all that matters is experience, their intended contrast is between what is inside and what is outside our heads, rather than an externally generated string of mind stuff on the one hand, and a our structured and ongoing inner lives on the other.

## VII

It may be difficult to keep the two readings apart. For there is a further problem with the machine. This problem is serious, and constitutes a powerful reason not to accept the original offer. But it is, I think, a problem with the details of that offer, and not a problem with machine life as such. It concerns Nozick's insistence that we should preprogramme our future experiences.

Machine life, in order to be attractive, has to meet certain conditions. It needs to be identity preserving. You must continue to exist, inside the machine. It needs, as well, to preserve what you care about – and that will likely involve at least much of memory and character – in your mental life. And it needs, third, to be convincing. You have to believe that life within the machine is life without. What will need to be true of machine life, if this condition is to be satisfied? What kind of existence, what series of experiences will be taken as real by someone like you, someone with your history, knowledge, memories, character? There will be several limitations. Given your general understanding of the world, experiences as of unaided flying, morphing of people into toads, the resurrection of the dead will appear to you, at least at the outset, as suspicious, and reason to doubt your world is real. Such things cannot happen. As important, there are things that can't happen to you. You are, as you know, nothing special. So seeming to rush up mountains, rival Proust as a novelist, become the love-interest of scores of young boys, will all strike you as odd.

There are difficulties here for preprogramming. No one wanting to be taken in by machine life should select for themselves experiences which, when they occur, will threaten the illusion. Is the lesson here simply that we need to select with care? There's a further, and deeper, difficulty. The experiences I have, in the real world, very much depend first on previous experiences, and second on my reactions and responses to these experiences. It's because I've enjoyed eight of Mahler's symphonies that I'll listen to the ninth; because I was bored silly on the beach that for my next holiday I'll go to the mountains. So for machine life to be at all attractive I'll want to preserve this ability to respond to my experiences, and to select or shape future experiences in light of this response. But then there are problems for detailed preprogramming. I choose to appear to climb Everest. But I do tend to become afraid. So once in the machine, and seemingly a mile from base camp, I want to turn round. Am I then to find that I'm somehow forced up the mountain, by some invisible power? It's difficult to imagine how anyone in this way captive to experience could continue to think their world real. The problem is skirted on the revised reading, of course, for I simply decide beforehand to set my fear aside. But then, as Nozick and I agree, climbing Everest is not something that I appear to do. I'm free of unwanted experiences at the cost of having any genuine wants, and genuine pleasures at all.

I don't want to exaggerate the extent of these difficulties. There will, of course, be some discontinuities between ordinary life and its successor in the machine. You'll need to forget entering the machine, for a start. There might be others. You want to seem to climb Everest. Perhaps you can, while acknowledging that this isn't straightforward, forget the physical condition that in your case makes it flatly impossible. Or you are, in reality, often painfully shy. Things are different in the machine. You welcome the change but are not deeply troubled by it. So certainly some alterations in psychology can be allowed. But they need both to be consonant with your sense of self, and not to threaten the illusion of reality. Nor need preprogramming disappear completely. You might insist that in five years you appear to meet the President of the United States, and then leave the scientists to work out the details, in light of intervening events. Or you might sketch the general and provisional shape of your future – it is to offer money, travel, gourmet dinners, fans – even while you retain the ability to tire of such offers, and opt for a different course. But detailed and long term preprogramming isn't going to work.

The problems here connect with a more general point, and a more general objection to machine life. James Griffin has insisted that we want control over our lives, and clearly believes that in wanting this there is reason not to enter the machine (Griffin 1986: 9). Here, allegedly, is an important respect in which machine life, even if valuable to some degree, will nevertheless fall short of many lives available outside. But this isn't altogether clear. Few of us have much control over our lives. Though we can influence such things, we can't control our health, our environment, the way others respond to us, and so on. Machine life, in contrast, is one that we can control. We'll have exactly the experiences we choose. If control is the issue, it would seem there's reason to go in. Yet this overlooks the difficulties just outlined. Assuming preprogramming, though we can now control our future experiences, we cannot then control anything at all. These difficulties can be circumvented: the scientists can know of your desires, and provide simulated experiences to satisfy them. You will appear to be in control of your life, and always to get what you want. Is this desirable? Would it be desirable if, in the real world, you are in control of your life, always getting just what you want? It might at first seem so, but on reflection most of us will have

doubts. Omnipotence, like immortality, would be boring. We want some control. But we want to some extent also to be subject to fortune. Griffin is wrong, then, to think, first that the machine is clearly second best with respect to control and, second that control over lives is unqualifiedly a good thing.<sup>21</sup>

## VIII

Near the end of his discussion of the experience machine Nozick says, ‘Perhaps what we desire is to live (an active verb) ourselves, in contact with reality’ (Nozick 1974: 45). We can drop ‘perhaps’. Almost all of us do indeed desire these things: we want both to be living our lives, thinking our thoughts, and somewhat accurately remembering earlier thoughts and the experiences; and also to be encountering the ordinary world in ordinary ways, and having most of our beliefs be true. These, though, are quite different desires. Both, for most of us, are more or less satisfied. But they might not be. We might be fully in control of our own thoughts, even while systematically deceived as to what is real. And we might be in the real world, encountering it in ordinary ways, yet with our reactions to and thoughts about it interfered with by scientists.

How does the satisfaction of these desires link with the experience machine? Under certain construals – either by opting directly for the revised reading, or if we are pressured to that via detailed preprogramming – we are both prevented from living our lives, thinking our thoughts, and kept apart from reality. But these are not the most natural or straightforward understandings of machine life. Hold to the standard reading, and while of course things only appear to be real the former desire continues to be satisfied. We have half of what we want.

Do we have reason to avoid the experience machine? As well as noting the desires we have we can ask whether these desires are rational. As I’ve said, the desire to be living our lives, thinking our thoughts, surely is. We have reason not to get into a situation that will in this sense cost us our autonomy. So, putting self-sacrifice cases aside, there are reasons to avoid some versions of the machine. Is the desire for reality similarly rational? It depends whether things being as they seem is in itself good for us. Hedonism, if that is a view simply about pleasure, is not easily defended. But wider construals of mental state views will fare better. I’ve suggested here only that the case against such views is not yet well made. But even if the desire for reality is rational there may still be reason to get in. Depending on circumstances, it may be our best option.

What do we learn from the experience machine? Nozick has insisted that thinking about it is instructive. But perhaps it is simply illustrative. We certainly don’t learn that autonomy or reality are things we want, and care about. And we don’t learn about the rationality, or otherwise, of these desires. Much of what we know and think about the good life, and pleasure, and the value of experience are things we know and think about already.<sup>22</sup>

Christopher Belshaw  
Department of Philosophy  
The Open University  
Milton Keynes, England MK7 6AA

c.d.belshaw@open.ac.uk

## REFERENCES

- Benson, J. (1999), *Environments, Ethics and Human Concern*. Milton Keynes: The Open University.
- Bradley, B. (2009), *Well-Being and Death*. Oxford: Oxford University Press.
- Brink, D. (1989), *Moral Realism and the Foundations of Ethics*. Cambridge: Cambridge University Press.
- Crisp, R. (2006), 'Hedonism Reconsidered', *Philosophy and Phenomenological Research*, 73: 619-645.
- Darwall, S. (1997), 'Self-Interest and Self-Concern', *Social Philosophy and Policy*, 14: 158-78.
- Donner, W. (1991), *The Liberal Self: John Stuart Mill's Moral and Political Philosophy*. Ithaca: Cornell University Press.
- Dworkin, R. (1993), *Life's Dominion: An Argument about Abortion and Euthanasia*. London: HarperCollins.
- Feldman, F. (2004), *Pleasure and the Good Life*. Oxford: Oxford University Press.
- Griffin, J. (1986), *Well-Being: Its Meaning, Measurement and Moral Importance*. Oxford: The Clarendon Press.
- Hanfling, O. (1987), *The Quest for Meaning*. Oxford: Blackwell.
- Lacey, A.R. (2001), *Robert Nozick*. Chesham: Acumen.
- Lemos, J. (2004), 'Psychological Hedonism, Evolutionary Biology, and the Experience Machine', *Philosophy of the Social Sciences*, 34: 506-526.
- Luper, S. (2009), *The Philosophy of Death*. Cambridge: Cambridge University Press.
- McMahan, J. (1988), 'Death and the Value of Life', *Ethics*, 90: 32-61.
- Nozick, R. (1974), *Anarchy, State, and Utopia*. New York: Basic Books.
- (1981), *Philosophical Explanations*. Cambridge, Mass.: Belknap, Harvard University Press.
- (1989), *The Examined Life*. New York: Simon and Schuster.
- Paul, J. (ed.) (1981), *Reading Nozick: Essays on Anarchy, State and Utopia*. Oxford: Blackwell.
- Rivera-Lopez, E. (2007), 'Are Mental State Welfarism and our Concern for Non-Experiential Goals Incompatible?', *Pacific Philosophical Quarterly*, 88: 74-91.
- (1998) 'What Does Nozick's Experience Machine Argument Really Prove?'  
<http://www.bu.edu/wcp/Papers/Valu/ValuRive.htm>
- Silverstein, M. (2000), 'In Defense of Happiness: A Response to the Experience Machine', *Social Theory and Practice*, 26: 279-300.
- Singer, P. (1981), 'The Right to be Rich or Poor' in ed. Paul (1981), 37-53.
- Sumner, L. W. (1996), *Welfare, Happiness and Ethics*. Oxford: The Clarendon Press.
- Tollefsen, C. (2003), 'Experience Machines, Dreams, and What Matters', *The Journal of Value Inquiry*, 37: 153-164.

---

## NOTES

<sup>1</sup> As well as those mentioned below see e.g. Stephen Darwall 1997, David Brink 1989. Notable among the few critics of Nozick is Peter Singer 1981, although more recently Roger Crisp, Eduardo Rivera-Lopez, and Ben Bradley have expressed doubts about the argument's strength

<sup>2</sup> See, respectively, John Benson 1999: 54; L.W. Sumner 1996: 98; Roger Crisp 2006: 620; Christopher Tollefsen 2003: 156; Oswald Hanfling 1987: 179.

<sup>3</sup> Nozick 1974: 42-45; 1981: 723 n.11; 1989: 104-8.

<sup>4</sup> This is the offer in Nozick 1974. In 1989: 107, he offers immersion for a day, a week, a year, or a lifetime. So it seems the period of immersion isn't essential to the case.

<sup>5</sup> Although some have appeared to think it is: 'Nozick invites us to imagine a machine capable of providing us only with pleasurable experiences', Silverstein 2000: 281. This is at least misleading – the machine is equally capable of providing only painful experiences.

<sup>6</sup> See here Wendy Donner 1991: 77-78, and A.R. Lacey 2001: 27.

<sup>7</sup> There might be further clarification of both hedonism and Nozick's target here. In a recent paper John Lemos 2004 insists that Nozick wants to refute psychological hedonism, or the view that as a matter of fact we are all pleasure seekers. And he does this, allegedly, by showing that not everyone would get into the machine. But a) there is more to desirable experience than pleasure, b) Nozick says that no one would get in, not merely that some would stay out, c) psychological hedonism, unlike its ethical counterpart, is implausible to begin with. And for much more on both the forms and feasibility of hedonism see Fred Feldman 2004.

<sup>8</sup> Michael Beaney made this point to me in conversation.

<sup>9</sup> Nozick 1974: 43. He allows the second point, but seems to think simulated contact is inferior to the real thing. We might wonder why. But where Beethoven and God are concerned, as there are restrictions only on sense experience, you can get the real thing. For if you can think in the machine then you can recall, internally, the music you know. So there remains, as the deaf Beethoven would attest, a way of getting Beethoven. And, as many religious people believe, both prayer and responses to prayer can go on within. So there's a way, also, of getting God. (But these communings will be less than free-flowing. He isn't to let on, presumably, that you are in a machine).

<sup>10</sup> What I am after here is the objective value of a person's life, insofar as it affects just that person. I'm not suggesting that if someone believes their life is valuable, then it is valuable. Some mental states may be more worthwhile than others, whatever their possessor thinks. But simply to talk of a life's value *tout court* is too wide: the mental state theorist might allow that my life too is of instrumental value, and that its non-mental features add to its value both for me and for others. Might he allow also, if he allows such value elsewhere, that my life could be of intrinsic value, and that this value too



---

depends on more than my mental states? Well, if he does allow this, and allows too that such intrinsic value is something I ought to care about and promote, then his position comes under considerable strain.

<sup>11</sup> See here Bradley's discussion, 2009: 9-18, where he usefully contrasts non-correspondence and correspondence theories of well-being; pure hedonism on the one hand, and a number of positions – true beliefism, achievementism, desire satisfactionism, truth-adjusted hedonism, on the other.

<sup>12</sup> Nozick allows there may be some such anxiety, but, 'What's a few moments of distress compared to a lifetime of bliss (if that's what you choose), and why feel any distress at all if your decision *is* the best one?', 1974: 43.

<sup>13</sup> Similarly in restricted cases. Perhaps you do wrong in betraying someone, or slandering them, but if what you do makes absolutely no difference to their experience, then it isn't at all clear that you can have harmed them. Similarly, too, in cases of alleged posthumous harm

<sup>14</sup> Nozick appears to acknowledge the difficulty here with his parenthetical and unanswered questions, '...*why* do we want to do the activities, rather than merely experience them?', 1974: 43.

<sup>15</sup> See Ronald Dworkin 1993: 72-4. Steven Luper perhaps muddles the two values here, claiming that a defect of machine life is that 'we will be deprived of other things that are good in themselves, such as love, friendship, accomplishments and so forth', Luper 2009: 91. Still, the way to avoid muddle isn't to insist that the different values are exclusive, and I make no such claim here.

<sup>16</sup> So even if we are wronged when subject to undiscovered betrayal, it's hard to see that we're harmed.

<sup>17</sup> See Tollefsen here, who appears to have this blob objection very much in mind: 'In the experience machine, it seems that what is missing is any act of the will. On Nozick's own understanding, the machine renders us utterly passive', Tollefsen 2003: 157. Yet this is something of a half way house. On the standard reading you will, in the machine, have beliefs and desires. There are acts of will there. On the revised reading there is neither will nor acceptance, even if, in momentary episodes, there might be the illusion of both.

<sup>18</sup> Consider again exiting the machine. Either you're no longer a person at all, or you remain just as you were when you entered, or you've been to some degree changed by your time inside. Nozick certainly discounts the first – when you come out you're able to consider reprogramming. And it's hard to believe your beliefs, memories etc remain intact but altogether unaltered. But if you are to some degree changed on leaving, it seems not unreasonable to suppose you are a kind of person while inside. Someone wants to improve their chess. They programme the machine to simulate a graded series of games. They leave after two years, and then play at Grand Master level. I see no reason to deny that shortly before exiting they were able to play chess at around this level. Someone else simply asks scientists to refigure and improve the chess playing areas of the brain. But again, if they can play when they're out, they can play when they're in. A fearful person wants to become more courageous. They

---

programme the machine to present them with spiders, mice, pit bulls, lions. On exiting their bravery is undeniable. I don't see why we shouldn't say both that they became and – in the later stages – were courageous while in the machine. Tollefsen is sceptical: 'genuine virtues are consequent on genuine acts... An agent who had merely the experience of courageous behaviour in a dream, or on the experience machine, could not expect that this would carry over as the virtue of courage in waking life', Tollefsen 2003: 157. Even though what he says about it is right, I doubt the dream analogy is apt. Consider instead procedures like aversion therapy, forms of brainwashing, or hypnosis. Even if claims about these are exaggerated, there is some carrying over of behaviour patterns.

<sup>19</sup> And then, of course, in contrast to Nozick's wording, it would be better to say that you *seem* to think, or are having an experience *as of* thinking, that you're reading or writing a novel.

<sup>20</sup> This uncertainty as to Nozick's position can be seen also by looking to the context in which discussion of the experience machine takes place. He is discussing utilitarianism, first in relation to animals, later to people, and insisting on its shortcomings. But what is utilitarianism? Is it concerned with, as Jan Narveson famously put it, making people happy or making happy people? Many of us are reasonably well-disposed to the former project, not so many to the latter. And the not so many will increase if it's suggested that a good world, and good lives, might contain just strings of disconnected happy moments. But it now seems that it's just that that the experience machine, on the revised reading, has to offer.

<sup>21</sup> If this isn't altogether fair to Griffin this may be because he is working with the revised reading of what the machine offers. Under that reading it is second best for control. See here also Nozick's discussion (in Nozick 1974) of the shortcomings of further machines that give us even more by way of control. There isn't the space to explore this here.

<sup>22</sup> My thanks to Alex Barber, Stephen Holland, Gary Kemp, Christian Piller, Carolyn Price, Tom Stoneham and referees of this journal for discussion and comment. An earlier version of this paper was presented at the SIFA (Italian Society for Analytic Philosophy) Conference, Genoa, 2004. My thanks to the organisers and the audience.