MIGHT KNOWLEDGE BE INSERTABLE? [[1]](#footnote-1)

SECTION I

INTRODUCTION

My question concerns whether we can coherently conceive of our having knowledge inserted into our minds, bypassing forms of knowledge acquisition hitherto relied on. As a first approximation of what I have in mind consider the last big exam that you studied for and the stack of books, and papers and other learning materials that you had to read, together with the exercises that you might have had to do to make sure you had understood and fully digested them. Then consider whether it would have been desirable to you to bypass this process altogether and advance straight to point of knowing by the use of some special technology. There are two reasons for raising these questions. First, insofar as technology develops apace, it is wise to start thinking about how we would use it before its uses are available to us lest we should unreflectively abuse it. Second, while having knowledge inserted might turn out to be physically impossible, its metaphysical possibility would still make reflecting on its desirability fruitful, since evaluating this will help us to better understand what it is that we value about education and about knowledge.

For instance, we might think that what we value about education is simply the good of having knowledge, and that the good of knowledge is the powers it affords us. Alternatively what we might value in education is the effortfulness of knowledge acquisition, or simply effortfulness in pursuit of its acquisition. In this paper, I provide a positive account of what it would mean to have knowledge inserted. I will reserve the task of arguing that it would often be a positive thing for us to have knowledge inserted for another occasion.

We should distinguish between practical knowledge and theoretical knowledge (or ‘knowledge how’ and ’knowledge that’), and while I propose to handle their cases somewhat separately, I argue that, for each, insertion is conceivable. Regarding the conceivability of inserting knowledge, we should be wary that not all scenarios that are *apparently* conceivable are coherent and so not all scenarios that are *apparently* conceivable are *actually* conceivable. That said, if a) we seem to be able to imagine the case, b) our seeming to be able to imagine it survives extended critical reflection, and c) we can cash out what it would consist in with plausible detail, then we have defeasible reason for thinking that inserting knowledge is in fact metaphysically possible. In section II, I discuss five objections to the thesis, so as to subject the case to extended critical reflection. In the rest of this section, as an example of what is meant by implanting or inserting knowledge, we briefly consider imaginary scenarios presented in the fictional film, *The Matrix* (1999). I then proceed to flesh out the scenario with some more plausible details to help make the case for its conceivability.

HOW TO IMAGINE INSERTING KNOWLEDGE

Our question is whether we might conceivably have knowledge inserted. Here I offer ways to understand how the insertion of beliefs might take place in a more concrete, but still speculative fashion.  What might such a processes look like? For a vivid and compelling example of what is meant by ‘inserting knowledge’, consider the following fictional scenario from The Wachowski Brothers’ film, *The Matrix*.[[2]](#footnote-2) A character named Neo is plugged into a machine for ‘combat training’ and ‘learns’ Jujitsu (or *inserts* a Jujitsu programme) in just a few seconds. ‘I know Kung Fu’ he announces, ‘show me’ replies Kung Fu expert, Morpheus. Sure enough, Neo is able to demonstrate Kung Fu skills rivalling those of Morpheus, skills that he never had before. There is a very distinct contrast between before and after Neo was plugged in here. By being plugged in, what Neo has bypassed is years and years of effortful practice in order to jump straight to the end state of what one would have acquired by those years of practice and without the same enormous opportunity costs that so much time spent would amount to.

The scenario in the Matrix seems to ask us to imagine human beings literally downloading digital information from a computer in the same way that computers do from one another. I do not wish to claim that that is something that we can imagine without also imagining human beings very different from ourselves, after all, humans do not store information in the same way that computers do. Nevertheless the example is useful for illustrating the general idea in play. What other ways do we have of cashing out the picture in more plausible detail? One intriguing possibility seems to be this. It seems that we must we admit that genetic ‘memory’ passes through our DNA. After all, it seems that infants come with cognitive equipment. For instance, the empirical evidence supports the view that an intuitive physics, containing certain core principles about solidity, continuity, cohesion, and property changes, is present as early in human infancy as can be tested for, “and the nature of the underlying representation is best characterized as primitive initial concepts that are elaborated and refined through learning and experience.”[[3]](#footnote-3) We know then, that genes can express themselves in the form of cognitive content, and we also know that genes can be mutated. Given this information, one might reasonably hope to insert quasi-innate cognitive content in a person by mutating their genes. Of course one might reply that one can acquire cognitive abilities or dispositions in this way, but deny that any cognitive content is transferred. This point is certainly controversial, but it may begin to persuade the reader that knowledge, or something close enough to it can be inserted.

Another apparent possibility is that of direct neuro-intervention – perhaps executed with the aid of Nanorobotics to reduce invasiveness of surgery – destroying, redistributing and rearranging connections between neurons, or perhaps inserting neurons that we might be able to either synthesize or grow. It seems that what knowledge, memory and other mental apparatus I have is (in part) a function of my neuronal landscape–of what neurons I have and of how they are connected. Furthermore, it seems that if one wants to add to what I know, one has to change this neuronal landscape. My contention is simply that if one can change neurons, while knowing what such changes will do, among the things that one can do is change what is known. While it is not brains but people that have beliefs and knowledge, one might be able to give a person a belief simply by rewiring their brains in these ways, and so too with skills, attitudes, and other sorts of things that would once have been transmitted through (e.g.) modelling, imitation, testimony, and trial and error.

One way to imagine this is the following. Suppose that in one possible world, I went off to university where I attended lectures, read books, engaged in discussions, read, wrote and rewrote papers, and came out, three years later, having a working knowledge of a lot of material and being, in a word, educated. Suppose now that in another possible world, before I went off to university, for a price, a shady neurosurgeon with knowledge of how I turn out in the first possible world, offered to save me all of that time and hassle and just bring about the end state. Such a scenario just consists in replacing my brain with an entirely new brain, and in the process one would simply be killing me and replacing me with a cognitive replica of myself from the first possible world. Obviously this is no good to me. Even if I somehow survived the transition, I would be lumbered with a lots of false memories about attending lectures, reading books and so on. But first the changes need not be so radical, and could be accumulated incrementally so that a sharp discontinuity between the pre-operation and post-operation persons need not result and render us numerically distinct. Second, what I could be given is the knowledge, without all of the false beliefs and other baggage. Quite how is down to the neurosurgeons to figure out. Knowing what networks to create and how to rewire brains so as to achieve the desired effects is a tall order, given that we cannot peer into alternate possible worlds, but not beyond the bounds of conceivability.

I have said that if a) we seem to be able to imagine the case of our inserting knowledge, b) our seeming to be able to imagine it survives extended critical reflection, and c) we can cash out what it would consist in with plausible detail, then we have defeasible reason for thinking that inserting knowledge is in fact metaphysically possible. In this section I have aimed to help the reader to have imagined the case of our inserting knowledge and to have cashed out the idea in some plausible detail. In Section II I will explain and consider a number of objections to it to see how it holds up to the requirement that it survive extended critical reflection. More immediately however, we must distinguish between three kinds of knowledge.

It is usual to distinguish between a few different kinds of knowledge which may not be reducible to one another and that might be exemplified by the following claims: 1) I know the streets of London, 2) I know the names of all the streets in London, and 3) I know my way around London. The first could be called knowledge by acquaintance, the next propositional (or theoretical) knowledge, and the last practical knowledge (or knowledge how). I take it that there is no question as to whether one can insert personal acquaintance, even if one can insert the impression of personal acquaintance. To be acquainted with an object is to have been around it. But inserting false memories of acquaintance is to give a mere illusion of acquaintance (one cannot insert a relation). I take it that there is more question about the conceivability of inserting practical knowledge, but what is most questionable is the possibility of inserting propositional knowledge, and I shall come to objections regarding this at the outset of the next section.[[4]](#footnote-4)

SECTION II

WHY IT IS NOT, NOT CONCIEVABLE TO INSERT KNOWLEDGE

In this section we shall consider four arguments denying that the insertion of knowledge can be coherently conceived. First, one might argue that inserting knowledge requires inserting justification, but that justification cannot be inserted. Second, one might contend that the possibility of inserting false beliefs undermines the possibility of self-trust, and so undermines the possibility of knowledge (Bernard Williams). Third, it might seem that inserted content has the wrong causal history to count as knowledge (Putnam and Davidson). Fourth, one might worry that any knowledge inserted would lack the appropriate connections to other mental items so as to count as knowledge, and would either a) amount to a tick, or b) simply remain inert. The first of these argument counts against the conceivability of inserting propositional knowledge in particular. The rest count against inserting either propositional knowledge or knowledge how. In the rest of this section I shall introduce and rebut each of these challenges.

OBJECTION 1: YOU CANNOT INSERT JUSTIFICATION

One might contend that although we might conceivably insert beliefs, we cannot conceivably insert knowledge. Unsurprisingly, I do not have a theory of what propositional knowledge consists in: the possibility of a simple Justified True Belief (JTB) account of propositional knowledge has been decisively undermined by Edmund Gettier’s well known counter examples.[[5]](#footnote-5) It is not clear how to amend the theory in order to account for propositional knowledge, and much reflection and argument has been committed to resolving the issue without yet generating consensus. Indeed, some theorists have defended a knowledge-first account which takes knowledge as a primitive concept, explaining other concepts in terms of it (for example, belief is understood as the result of a failed attempt at knowledge). For others, the hope remains of supplementing JTB with some additional ingredient to produce a successful reductive identification. All the same, it seems that JTB is necessary to (if not sufficient for) propositional knowledge.[[6]](#footnote-6)   Supposing that justified, true beliefs are essential components of knowledge, if we cannot insert them, then we cannot insert knowledge.

In what follows, I will have in mind beliefs in this sense: mental representations of states of affairs that one represents as being true.[[7]](#footnote-7) Such representations may be fine or coarse grained: my daughter may believe that daddy is a doctor without having a very fine grained notion of what that is. Such representations may be more model-like than sentence-like: I may have a single representation of my bicycle that I can partially express with countless individual propositions, rather than countless sentence-like beliefs about my bike. By justification, I have in mind whatever raises the odds that the proposition believed is so much more likely to be true than to be false that it is no longer reasonable for the person who knows it to suspend judgement between the two. Inserted beliefs that can be said to count as knowledge however, must, if JTB is a necessary (if not sufficient) condition of knowledge, consist in something more than representations of states of affairs that one represents as being true. They must also boast justification. However, it may be objected, it is not possible to simply so insert justification in the same way as it is possible to insert beliefs, without starting out on an infinite regress of justification requiring further inserted beliefs.[[8]](#footnote-8)

How should we respond? One reasonable response, albeit rather limited in its application, is to observe that it seems in the case of *a priori* knowledge we could simply insert the justification for the belief along with the belief. For instance, the knowledge that a triangle has three sides, requires nothing more than introspection on the relevant concept. The principles of Modus Tollens and Modus Ponens could equally be inserted. However, a second, more powerful, response start with the observation that one form of justification consists in mere deference to reliable, or trustworthy testimony. Given this, one could start the journey to knowledge by inserting a set of beliefs from a source that one took to be authoritative anyway. These beliefs would become knowledge when combined with the belief that they can be traced to the source that one took to be authoritative (without any belief insertion). This indication of origin together with one’s pre-existing trust of that source allows for rational belief and thereby knowledge in the case that it is true (and some other Gettier-case defying condition holds). It seems then that issues regarding the insertability of justification for belief pose no problem in principle to the insertability of knowledge. Let us now consider an altogether different style of objection that claims that the possibility of inserting false beliefs undermines the possibility of having knowledge at all, and thereby undermines the possibility of having inserted knowledge.

OBJECTION 2: INSERTING FALSE BELIEFS UNDERMINES SELF-TRUST

It seems that the ability to insert beliefs would give us the option to voluntarily select our belief sets on non-evidential grounds, it would enable us to believe false things in exchange for money, or because believing such things would make us happy. Bernard Williams has argued that “it is not a contingent fact that I cannot bring it about, just like that, that I believe something”. [[9]](#footnote-9) Why is this? Williams’ worry is that if we can decide to believe anything at all at will, and not on epistemic, but on practical grounds (e.g. in return for a $500,000,000 reward), this may undermine the possibility of belief *tout court*. If I insert false beliefs for money, say, it may be that I begin to undermine the warrant for any belief that I form, and that I may undermine any conception of myself as a truth-seeker who can believe what he tells himself. Even if I pay to have such worries erased, they will only emerge again at every juncture whereat I consider exchanging belief for money. I then become an object of my own scepticism. How should we respond to this line of worry? We can observe that first of all, things need not play out in this way at all. So long as I do not entertain such temptations seriously, I do not become a victim of such scepticism. Second we can rehearse the familiar response to the Cartesian method of doubt that although we may have grounds to doubt any particular one of our beliefs, we cannot take this to be grounds to doubt all of our beliefs altogether. Thus the possibility of inserting false beliefs need not open the gates to any thorough going scepticism, and certainly not to any beyond those to which we are already vulnerable. The third objection which we shall now consider contends that inserted ‘knowledge’ would have the wrong kind of causal history to count as any such thing.

OBJECTION 3: INSERTED ‘KNOWLEDGE’ HAS THE WRONG CAUSAL HISTORY

Consider a thought experiment described by Derek Parfit:

Here on Earth, I enter the *Teletransporter*. When I press some button, a machine destroys my body, while recording the exact states of all my cells. This information is sent by radio to Mars, where another machine makes, out of organic materials, a perfect copy of my body. The person who wakes up on Mars seems to remember living my life up to the moment when I pressed the button, and is in every other way just like me.[[10]](#footnote-10)

Call these individuals Earth John and Mars John. Now ask yourself the following question: what does Mars John know? Parfit says that Mars John “seems to remember living [Earth John’s] life up to the moment when [he] pressed the button.” That is surely a natural thought. Let’s suppose that Mars John also believes that he is numerically identical with Earth John. Let’s further suppose that Mars John meets Earth John’s Mother. Now let’s parse out the sorts of things that Mars John might be expected to know with more specific questions.

Question One: Does Mars John know that this is his own mother? Answer: No! After all, he does not even have a mother since he is a replica that was synthesized in a machine. It would seem that Mars John has a false belief that this is his own mother.

Question Two: Does Mars John know Earth John’s mother (knowledge by acquaintance)? Answer: Mars John is not acquainted with Earth John’s mother, so although he might seem to know her, he does not. After all, he has not spent any time in her company, or had any previous direct contact with her. To help illustrate this point, consider the following example from Timothy Williamson:

The first time I felt my unborn daughter kicking with my hand, what moved me most was the feeling that I was having singular thoughts directly about her, not just as ‘the baby, whoever it is, inside my wife’; in Bertrand Russell’s terminology, it was knowledge by acquaintance, not by description.[[11]](#footnote-11)

If Timothy Williamson were to be cloned, his clone would not have felt his unborn daughter kicking with his hand, and so would not be acquainted with her, and could not have singular thoughts directly about her.

Question Three: Does Mars John know that he has a brain (propositional knowledge)? Answer: he seems to believe he has a brain, he certainly does have a brain, but does he know it? It is a belief that was generated (along with all his other, often false, beliefs) by the machine and not by the usual processes by which one ordinarily come to believe that one has a brain.

Question Four: Does Mars John know that he is on mars (propositional knowledge)? Answer: he seems to think that he is on Mars, and he is right that he is on Mars. Depending on what one’s theory of knowledge consists in, it at least seems prima facie plausible that he knows that he is on Mars. He opens his eyes (for the first time) believing that he has just pressed the button. He believes that if he has pressed the button, and nothing happens, this means he must be on Mars, and indeed, nothing seems to happen, so he concludes that he is on Mars. But he has not pressed the button, he merely thinks he has.

Question Five: Does Mars John know how to walk (knowledge how)? Answer: it might seem pretty obvious that he knows how to walk. However, one might object that even this ‘knowledge’ needs a particular causal history or it cannot count as knowledge at all.

If that’s right, it may turn out that no matter how good we get at brain surgery, and no matter how closely we can get a patient’s brain to approximate that of someone who knows, they will never produce knowledge because the causal history is all wrong. Donald Davidson, following Hilary Putnam, argues to this effect:

“aspects of the natural history of how someone learned the use of a word necessarily make a difference to what the word means. It seems to follow that two people might be in physically identical states, and yet mean different things by the same words.”[[12]](#footnote-12)

But “If meanings ain't in the head, then neither, it would seem, are beliefs and desires and the rest” (KOOM, 443), this is because beliefs and desires have the same kind of intentional content as meanings. Clearly Davidson thinks that causal history is very significant for mental items to have the intentional content that they do. He asks us to consider this scenario:

Suppose lightning strikes a dead tree in a swamp; I am standing nearby. My body is reduced to its elements, while entirely by coincidence (and out of different molecules) the tree is turned into my physical replica. My replica, The Swampman, moves exactly as I did; according to its nature it departs the swamp, encounters and seems to recognize my friends, and appears to return their greetings in English. It moves into my house and seems to write articles on radical interpretation. No one can tell the difference. (KOOM, 443)

But there is a difference. My replica doesn’t recognize my friends; it can’t recognize anything, since it never cognized anything in the first place. It can’t know my friends’ names (though of course it seems to), it can’t remember my house. It can’t mean what I mean by ‘house’, for example, since the sound ‘house’ it makes was not learned in a context that would give it the right meaning—or any meaning at all. Indeed, I don’t see how my replica can be said to mean anything by the sounds it makes, nor to have any thoughts’ (KOOM, 444)

I should emphasize that I am not suggesting that an object accidentally or artificially created could not think; The Swampman simply needs time in which to acquire a causal history that would make sense of the claim that he is speaking of, remembering, identifying, or thinking of items in the world. (I return to this point later) (KOOM, fn 4, 456)

How might we respond to this line of criticism? Imagine that you saw a video recording of the episode in the swamp and decided to show it to The Swampman. The Swampman might seem to be surprised. He might even seem to be upset. I suggest that it beggars belief to deny that he is upset or at least *as good as* upset. We have to try to make sense of his reaction to the video, and the way it makes most sense to do so is by attributing to him the property of sadness together with the explanation that he has just found out an unsettling fact which tells him that he is not who he thinks he is. Adding scare quotes around ‘sad’ and ‘belief’ and so forth beings to look pretty vacuous at this point. If the qualitative (rather than historical) character of his mental states in no way differs from that of more normally developed creatures, then it’s as good as one another. Insisting that The Swampman doesn’t know but merely ‘knows’ looks pretty pedantic. The Swampman *as good as* knows.[[13]](#footnote-13) The ideas that I am appealing to here are their functional similarity, and how they seem ‘from the inside’.

To develop this point in more detail, one could draw on two-dimensional semantics to rescue The Swampman’s mental content. For David Chalmers the sentence “Water is H2O” has a primary and secondary intension. The primary intension of a concept allows us to determine what its referent is in this, our actual world. However, that primary intension could be satisfied by an altogether different referent in a counterfactual world. The secondary intension of concept is just whatever the referent of the concept happens to be in this, the actual world. Roughly, the primary intension of ‘Water’ “picks out the dominant clear, drinkable liquid in the oceans and lakes,” something that could equally be satisfied by H2O or some other chemical composition if that were the dominant clear, drinkable liquid in the oceans and lakes.[[14]](#footnote-14) Since the referent of the concept ‘Water’ just so happens to be H2O, H2O is the secondary intension of that concept. While The Swampman’s secondary intentions would require “time in which to acquire a causal history that would make sense of the claim that he is speaking of … identifying, or thinking of items in the world,” it seems that his primary intentions would require no such time to acquire meaning. Indeed, they might be wholly satisfied by quite different items in a counterfactual world.

For these reasons then, I do not take the Swampman considerations to rule out the possibility of inserting knowledge or something close enough to it. There are two other arguments, or two versions of the same argument left to consider.

OBJECTION 4A: INSERTED ‘KNOWLEDGE’ WOULD BE ‘TICK-LIKE’

Another interesting scenario described by Daniel C. Dennett might be thought to hold the germ of a problem for the thesis I am defending. “Suppose we have entered the age of neurocryptography,” Dennett says, “and it has become possible for ‘cognitive micro-neurosurgeon’ to do a bit of relevant tinkering and *insert* a belief into a person’s brain.”[[15]](#footnote-15) Suppose that Tom has the belief “I have an older brother living in Cleveland” inserted, and that it is false. This might play out in two ways, says Dennett: one in which Tom’s rationality is left intact, and one in which it is damaged. Were Tom asked whether he has a brother, he would answer that he does, and that he lives in Cleveland. Were he asked anything more about his brother, he might realise that the belief was false and eliminate it (after all, he also believes that he is an only child, and he has no recollection of his brother). On the other hand, he might maintain that he has an older brother living in Cleveland (“in the nature of a tick”), and that he is also an only child, in which case “his frank irrationality would disqualify him as a believer.”[[16]](#footnote-16) Worse, perhaps, “pathologically, the brain will surround the handiwork with layers of pearly confabulation,” such as a made-up name and what not. If Tom is to remain rational, in order for the belief to be sustained, it cannot be inserted in isolation, or without simultaneously eradicating other beliefs. Here the notion of belief is illuminated: beliefs must function in particular way, they must be susceptible to creation and annihilation in light of evidence, and must be able to affect the promotion and annihilation of other beliefs. In particular, it is the content of these beliefs that makes further creation and annihilation appropriate through logical relations: if I regard myself as having no brother, I cannot rationally think I have a brother from Cleveland. Rather than rendering the prospect of our having knowledge inserted incoherent, this example merely places the further constraints on knowledge that it be appropriately integrated among our wider sets of beliefs: poised to eliminate those which it contradicts, for instance. In the next section we consider a connected problem about how inserted knowledge is to be properly integrated among affective items.

OBJECTION 4B: ANY INSERTED ‘KNOWLEDGE’ WOULD BE INERT

In *Descartes’ Error*, neuroscientist Antonio Damasio presents cases of patients who have sustained damage to the Ventromedial region of their prefrontal cortices, without sustaining any intellectual damage (e.g. to their capacities for memory, attention, language). It turns that such patients out consistently act in what we would consider to be dramatically irrational ways. He contends that it is because patients who have sustained this kind of brain damage lack ‘somatic markers’ – bodily feelings such as anxiety, nausea and disgust that are cued by emotions – that they fail to translate into action what they can recognize intellectually as being best. This he called the Somatic Marker Hypothesis, according to which:

emotions marked certain aspects of a situation, or certain outcomes of possible actions. Emotion achieved this quite overtly, as in a “gut feeling,” or covertly, via signals occurring below the radar of our awareness.[[17]](#footnote-17)

If we were to re-wire neurological connections so as to create knowledge, the lesson from Damasio’s case studies seems to be that we ought to wire affect up with the knowledge if it is ever going to be deployed sensibly. Indeed, following philosopher of education Michael Hand, we ought to acknowledge that items of knowledge cannot be hived off from the affective responses that they render appropriate:

cognition and affect are not at all easy to separate: an integral part of coming to understand the facts, theories, texts and narratives that make up the cognitive content of the curriculum is coming to feel their interest and excitement, their inspiration or disenchantment, their nobility, injustice, comedy or tragedy.[[18]](#footnote-18)

The sort of understanding that Hand has in mind here is appreciative, which can be distinguished from more formal and disinterested forms of understanding. For instance, someone may be able to answer questions about the holocaust correctly, to reason well about it from the evidence, and yet take a disinterested, merely professional view of it. Here, it would seem that there was a major failing in their appreciative understanding, even if their more formal understanding was outstanding.

CONCLUSION

While the scenarios from our introduction might seem fanciful, we ought not to rule out the possibility without argument. Although the onus may be on those that propose knowledge’s insertability as a possibility to flesh a coherent account of what it is that is to be affirmed or denied, and perhaps concocting a phrase such as ‘inserting knowledge’, and pointing to a few fictional scenarios will not serve this purpose. Indeed, we should be wary that not all imaginable scenarios are coherent and so not coherently conceivable, which is what seems to be required of the notion of conceivability. For instance, I might be able to imagine myself going back in time and killing my parents before they had me, but we cannot coherently conceive of that, for I would not then have been born in the first place.[[19]](#footnote-19) On the other hand we might suggest that, while wary that the scenario might prove inconceivable, the fact that we seem to be able to conceive of it gives us a defeasible reason for thinking that it is conceivable.  Many people take themselves to be able to imagine a bicycle, but, on closer inspection, when asked to draw the bicycle, they find themselves unable to do so (not in terms of making it look pretty, but simply in terms of including the rights sorts of components in the right sort of relation to make something like a functional bike, or, one might say, a bike). They might take themselves to be riding it and operating it, but they could not possibly be performing those actions on that piece of equipment, because it has, e.g. no breaks, no gears, and so on. Alex Carruth has argued that something, although apparently conceivable, could in fact prove inconceivable:

Here’s an analogy, think of a mechanical clock, indeed, an exact duplicate of a mechanical clock you’re acquainted with. Can you conceive of the duplicate’s hands running anti-clockwise, rather than clockwise, or not running at all? You certainly could form a mental picture of the clock and say ‘and the hands run backwards’. But under close inspection, it’s not clear one could maintain this picture under scrutiny without making some change to the clock — say by rearranging the gears, or changing the direction of the motion imparted by the motor.[[20]](#footnote-20)

Notwithstanding these considerations, I have contended that if a) we seem to be able to imagine something, b) our seeming to be able to imagine it survives extended critical reflection, and c) we can cash out what it would consist in with plausible detail, then we have defeasible reason for thinking that that thing is in fact metaphysically possible. In Section I, I attempted to motivate the view that we at least seem to be able to imagine inserting knowledge, and provide some plausible detail to flesh out that imagining further. In Section II, I attempted to defend the conceivability of these case from a series of critiques, showing that it can survive extended critical reflection by appraising a number of arguments that it is not coherently conceivable. These included arguments that inserting knowledge requires inserting justification, but that justification cannot be inserted, that the possibility of inserting false beliefs undermines the possibility of self-trust which is crucial to knowledge *tout court*, that inserted content has the wrong causal history to count as knowledge, and that inserted knowledge would lack the appropriate connections to other mental items in order to count as knowledge. We have answered each of these objections in turn and thus far it seems that we seem to be able to imagine the insertion of knowledge, that what it would consist in can be cashed out with plausible detail, and that the notion can survive extended critical reflection. In this paper, I have not had space to discuss the desirability of having knowledge uploaded. As final remark however, I think that we should regard the fact that these questions about desirability of inserting knowledge are compelling as adding further weight to the claim that they are about something coherently conceivable. Witness the hapless critic who says: knowledge insertion is an incoherent idea, and if it were coherent is would be a disastrous undertaking. Making sense of the last accusation would require rejecting the first accusation, and who can resist having opinions about the second claim?

1. I would like to thank David Aldridge, Meira Levinson, Simon Palmer, and Bryan Warnick for helpful conversations and Alex Carruth, Guy Longworth, and David Rowthorn commenting on drafts. [↑](#footnote-ref-1)
2. The Wachowski Brothers, *The Matrix* (1999, Warner Bros).   [↑](#footnote-ref-2)
3. Susan J. Hespos and Kristy van Marle (2012) ‘Physics for infants: characterizing the origins of knowledge about objects, substances, and number’ WIREs Cognitive Science, 3: pp. 19–27. [↑](#footnote-ref-3)
4. For an influential attempt to reduce knowledge-how to instances of knowledge-that, see Stanley, Jason, and Timothy Williamson, 2001, “Knowing how”, Journal of Philosophy, 92: 411–444. [↑](#footnote-ref-4)
5. Gettier, Edmund L. ‘Is Justified True Belief Knowledge?’ *Analysis* 23.6 (1963): 121-123. A familiar Gettier-style case is that of a person forming a true belief that it is 12pm on the basis of observing a reliable clock that, unbeknownst to himself, stopped 12 hours ago. Here the elements of justification, truth, and belief are all present, but still the subject seems not to know that it is 12pm. [↑](#footnote-ref-5)
6. Some deny this, claiming that either truth, justification or indeed, belief are not necessary to belief. [↑](#footnote-ref-6)
7. For a detailed discussion and literature review of theories of belief, see: Schwitzgebel, Eric, "Belief", The Stanford Encyclopedia of Philosophy (Summer 2015 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/sum2015/entries/belief/>. [↑](#footnote-ref-7)
8. An alternative response might be this: look: it’s possible to insert justified true beliefs, but not justification, but that one need not have access to the justification anyway. It is my believing x that has to be justified, and where am I to find that justification? Can I have it implanted? And if I can, won’t that further justification need to be implanted along with it to justify it? This seems to amount to the same problem that exists for knowledge whether it is inserted or not! [↑](#footnote-ref-8)
9. Williams, B. ‘Deciding to Believe,’ in his *Problems of the Self* (Cambridge: Cambridge University Press, 1970), p. 148 [↑](#footnote-ref-9)
10. It seems to me that the copy of me made on Mars is exactly that: a copy, I would have been destroyed. This can be brought out by considering the following case due to Bernard Williams. Imagine that when Earth John is destroyed, two Mars Johns are created. I cannot be numerically identical with both of them, for they cannot be numerically identical with each other. But since they have as much in common with me as one another, if I am one of them, I must be the other also, but that is absurd. Therefore I cannot be identical with either of them. Williams, Bernard, ‘Personal Identity and Individuation’, Proceedings of the Aristotelian Society, New Series, Vol. 57 (1956 - 1957), pp. 229-252 [↑](#footnote-ref-10)
11. Cliff Sosis ‘What is it like to be a philosopher?’ An interview with Timothy Williamson Available @ <http://www.whatisitliketobeaphilosopher.com/timothy-williamson/> [Last accessed 4/4/2017] [↑](#footnote-ref-11)
12. Donald Davidson (1987) ‘Knowing One's Own Mind’, (Hereafter KOOM) Proceedings and Addresses of the American Philosophical Association, Vol. 60, No.3, pp. 441-458, p. 433 [↑](#footnote-ref-12)
13. There is one point on which I hesitate: is Mars John’s ‘knowledge’ of my mother as good as Earth Johns knowledge of my mother? Perhaps, or perhaps not. Suppose Earth John has died, Earth John’s mother and Mars John, both realize that Mars John is numerically non-identical with Earth John. What will they mean to each other? One way to imagine it is as an awkward scenario. Perhaps Earth John’s mother adopts Mars John as a surrogate, and perhaps Mars John is happy with this. There may be an awkward oscillation between supposing Mars John to be Earth me, and supposing him to be something as good, perhaps made as good by the absence of Earth John. The case is similar to that in which a heart donor does, and the man who takes the heart takes on a son-like mantel for the mother of the heart donor. [↑](#footnote-ref-13)
14. David Chalmers (1996) *The Conscious Mind* (Oxford: Oxford University Press), p.64 [↑](#footnote-ref-14)
15. Dennett, D.C. *Intuition Pumps and Other Tools for Thinking* (New York: WW. Norton & Company, 2013), p. 65. [↑](#footnote-ref-15)
16. Dennett, D.C. *Intuition Pumps and Other Tools for Thinking* (New York: WW. Norton & Company, 2013), p. 66. [↑](#footnote-ref-16)
17. *Descartes' Error: Emotion, Reason, and the Human Brain*, London, Vintage (2006), p. xviii [↑](#footnote-ref-17)
18. Michael Hand, ‘Should We Promote Patriotism in Schools?’ *Political Studies* 59(2) (2011), pp. 328–347, p. 330. [↑](#footnote-ref-18)
19. Lowe, E.J. *Locke on Human Understanding* (London: Routledge, 1995)  [↑](#footnote-ref-19)
20. # Alex Carruth, Imagining zombies, <https://blog.oup.com/2016/03/imagining-zombies-philosophy/>

    [↑](#footnote-ref-20)