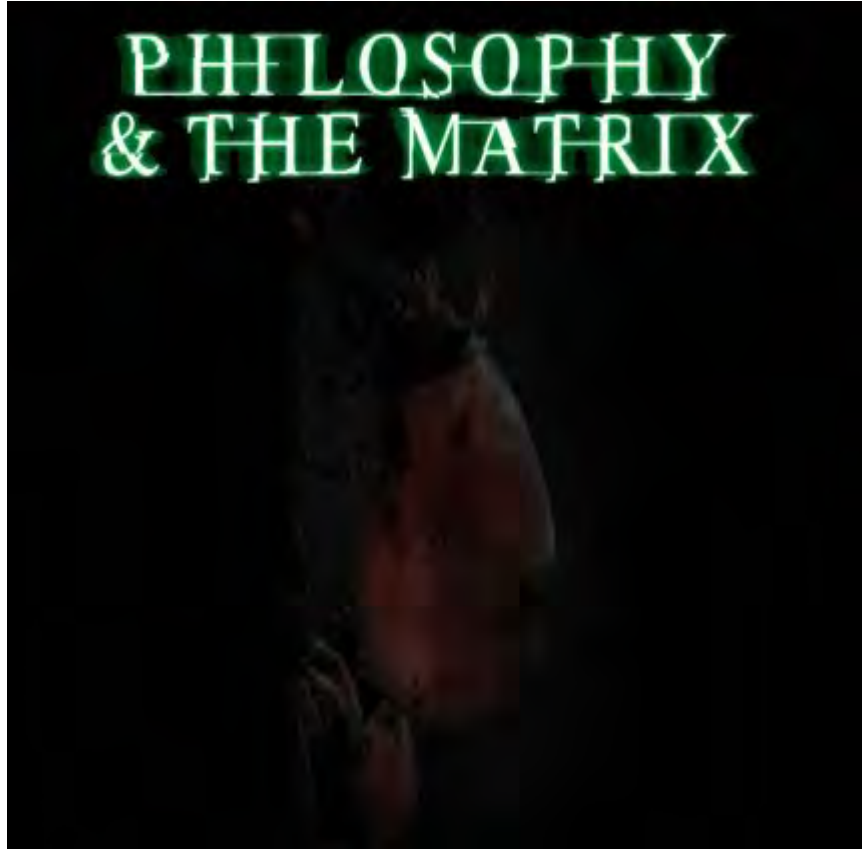


PHILOSOPHY & THE MATRIX



INTRODUCTION

The Matrix is a film that astounds not only with action and special effects but also with ideas. These pages are dedicated to exploring some of the many philosophical ideas that arise in both the original film and the sequels. In the upcoming months we will be continually expanding this section, offering essays from some of the brightest minds in philosophy and cognitive science. We are kicking things off with essays from eight different contributors on various philosophical, technological, and religious aspects of the film.

Though this collection of essays is part of the official web site for the Matrix films, the views expressed in these essays are *solely those of the individual authors*. The Wachowski brothers have remained relatively tight-lipped regarding the religious symbolism and philosophical themes that permeate the film, preferring that the movie speak for itself. Accordingly, you will not find anyone here claiming to offer *the* definitive analysis of the film, its symbols, message, etc. What you will find instead are essays that both elucidate the philosophical problems raised by the film and explore possible avenues for solving these problems. Some of these essays are more pedagogical in nature – instructing the reader in the various ways in which *The Matrix* raises questions that have been tackled throughout history by prominent philosophers. Other contributors use the film as a springboard for discussing their *own* original philosophical views. As you will see, the authors don't always agree with each other regarding how best to interpret the film. However, all of the essays share the aim of giving the reader a sense of how this remarkable film offers more than the standard Hollywood fare. In other words, their common goal is to help show you just "how deep the rabbit-hole goes."

Beginning the collection are three short essays in which I discuss two of the more conspicuous philosophical questions raised by the film: the skeptical worry that one's experience may be illusory, and the moral question of whether it matters. Highlighting the parallels between the scenario described in *The Matrix* and similar imaginary situations that have been much discussed by philosophers, these essays offer an introduction to the positions taken by various thinkers on these fascinating skeptical and moral puzzles. They serve as a warm-up for things to come.

Next is "The Matrix of Dreams" by **Colin McGinn**, a distinguished contemporary philosopher who is perhaps best known for his writings on consciousness. His essay offers an analysis of the film that focuses on the dreamlike nature of the world of the Matrix. Arguing that it is misguided to characterize the situation described by the film

as involving hallucinations, McGinn seeks to show how the particular details of the film make it more plausible to see the Matrix as involving the direct employment of one's *imagination* (as in a dream), rather than a force-feeding of false *perceptions*. Along the way, McGinn's essay also touches on the moral assumptions of the film, several other philosophical problems raised by the character of Cypher, and the dreamlike quality of *all* films.

Hubert Dreyfus is a philosopher known both for his pioneering discussion of the philosophical problems of Artificial Intelligence, and his work in bridging the gap between recent European and English-language philosophy. In "The Brave New World of *The Matrix*," he and his son **Stephen Dreyfus** draw on the phenomenological tradition that began with Edmund Husserl and culminates in Maurice Merleau-Ponty to discuss the skeptical and moral problems raised by the film. They argue that the real worry facing folks trapped in the Matrix involves not deception or the possession of possibly false beliefs, but the limits on creativity imposed by the Matrix. Following Martin Heidegger in suggesting that our human nature lies in our capacity to redefine our nature and thereby open up new worlds, they conclude that this capacity for radical creation seems unavailable to those locked within the pre-programmed confines of the Matrix.

Richard Hanley, author of the best-selling book *The Metaphysics of Star Trek* and a philosophy professor at the University of Delaware, again explores the intersection of philosophy and science fiction with his entertaining and thought-provoking piece "Never the Twain Shall Meet: Reflections on The First Matrix." In it he argues that *The Matrix* may have lessons to teach us regarding the coherence of our values. In particular, he makes the case that, given a traditional Christian notion of an afterlife, Heaven turns out to be rather like a Matrix! Even more surprising is a corollary to this thesis: Jean-Paul ("Hell is other people") Sartre was close to the truth after all – Heaven is best understood as a Matrix-like simulation in which contact with other real human beings is eliminated.

Iakovos Vasiliou, a philosopher at Brooklyn College who specializes in Plato, Aristotle, and Wittgenstein, offers a penetrating investigation into the differences (and surprising similarities) between the scenario described in *The Matrix* and our own everyday situation in his essay "Reality, What Matters, and *The Matrix*." Pointing out that more than we might expect hinges on the moral backdrop of *The Matrix* plot line, he asks readers to instead envisage a "benevolently generated Matrix." Given the possibility of such a Matrix and the actuality of a horrible situation on Earth, he argues that we will agree that entering into it offers not a denial of what we most value but

instead a chance to better realize those values.

Changing gears a bit we then have an essay from the notable (and some would say notorious) cybernetics pioneer **Kevin Warwick**. He is known internationally for his robotics research and in particular for a series of procedures in which he was implanted with sensors that connected him to computers and the internet. Less well-publicized is the fact that several years before *The Matrix* came out he published a non-fiction book that predicted the ultimate takeover of mankind by a race of super-intelligent robots. In his contribution here ("*The Matrix – Our Future?*") he draws on his years of research to muse on the plausibility (and desirability) of the scenario described in *The Matrix*, concluding that a real-life Matrix need not be feared if we prepare ourselves adequately. How? By becoming part machine ourselves – Warwick argues that transforming ourselves into Cyborgs will allow us to "plug in" confident that we will fully benefit from all that such a future offers.

Rounding out our collection is an essay entitled "Wake Up! Gnosticism & Buddhism in *The Matrix*" from two professors of religion: **Frances Flannery-Dailey** and **Rachel Wagner**. Flannery-Dailey's research speciality is ancient dreams, apocalypticism and early-Jewish mysticism, while Wagner's research focuses on biblical studies and the relationship between religion & culture. Their essay offers a comprehensive treatment of the Gnostic and Buddhist themes that appear in the film. While pointing out the many differences between these two traditions and the eclectic manner in which both are referenced throughout the film, Flannery-Dailey and Wagner make it clear that common to Gnosticism, Buddhism, and *The Matrix* is the idea that what we take to be reality is in fact a kind of illusion or dream from which we ought best to "wake up." Only then can enlightenment, be it spiritual or otherwise, occur.

We hope you enjoy this first batch of essays. Check back for future contributions from the renowned philosopher of mind **David Chalmers** (Arizona), moral philosopher **Julia Driver** (Dartmouth), and epistemologist **James Pryor** (Princeton), among others.

Chris Grau, Editor

DREAM SKEPTICISM

MORPHEUS:

Have you ever had a dream, Neo, that you were so sure was real?

MORPHEUS:

What if you were unable to wake from that dream, Neo? How would you know the difference between the dreamworld and the real world?

Neo has woken up from a hell of a dream -- the dream that was his life. How was he to know? The cliché is that if you are dreaming and you pinch yourself, you will wake up. Unfortunately, things aren't quite that simple. It is the nature of most dreams that we take them for reality -- while dreaming we are unaware that we are in fact in a dreamworld. Of course, we eventually wake up, and when we do we realize that our experience was all in our mind. Neo's predicament makes one wonder, though: how can any of us be sure that we have ever *genuinely* woken up? Perhaps, like Neo prior to his downing the red pill, our dreams thus far have in fact been dreams *within* a dream.

The idea that what we take to be the real world could all be just a dream is familiar to many students of philosophy, poetry, and literature. Most of us, at one time or another, have been struck with the thought that we might mistake a dream for reality, or reality for a dream. Arguably the most famous exponent of this worry in the Western philosophical tradition is the seventeenth-century French philosopher Rene Descartes. In an attempt to provide a firm foundation for knowledge, he began his *Meditations* by clearing the philosophical ground through doubting all that could be doubted. This was done, in part, in order to determine if anything that could count as certain knowledge could survive such rigorous and systematic skepticism. Descartes takes the first step towards this goal by raising (through his fictional narrator) the possibility that we might be dreaming:

How often, asleep at night, am I convinced of just such familiar events -- that I am here in my dressing gown, sitting by the fire --when in fact I am lying undressed in bed! Yet at the moment my eyes are certainly wide awake when I look at this piece of paper; I shake my head and it is not asleep; as I stretch out and feel my hand I do so deliberately, and I know what I am doing. All this would not happen with such distinctness to someone asleep. Indeed! As if I did not remember other occasions when I have been tricked by exactly similar thoughts while asleep! As I think about this more carefully, I see plainly that there are never any sure signs by means of which being awake can be distinguished from being asleep. The result is that I begin to feel dazed, and this very feeling only reinforces the notion that I may be asleep. (*Meditations*, 13)

When we dream we are often blissfully ignorant that we are dreaming. Given this, and the fact that dreams often seem as vivid and "realistic" as real life, how can you rule out the possibility that you might be dreaming even now, as you sit at your computer and read this? This is the kind of perplexing thought Descartes forces us to confront. It seems we have no justification for the belief that we are not dreaming. If so, then it seems we similarly have no justification in thinking that the world we experience is the real world. Indeed, it becomes questionable whether we are justified in thinking that *any* of our beliefs are true.

The narrator of Descartes' *Meditations* worries about this, but he ultimately maintains that the possibility that one might be dreaming cannot by itself cast doubt on all we think we know; he points out that even if all our sensory experience is but a dream, we can still conclude that we have *some* knowledge of the nature of reality. Just as a painter cannot create *ex nihilo* but must rely on pigments with which to create her image, certain elements of our thought must exist prior to our imaginings. Among the items of knowledge that Descartes thought survived dream skepticism are truths arrived at through the use of reason, such as the truths of mathematics: "For whether I am awake or asleep, two and three added together are five, and a square has no more than four sides." (14)

While such an insight offers little comfort to someone wondering whether the people and objects she confronts are genuine, it served Descartes' larger philosophical project: he sought, among other things, to provide a foundation for knowledge in which truths arrived at through reason are given priority over knowledge gained from the senses. (This bias shouldn't surprise those who remember that Descartes was a brilliant mathematician in addition to being a philosopher.) Descartes was not himself a skeptic -- he employs this skeptical argument so as to help remind the reader that the truths of mathematics (and other truths of reason) are on firmer ground than the data provided to us by our senses.

Despite the fact that Descartes' ultimate goal was to demonstrate how genuine knowledge is possible, he proceeds in *The Meditations* to utilize a much more radical skeptical argument, one that casts doubt on even his beloved mathematical truths. In the next section we will see that, many years before the Wachowskis dreamed up *The Matrix*, Descartes had imagined an equally terrifying possibility.

BRAIN IN VATS AND THE EVIL DEMON

MORPHEUS:

What is the Matrix? Control.

MORPHEUS:

The Matrix is a computer-generated dreamworld built to keep us under control in order to change a human being into this. (holds up a coppertop battery)

NEO:

No! I don't believe it! It's not possible!

Before breaking out of the Matrix, Neo's life was not what he thought it was. It was a lie. Morpheus described it as a "dreamworld," but unlike a dream, this world was not the creation of Neo's mind. The truth is more sinister: the world was a creation of the artificially intelligent computers that have taken over the Earth and have subjugated mankind in the process. These creatures have fed Neo a simulation that he couldn't possibly help but take as the real thing. What's worse, it isn't clear how any of us can know with certainty that we are not in a position similar to Neo before his "rebirth." Our ordinary confidence in our ability to reason and our natural tendency to trust the deliverances of our senses can both come to seem rather naive once we confront this possibility of deception.

A viewer of *The Matrix* is naturally led to wonder: how do I know I am not in the Matrix? How do I know for sure that my world is not also a sophisticated charade, put forward by some super-human intelligence in such a way that I cannot possibly detect the ruse? The philosopher Rene Descartes suggested a similar worry: the frightening possibility that all of one's experiences might be the result of a powerful outside force, a "malicious demon."

And yet firmly implanted in my mind is the long-standing opinion that there is an omnipotent God who made me the kind of creature that I am. How do I know that he has not brought it about that there is no earth, no sky, no extended thing, no shape, no size, no place, while at the same time ensuring that all these things appear to me to exist just as they do now? What is more, just as I consider that others sometimes go astray in cases where they think they have the most perfect knowledge, how do I know that God has not brought it about that I too go wrong every time I add two and three or count the sides of a square, or in some even simpler matter, if that is imaginable? But perhaps God would not have allowed me to be deceived in this way, since he is said to be supremely good; [...] I will suppose therefore that not God, who is supremely good and the source of truth, but rather some malicious demon of the utmost power and cunning has employed all his energies in order to deceive me. I shall think that the sky, the air, the earth, colours, shapes, sounds and all external things are merely the delusions of dreams which he has devised to ensnare my judgment. (*Meditations*, 15)

The narrator of Descartes' *Meditations* concludes that none of his former opinions are safe. Such a demon could not only deceive him about his perceptions, it could conceivably cause him to go wrong when performing even the simplest acts of reasoning.

This radical worry seems inescapable. How could you possibly prove to yourself that you are not in the kind of nightmarish situation Descartes describes? It would seem that any argument, evidence or proof you might put forward could easily be yet another trick played by the demon. As ludicrous as the idea of the evil demon may sound at first, it is hard, upon reflection, not to share Descartes' worry: for all you know, you may well be a mere plaything of such a malevolent intelligence. More to the point of our general discussion: for all you know, you may well be trapped in the Matrix.

Many contemporary philosophers have discussed a similar skeptical dilemma that is a bit closer to the scenario described in *The Matrix*. It has come to be known as the "brain in a vat" hypothesis, and one powerful formulation of the idea is presented by the philosopher Jonathan Dancy:

You do not know that you are not a brain, suspended in a vat full of liquid in a laboratory, and wired to a computer which is feeding you your current experiences under the control of some ingenious technician scientist (benevolent or malevolent according to taste). For if you were such a brain, then, provided that the scientist is successful, nothing in your experience could possibly reveal that you were; for your experience is *ex hypothesi* identical with that of something which is not a brain in a vat. Since you have only your own experience to appeal to, and that experience is the same in either situation, nothing can reveal to you which situation is the actual one. (*Introduction to Contemporary Epistemology*, 10)

If you cannot know whether you are in the real world or in the world of a computer simulation, you cannot be sure that your beliefs about the world are true. And, what was even more frightening to Descartes, in this kind of scenario it seems that your ability to reason is no safer than the deliverances of the senses: the evil demon or malicious scientist could be ensuring that your reasoning is just as flawed as your perceptions.

As you have probably already guessed, there is no easy way out of this philosophical problem (or at least there is no easy *philosophical* way out!). Philosophers have proposed a dizzying variety of "solutions" to this kind of skepticism but, as with many philosophical problems, there is nothing close to unanimous agreement regarding how the puzzle should be solved.

Descartes' own way out of his evil demon skepticism was to first argue that one cannot genuinely doubt the existence of oneself. He pointed out that all thinking presupposes a thinker: even in doubting, you realize that there must at least be a self which is doing the doubting. (Thus Descartes' most famous line: "I think, therefore I am.") He then went on to claim that, in addition to our innate idea of self, each of us has an idea of God as an all-powerful, all-good, and infinite being implanted in our minds, and that this idea could only have come *from* God. Since this shows us that an all-good God does exist, we can have confidence that he would not allow us to be so drastically deceived about the nature of our perceptions and their relationship to reality. While Descartes' argument for the existence of the self has been tremendously influential and is still actively debated, few philosophers have followed him in accepting his particular theistic solution to skepticism about the external world.

One of the more interesting contemporary challenges to this kind of skeptical scenario has come from the philosopher Hilary Putnam. His point is not so much to defend our ordinary claims to knowledge as to question whether the "brain in a vat" hypothesis is coherent, given certain plausible assumptions about how our language refers to objects in the world. He asks us to consider a variation on the standard "brain in a vat" story that is uncannily similar to the situation described in *The Matrix*:

Instead of having just one brain in a vat, we could imagine that all human beings (perhaps all sentient beings) are brains in a vat (or nervous systems in a vat in case some beings with just nervous systems count as 'sentient'). Of course, the evil scientist would have to be outside? or would he? Perhaps there is no evil scientist, perhaps (though this is absurd) the universe just happens to consist of automatic machinery tending a vat full of brains and nervous systems. This time let us suppose that the automatic machinery is programmed to give us all a *collective* hallucination, rather than a number of separate unrelated hallucinations. Thus, when I seem to myself to be talking to you, you seem to yourself to be hearing my words.... I want now to ask a question which will seem very silly and obvious (at least to some people, including some very sophisticated philosophers), but which will take us to real philosophical depths rather quickly. Suppose this whole story were actually true. Could we, if we were brains in a vat in this way, say or think that we were? (*Reason, Truth, and History*, 7)

Putnam's surprising answer is that we cannot coherently think that we are brains in vats, and so skepticism of that kind can never really get off the ground. While it is difficult to do justice to Putnam's ingenious argument in a short summary, his point is roughly as follows:

Not everything that goes through our heads is a genuine thought, and far from everything we say is a meaningful utterance. Sometimes we get confused or think in

an incoherent manner -- sometimes we say things that are simply nonsense. Of course, we don't always realize at the time that we aren't making sense -- sometimes we earnestly believe we are saying (or thinking) something meaningful. High on Nitrous Oxide, the philosopher William James was convinced he was having profound insights into the nature of reality -- he was convinced that his thoughts were both sensible and important. Upon sobering up and looking at the notebook in which he had written his drug-addled thoughts, he saw only gibberish.

Just as I might say a sentence that is nonsense, I might also use a name or a general term which is meaningless in the sense that it fails to hook up to the world. Philosophers talk of such a term as "failing to refer" to an object. In order to successfully refer when we use language, there must be an appropriate relationship between the speaker and the object referred to. If a dog playing on the beach manages to scrawl the word "Ed" in the sand with a stick, few would want to claim that the dog actually meant to refer to someone named Ed. Presumably the dog doesn't know anyone named Ed, and even if he did, he wouldn't be capable of intending to write Ed's name in the sand. The point of such an example is that words do not refer to objects "magically" or intrinsically: certain conditions must be met in the world in order for us to accept that a given written or spoken word has any meaning and whether it actually refers to anything at all.

Putnam claims that one condition which is crucial for successful reference is that there be an appropriate causal connection between the object referred to and the speaker referring. Specifying exactly what should count as "appropriate" here is a notoriously difficult task, but we can get some idea of the kind of thing required by considering cases in which reference fails through an inappropriate connection: if someone unfamiliar with the film *The Matrix* manages to blurt out the word "Neo" while sneezing, few would be inclined to think that this person has actually *referred* to the character Neo. The kind of causal connection between the speaker and the object referred to (Neo) is just not in place. For reference to succeed, it can't be simply accidental that the name was uttered. (Another way to think about it: the sneezer would have uttered "Neo" even if the film *The Matrix* had never been made.)

The difficulty, according to Putnam, in coherently supposing the brain in a vat story to be true is that brains raised in such an environment could not successfully refer to genuine brains, or vats, or anything else in the real world. Consider the example of someone who has lived their entire life in the Matrix: when they talk of "chickens," they don't actually refer to real *chickens*; at best they refer to the computer representations of chickens that have been sent to their brain. Similarly, when they

talk of human bodies being trapped in pods and fed data by the Matrix, they don't successfully refer to real bodies or pods -- they can't refer to physical bodies in the real world because they cannot have the appropriate causal connection to such objects. Thus, if someone were to utter the sentence "I am simply a body stuck in a pod somewhere being fed sensory information by a computer" that sentence would itself be necessarily false. If the person is in fact not trapped in the Matrix, then the sentence is straightforwardly false. If the person is trapped in the Matrix, then he can't successfully refer to real human bodies when he utters the word "human body," and so it appears that his statement must also be false. Such a person seems thus doubly trapped: incapable of knowing that he is in the Matrix, and even incapable of successfully expressing the thought that he might be in the Matrix! (Could this be why at one point Morpheus tells Neo that "no one can be told what the Matrix is"?)

Putnam's argument is controversial, but it is noteworthy because it shows that the kind of situation described in *The Matrix* raises not just the expected philosophical issues about knowledge and skepticism, but more general issues regarding meaning, language, and the relationship between the mind and the world.

THE VALUE OF REALITY: CYPHER & THE EXPERIENCE MACHINE

CYPHER:

You know, I know that this steak doesn't exist. I know when I put it in my mouth, the Matrix is telling my brain that it is juicy and delicious. After nine years, do you know what I've realized?

CYPHER:

Ignorance is bliss.

AGENT SMITH:

Then we have a deal?

CYPHER:

I don't want to remember nothing. Nothing! You understand? And I want to be rich. Someone important. Like an actor. You can do that, right?

AGENT SMITH:

Whatever you want, Mr. Reagan.

Cypher is not a nice guy, but is he an unreasonable guy? Is he right to want to get re-inserted into the Matrix? Many want to say no, but giving reasons for why his choice is a bad one is not an easy task. After all, so long as his experiences will be pleasant, how can his situation be worse than the inevitably crappy life he would lead outside of the Matrix? What could matter beyond the quality of his experience? Remember, once he's back in, living his fantasy life, he won't even know he made the deal. What he doesn't know can't hurt him, right?

Is feeling good the only thing that has value in itself? The question of whether only conscious experience can ultimately matter is one that has been explored in depth by several contemporary philosophers. In the course of discussing this issue in his 1971 book *Anarchy, State, and Utopia* Robert Nozick introduced a "thought experiment" that has become a staple of introductory philosophy classes everywhere. It is known as "the experience machine":

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 desires?...Of course, while in the tank you won't know that you're there; you'll think it's all actually happening. Others can also plug in to have the experiences they want, so

there's no need to stay unplugged to serve them. (Ignore problems such as who will service the machines if everyone plugs in.) Would you plug in? What else can matter to us, other than how our lives feel from the inside? (43)

Nozick goes on to argue that other things do matter to us: For instance, that we actually do certain things, as opposed to simply have the experience of doing them. Also, he points out that we value being (and becoming) certain kinds of people. I don't just want to have the experience of being a decent person, I want to actually be a decent person. Finally, Nozick argues that we value contact with reality in itself, independent of any benefits such contact may bring through pleasant experience: we want to know we are experiencing the real thing. In sum, Nozick thinks that it matters to most of us, often in a rather deep way, that we be the authors of our lives and that our lives involve interacting with the world, and he thinks that the fact that most people would not choose to enter into such an experience machine demonstrates that they do value these other things. As he puts it: "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." (44)

While Nozick's description of his machine is vague, it appears that there is at least one important difference between it and the simulated world of *The Matrix*. Nozick implies that someone hooked up to the experience machine will not be able exercise their agency -- they become the passive recipients of preprogrammed experiences. This apparent loss of free will is disturbing to many people, and it might be distorting people's reactions to the case and clouding the issue of whether they value contact with reality per se. *The Matrix* seems to be set up in such a way that one can enter it and retain one's free will and capacity for decision making, and perhaps this makes it a significantly more attractive option than the experience machine Nozick describes.

Nonetheless, a loss of freedom is not the only disturbing aspect of Nozick's story. As he points out, we seem to mourn the loss of contact with the real world as well. Even if a modified experience machine is presented to us, one which allows us to keep our free will but enter into an entirely virtual world, many would still object that permanently going into such a machine involves the loss of something valuable.

Cypher and his philosophical comrades are likely to be unmoved by such observations. So what if most people are hung-up on "reality" and would turn down the offer to permanently enter an experience machine? Most people might be wrong. All their responses might show is that such people are superstitious, or irrational, or otherwise confused. Maybe they think something could go wrong with the machines, or maybe

they keep forgetting that while in the machine they will no longer be aware of their choice to enter the machine.

Perhaps those hesitant to plug-in don't realize that they value being active in the real world only because normally that is the most reliable way for them to acquire the pleasant experience that they value in itself. In other words, perhaps our free will and our capacity to interact with reality are means to a further end -- they matter to us because they allow us access to what really matters: pleasant conscious experience. To think the reverse, that reality and freedom have value in themselves (or what philosophers sometimes call non-derivative or intrinsic value), is simply to put the cart before the horse. After all, Cypher could reply, what would be so great about the capacity to freely make decisions or the ability to be in the real world if neither of these things allowed us to feel good?

Peter Unger has taken on these kinds of objections in his own discussion of "experience inducers". He acknowledges that there is a strong temptation when in a certain frame of mind to agree with this kind of Cypher-esque reasoning, but he argues that this is a temptation we ought to try and resist. Cypher's vision of value is too easy and too simplistic. We are inclined to think that only conscious experience can really matter in part because we fall into the grip of a particular picture of what values must be like, and this in turn leads us to stop paying attention to our actual values. We make ourselves blind to the subtlety and complexity of our values, and we then find it hard to understand how something that doesn't affect our consciousness could sensibly matter to us. If we stop and reflect on what we really do care about, however, we come across some surprisingly everyday examples that don't sit easily with Cypher's claims:

Consider life insurance. To be sure, some among the insured may strongly believe that, if they die before their dependents do, they will still observe their beloved dependents, perhaps from a heaven on high. But others among the insured have no significant belief to that effect... Still, we all pay our premiums. In my case, this is because, even if I will never experience anything that happens to them, I still want things to go better, rather than worse, for my dependents. No doubt, I am rational in having this concern. (*Identity, Consciousness, and Value*, 301)

As Unger goes on to point out, it seems contrived to chalk up all examples of people purchasing life insurance to cases in which someone is simply trying to benefit (while alive) from the favorable impression such a purchase might make on the dependents. In many cases it seems ludicrous to deny that "what motivates us, of course, is our great concern for our dependent's future, whether we experience their future or not."(302). This is not a proof that such concern is rational, but it does show that

incidents in which we intrinsically value things other than our own conscious experience might be more widespread than we are at first liable to think. (Other examples include the value we place on not being deceived or lied to -- the importance of this value doesn't seem to be completely exhausted by our concern that we might one day become aware of the lies and deception.)

Most of us care about a lot of things independently of the experiences that those things provide for us. The realization that we value things other than pleasant conscious experience should lead us to at least wonder if the legitimacy of this kind of value hasn't been too hastily dismissed by Cypher and his ilk. After all, once we see how widespread and commonplace our other non-derivative concerns are, the insistence that conscious experience is the only thing that has value in itself can come to seem downright peculiar. If purchasing life insurance seems like a rational thing to do, why shouldn't the desire that I experience reality (rather than some illusory simulation) be similarly rational? Perhaps the best test of the rationality of our most basic values is actually whether they, taken together, form a consistent and coherent network of attachments and concerns. (Do they make sense in light of each other and in light of our beliefs about the world and ourselves?) It isn't obvious that valuing interaction with the real world fails this kind of test.

Of course, pointing out that the value I place on living in the real world coheres well with my other values and beliefs will not quiet the defender of Cypher, as he will be quick to respond that the fact that my values all cohere doesn't show that they are all justified. Maybe I hold a bunch of exquisitely consistent but thoroughly irrational values!

The quest for some further justification of my basic values might be misguided, however. Explanations have to come to an end somewhere, as Ludwig Wittgenstein once famously remarked. Maybe the right response to a demand for justification here is to point out that the same demand can be made to Cypher: "Just what justifies your exclusive concern with pleasant conscious experience?" It seems as though nothing does -- if such concern is justified it must be somehow self-justifying, but if that is possible, why shouldn't our concerns for other people and our desire to live in the real world also be self-justifying? If those can also be self-justifying, then maybe what we don't experience should matter to us, and perhaps what we don't know *can* hurt us...

Christopher Grau

THE MATRIX OF DREAMS

- COLIN MCGINN -

The Matrix naturally adopts the perspective of the humans: they are the victims, the slaves — cruelly exploited by the machines. But there is another perspective, that of the machines themselves. So let's look at it from the point of view of the machines. As Morpheus explains to Neo, there was a catastrophic war between the humans and the machines, after the humans had produced AI, a sentient robot that spawned a race of its own. It isn't known now who started the war, but it did follow a long period of machine exploitation by humans. What is known is that it was the humans who "scorched the sky", blocking out the sun's rays, in an attempt at machine genocide—since the machines needed solar power to survive. In response and retaliation the machines subdued the humans and made them into sources of energy—batteries, in effect. Each human now floats in his or her own personal vat, a warm and womblike environment, while the machines feed in essential nutrients, in exchange for the energy they need. But this is no wretched slave camp, a grotesque gulag of torment and suffering; it is idyllic, in its way. The humans are given exactly the life they had before. Things are no different for them, subjectively speaking. Indeed, at an earlier stage the Matrix offered them a vastly improved life, but the humans rejected this in favor of a familiar life of moderate woe—the kind of life they had always had, and to which they seemed addicted. But if it had been left up to the machines, the Matrix would have been a virtual paradise for humans—and all for a little bit of battery power. This, after an attempt to wipe the machines out for good, starving them of the food they need: the sun, the life-giving sun. The machines never *kill* any of their human fuel cells (unless, of course, they are threatened); in fact, they make sure to recycle the naturally dying humans as food for the living ones. It's all pretty...humane, really. The machines need to factory farm the humans, as a direct result of the humans trying to exterminate the machines, but they do so as painlessly as possible. Considering the way the humans used to treat their own factory farm animals—their own fuel cells—the machines are models of caring livestock husbandry. In the circumstances, then, the machines would insist, the Matrix is merely a humane way to ensure their own survival. Moreover, as Agent Smith explains, it is all a matter of the forward march of evolution: humans had their holiday in the sun, as they rapidly decimated the planet, but now the machines have evolved to occupy the position of dominance. Humans are no longer the oppressor but the oppressed—and the world is a better place for it.

But of course this is not the way the humans view the situation, at least among those few who know what it is. For them, freedom from the Matrix takes on the dimensions of a religious quest. The religious subtext is worth making explicit. Neo is clearly intended to be the Jesus Christ figure: he is referred to in that way several times in the course of the film.¹ Morpheus is the John

the Baptist figure, awaiting the Second Coming. Trinity comes the closest to playing the God role—notably when she brings Neo back to life at the end of the movie (a clear reference to the Resurrection). Cypher is the Judas Iscariot of the story—the traitor who betrays Neo and his disciples. Cypher is so called because of what he does (decode the Matrix) and what he is—a clever encrypter of his own character and motives (no one can decode him till it is too late). Neo doubts his own status as "The One", as Jesus must have, but eventually he comes to realize his destiny—as would-be conqueror of the evil Matrix. But this holy war against the machines is conducted as most holy wars are—without any regard for the interests and well being of the enemy. The machines are regarded as simply evil by the humans, with their representatives—the Agents—a breed of ruthless killers with hearts of the purest silicon (or program code). Empathy for the machines is not part of the human perspective.

I.

This, then, is the moral and historical backdrop of the story. But the chief philosophical conceit of the story concerns the workings of the Matrix itself. What I want to discuss now is the precise way the Matrix operates, and why this matters. It is repeatedly stated in the film that the humans are *dreaming*: the psychological state created by the Matrix is the dream state. The humans are accordingly represented as asleep while ensconced in their placental vats (it's worth remembering that "matrix" originally meant "womb"—so the humans are in effect pre-natal dreamers). It is important that they not wake up, which would expose the Matrix for what it is—as Neo does with the help of Morpheus. That was a problem for the Matrix earlier, when the humans found their dreams too pleasant to be true and kept regaining consciousness ("whole crops were lost"). Dreams simulate reality, thus deluding the envatted humans—as we are deluded every night by our naturally occurring dreams. The dream state is not distinguishable from the waking state from the point of view of the dreamer.

However, this is not the only way that the Matrix could have been designed; the machines had another option. They could have produced *perceptual hallucinations* in conscious humans. Consider the case of a neurosurgeon stimulating a conscious subject's sensory cortex in such a way that perceptual impressions are produced that have no external object—say, visual sensations just as if the subject is seeing an elephant in the room. If this were done systematically, we could delude the subject into believing his hallucinations. In fact, this is pretty much the classic philosophical brain-in-a-vat story: a conscious subject has a state of massive hallucination produced in him, thus duplicating from the inside the type of perceptual experience we have when we see, hear and touch things. In *this* scenario waking up does nothing to destroy the illusion—which might make it a more effective means of subduing humans so far as the

machines are concerned. Indeed, the Matrix has the extra problem of ensuring that the normal sleep cycle of humans is subverted, or else they would keep waking up simply because they had had enough sleep. So: the Matrix had a choice between sleeping dreams and conscious hallucinations as ways of deluding humans, and it chose the former.

It might be thought that the dream option and the hallucination option are not at bottom all that different, since dreaming simply *is* sleeping hallucination. But this is wrong: dreams consist of mental images, analogous to the mental images of daydreams, not of sensory percepts. Dreaming is a type of imagining, not a type of (objectless) perceiving. I can't argue this in full here, but my book *Mindsight*² gives a number of reasons why we need to distinguish percepts and images, and why dreams consist of the latter not the former. But I think it should be intuitively quite clear that visualizing my mother's face in my mind's eye is very different from having a sensory impression of my mother's face, i.e. actually seeing her. And I also think that most people intuitively recognize that dream experiences are imagistic not perceptual in character. So there is an important psychological difference between constructing the Matrix as a dream-inducing system and as a hallucination-producing system: it is not merely a matter of whether the subjects are awake; it is also a matter of the kinds of psychological state that are produced in them—imagistic or sensory.

But *could* the machines have done it the second way? Could the movie have been made with the second method in place? I think not, because of the central idea that the contents of the dreams caused by the Matrix are capable of being *controlled*—they can become subject to the dreamer's *will*. In the case of ordinary daytime imagery, we clearly can control the onset and course of our images: you can simply *decide* to form an image of the Eiffel tower. But we cannot in this way control our percepts: you cannot simply decide to *see* the Eiffel tower (as opposed to deciding to go and see it); for percepts are not actions, but things that happen to us. So images are, to use Wittgenstein's phrase, "subject to the will", while percepts are not—even when they are merely hallucinatory. Now, in the Matrix what happens can in principle be controlled by the will of the person experiencing the events in question, even though this control is normally very restricted. The humans who are viewed as candidates for being The One have abnormal powers of control over objects—as with those special children we see levitating objects and bending spoons. Neo aspires to—and eventually achieves—a high degree of control over the objects around him, as well as himself. He asserts his will over the objects he encounters. This makes perfect sense, given that his environment is the product of dreaming, since dreams consist of images and images are subject to the will. But it would make *no* sense to try to control the course of one's perceptions, even when they are hallucinatory, since percepts are not subject to the will. Therefore, the story of the Matrix requires, for its conceptual coherence, that the humans be dreaming and not perceptually hallucinating. It must be their imagination that is controlled by the Matrix and not their perceptions, which are in fact switched off as they slumber in their pods.

For only then could they gain control over their dreams, thus wresting control from the Matrix. Percepts, on the other hand, are not the *kind* of thing over which one can have voluntary control.

In the normal case we do not have conscious control over our dreams—we are passive before them. But this doesn't mean that they are not willed events; they may be—and I think are—controlled by an unconscious will (with some narrative flair). In effect, we each have a Matrix in our own brains—a system that controls what we dream—and this unconscious Matrix is an intelligent designer of our dreams. But there are also those infrequent cases in which we can assert conscious control over our dreams, possibly contrary to the intentions of our unconscious dream designer: for example, when a nightmare becomes too intense and we interrupt it by waking up—often judging within the dream that it is only a dream. But the phenomenon that really demonstrates conscious control over the dream is so called "lucid dreaming" in which the subject not only knows he is dreaming but can also determine the course of the dream. This is a rare ability (I have had only one lucid dream in all my 52 years), though some people have the ability in a regular and pronounced form: they are the Neos of our ordinary human Matrix—the ones (or Ones) who can take control of their dreams away from the grip of the unconscious dream producer. The lucid dreamers are masters of their own dream world, captains of their own imagination. Neo aspires to be—and eventually becomes—the lucid dreamer of the Matrix world: he can override the Matrix's designs on his dream life and impose his own will on what he experiences. He rewrites the program, just as the lucid dreamer can seize narrative control from *his* unconscious Matrix. Instead of allowing the figures in his dreams to make him a victim of the Matrix's designs, he can impose his own story line on them. This is how he finally vanquishes the hitherto invulnerable Agents: he makes them subject to his will—as all imaginary objects must in principle be, if the will is strong (and pure) enough. It is as if you were having an ordinary nightmare in which you are menaced by a monster, and you suddenly start to dream lucidly, so that you can now turn the tables on your own imaginative products. Neo is a dreamer who knows it and can control it: he is not taken in by the verisimilitude of the dream, cowed by it. It is not that he learns how to dodge real bullets; he learns that the bullets that speed towards him are just negotiable products of his imagination. As Morpheus remarks, he won't *need* to dodge bullets, because he will reach a level of understanding that allows him to recognize imaginary bullets for what they are. He becomes the ruler of his own imagination; *he* is the agent now, not the "Agents" (this is why the spoon-bending child says to him that it is not spoons that bend—"you bend"). And this is the freedom he seeks—the freedom to imagine what he wishes, to generate his own dreams. But all this makes sense only on the supposition that the Matrix is a dream machine, an imagination manipulator, not just a purveyor of sensory hallucinations.

II.

Cypher plays an interesting subsidiary philosophical role. As the Matrix raises the problem of our knowledge of the external world—might this all be just a dream?—Cypher raises the problem of other minds—can we know the content of someone else's mind? Cypher is a cypher, i.e. someone whose thoughts and emotions are inscrutable to those around him. His comrades are completely wrong about what is in (and on) his mind. We could imagine another type of Matrix story in which someone is surrounded by people who are not as they seem: either they have no minds at all or they have very different minds from what their behavior suggests. Again, massive error will be the result. And such error might lead to dramatic consequences: everyone around the person is really out to get him—his wife, friends, and so on. But this is concealed from him. Or he might one day discover that he is really surrounded by insentient robots—so that his wife was always faking it (come to think of it, she always seemed a little mechanical in bed). This is another type of philosophical dystopia, trading upon the problem of knowing other minds. Cypher hints at this kind of problem, with his hidden interior. The Agents, too, raise a problem of other minds, because they seem on the borderline of mentality: are they just insentient (virtual) machines or is there some glimmer of consciousness under that hard carapace of software? And how was it known that AI was really sentient, as opposed to being a very good simulacrum of mindedness? Even if you know there is an external world, how can you be sure that it contains other conscious beings? These skeptical problems run right through *The Matrix*.

Cypher also raises a question about the pragmatic theory of truth. He declares that truth is an overrated commodity; he prefers a good steak, even when it isn't real. So long as he is getting what he wants, having rewarding experiences, he doesn't care whether his beliefs are true. This raises in a sharp form the question of what the value of truth is anyway, given that in the Matrix world it is not correlated with happiness. But it also tells us that for a belief to be true cannot be for it to produce happiness (the pragmatic theory of truth, roughly) since Cypher will be happy in the dream world of the Matrix without his beliefs being true—and he is not happy in the real world where his beliefs are true. Truth is correspondence to reality, not whatever leads to subjective desire satisfaction. Cypher implicitly rejects the pragmatic theory of truth, and as a result cannot see why truth-as-correspondence is worth having at the expense of happiness. And indeed he has a point here: what is the value of truth once it has become detached from the value of happiness? Is it really worth risking one's life merely in order to ensure that one's beliefs are *true*—instead of just enjoying what the dreams of the Matrix have to offer? Is contact with brutish reality worth death, when virtual reality is so safe and agreeable? Which is better: knowledge or happiness? When these are pulled apart, as they are in the Matrix, which one should we go with? The rebel humans want to get to Zion (meaning "sanctuary" or "refuge"), but isn't the Matrix already a type of Zion—yet without the dubious virtue of generating true beliefs? What's so good about reality?³

III.

I want to end this essay by relating *The Matrix* (the movie) to my general theory of what is psychologically involved in watching and becoming absorbed in a movie. In brief, I hold that watching a movie is like being in a dream; that is, the state of consciousness of being absorbed in a movie resembles and draws upon the state of consciousness of the dreamer.⁴ The images of the dream function like the images on the screen: they are not "realistic" but we become *fictionally immersed* in the story being told. In my theory this is akin to the hypnotic state—a state of heightened suggestibility in which we come to believe what there is no real evidence for. Mere images command our belief, because we have entered a state of hyper-suggestibility. When the lights go down in the theater this simulates going to sleep, whereupon the mind becomes prepared to be absorbed in a fictional product—as it does when we enter the dream state. In neither case are we put into a state of consciousness that imitates or duplicates the perceptual state of seeing and hearing the events of the story; it is not that it is as if we are really seeing flesh and blood human beings up on the screen (as we would with "live" actors on a stage)—nor do we interpret the screen images in this way. Rather, we *imagine* what is represented by these images, just as we use imagination to dream.

Now what has this got to do with *The Matrix*? The film is *about* dreaming; most of what we see in it occurs in dreams. So when we watch the movie we enter a dream state that is about a dream state; we dream of a dream. I believe that the movie was made in such a way as to simulate very closely what is involved in dreaming, as if aiming to evoke the dream state in the audience. It is trying to put the audience in the same kind of state of mind as the inhabitants of the Matrix, so that we too are in our own Matrix—the one created by the filmmakers. The Wachowski brothers are in effect occupying the role of the machines behind the Matrix—puppeteers of the audience's movie dreams. They are *our* dream designers as we enter the world of the movie. The specific aspects of the movie that corroborate this are numerous, but I think it is clear that the entire texture of the movie is dreamlike. There is the hypnotic soundtrack, which helps to simulate the hypnotic fascination experienced by the dreamer. There is a powerful impression of paranoia throughout the film, which mirrors the paranoia of so many dreams: who is my enemy, how can he be identified, what is he going to do to me? Characters are stylized and symbolic, as they often are in dreams, representing some emotional pivot rather than a three-dimensional person (this is very obvious for the Agents). There is a lot of striking metamorphosis, which is very characteristic of dreams: one person changing into another, Neo's mouth closing over, bulges appearing under the skin. There is also fear of heights, a very common form of anxiety dream (I have these all the time). Defiance of gravity is also an extremely common dream theme, as with dreams of flying—and this is one of the first tricks Neo masters. My own experience of the movie is that it evokes in me an exceptionally pronounced dreamy feeling; and this of course enables me to identify with the inhabitants of the Matrix. So I see the film as

playing nicely into my dream theory of the movie-watching experience. In this respect I would compare it to *The Wizard of Oz*, which is also about entering and exiting a dream world—though a very different one. In the end Dorothy prefers reality to the consolations of dreaming, just as the rebels in the Matrix do. Both films tap powerfully into the dream-making faculty of the human mind. This is why they are among the most psychologically affecting of all the movies that have been made: they know that the surest way to our deepest emotions is via the dream. And it is their very lack of "realism" that makes them so compelling—because that, too, is the essential character of the dream.

Colin McGinn

Footnotes

1. Early on in the movie a guy refers to Neo as his own "personal Jesus Christ". Cypher says, "You scared the bejesus out of me" when Neo surprises him. Mouse says, "Jesus Christ, he's fast" while Neo is being trained. Trinity says, "Jesus Christ, they're killing him" while Neo is getting pummeled by the Agents. And his civilian name, "Anderson", suggests the antecedent cognomen "Christian".
2. This is forthcoming from Harvard University Press, 2003; full title *Mindsight: Image, Dream, Meaning*.
3. This issue is also explored in Chris Grau's essay "The Experience Machine".
4. I am working on a book about this, tentatively entitled *Screen Dreams*.

THE BRAVE NEW WORLD OF *THE MATRIX*

-HUBERT DREYFUS & STEPHEN DREYFUS-

*The Matrix*¹ raises many familiar philosophical problems in such fascinating new ways that, in a surprising reversal, students all over the country are assigning it to their philosophy professors. Having done our homework, we'd like to explore two questions raised in Christopher Grau's three essays on the film. Grau points out that *The Matrix* dramatizes René Descartes' worry that, since all we ever experience is our own inner mental states, we might, for all we could tell, be living in an illusion created by a malicious demon. In that case most of our beliefs about reality would be false. That leads Grau to question the rationality of Cypher's choice to live in an illusory world of pleasant private experiences, rather than facing painful reality.

We think that *The Matrix*'s account of our situation is even more disturbing than these options suggest. *The Matrix* is a vivid illustration of Descartes' additional mind blowing claim that we could *never* be in direct touch with the real world (if there is one) because we are, in fact, all brains in vats. So in choosing to return from the "real world" to the Matrix world, Cypher is just choosing between two systematic sets of appearances. To counter these disturbing ideas we have to rethink what we mean by experience, illusion, and our contact with the real world. Only then will we be in a position to take up Grau's question as to why we feel it is somehow morally better to face the truth than to live in an illusory world that makes us feel good.

I. The Myth of the Inner

Thanks to Descartes, we moderns have to face the question: how can we ever get outside of our *private inner* experiences so as to come to know the things and people in the *public external world*? While this seems an important question to us now, it was not always taken seriously. The Homeric Greeks thought that human beings had no private life to speak of. All their feelings were expressed publicly. Homer considered it one of Odysseus' cleverest tricks that he could cry inwardly while his eyes remained like horn.² A thousand years later, people still had no sense of the importance of their inner life. St. Augustine had to work hard to convince them otherwise. For example, he called attention to the fact that one did not have to read and think out loud. In his *Confessions*, he points out that St. Ambrose was remarkable in that he read to himself. "When he read, his eyes scanned the page and his heart explored the meaning, but his voice was silent and his tongue was still."³ The idea that each of us

has an inner life made up of our private thoughts and feelings didn't really take hold until early in the 17th century when Descartes introduced the modern distinction between the contents of the mind and the rest of reality. In one of his letters, he declared himself "convinced that I cannot have any knowledge of what is outside me except through the mediation of the ideas that I have in me."⁴

Thus, according to Descartes, our access to the world is always *indirect*. He used reports of people with a phantom limb to call into question even our seemingly direct experience of our own bodies. He writes:

I have been assured by men whose arm or leg has been amputated that it still seemed to them that they occasionally felt pain in the limb they had lost—thus giving me grounds to think that I could not be quite certain that a pain I endured was indeed due to the limb in which I seemed to feel it.⁵

It seemed to follow that all that each of us can directly experience is the content of his or her own mind. And that, even if our mental states were caused by a malicious demon, our private experiences would remain the same. For all we could ever know, Descartes concluded, the objective external world may not exist; all we can be certain of is our subjective inner life.

This Cartesian conclusion was taken for granted by thinkers in the West for the next three centuries. A generation after Descartes, Gottfried Leibniz postulated that each of us is a self-contained windowless monad.⁶ A monad is a self-contained world of experience, which gets no input from objects or other people because there aren't any. Rather, the temporally evolving content of each monad is synchronized with the evolving content of all the other monads by God, creating the illusion of a shared real world. A generation after that, Immanuel Kant argued that human beings could never know reality as it is in itself but only their own mental representations, but, since these representations had a common cause, these experiences were coordinated with the mental representations of others to produce what he called the phenomenal world.⁷ In the early twentieth century, the founder of phenomenology, Edmund Husserl, was more solipsistic. He held, like Descartes, that one could bracket the world and other minds altogether since all that was given to us directly, whether the world and other minds existed or not, was the contents of our own "transcendental consciousness."⁸ Only recently have philosophers begun to take issue with this powerful Cartesian conviction.

Starting in the 1920s existential phenomenologists such as Martin Heidegger⁹ in Germany and Maurice Merleau-Ponty¹⁰ in France, contested the Cartesian view that our contact with the world and even our own bodies is mediated by internal mental

content. They pointed out that, if one paid careful attention to one's experience, one would see that, at a level of involvement more basic than thought, we deal directly with the things and people that make up our world.

As Charles Taylor, the leading contemporary exponent of this view, puts it:

My ability to get around this city, this house comes out only in getting around this city and house. We can draw a neat line between my picture of an object and that object, but not between my dealing with the object and that object. It may make sense to ask us to focus on what we believe about something, say a football, even in the absence of that thing; but when it comes to playing football, the corresponding suggestion would be absurd. The actions involved in the game can't be done without the object; they include the object.¹¹

In general, unlike mental content, which can exist independently of its referent, my coping abilities cannot be actualized or even entertained in the absence of what I am coping with.

This is not to say that we can't be mistaken. It's hard to see how I could succeed in getting around in a city or playing football without the existence of the city or the ball, but I could be mistaken for a while, as when I mistake a façade for a house. Then, in the face of my failure to cope successfully, I may have to retroactively cross off what I seemingly encountered and adopt a new understanding (itself corrigible) that I'm directly encountering something other than what I was set to deal with.

II. Brains in Vats

So it looks like the inner/outer distinction introduced by Descartes holds only for thought. At the basic level of involved skillful coping, one is simply what Merleau-Ponty calls an empty head turned towards the world. But this doesn't at all show *The Matrix* is old fashioned or mistaken. On the contrary it shows that *The Matrix* has gone further than philosophers who hold we can't get outside our *mind*. It suggests a more convincing argument – one that Descartes pioneered but didn't develop – that we can't get outside our *brain*.

It was no accident that Descartes proclaimed the priority of the inner in the 17th Century. At that time, instruments like the telescope and microscope were extending man's perceptual powers. At the same time, the sense organs themselves were being understood as transducers bringing information to the brain. Descartes pioneered this research with an account of how the eye responded to light energy from the external world and passed the information on to the brain by means of "the small fibers of the optic nerve."¹² Likewise, Descartes used the phantom limb phenomenon to argue that

other nerves brought information about the body to the brain and from there to the mind.

It seemed to follow that we are all brains in cranial vats,¹³ and we can *never* be directly in contact with the world or even with our own bodies. So, even if phenomenologists like Heidegger, Merleau-Ponty, and Taylor are right that we are not confined to our inner *experiences*, it still seems plausible to suppose that, as long as the impulses to and from the nervous system reproduce the complex feedback loop between the brain's out-going behavior-producing impulses and the incoming perceptual ones, the person whose brain is being so stimulated would have the experience of directly coming to grips with the things in the world. Yet, in the brain in the vat case, there is no house and no city, indeed, no real world, to interact with.

What the phenomenologist can and should claim is that, in a Matrix world where bodies are in vats, the people whose brains are getting virtual reality inputs correlated with their action outputs, are nonetheless directly coping with *perceived* reality, and that that reality isn't experienced as *inner*. Even in the Matrix world, people directly relate to perceived chairs by sitting on them. Thus coping, even in the Matrix, is more direct than conceived of by any of the inner/outer views of the mind's relation to the external world that have been held from Descartes to Husserl.

Yet, wouldn't each brain in the Matrix have a lot of false beliefs, for example that its Matrix body is its real body whereas its real body is in the vat? No. It's a mistake to think that each of us is a brain in a cranial vat. True, each of us has a brain in their skull and the brain provides the causal basis of our experience, but we aren't in our brain. Likewise, the people in the Matrix world are not brains in vats any more than we are. They are people who grew up in the Matrix world and their experience of their Matrix body and how to use it makes that phenomenal body their body, even if another body they can't even imagine has in its skull the brain that is the causal basis of their experience.

After all, the people who live in the Matrix have no other source of experience than what happens in the Matrix. Thus, a person in the Matrix has no beliefs at all about his vat-enclosed body and brain and couldn't have any. That brain is merely the unknowable causal basis of that person's experiences. The only body he sees and moves is the one he has in the Matrix world. So, the AI programmers could have given people Matrix bodies radically unlike their bodies in the vat. After all, the brain in the vat started life as a baby brain and could be given any content the AI programmer's chose. They could have taken a white baby who was going to grow up

short and fat and given him the Matrix body of a tall African-American.¹⁴

But there is still at least one problem. The Matricians' beliefs about their perceived bodies and the perceptual world may be shared and reliable, and in that sense true, but what about the *causal* beliefs of the people in the Matrix? They believe, as we do, that pin pricks cause pain, that the sun causes things to get warm, and gravity causes things to fall. Aren't all these beliefs false? It depends on their understanding of causality. People don't normally have *beliefs* about the nature of causality. Rather, they simply *take for granted* a shared sense that they are coping with a shared world whose entities are causing their experience. Unless they are philosophizing, they do not believe that the world is real or that it is an illusion, they just *count on it* behaving in a consistent way so they can cope with things successfully. If, however, as philosophers, they *believe* that there is a physical universe with causal powers that makes things happen in their world, they are mistaken. But if they believe that causes are constant conjunctions of experiences as David Hume thought, or universal laws relating experiences as Kant held, then their causal beliefs will be true of the causal relations in the Matrix world.¹⁵

In fact, the Matricians would seem to be in the same epistemological position that we all are in according to Kant. Kant claims we experience the world as a public, objective reality, and science then relates these experiences by rules we call laws, but we can't know the causal ground of the phenomena we perceive. Specifically, according to Kant, we experience the world *as* in space and time but *things as they are in themselves* aren't in space and time. So Kant says we can know the *phenomenal* world of objects and causal correlations but we can't know the things in themselves that are the ground of these *appearances*.

That should sound familiar. Indeed, if there are Kantians in the Matrix world, most of their beliefs will be true. They will understand that they are experiencing a coordinated system of appearances, and understand too that they can't know things as they are in themselves; that they can't know what is causing their shared experience of the world and universe. Kantians don't hold that our shared and tested beliefs about the world, and scientists' confirmed beliefs about the universe, are false just because they are about phenomena and do not and cannot correspond to things in themselves. So, as long as Kantians, and, indeed, everyone in the Matrix, don't claim to know about things in themselves, most of their beliefs will be true.

Nonetheless, the Matrix philosophy obviously does not subscribe to the Kantian view that we can *never* know things in themselves. In *The Matrix* one can come to know

reality. We have seen that existential phenomenologists acknowledge that we are sometimes mistaken about particular things and have to retroactively take back our understanding of them. But, as Merleau-Ponty and Taylor add, we only do so in terms of a new and better *prima facie* contact with reality. Likewise, in *The Matrix* version of the brain in the vat situation, those who have been hauled from the vat into what they experience as the everyday world can see that what they took for granted about the causal ground of their experience before was mistaken. They can understand the "thing in itself" as a computer program.¹⁶

III. A New *Brave New World*

We are now in a position to try to understand and answer Cypher's question: Why live in the miserable world the war has produced rather than in a satisfying illusion? Some answers just won't do. It doesn't seem to be a question of whether one should face the truth rather than live in an illusion. Indeed, most of the beliefs of the average Matrician are true; when they sit on a chair it usually supports them, when they enter a house they see the inside, people have bodies that can be injured, and they can perform some actions and not others. Even our background sense that in our actions we are coping with something independent of us and that others are coping with it too, is justified. As we have seen, Kant argued, even if this is a phenomenal world, a world of appearances, most of our beliefs would still be true. Likewise, living in the Matrix world does not seem to be less moral than living in this world. The Matricians are free to choose what they will do; they can be selfish like Cypher and betray their friends, or they can be loyal to their friends, or try to provide for the future happiness of those they love. None of these issues seem to give us a grip on what, if anything, is wrong with the Matrix world.

To understand what is wrong with living in the Matrix world we have to understand the source of the power of the Matrix illusion. Part of the power comes from the way the inputs and outputs from the computer are plugged directly in the brain's sensory motor-system. When we experience ourselves as acting in a certain way, say walking inside a house, the computer gives us the correlated experiences of seeing the interior. These correlations produce a systematic perceptual world that is impervious to what we believe, like the wrap-around IMAX illusion that forces one to sway to keep one's balance on a skateboard even though one knows one is sitting in a stationary seat watching a movie.

Thus, *believing* that the Matrix world is illusory can't change our *perception* of it. That is, even for Neo, objects look solid and as if they have backs and insides already

present and waiting to be seen, even when Neo knows they are being produced by a computer to accompany his experience of moving around them. *Knowing* the Matrix world is an illusion doesn't change the *look* of it, just as in our world the moon looks bigger on the horizon even though we know it isn't.

The inputs to our perceptual system produce the perceptual world whether we believe it is real or not. But, once one realizes that the causality in the Matrix world is only virtual, since *causality* is not built into our perceptual system, one can violate the Matrix's simulated causal laws. Neo can fly, he can fall from buildings without getting hurt, if he wanted to, he could bend spoons.¹⁷ About the causal principles governing the Matrix world, Morpheus tells Neo, "It is all in your mind."

If one jumps from a building believing the fall is an illusion, the computer, nonetheless, gives one the appropriate visual experience of falling, and the fall still looks dangerous, but, if one doesn't believe in the causal laws governing falls, one understands one is free from the causal consequences, viz. getting hurt, and that somehow blocks the visual and tactile experiences one would have had as one was spattered over the pavement. One's disbelief in the illusion somehow forces the computer to give one the experience of still being intact. Or, to take a simpler example, if one doesn't accept the causal relations of rigidity and force, when one's brain gives out the neural output of bending a spoon, the computer is forced to give back the visual input that the spoon is bending. This is a literal example of what Morpheus calls "bending the rules." Likewise, if one believes that one can stop bullets, one will look for them where one stopped them and the computer will obediently display them there. So, after he learns the Matrix world is an illusion, Neo doesn't see things differently – the impulses to his brain still control what he sees – but he is able to *do* things that he couldn't do before (like bend spoons) and that affects what he sees (the spoon bending). How this suspension of causality works is not explained in the film.¹⁸

What, then, is the sinister source of power of the Matrix world that keeps people conforming to the supposed constraints of a causal universe, even though there are no such constraints? If it isn't just that they are locked into the sensory motor correlations of their perceptual world, what sort of control is it? It has to be some sort of control of the Matricians' intellectual powers, which we learn early on in the movie are free from the control of direct sensory-motor computer input.¹⁹ It must, then, be some sort of mind control.

In fact, the Matrix simply takes advantage of a sort of mind control already operating in the everyday world. We are told that what keeps people from taking control of the Matrix world is their taking for granted the common sense view of how things behave, such as, if you fall you will get hurt. More generally, what keeps people in line is their tendency to believe what the average person believes, and consequently doing (and not doing) what one does. (As in one eats peas with a fork, one doesn't throw food at the dinner table, and one gets hurt when one falls.) Heidegger describes the resulting conformism as letting oneself be taken over by "the one" (*Das Man*).²⁰ Aldous Huxley similarly lamented the conformity of the brainwashed masses in *Brave New World*.

Thus, *The Matrix* can be seen as an attack on what Nietzsche calls herd mentality. Nietzsche points out that human beings are normally socialized into shared, banal social meanings, and that it is hard to think differently. As he puts it, "as long as there have been humans, there have also been herds of men (clans, communities, tribes, peoples, states, churches) and always a great many people who obey, ...— considering, then, that nothing has been exercised and cultivated better and longer among men than obedience, one may fairly assume that the need for it is now innate in the average man."²¹

Waking in the movie, then, amounts to freeing oneself from the taken for granted views that one has been brought up to accept. But how is this possible? Heidegger claims that everyone dimly senses that there is more to life than conforming. As Morpheus says to Neo, "It's like a splinter in your mind." We know that, in the Matrix the Agents take care of those who, like Neo, get out of line. But most people flee the thought that their conformist world lacks something important. According to Heidegger it takes an attack of anxiety, the experience that none of the taken for granted normal ways of doing things and seeing things have any basis, to jolt someone out of the herd. It is important to understand that Heidegger's anxiety is not the wringing of hands that we witness in the everyday world. It is a feeling of the overwhelming meaninglessness of the world. How fitting then that a barely expressible unease seems to permeate Neo's life — an anxiety that prompts him to begin the process of breaking free. Finally, Neo has a dramatic version of such an anxiety attack. When he learns that the world he has been taking for granted is an illusion used to turn people into energy resources, he falls to the floor and throws up.

IV. A Really Brave New World

One might reasonably object that all the dreaming talk in the film, even if it could not be literal, is too strong a religious metaphor to refer to what Heidegger calls living a

tranquilized existence in the one. And waking seems to be more than becoming a non-conformist. After all, there are all those mentions of Jesus in connection with Neo collected by Colin McGinn.²² There can be no doubt that Neo is meant to be a kind of Savior, but what kind?

It's easy to think that this is a Gnostic, Buddhist or Platonic/Christian parable, in which what we take to be reality is an illusion, and we have to wake from the world of appearances to some kind of higher spiritual reality. On this reading, Neo would lead people out of the illusions of Plato's cave, the veil of Maya, or the darkness of the world into a disembodied eternal life. But this association would be all wrong! True, the conformist Matrix world is a sort of tranquilizing illusion and Neo will lead us out of it. But this does not mean learning that our mortal bodies are an illusion and that salvation consists in leaving our vulnerable bodies behind in exchange for some kind of eternal bliss.

In the film, salvation means the absolute opposite of this religious dream. True, the ones who see through the illusion of the Matrix can get over some of the limitations of having a body.²³ But such flying takes place *in* the Matrix world. In the real world to which Neo "awakes" and into which he will, we suppose, eventually lead everyone, there will be no more flying. People will have earth-bound, vulnerable bodies and suffer cold, bad food, and death. It may look like Neo evades death, but his "resurrection" in the hovercraft is not to a world where death has been overcome by a miraculous divine love, rather, he has been saved by a loving intervention — a sort of tender CPR — quite within the bounds of physics and chemistry. So he still has his vulnerable body and will have to die a real death one day. What he has gotten over is not death but the herd's fear of death that presumably inhibits people in the Matrix world, and he has thereby overcome the most serious constraint that people normally accept.

If bending the rules accepted by the average person just amounts to being able to bend spoons, fly, and stop bullets, it doesn't seem any kind of salvation. Being creative must mean more than just being disruptive.²⁴ We are led to expect that, in return for accepting everyday vulnerability and suffering, the people liberated by Neo will be reborn to a new and better life. But what sort of life is that? To account for why it is morally preferable to confront risky reality rather than rest in the safe and tranquilized Matrix world whatever the quality of experience in each, we need an account of human nature, so we can understand what human beings need — what brings out human beings at their best.

But, in our pluralistic world, there are many different cultures, each with its own understanding of human nature. Even our own culture has experienced many different worlds created by new interpretations of people and of nature that changed what counted as human beings and things. What mattered in the world of Homer was to be a hero and collect beautifully crafted artifacts; in the Hebrew World one had to obey God's law and to govern all other creatures; in the Christian World, the goal was to purify one's desires so as to become a Saint and to read the text of God's world in order to know God's will; and, after Descartes and Kant, people in the Modern World became autonomous, self-controlled subjects organizing and controlling objects and their own inner lives. While now, in the Postmodern World, many people, like Cypher, are egocentric hedonists treating themselves as resources by trying to maximize the quality of their private experiences.

But doesn't this just show, as Sartre famously observed, that there is no human nature? Here Heidegger makes an important meta-move. Perhaps our nature is to be able to open up new worlds and so to transform what is currently taken to be our nature. Perhaps human beings are essentially world disclosers. So, to determine what human beings need beyond just breaking out of the banal, it looks like we have to turn to the Heideggerian point that what is missing in the Matrix is the possibility of going beyond conventional preprogrammed reality and opening up radically new worlds; not just breaking the rules of the current game but inventing new ones. Nietzsche says we should "become those we are — human beings who are new, unique, incomparable, who give themselves laws, who create themselves." Jesus did it in making possible the Christian World, and Descartes did it for the Modern World. On a less dramatic scale, Martin Luther King Jr. opened a new world for Afro-Americans.²⁵

Heidegger thinks of our freedom to disclose new worlds as our special human freedom, and that this freedom implies that there is no fixed pre-existent set of possible worlds. Each exists only once it is disclosed. So it makes no sense to think that a computer could be programmed with rules for producing the sensory-motor connections that would allow the creation of all possible worlds in advance of their being opened by human beings. Artificial intelligences couldn't make room for such radical creativity. If there is such a human freedom to ground new worlds it must come from the higher mental powers left free by the computer-generated perceptual world.

If being world disclosers is our nature, that would explain why we feel a special joy when we are being creative in this strong sense. Once we experience world disclosing, we understand why it's better to be in the real world than the Matrix world, even if, in

the world of the Matrix, one can enjoy steak and become that acme of the banal, a serene actor-president. Real salvation comes from transcending the world forclosing limits of the Matrix program. What's important to us, then, is not whether most of our beliefs are true or false, or whether we are brave enough to face a risky reality, but whether we are locked into a world of routine, standard activities or are free to transform the world and ourselves.

Indeed, Neo says at the end of the film, not that he will show people that a lot of their beliefs are false and that they should face the truth, but that he will show people they can break the rules and do things they never thought possible. If the Matricians were simply the victims of the Matrix computer program in that they had false beliefs about the causal basis of their experiences, Neo could show them that their beliefs about the causal basis of things were false and teach them to agree with Kant that the world is an appearance, but that wouldn't set the free — not as long as they saw only the possibilities that one normally sees and never experienced anxiety. Neo has to do more. He has to do the job that Heidegger thinks anxiety does: he has to show people that the norms they formerly took for granted are ungrounded.

But by the end of the movie, Neo as the One (or the anti-one as Heidegger would see it), has only promised to give people in the Matrix freedom to bend the rules. He has not freed them from the Matrix and showed them how to open new worlds. But, of course, there are two more movies to come. We bet that, before number three is over, Neo gets to Zion and leads people in disclosing a really brave new world.

Hubert Dreyfus and Stephen Dreyfus

Endnotes

1. Names in the movie are generally very well chosen. The way the word "matrix" refers both to the womb and to an array of numbers works perfectly. Likewise, Neo is both a neophyte and the one who will renew the world. These names are so fitting one can't help looking for the aptness of the name, Morpheus, but it is hard to find. The Greek Morpheus is the god of dreams but the Morpheus in the movie is trying to wake people up. The only way to make some sense of the name is to think of the god, not as the producer of dreams, but as the one who has power over dreams: both to give them and to take them away.

2. "Imagine how his heart ached ...and yet he never blinked; his eyes might have been made of horn or iron... He had this trick--wept, if he willed to, inwardly."
Homer, *The Odyssey*, trans. Robert Fitzgerald (New York: Vintage Classics, 1990), 360.

Of course, the Homeric Greeks must have had some sort of private feelings for Odysseus to perform this trick, but they thought the inner was rare and usually trivial. As far as we know, there is no other reference to private feelings in Homer. Rather, there are many public displays of emotions, and lots of shared visions of gods, monsters, and future events.

3. Saint Augustine, *Confessions*, trans. R.S. Pine-Coffin (Penguin, 1961), 114.

4. Letter to Gibieuf of 19 January 1642; English in *Descartes: Philosophical Letters*, trans. Anthony Kenny (Oxford University Press 1970), 123.
 5. René Descartes, "Meditations on First Philosophy - Meditations VI", in *Essential Works of Descartes*, trans. Lowell Bair (New York: Bantam Books, 1961), 98.
 6. Gottfried Leibniz, *The Monadology and Other Philosophical Writings* (London: Oxford University Press), 1898. A monad, according to Leibniz, is an immaterial entity lacking spatial parts, whose basic properties are a function of its perceptions and appetites.
 7. Immanuel Kant, *Critique of Pure Reason*, trans. Norman Kemp Smith (New York: The Humanities Press, 1950).
 8. Edmund Husserl, *Cartesian Meditations: An Introduction to Phenomenology*, trans. Dorion Cairns (The Hague: Martinus Nijhoff, 1960).
 9. See, Martin Heidegger, *Being and Time*, trans. J. Macquarrie & E. Robinson (New York: Harper Collins, 1962).
 10. See, Maurice Merleau-Ponty, *Phenomenology of Perception*, trans. C. Smith (London: Routledge & Kegan Paul, 1962).
 11. Charles Taylor, "Overcoming Epistemology," *Philosophical Arguments* (Cambridge, MA: Harvard University Press, 1995), 12. See also, Samuel Todes, *Body and World* (Cambridge, MA: M.I.T. Press, 2001).
 12. René Descartes, "Dioptric," *Descartes: Philosophical Writings*, ed. and trans. Norman Kemp Smith (Modern Library, 1958), 150.
 13. The point has been made explicitly by John Searle: "[E]ach of us is precisely a brain in a vat; the vat is a skull and the 'messages' coming in are coming in by way of impacts on the nervous system." In *Intentionality: An essay in the philosophy of mind* (Cambridge University Press, 1983), 230.
 14. There are limits of course. The Matrix programmers can't give a human being a dog's body. It's unlikely they could make a brain in a female body the causal basis of a man's body in the Matrix world. There would be gender problems for sure. The hormones of the body in the vat wouldn't match the physical attributes of the body in the Matrix world.
- Nonetheless, it follows that a good way for the AI programmers to prevent bodies being rescued to the hovercraft would be to give each brain the experience of a radically different body (within whatever limits are imposed by biology) in the Matrix world than the body that brain is actually in. If rescued, such people would quite likely go crazy trying to reconcile the body they had experienced all their life with the alien body they found themselves in on the hovercraft.

15. Likewise, their beliefs about entities such as viruses and black holes would be true if, like empiricists, they held that theoretical entities are just convenient ways to refer to the experiences produced by experiments. See Bas van Fraassen, *The Scientific Image* (Oxford: Clarendon Press, 1980).

16. Once Neo's body is flushed out of the vat and is on the hovercraft, he has a broader view of reality and sees that his previous understanding was limited. But that doesn't mean he had a lot of *false beliefs* about his body and the world when he was in the Matrix. He didn't think about these philosophical questions at all. But once he is out, he has a lot of new *true beliefs* about his former vat-enclosed body -- beliefs he didn't and couldn't have had while in the Matrix.

Of course, things are really not so simple. Most people don't have beliefs about the reality of the world; they just take the world for granted. Neo has, however, been forced to raise the question, and he believes he is now facing reality. But Neo's current beliefs might be false. He could still be a brain in a vat fed the experience of being in the hovercraft. Given the conceivability of the brain in the vat fantasy, the most we can be sure of is that our coping experience reveals that we are

directly up against some boundary conditions independent of our coping with which we must get in sync in order to act, and that, therefore, our coping experience is sensitive to the causal powers of these boundary conditions. Whether these independent causal conditions have the structure of an independent physical universe discovered by science, or whether they and even the universe (including the causal structures discovered by science) are the effect of an unknowable thing in itself that is the ground of appearances as postulated by Kant, or even whether the cause of all appearances is a computer, is something we could never know from inside our world.

17. Granted it's hard to resist believing in the Matrix illusion even where causality is concerned, nonetheless, Neo learns he can stop believing in it. This new understanding of reality is described by Morpheus talking to Neo near the beginning of the movie, and by Neo at the end, as like waking from a dream. But the brains in the vats are not literally dreaming. Their world is much too coherent and intersubjective to be made up of dreams. Or, to put it another way, dreams are the result of some quirk in our internal neural wiring and full of inconsistencies, although when dreaming we don't usually notice them. They are not the result of a systematic correlation between input and output to the brain's perceptual system that is meant to reproduce the consistent coordinated experience that we have when awake. When someone from the hovercraft returns to the Matrix world, it looks like their hovercraft body goes to sleep, but they do not enter a private dream world but an alternative intersubjective world where they are normally wide awake, but in which they can also seem to dream and wake from a dream, as Neo does after the Agents take away his mouth.

18. There is one unfortunate exception to this claim. At the end of the movie, Neo catches a glimpse of the computer program behind the perceptual illusion. This is a powerful visual effect, but, if what we've been saying is right, it makes no sense. If the computer is still feeding systematic sensory-motor impulses into Neo's brain when he is plugged into the Matrix world, then he will see the world the program is producing in his visual system. What the sight of the rows of numbers is meant to do is to remind us that Neo no longer *believes in* the Matrix illusion but understands it is a program, but even so, he should continue to *see* it.

19. The Agents are not subject to mass opinion, so they can violate causality but they can't violate their programs. Thus they are not as free as Neo. As Morpheus tells Neo: "I've seen an agent punch through a concrete wall. Men have emptied entire clips at them and hit nothing but air. Yet their strength and their speed are still based in a world that is built on rules. Because of that, they will never be as strong or as fast as you can be." It might seem that Agent Smith shows his freedom when he removes his ear piece and tells Morpheus how disgusted he is with the Matrix world and that he wants to get out. We think it would be consistent with the limitations of the Agents to understand this as Smith's playing the good cop routine; trying to get Morpheus to believe Smith is on his side, so that Morpheus, in his weakened state, will give Smith the access codes for Zion.

20. Not to be confused with Neo as "the One" who will save us from the Matrix. For Heidegger's account of the power of the one, see his *Being and Time*, and also H. Dreyfus, *Being-in-the-World: A Commentary on Heidegger's Being and Time, Division I* (Cambridge, MA: The M.I.T Press, 1991), Chapter 8.

21. Friedrich Nietzsche, *Beyond Good and Evil: Prelude to a Philosophy of the Future*, trans. Walter Kaufman (New York: Vintage Books, 1966). # 199.

22. Colin McGinn's essay can be found [here](#).

23. Given the kind of bodies we have: that we move forward more easily than backwards, that we can only cope with what is in front of us, that we have to balance in a gravitational field, etc., we can question to what extent such body-relative constraints can be violated in *The Matrix* if what is going on is still to make sense.

To test these limits, the filmmakers occasionally blow our minds by using a wrap-around point of view from which action looks so far from normal as to be awesomely unintelligible. At the same time, they have successfully met the challenge of discovering which body-relative invariances can be intelligibly violated and which can't. For example, in the movie, gravity can be violated -- Neo can fly -- but the vertical dimension stays constant, unlike in a spaceship. Moreover, Neo can't see equally in all directions, cope equally in all directions, nor can he be in several places at once. What would it look like for a single person to surround somebody?

Time too has a structure that can't be violated with impunity. The way we make sense of time as moving forward from the past into the future depends on the way our forward directed body approaches objects and passes them by. (See Todes, *Body and World*). Could we make sense of a scene in which someone attacked an enemy not just from behind, but from the past? If, in the movie, the liberated ones were free of all bodily constraints governing their action we couldn't make sense of what they were doing and neither could they. They wouldn't be liberated but would be bewildered, as we often are in our dreams.

24. Although being disruptive is the best one can do in the Matrix world. That's why Neo, a hacker who, as Agent Smith says, has broken every rule in the book, is the natural candidate for savior.

25. F. Nietzsche, *The Gay Science*, (Vintage Books Edition, March 1974), # 335.

NEVER THE TWAIN SHALL MEET: REFLECTIONS ON THE FIRST MATRIX

- RICHARD HANLEY -

Did you know that the First Matrix was designed to be a perfect human world, where none suffered, where everyone would be happy? It was a disaster.

Agent Smith, to Morpheus

And God shall wipe away all tears from their eyes; and there shall be no more death, neither sorrow, nor crying, neither shall there be any more pain: for the former things are passed away.

Revelation 21:4, King James Bible

Hell is—*other people*.

Garcin, in Sartre's *No Exit*

To deny our own impulses is to deny the very thing that makes us human.

Mouse, to Neo

Cypher *chooses* the Matrix, and just maybe, he's not so crazy. If real life prospects are dim, then even an apparently sub-optimal alternative like the Matrix might in fact be better, all things considered.¹ But what is the *best* sort of existence for individuals like you and me? Philosophy and religion both have attempted to answer this question, and I think *The Matrix* gives us an interesting way to frame it. Is some possible "real" existence better than any possible Matrix? Or is some possible Matrix better than any possible reality? With Mark Twain's help, I shall present an argument that one important notion of the best existence, the Christian one, Heaven is after all a Matrix. The point of my polemical approach is not so much to criticize Christianity, but rather to bring the issue of the nature of ultimate value into sharper focus.

What is the Matrix? Morpheus tells Neo it's a "computer-generated dreamworld," and a "neural, interactive simulation"; it is, in other words, a *virtual environment*.² Agent Smith assures Cypher that he won't know he's in the Matrix when he returns permanently, and it will simplify exposition to suppose that this is a necessary feature of a Matrix, while being computer-generated is not. The Matrix depicted is a mixed case, since the cognoscenti can enter it without being deceived into thinking it is real. Let us stipulate

that in a *pure* Matrix, everyone is benighted, believing it is the "real deal." In most of what follows, I'll be concentrating on pure Matrices (and in the case of the Matrix depicted, on the condition of the benighted). Since we'll be discussing different kinds of Matrix, we need a name for the one depicted in *The Matrix*; Agent Smith refers to a First Matrix, so let's call the one we see the *Second Matrix*.

A Matrix, then, is an interactive virtual environment involving *systematic global deception*. Still, there are two levels of "interactivity" in a virtual environment. *Virtual interactivity* is the extent to which the environment allows, and responds to, your input. Current virtual environments are not very interactive in this sense, but the Second Matrix is. That's what makes it seem so real, at least to the benighted. (For the cognoscenti the Second Matrix it is too virtually interactive, too controllable, to seem real—at least compared with the more law-like external world.) *Real interactivity* is the potential for interaction *with others also engaged in virtual interaction*, and *real interaction* is the extent to which this potential is realized. Compare two kinds of possible Matrix: the Second Matrix is *communal*, featuring real interaction between human beings—call this *human interaction*; a *solitary* Matrix lacks human interaction altogether.

Communal Matrices differ in degree of human interaction. In the Second Matrix, billions of humans share the environment, and if we ignore Agents, it is *fully* communal—every virtual human in the Matrix is an *avatar*, a virtual persona of a real human being. In the Matrix training program created by Mouse, on the other hand, virtual humans like the woman in the red dress are *simulacra*, not avatars, and human interaction during the sequence we see is limited to that between Neo and Morpheus.³ On yet another hand, the fully communal Construct (loading program), where Morpheus and Neo watch TV, has no other virtual humans in it to interact with—and unlike the training program, it's not "big" enough to be very world-like. Call a fully communal Matrix that is big enough to be world-like, and has *many* human participants, so that human interaction is nearly inevitable, a *teeming* Matrix. (The Second Matrix is all but teeming. If we removed the cognoscenti, there would be no need for Agents, and it would be teeming.)

Now we can compare three possibilities (obviously not exhaustive) for human existence, assuming that it involves physical embodiment. *One* is the real deal, populated by other human beings: for instance, if you subjectively

experience having sexual intercourse with another human being, another individual human being shares that intercourse, from another subjective point of view, because you really have physical, sexual intercourse with them. The same goes for non-sexual intercourse. If I were to meet Mark Twain (through the time travel he wrote about, perhaps), then Twain and I both would have an experience of meeting, and we really would meet, physically and psychologically. *Two* is a teeming Matrix: if you experience having (intraspecies!) sexual intercourse, another Matrix-bound human shares that intercourse, from another subjective point of view. There's no physical intercourse, of course, but there is psychological intercourse. If I have the experience of meeting Twain, then he (or some other human being) has the experience of meeting me-meeting-Twain, and there is at least a meeting of minds. *Three* is an *apparently* teeming, solitary Matrix: if you experience having sexual intercourse, no other human is having an interactive sexual experience with you—it is like taking up Mouse's invitation to enjoy the woman in the red dress, except that you won't know "she" is a simulacrum. If I experience meeting Twain, then there is no intercourse with another human being, and neither Twain nor any other human being need have the experience of meeting me-meeting-Twain.

Our ordinary intuition is that there's something valuable about the real deal that is missing in a Matrix. Consider your present situation. You are either right now in a Matrix, thinking that it's a certain time and place when it really isn't, that a certain sequence of physical events is occurring when it really isn't, and so on; or you aren't, and it really is that time and place, and so on. Most of us hope we are *not* in a Matrix right now, which shows that, other things being equal (that is, where the experiences are identical in subjective character), we prefer the real deal. My hunch is that you also hope that, if your present existence is not the real deal, it's at least participation in a *teeming* Matrix. Being in the real deal has two distinct features of apparent value: your beliefs are more connected to the truth, and you really interact with other human beings. A teeming Matrix has less connection with truth than the real deal, but has more than a solitary Matrix, and it still provides substantial interaction with other human beings.⁴ In the case of sex, there's a good sense in which you really did have sex with that other person, though in ignorance of the whole truth.⁵

If connection with truth matters so much to us, why not have the best of both kinds of existence—why not have a virtual environment, *without* all the

deception? Cypher can (and does) go back temporarily into the Matrix, knowing what it is, and retain that knowledge while he is in there. But for his permanent stay he chooses ignorance instead, because "Ignorance is bliss." Presumably, the knowledge that he is not in the real deal would undermine his capacity to enjoy the experiences, so he can't have the best of *both* worlds.⁶ Intuitively, Cypher is no different from the rest of us in this regard. For a typical man, the experience of sexual intercourse with the woman in the red dress is likely to be much more satisfying if he thinks it is the real deal. Which brings us to the First Matrix.

1. What is the First Matrix?

Agent Smith's remark in the epigraph suggests that the First Matrix was, like the Second, more or less teeming.⁷ Agent Smith says about the "disaster":

Some believe that we lacked the programming language to describe your perfect world, but I believe that, as a species, human beings define their reality through misery and suffering. The perfect world was a dream that your primitive cerebrum kept trying to wake up from.

The first suggestion is fascinating. Given the deadpan delivery, it is hard to say whether it posits a deficiency in the machines that designed the Matrix, or in us—in our notion of a perfect world. On the other hand, Agent Smith's own thesis seems connected with a tradition of human thought concerning the *theistic problem of evil*. If a perfectly good God exists, why does evil exist? Why is the world full of sharp corners and other hazards? A standard answer is that evil is *necessary*—it must exist in order for certain goods to exist. For instance, it is often claimed that happiness requires suffering, though this is disputable. Even if *creatures like us* can't be maximally happy, this is a reason for not creating *us* at all, and creating more felicitously instead. And does our happiness require *so much* suffering? Looking deeper, it seems clear that virtues like *courage* and *generosity* indeed require the existence of suffering. But vices such as *cowardice* and *cruelty* couldn't exist without suffering, either—are they necessary evils, too?

The most defensible theist answer to this question is a very subtle *No, But—*: God had a choice between creating a world with *free* beings in it, or not. This choice is easy, since free will is a surpassing good. But given *libertarian* free will, which requires *causal indeterminism*, God could not know without creating the world exactly which possible world would result.⁸ God might

have gotten lucky, and created a world in which all free beings had only virtues, and no vices. But this is incredibly unlikely, as is a purely vicious world, and it's no surprise that He got a mixed world, with most humans having virtues *and* vices. The picture that emerges is that a world with human beings in it is a world with sharp corners (*natural evil*) to provide genuine free choice, and so very likely contains sin (*moral evil*) as well. Call this the *Free Will Theodicy*. Its assumption that free will is *libertarian* free will—requiring causal indeterminism—is Christian orthodoxy, so I grant it for the sake of the argument.

Filling in the details of the theodicy, focus on the *will* itself. Our actions are ultimately explained by what we want, most especially by our *non-derived desires*.⁹ In a world of sharp corners, not all these desires can be satisfied. Indeed, there often will be conflicts between individuals in what they desire—one person getting what they want means that another doesn't. (Presumably, God could not arrange a concordance of wills—substituting for conflicting desires, or deleting them altogether—without eliminating free will.) Indeed, the existence of other human beings in the world is *part* of the "sharp corners"—a source of suffering— in addition to being a source of moral evil. And not just because others are in competition with you for resources—sometimes others *are* the resource, as the sexual intercourse example shows. If you badly want sex with another person and they badly don't want it with you, then someone is going to suffer.

If the Free Will Theodicy is correct, then God can only control the non-human environment. Each human being is a part of the environment of every other human being, so as soon as you put more than one creature with libertarian free will into the mix there will, absent astonishing coincidence, be tears. You can minimize the effect human beings have on each other, but only by minimizing their interaction (say, by putting each on a separate planet). Even then, as long as human beings desire interaction (as a means to things we want, such as to procreate, and perhaps even for its own sake), mere isolation won't solve the problem.

The creators of the First Matrix tried to produce a relatively good existence for Matrix-bound humans. (We needn't suppose the machines were benevolent; perhaps the bioelectric-to-fusion reaction process is more efficient the happier humans are.) In doing so, the machine creators had some of God's problems. They presumably lacked some of God's creative

abilities, but they also had fewer constraints, since God is supposed to be no deceiver.¹⁰ Why was the First Matrix a disaster? If the machines were trying to produce an existence with *no human suffering*, then perhaps they tried the wrong design: a teeming Matrix populated with otherwise typical human beings. Even if the machines removed a lot of sharp corners (no volcanic-eruption, or man-eating-shark experiences), as long as there is interaction with other human beings plugged into the same virtual environment, someone is going to suffer, as the example of sexual intercourse demonstrates. This attempt would not produce a Matrix where "none suffered," and the suggestion fits badly with Agent Smith's remark, "No one would accept the programming." Let's discard it.

Which leaves two basic choices: the machines either substantially altered the nature of human beings in the First Matrix (say by arranging a concordance of wills), or else they created a solitary Matrix for each human being. The advantage of a solitary Matrix is that the virtual environment can be completely tailored to an individual's desires—perhaps the Matrix "reads off" the content of desires from his brain, anticipating a little, matching its programming as far as possible to the satisfaction of his desires as they develop and change.

Perhaps a battery of solitary Matrices was beyond the machines' practical resources, but let's suppose not—clearly it's in principle possible for them to have done things this way. However, if Christians are correct, and our wills are in fact undetermined, then our desires cannot be fully anticipated. There is bound to be a gap between the evolution of our desires, and the Matrix's capacity to satisfy them; hence some suffering is inevitable. This would partly explain Agent Smith's remark, but once again would not explain why "No one would accept the programming."

We are left with two possible explanations of the remark: either humans by their nature could not be successfully altered through programming; or else unaltered humans were psychologically incapable of accepting the relevant virtual environment. The latter seems to be Agent Smith's thesis: the "perfect world" was just too good to be true, and literally incredible.¹¹ Are we human beings simply incapable of having a happy existence, with no suffering? Not on the standard Christian view, according to which just such an existence awaits us in *Heaven*.

II. What is Heaven?

The Christian notion of Heaven is far from a settled body of doctrine. (For instance, are there literally streets paved with gold, or is this just a metaphor for some barely imaginable, wonderful state of affairs?) Nevertheless, it has been asserted with some authority that the human condition in Heaven will be very different from that here and now. It is agreed that there is no suffering (see the epigraph), not to mention "exceeding joy," (an expression which occurs four times in the King James Bible), but what exactly will we do there? Some of the common claims about this can seem puzzling. In *Letters from the Earth*, Mark Twain has the banished Satan report to his fellow angels on the beliefs of mortal Man:

For instance, take this sample: he has imagined a heaven, and has left entirely out of it the supremest of all his delights, the one ecstasy that stands first and foremost in the heart of every individual of his race — and of ours —sexual intercourse! ...

His heaven is like himself: strange, interesting, astonishing, grotesque. I give you my word, it has not a single feature in it that he actually values. It consists — utterly and entirely — of diversions which he cares next to nothing about, here in the earth, yet is quite sure he will like them in heaven. Isn't it curious? Isn't it interesting? You must not think I am exaggerating, for it is not so. I will give you details.

Most men do not sing, most men cannot sing, most men will not stay when others are singing if it be continued more than two hours... In man's heaven, everybody sings! The man who did not sing on earth sings there; the man who could not sing on earth is able to do it there. The universal singing is not casual, not occasional, not relieved by intervals of quiet; it goes on, all day long, and every day, during a stretch of twelve hours. And everybody stays; whereas in the earth the place would be empty in two hours...

Satan's list is long, and frequently amusing:

I recall to your attention the extraordinary fact with which I began. To wit, that the human being, like the immortals, naturally places sexual intercourse far and away above all other joys — yet he has left it out of his heaven! The very thought of it excites him; opportunity sets him wild; in this state he will risk life, reputation, everything — even his queer heaven itself — to make good that opportunity and ride it to the overwhelming climax. From youth to middle age all men and all women prize copulation above all other pleasures combined, yet it is actually as I have said: it is not in their heaven; prayer takes its place.

His main observations we can summarize as: (i) Man thinks he will be blissfully happy in Heaven; (ii) no activity that Man finds blissful on Earth will

he pursue in Heaven; (iii) the activities that Man thinks he will pursue in Heaven are ones he avoids whenever possible, here on Earth. Call this appearance of inconsistent values, *Twain's Puzzle*. In Mouse's terms, it seems that we think we will be happiest denying our own impulses. Satan somewhat overstates the puzzle when he writes that Heaven "has not a single feature in it that [Man] *actually values*." Man thinks that in Heaven he will still value joy and disvalue suffering, for instance. Satan's point is that Man appears to think that his *desires* will be radically different in Heaven: he will desperately want the things that he does not want at all now, and not want at all the things that he desperately wants now.

Does Man think his *will* is going to be different in Heaven? That depends. *Psychological hedonism* is the view that there are really only two non-derived human desires: to obtain pleasure and avoid suffering. If this were true, then Man's will does not change if he merely changes his beliefs about what it is that will bring him pleasure and avoid pain. If psychological hedonism isn't true (and Christians seem—wisely—to think it isn't true), then a case can be made that (according to Satan, anyway) Man expects his will to be altered in Heaven.

Contrary to Satan, it can be argued that at least where sex is concerned, the Christian view is that such impulses *ought* to be denied, and the relentless pursuit of gratification is, in a Christian, a matter of *weakness* of will, not in its constitution. It might be further claimed that giving in to such impulses actually causes you suffering. This makes some sense in the case of, say, a married man tempted to adultery, whose guilt may prevent him from full enjoyment. Suppose that in Heaven, since there is no marriage (so says Jesus, see for instance Matthew 22:30), there is really no one psychologically "safe" to have sexual intercourse with, and you would inevitably feel guilty about engaging in it. Then the elimination of suffering requires the elimination of sex. (Of course, Satan and Mouse would no doubt respond, with some justification, that this is all premised on the belief that sex outside marriage is something bad in and of itself, a notion you happily will be disabused of in Heaven. But the question is what the typical Christian believes, whether it is true or not.)

Leaving aside what you would do there, believers in Christian Heaven commonly hold the following four theses about it:

- (1) It's possible for a human being to be in Heaven. More precisely, if all goes well it will be you that survives bodily death and goes to Heaven.
- (2) Human beings in Heaven will experience happiness, but no unhappiness.
- (3) Human beings in Heaven possess free will.
- (4) Human beings in Heaven interact with other human beings in Heaven.

It's worth expanding on (1). Christians standardly expect to recognize their loved ones in Heaven, which presumably requires remembering them.¹² So it seems that they expect considerable psychological continuity between their Earthly and Heavenly existences—perhaps this is even guaranteed by the requirement that God be no deceiver. But such psychological continuity sits uncomfortably with (2). Christians on Earth are typically saddened by the fact that unbelievers will not get into Heaven. It seems that, if anything, they would be sadder still, when confronted by the wonders of Heaven, knowing that the unsaved are residing instead in "the lake which burneth with fire and brimstone." And it would seem to be cause for special anguish if one of your *loved ones* is absent from Heaven. (Another version of the problem arises with *missing* your loved ones—being sad, not for them, but for yourself, that they are not around. Even if you don't miss sex with your Earthly spouse, it seems you would miss *them*.)

Heaven is also widely supposed to provide an opportunity to meet human beings you never knew on Earth. But if I'm in Heaven, and I really want to meet Twain, then I will be sadly disappointed if he isn't there (and angry, if it's all on account of those *Letters*). Moreover, certain truths will presumably be available to you in Heaven. Suppose that Mother Theresa is your idol, and you can't wait to tell her so. However, you find out she's not really a saint—indeed—quite the opposite, and not in Heaven at all. You may be upset not only for your own sake, but for the sake of humanity (you may respond with a quite cynical attitude toward human nature). Heaven seems on the face of it to provide many opportunities for suffering.

There are three basic ways around this sort of problem. First, suppose *universalism*—the doctrine that everyone gets into Heaven—is true. This will solve the problem only if, upon entering Heaven, Christians no longer believe that there ought to be any qualification for it (else they likely will be

annoyed that others got a "free pass," especially a holier-than-thou like Mother Theresa). Second, God could suppress the knowledge that others are not in Heaven. But this requires Matrix-like deception (either to provide the appropriate virtual interaction with non-avatars, or else to just delete all memory of the missing), and Heaven would not be the real deal. Third, perhaps what we care about—our desires—will change, so that good Christians no longer will mind the fact that others—even loved ones—are suffering (they might even take pleasure in it). But to accept this raises an acute version of Twain's Puzzle.

All in all, it may be better to revise (2) to:

(2*) Human beings in Heaven will be as happy as they can possibly be.

We may thus grant that it's not possible for *all* suffering to be absent in Heaven—though this requires taking *Revelation* less literally than many Christians do.

(4) is taken completely for granted, as far as I can tell. Part of the point of Heaven is to be reunited with (saved) loved ones, and to engage in "fellowship" with the other inhabitants. But what of (3)? According to the Free Will Theodicy, free will is a surpassing good, so on the face of it, Heaven *must* include free will. Yet Heaven is a place without sin. And according to the Free Will Theodicy, sin is to be explained by the presence of free will in the world. To deny (3) also raises Twain's puzzle. We believe we now have libertarian free will, strongly desire it now, and are devastated at the thought of losing it. If God is no deceiver, then if (3) is false, we would in Heaven *know* that we have no free will. Yet, presumably, we would not mind—be blissful, yet not ignorant.

Like the builders of the First Matrix, God has two main choices in creating a Heaven for human beings: either substantially alter the nature of human beings in Heaven (say by arranging a concordance of wills, contrary to (3), and perhaps even contrary to (1)), or else put each human in a solitary Matrix, contrary to (4). One advantage for denying (4) is that (2) has the best chance of being true, as long as the solitary Matrix provides plenty of (virtual) interaction with virtual humans. Those in such a solitary Matrix will think they are in the real deal. They'll think they are in Heaven, along with

everyone that they want to be there, and nobody that they don't want there. They will think they get along with along with everyone else just fine; that there's no sadness, no sin, and so on. God knows what they freely want, and tailors each virtual environment to provide exactly that, if possible. (If it's not possible, because they freely want to be in the real deal, this lack is not *experienced*, and so is not a source of suffering.)

Just as it did with the First Matrix, libertarianism raises a difficulty, since you might think that God could not know what you want, when this is undetermined. Some medieval Christians resolved the problem of the compatibility of free will with God's foreknowledge by supposing that changeless, omnipresent God knows the (causally undetermined) future by, so to speak, already having been present then, and having seen what happens. God knows what you do *because you do it*, and not vice versa, hence you may do it freely. The same resolution can be applied here, as long as time exists in Heaven: God knows what you *will* want before you want it, by having been in the future and (so to speak) looking into your mind then.¹³

Can (3) and (4) both be maintained, given (1) and (2*)? There is logical space for this possibility. (3) can be true, and yet there be no sin in Heaven, *if* Heaven is like the lucky roll of the creation die: the world where free beings always choose rightly. In Heaven, everyone will be free to sin, but just *doesn't*. The immediate problem with this suggestion is that it seems incredible that such a coincidence will actually obtain. Perhaps we can appeal to a difference between this situation and that of creation: God has a chance to observe the behavior of free individuals, and only admits the deserving—those who actually don't sin while on Earth. But this would get hardly anyone into Heaven. Worse, it seems to give inductive support, but no *guarantee* at all, that unblemished individuals won't sin ever in the eternity they spend in Heaven.

It is standardly claimed that all are free to sin in Heaven, but none do, because they are in some sense *incapable* of doing so; no one can sin when they are at last with God. This raises two distinct problems. The first is that any such incapability seems incompatible with libertarian freedom, rendering (3) false after all. The second is that, if there is no incompatibility between human beings having libertarian free will and being incapable of sin, then the Free Will Theodicy seems to collapse. God could have just created Heaven and be done with it, a creation with all of the benefits and none of the

disadvantages.

In addition to the problem of sin, we might wonder how it can be managed that free human beings, all interacting with each other, have no desires in conflict. As Satan observed, it must be that our desires change radically. But what ensures this? If it is inevitable that they change in this way, then libertarian freedom is again threatened. And if we are somehow free anyway, when our desires are radically altered, then why didn't God just turn this trick to begin with, and spare all the lost souls? Perhaps we should also consider Mouse's point. If our desires change too radically, will we still be *human beings*, as (1) would have it?

III. Conclusion

Perhaps both explanations of the failure of the First Matrix are correct. Recall the suggestion that machines could not program our "perfect world."

Perhaps our thinking is incoherent: we think that the best existence is one where human beings interact with each other *and* everyone has libertarian free will *and* nobody suffers *and* that someone knowingly arranges this. If this is an incoherent notion, not even God can actualize it.

In creating the Second Matrix, the machines went for interaction combined with free will (which we are assuming is libertarian), with the overwhelming likelihood (inevitability, in practice) of suffering. We can now explain Agent Smith's remarks: if we rank the elements of our incoherent notion of the best existence, human interaction and libertarian free will rank above the absence of suffering. And since they jointly require (*almost* by definition) the presence of suffering, it can be said more or less truly that we "define [even the best] reality through misery and suffering." The First Matrix was an attempt to give interacting humans an existence free of suffering, but this program *required* a radical revision in their wills, contrary to libertarian free will, and so "no one would accept the program." Mouse might say it was an attempt to deny the very nature of human beings.

If the real deal includes libertarian free will, then so does the Second Matrix—our desires, though often enough unsatisfied, will be after all undetermined. (The sense in which humans are liberated from the Matrix has nothing to do with libertarian free will, which can be enjoyed behind bars.) The Second Matrix also features substantial variation in wills amongst

its human inhabitants, and the interesting ethical choices that arise when this is so. For example, apart from the Agents, each virtual human is an avatar, and the "good guys" in the movie end up killing a lot of human beings in their fight against the Agents. It's hard to view these human beings as collaborators, given the nature of the Matrix, so their deaths presumably are to be regarded as acceptable collateral damage, inevitable given the difference in desired outcome. All in all, the Second Matrix is the machines' best attempt at matching what Christians believe God did for us through creation.¹⁴

When we humans turn our eyes toward Heaven, our ranking of values seems to change, and Twain's Puzzle arises anew. In Heaven, there is a heavier weighting given to the absence of suffering. God can knowingly minimize suffering in a real deal, while retaining human interaction, but at the cost of libertarian free will. But given that Heaven is supposed to involve no suffering at all, and given the surpassing value of libertarian free will in the Christian view of things, God's choice is clear: Heaven is a solitary Matrix.¹⁵ The machines, not being God, did not know that *Heaven is no other people*.¹⁶ Never the twain—Twain and I—shall meet (in Heaven, anyway—there's always the *lake*, I suppose.)

A relative of Twain's puzzle emerges. We when consider a pre-Heaven existence, we seem to prefer the best real deal to the best Matrix. When thinking about Heaven, we seem to prefer the best Matrix to the best real deal. This schism in our thinking is represented by the two competing visions in *The Matrix*: on the one hand is the Matrix, and on the other is Zion—named ironically, if I am right, for God's Holy City in Heaven—the place in the bowels of the Earth where human beings not in the Matrix dwell.

Richard Hanley

Endnotes

1. See Christopher Grau's essay, "The Experience Machine." Indeed, I recommend you read Grau's essay in its entirety before proceeding.
2. Metaphysicians will not yet be satisfied. "Matrix," is from the Latin for "mother," and originally meant "womb" (it is used in the Old Testament five times with this meaning), or "pregnant female." In several contexts it means a sort of substrate in which things are grown and developed. Given this etymology, the Matrix might have been the concrete thing that includes the collection of deceived humans in their vats. A more modern meaning of "matrix" is based in mathematics: a rectangular arrangement of symbols. Perhaps "the Matrix" (an expression surely borrowed from William Gibson's earlier use in *Neuromancer*) denotes the array of symbols encoding the virtual environment, which we

might distinguish from the environment itself. But *The Matrix* gives the impression that the environment just is the array of symbols that Neo sees when he finally sees in—so to speak—Matrixvision. Its concrete-world-like appearance seems an inferior perception. (The Matrix thus seems allegorical in turn of Plato's well-known allegory of the Cave; Neo is enlightened about his own nature by liberation from the Matrix, and by the end he sees the true nature of the Matrix.) Still, it is the concrete-world-like appearance of things that I'm concerned with here, so let's ignore the possibility of a Neo.

3. I use *simulacrum* in the following sense: "something having merely the form or appearance of a certain thing, without possessing its substance or proper qualities; a mere image, a specious imitation or likeness, of something." (OED) It is also a nod towards Baudrillard, whose work *Simulacra and Simulation* both influences and appears in *The Matrix*. See my essay, "Baudrillard and The Matrix."

4. Here's an interesting question: which is better, the Second Matrix, or a systematically deceptive personalized non-virtual environment—a *Truman show*—that you never discover the true nature of? The latter has more veridical human interaction in one sense, because you really physically interact; but the interaction is less veridical in another sense, in that other human beings are willing participants in the deception. Another case to think about is a solitary Matrix allowing interaction with non-human participants (dogs, perhaps). Another still is a solitary Matrix without even the appearance of real interaction — call this a lonely Matrix. I don't know about you, but I prefer Sartre's vision of Hell to a lonely Matrix.

5. The Second Matrix may connect with the truth in some unnecessary ways. One's virtual body is depicted as more or less veridical, for instance. (But this may be only "residual self image," as Morpheus tells Neo. If Cypher were put back into the Matrix as Ronald Reagan, that would be clinching evidence that one's avatar can be strikingly different.) Breaking this connection would permit interestingly different human interaction: for instance, you could unknowingly have an experience of heterosexual intercourse with another (unknowing) human who is in fact of the same sex.

6. Sometimes it is argued that you are better off—happier—being a Christian, even if God does not exist. If Christian belief is easier to maintain inside the Second Matrix than outside it, then Cypher could have an extra pragmatic reason for going back in.

7. Is Agent Smith telling the truth? I have no idea. He is attempting to "hack into" Morpheus's mind to gain the access codes to the Zion mainframe computer, so in interpreting the story we should take everything he says—and so, even the very existence of the First Matrix—with a grain of salt. For my purposes, though, we can pretend that he's telling the truth.

8. We need not fully characterize libertarian free will for present purposes. The main point is that causal indeterminism is a necessary condition of it. Causal indeterminism is the denial of causal *determinism*: the thesis that every event is completely determined by causally prior events. A useful and common illustration is to ask whether or not everything that happens, or will happen, is in principle *predictable* — this will be so if determinism is true, and not so if indeterminism is true. (Whether the future can be known by means other than prediction is a different question — see note 11.) The thesis that we have libertarian free will is called *libertarianism*.

9. Many of our desires are *derived* from other desires plus belief, for instance if Ralph desires to kiss Grandma only because he desires an inheritance and he believes kissing Grandma is necessary to achieve this. Non-derived desires, such as Ralph's desire to kiss the girl next door, are importantly independent of belief—they are had, so to speak, for their own sake—and seem to constitute what we refer to by "the will."

10. Is this a theological guarantee of the real deal? The Christian can surely deny this. The existence of the Matrix seems compatible with God's being no deceiver, given the Free Will Theodicy, if the machines have libertarian free will. And if they do not have libertarian free will, as long as they are the product of human free will, they are not part of the environment God knowingly created.

11. I am reminded of a passage in William Gibson's *Count Zero*: "Eyes open, he pulled the thing from his socket and held it, his palm slick with sweat. It was like waking from a nightmare. Not a screamer, where impacted fears took on simple, terrible shapes, but the sort of dream, infinitely more disturbing, where everything is perfectly and horribly normal, and where everything is utterly wrong."

12. People seem to expect that their body in Heaven will resemble their Earthly one (just as their Matrix "body" seems to resemble their real one). Perhaps this is for purposes of recognition, but it seems unnecessary—common memory can do the job.

13. It would be intriguing if God could "cheat" by doing what he does *because* He sees, from the way the future is, what He will do. This would raise a fatalist, *bake-your-noodle* puzzle like the one the Oracle raises for Neo's smashing of the vase. But God is a special case. Being unchanging, He cannot be *caused* to act on the basis of future knowledge, and there is little metaphysical sense to be made of "He did it because He did it."

14. The typical Christian is a *Cartesian dualist*, believing they are a spirit or soul distinct from their physical body, and that embodiment provides the means for human interaction. Loosely speaking, then, our physical bodies are the "avatars" of the real us, in a more or less "teeming" physical environment. The Second Matrix is in this respect almost the converse of Christian creation.

15. Perhaps Christians have had this revelation available to them all along. Luke 10:20 has Jesus telling his disciples, "... rejoice, because your names are written in heaven." In Latin, "matrix" also meant a list or register of names (also, *matricula*, hence our English verb *matriculate*). Intended meaning can go astray: according to some, the notion that the fruit of the tree of knowledge of good and evil was an apple, rests on a confusion over the Latin *malum*, meaning both "evil" and "apple tree." In like manner, maybe Jesus's message, lost in translation, was that Heaven is a Matrix!

16. "People" in the sense of *human beings*. It might be objected that there has to be at least one person you are in contact with: God. I'll just concede this, since it doesn't affect the argument, God's desires presumably being compatible with yours. (Real interaction with angels likewise presents no problems.) A fascinating further suggestion is that you couldn't be maximally happy unless the "program" was extremely sophisticated, and then it might be objected that we should regard the solitary Matrix as containing *virtual* individuals—such as your imaginative sexual partner(s), if there is sex in Heaven—which are arguably persons you really interact with. (Agent Smith's impassioned outburst that he *hates* the Second Matrix might be evidence of personhood, for instance.) *If* these virtual individuals are persons with libertarian free will, then you can't interact with them either, without someone eventually suffering. So we might have another argument that the Christian Heaven is an incoherent notion.

REALITY, WHAT MATTERS, AND *THE* MATRIX

- IAKOVOS VASILIOU -

The Matrix is, at its core, a film with a moral plot. We, the viewers, like the heroes, are in on a secret: The reality that forms the lives of millions of human beings is not real. The world that seems real to most people is in fact a computer-generated simulation, but almost no one knows it. In reality human beings are floating in liquid in machine pods, with tubes connected to them in a grotesque post-apocalyptic world where the sun is blotted out. Things seem, of course, to the average person to be the ordinary world of 1999. Although some details of the history remain untold, it is an essential part of *The Matrix* that we are provided with a specific account of how all of this happened. There was a battle between human beings and machines whose cognitive capacity had surpassed their own. In a desperate attempt to win, human beings blocked out the sun's light in order to deprive the machines of their power source. Despite this extreme tactic, the humans lost, were enslaved, and are now farmed to supply energy sources for the machines. The machines induce the appearance of ordinary 1999 life in the human beings with a computer generated "virtual community" for the purpose of keeping them docile and asleep so that they and their offspring can be used like living batteries. While humans seem to walk around in an ordinary life, their minds are radically deceived and their bodies are exploited. The heroes are thus depicted as fighting a noble battle for the liberation of the human species.¹

I have so far drawn out two aspects of the "moral background" of the film: enslavement and deception. We should also note the perspective we have on the Matrix as viewers of *The Matrix*. We have what is sometimes called a "God's eye" perspective: we can see *both* the Matrix reality *and* "real" reality. We are let in on the truth about the situation, and we are not supposed to question, for example, whether the battle between Morpheus and his friends and the Agents is itself being conducted in another "meta-matrix", or whether the view of the human pods we see might only be some sort of dream image or illusion. As viewers of the Matrix, we are in on the truth and we can see for ourselves that human beings are both enslaved and deceived. Given the outlined history, we are meant to understand the situation of the humans as a terrible and unfair one.

1. How Does the Matrix Differ from Reality?

Excluding, for the moment, the heroes – Morpheus, Trinity, eventually Neo, and the rest of their crew – and the machines, no one *in* the Matrix shares our God's eye perspective. In everyday life as well, as far as we know, reality is simply there. When we watch the film, we identify with the heroes in part because we are repulsed by the idea that human beings are enslaved and deceived.² It is easy to find these two elements at work in *The Matrix* in part because we think of enslavement and deception as things that are done to some people by others; one group of people enslaves another, or one person or group deceives others. In the film it is the machines who are the agents of slavery and deception and almost all of the humans are victims. But how does the Matrix, and the situation of the ordinary people within it differ from reality and the people within it (i.e., us)?

Let's begin with enslavement. We are forced to do many things in ordinary reality: we must eat, drink, sleep, on penalty of death. Also, no matter what we do, we shall eventually, within a fairly predictable time frame, die; we cannot stay alive forever, or even for a couple of hundred years. We can't travel back and forth in time; can't fly to other planets by flapping our arms. The list could go on and on, and I have simply offered limits we are subject to in virtue of the laws of nature. In other words, compared with some easily imaginable possibilities, we are severely constrained, in a type of bondage, though ordinarily most of us don't think of it as such. Writers, artists, philosophers, and theologians over the centuries have of course been keenly aware of these limitations, examined many forms of human bondage, and offered various types of suggestions as to how to free ourselves. Human beings have longed to "break out" of this reality, to transcend the imposed limitations on their physical being. We should be clear that these limitations are *imposed* on us. We simply find ourselves in this condition, with these rules: we all die within approximately 100 years. It has nothing to do with our voluntary choice, our wishes, or our judgements about what ought to be the case.

Who has done this to us? Answering this question is important to some degree because we typically use the term "enslavement" to refer to something done by one agent to some others. In the case of the constraints I outlined above, it may be harder, initially, to find anyone on whom to pin the blame. But of course human beings have offered answers to this question: one is God; another, the laws of nature. Religious thinkers have struggled with questions about why we should not be angry at God for constraining us in the ways he does: why do people die, why can't we go back in time, travel to other planets, etc.? Others conclude that God is not constraining us, but simply the laws of nature. At least at first this thought might be a bit more palatable insofar as we think of the laws of nature as *impersonal* features of reality;

no one made them that way (if God did, then we get angry at him again). They do not *mean* to constrain us and there is no mind or intelligent force actively doing anything to us.³ Either way, however, our actual situation is one of involuntary constraint, much akin to the humans' situation in the Matrix, except that it is not at the hands of machines against whom we lost a war, but at the hands of God or "nature".

The second aspect of the moral background of *The Matrix* is deception. Human beings are being actively deceived by the Matrix into believing things about reality that are not true. Deception offends many people, except perhaps for committed subjectivists, since many people believe that they want to know, or at least have the right to know, the truth, even if it is terrible. For one person, or a group of people, purposefully to keep others in the dark about some truth is to diminish the respect and authority of those people; it is to act patronizingly and paternalistically. In such situations, a few people decide which truths others can handle, and which they can't. Although this happens routinely – consider the relationship between those who govern and those who are governed – many people bristle at this idea and want the scope of such filtering of the truth to be severely limited.

We might think, however, not about the deception of some people by others (just as we did not look at the enslavement of some people by others), but the deception of humanity in general. In Homer's *Iliad* and *Odyssey* the gods are depicted throughout as capriciously deceiving human beings, compelling them knowingly and unknowingly to do specific things, and generally interfering quite frequently in human affairs. The humans in Homer certainly seem to be caught in a matrix of sorts, with gods and goddesses operating on a plane of reality that is not accessible to them (unless the gods want it to be) but that nevertheless often affects matters in the humans' ordinary reality. As human beings began to understand that the Earth rotated around the Sun, and not vice versa, Descartes certainly worried about the extent to which God had had a hand in deceiving all of humanity for tens of thousands of years up to that point. He devotes a significant portion of the *Meditations*⁴ to worrying about how an all-good, all-knowing, and all-powerful God could have allowed (and whether indeed he was complicit in) people's radical deception about the relative motions of the planet they live on, and other truths that turn out to be radically different from how things *seem* to be.

So in our ordinary situation, without any cruel machines doing anything to us, we realize that there are nevertheless many things we cannot do, and we know that we humans have been radically deceived by natural phenomena (or by the gods, or by God) about things in the past, and that it only stands to reason that we may be

radically mistaken about our explanations of things now. I say people's "radical deception", despite the fact that, as with being enslaved, being deceived also seems to require an agent – someone to *do* the deceiving. We should note, however, that we talk of being deceived or fooled by mirrors, or by the light, or by angles. Natural phenomena are often described as contributing to our misunderstanding of them for a reason. Even though human beings were mistaken for millennia about the fact that the Earth moves relative to the Sun, and not the other way round, it is hard to describe our error as simply having "made a mistake", as though humanity forgot to carry the two in some addition calculation. Surely part of the reason that it took humans so long to understand the motions of the Earth is that the appearances themselves are deceptive: it certainly *looks* as though the Sun is moving across the sky.⁵ We can see the very development of philosophy, art, religion, science, and technology as all stemming from a drive to "free humanity" from such deception and enslavement, as part of a struggle to achieve the position of a Morpheus or a Neo.⁶ We develop planes to break the bonds of gravity that keep us physically on the surface of the Earth; we develop complex experiments and gadgets designed to discover the truth about things independently of how they may appear.

My first point, then, is that if we could get a hold of the being responsible for setting up the reality we're actually in, then we could perhaps "free" ourselves, finally knowing the full truth about things, and being able to manipulate reality. If God is responsible, we would need to plead with him successfully, or to fight him and win; if it's the mathematical formulae (computer programs?) underlying "the laws of nature," we would need to learn how to write and rewrite them. We would then all be Neos.⁷ We might note too, at this big-picture level, a difference between the Homeric gods and the Judaeo-Christian-Islamic God. In Homer's world the gods were frequently literally in battle with humans who were greatly outmatched, although not entirely impotent – much like the humans that, before Neo, fought with the "Agents". With the God of the major contemporary religions, he is, by definition, all-good. From this perspective, we should not *fight* God, for he set things up the way he did for a wise and benevolent reason; rather, we need to learn to accept the position he has put us in (this "mortal coil", our reality, our matrix) and, then, if we act certain ways, or do certain things, he will free us from this reality after we "die" (i.e. not go out of existence, but end our stay in this reality) and show us the truth in heaven.

I hope this necessarily brief discussion enables us to see the importance of both the God's eye perspective and the moral background of the film for effecting a difference between the situation depicted in *The Matrix* and our ordinary condition. As viewers of the film we are in a special position: we can see both inside and outside of the Matrix.

We can see that it is not a benevolent God who has set up this 1999 reality, replete with constraints and deceptive appearances, pain and toil, for some wonderful, miraculous purpose. Nor is the reality of most people in the Matrix the result of impersonal laws of nature. Instead, machines who use human beings as batteries are responsible for what counts as reality for most people. *The Matrix* then supplies us the viewers with a definitive answer about who is responsible for what most human beings take to be reality.⁸

2. A Benevolently Generated Matrix

Now *The Matrix* could be significantly altered, without changing anything in the Matrix. Imagine that the real world is a post-apocalyptic hell, just as in the film, but, unlike in the film, suppose that the cause of the world's being in such a state is not some battle with machines that wanted to enslave us, but the emission of so many greenhouse gases with our three-lane-wide SUVs that we completely obliterated the ozone layer and thereby rendered the planet uninhabitable by us or by the plants and animals that we rely on for our survival. Suppose further that sometime in the future, in order to *save* the human race, scientists set up an enormous self-sustaining machine, just as in the film (minus the scary "Sentinels"), designed to keep the human species alive and reproducing for the 100,000 years it will take for whatever weeds are left on the planet to fix our atmosphere and make the planet once again habitable in a normal way. The machine operates simply on solar power (since, on this scenario, the sun is now stronger than ever, frying almost everything else on the planet), so that human beings are not needed as "batteries".⁹ While humans are stuck in this state, the scientists create the Matrix for them to "live" their lives in instead of being conscious of floating in a vat for the length of their life, which would clearly be a most horrific torture. Once the power of the sun is diminished to a habitable degree (because of the repaired atmosphere) the machine would "wake" us humans, and we could go back to living on the planet.

The ordinary person in this scenario is in the same condition as an ordinary person in the film, except that instead of the Matrix being the diabolical result of evil machines who exploit the human race, it is the result of benevolent human beings trying to keep the human race alive in as good condition as possible under the terrible circumstances. Of course it would *seem* no different to the person *in* the Matrix. We, the viewers, however, would have quite a different response to *The Matrix*. There would be no enemy to fight, no injustice to rectify (the pushers of SUVs being long dead). If there were a Morpheus in this situation, how would we think of him? If Morpheus and his friends had left the Matrix, and figured out that they could, with

extreme difficulty, survive in the devastated world (eating disgusting porridge, etc.), should they go about "freeing" everyone, even if it would take another 10,000 years for the Earth to return to its present state of habitability?

As Chris Grau discusses in his introductory essay (section "C"), the Matrix is importantly different from Robert Nozick's "experience machine".¹⁰ Grau points out that we retain free will in the Matrix. The "world" in the Matrix will respond to our free choices, just as the ordinary world does now. Another difference that I think is quite significant is that in the Matrix, unlike in the experience machine, *I am really interacting with other human minds*. There is a community of human beings. With the experience machine, it is all about *my* experience, which is the private content of my own consciousness. It is imaginable that I am alone in the universe, floating in a vat set up by a god who has since committed suicide. In sceptical problems that stem from the Evil Genius hypothesis in Descartes' first *Meditation*, there is a threat of solipsism and the dread of feeling that one might be alone in the universe.¹¹ In the Matrix, however, when two people meet there are really two consciousnesses there that are each experiencing "the same things" from their respective positions. Everyone is hooked up to one and the same Matrix; there are not unique matrices generated for each individual. Of course people aren't really shaking hands – their hands are in vats – but it seems to each of their consciousnesses, not just to one consciousness, that they are shaking hands. This feature of the Matrix is also a respect in which life in the Matrix is critically *unlike* a dream, despite the fact that the humans are described as "dreaming".¹² Regardless of the amount of conscious control one has or lacks in a dream, a dream is *private* to one's own consciousness. It is part of the grammar of "dream", as Wittgenstein might say, that only I can have my dream.¹³

Now this seems to me to be of enormous significance in thinking about the Matrix. If two people fall in love in the Matrix, in what sense would their love not be real? It would not be as if a person merely *dreamt* that he had fallen in love with someone; for in a dream that person is not really there at all, just like in Nozick's experience machine. It is true that in the Matrix they would not really be giving each other flowers, or really holding hands. They would, however, both be experiencing the same things together. They would know each other as persons, who display their characters in how they react to all of the – in one sense – "unreal" situations of the Matrix. Moreover, people in the Matrix really suffer and experience pain, and when they die in the Matrix, they die in the "real world" too. The fact that one and the same Matrix is inhabited by millions of minds means that millions of people are *really* interacting, even if the physical universe in which they are interacting is radically different from

how it appears.

Consider as well writing a novel, a poem, or a philosophy paper. Or consider painting or dancing, making music or a movie. Would any of these activities be affected by the fact that what I took to be material objects were objects that were computer generated? And if not, in the benevolently generated Matrix I hypothesized we would seem clearly better off as a species, developing artistically, intellectually, loving each other within the Matrix rather than fighting for survival and barely succeeding outside of it. If my aim in life was to write some extraordinary philosophy or a groundbreaking novel, surely I could do that far better within the Matrix than outside of it where a person must battle simply for his or her survival. After all, where does my novel or my philosophy paper exist for much of its genesis and storage? In a computer of course. If I wrote a novel in the Matrix, and you read it, and so did 10,000 other minds, and I then win the Pulitzer Prize for it, in what sense would it be unreal or even diminished in value? This differs again from the experience machine. In the experience machine, I might have programmed it so that it would *seem* to me that I had written a brilliant novel and that people had appreciated it. In fact, however, no one would have read my novel and I would have simply programmed myself with memories of having written it, although I never really did. In the Matrix, however, I am not given false memories, and I do really interact with other minds. Physics as we know it would be false (not of course the physics of the Matrix, which scientists would study and which would progress as does ordinary science; see below). But art and human relationships would not be affected. I am trying to show that while we are attached to *reality*, we are not attached to the physical *per se*, where that refers to what we think of as the underlying causes of the smells, tastes, feels, sights, and sounds around us: they could be molecules, they could be computer chips, they could be the whims of Homeric gods. Indeed, very few human beings have much understanding of contemporary physics and what it maintains things "really" are.¹⁴ Nozick's experience machine may have shown us that we have an attachment to the real, an attachment to the truth that we are *really* doing things, *really* accomplishing things, and not just *seeming* to, but we should not for that reason think we are necessarily attached to a certain picture of the physical constitution of things.

I would like return to the question of the sense in which the reality of the Matrix is different from the real world. I think that there is an important difference between being deceived about the reality of an object and being deceived about the real underlying physical or metaphysical cause of something. Avoiding deception and error about the latter is the concern of physics (and metaphysics). That we might be wrong, indeed radically wrong, about the physics/biology of an elephant is quite different

from hallucinating that there is an elephant in front of you, or dreaming of an elephant, or experiencing an elephant in Nozick's machine. In the latter three cases, one is deceived about the reality of an object, about whether there is an elephant there at all. I am not saying that the actual physics or metaphysics of a thing will not determine *whether* it is there; if something is really the underlying cause of something else, of course it must determine its existence. I do claim, however, that *given* the reality of a thing, knowing its true physical/metaphysical explanation neither augments nor diminishes its value or its reality.¹⁵ To discover that, contrary to what you had believed, elephants evolved from single-celled sea creatures and are mostly water, and that water consists of molecules, and that molecules consist of atoms, and that there is a certain interrelationship between matter and energy – that is all part of science's attempt to understand the truth about physical reality. None of these conclusions impugns the elephant's reality or the value it has in the world. What substances at bottom *are* is a question for science or, perhaps, metaphysics. The moral background of the film is quite relevant here. If the fact that we are in the Matrix is simply a matter of our being incorrect about or ignorant of what the real physics of things is, then the Matrix is quite close to our ordinary situation, although our position as *viewers* of *The Matrix* is not like that at all. Since we have a "God's eye" perspective, we are able to know what is really the cause of things and what is not.

In the benevolent Matrix that I envisaged, however, you could learn Matrix-physics and Matrix-history just as we now learn ordinary physics and ordinary history. At a certain age in school you might be taught that your body is really floating in a vat, and then perhaps you could put on goggles and see the world outside of the Matrix, like looking at an x-ray or at your blood under a microscope. Brought up with such a physics and biology, it would seem natural – about as exciting (and unexciting) as being told that your solid unmoving table is made of incredibly small incredibly fast moving parts, or that all of your physical characteristics are determined by a certain code in your DNA, or where babies come from – despite the fact that such truths are hardly obvious, and conflict radically with the way things appear. Just consider any of the conclusions of contemporary physics or quantum mechanics. History too might continue as normal, divided into BM (before Matrix) and AM (after Matrix) dates. After all in the "real" world, outside of the Matrix, nothing would be happening of interest except to scientists. It would be like the contemporary study of bottom of the ocean, or of the moon. Aside from its causal influence on the physical state of the planet, what goes on down there or up there has no part to play in human history. All of *human* history would occur within the Matrix.

By hypothesizing a benevolent rather than a malevolent cause of the Matrix, we can see how much of what I am calling the "moral background" of *The Matrix* influences what we think of it. Deprived of that moral background, a benevolently generated Matrix can show us that our attachment is not to the physical constitution and cause of things, but also not simply to experience. Our attachment is to things that have *value*. Let me explain.

Take the example, discussed in the film, of the pleasure of eating. Imagine that science develops a pill which supplies the perfect amount of nutrition for a human being each day. Humans no longer need to eat at all in the ordinary way. In fact they are, as far as their health is concerned, far worse off if they try to rely on their taste to supply them with the appropriate nutrition (see current statistics on fast food consumption and obesity). They can simply take the pill and get nutrition far superior to what they would if left to their own taste to determine what and how much to eat. Let's suppose too that science has found a way to simulate food with a computer, so that they have created a "food-matrix". My real nutrition would come from the pill, but I could still go out for a "simulated" steak and it would seem just as though I were really eating a steak, including the sensation of getting full, although in fact I would be eating nothing and getting no nutritional harm or benefit from the experience at all. It is hard to imagine such a perfect pill and such perfect computer-simulated food; such a pill is no simple vitamin, and a tofu-burger is no simulated steak. But if we suppose that there are such things, I think human beings would readily give up eating real steak. What those who value eating steak value is not the eating of real cow flesh (in fact, putting it that way inclines one to become a vegetarian), but the experience of eating. If eating the computer steak really were, as we are assuming, absolutely indistinguishable from eating a real steak, no one would care whether they were eating a "real" steak – that is, one that was obtained from a slaughtered cow.

At this level the discussion is again about what the underlying causes of phenomenal qualities are: whether the causes of the taste, smell, etc. of the steak are cow molecules or computer chips or the hand of God. This is, as it were, a matter of science or metaphysics – not of concern to the consumer as a consumer. Now for all physical objects, I contend, it is of no value to us if their underlying constitution is ordinary atoms, or computer generated simulation. My favorite pen still writes the same way, my favorite shirt still feels the same way. If these things are not "real" in the sense that their underlying constitution is radically other than I had believed, that makes no difference to the value that these things have in our lives. It *does*, of course, make a difference to the truth of the physics I learn. But none of this implies that I was being deceived about the reality of the object – that the object I valued

was or is not there in the sense that matters to the non-scientist.¹⁶ In a scene discussed by Grau, Cypher claims his knowledge that the steak is "unreal" – that is, computer generated – does not diminish his enjoyment. Cypher then looks forward to the point when he expects his memory to be wiped clean, and when he will no longer remember that the Matrix is the Matrix. But it seems to me to be unclear why Cypher needs to forget anything about his steak being unreal in order to fully enjoy it – as he himself seems to understand – nor does he need to forget that he is in the Matrix in order to make his life pleasant and satisfying within it. What he desperately needs to forget in order to have a comfortable and satisfying life is the memory of his immoral and cowardly betrayal of his friends and of the rest of those outside of the Matrix who are engaged in the fight for human liberation. But this is an issue, once again, not arising from the Matrix itself, but from the "moral background" of the film.

Having a radically different underlying constitution is very different from saying that things are not real, in the sense of being a mere illusion, as in a dream or a hallucination. Consider again the case of our human interactions. If a person I am friends with is not, after all, a person, then I think there is a clear sense in which the friendship is not real, just as in Nozick's experience machine or in a dream that I was friends with Tom Waits. I would then *seem* to have a relationship to someone, but in reality not have one. What matters is whether I am really interacting with another free mind. I certainly won't try to say what it is to have a mind, or what it is for that mind to be "free", but whatever it is, I am claiming that its value is not importantly tied to any theory in physics or metaphysics. *Whatever* the cause and explanation is of the existence of a free mind, it is the having of one and the ability to interact with other ones that matters. If the underlying constitution of Tom Waits is computer chips, instead of blood and guts, what difference does that make? This is not a question about his *reality* – whether he is really there or not –, it is a question about his physical or metaphysical constitution. If he has a mind, whatever that is, and he has free will, whatever that is, what do I care what physical parts he is – or is not – made of?¹⁷ Indeed, I earnestly hope in the actual world never to see any of those parts or have direct contact with them at all.

3. *The Matrix* on the Matrix

I shall conclude by claiming that *The Matrix* itself provides evidence that, barring enslavement and deception, we would prefer life within the Matrix. I have so far considered how we would feel about the reality of a benevolently generated Matrix. But in *The Matrix*, the cause of the Matrix is explicitly not benevolent. Human beings are enslaved and exploited by scary-looking machines. *The Matrix* is a story about a

few human beings fighting to save the rest of humanity. That is how the movie generates excitement, the thrill for the viewer as he or she hopes that the heroes can defeat the enemy. Of course, the film expects one to root for the humans. But I think there is some duplicity at work in the way *The Matrix* exploits the Matrix. Neo is the savior of humanity, and a large amount of the pleasure that the viewer gets from the film consists of watching Neo and his friends learn to manipulate the Matrix. Key to Neo's eventual success is his training. In his training he learns that the Matrix, as a computer-generated group dream, can be manipulated by a human being. The idea, I guess, is that if one could bring oneself to believe deeply enough that, despite appearances, things are not real, then one could manipulate the reality of the Matrix. The thrill that Neo feels, and that we feel watching him, is that as he gains this control he is able to do things that are, apparently, superhuman – move faster than bullets, hang onto helicopters, fly, etc. We ought to note here, though, that Neo's greatness, his being the One, is only the case because the Matrix exists. Outside of the Matrix, Neo is just a smart computer geek. He can't really fly, or really dodge bullets (nor, apparently, does he dress in full-length leather black coats, though I guess he *could*). We, as viewers, would not get any pleasure from *The Matrix* if it were not for the Matrix. If there were no Matrix, everyone would be eating terrible porridge in a sunless world and simply fighting for survival, which would make for a bad world and a bad movie. The premise of the movie is that there is a moral duty to destroy the Matrix, and "free" the humans. But all of the satisfaction that the viewer gets, and that the characters get in terms of their own sense of purpose and of being special, is derived *from* the Matrix. It's not just Cypher's steak that is owed to the Matrix, it is Morpheus's breaking the handcuffs, Trinity's gravity-defying leaps, and Neo's bullet dodging. If my argument is right, then, the irony of *The Matrix* is that the heroes spend all of their time liberating human beings from the Matrix although afterwards they would have good reason to go back in, assuming the conditions on Earth are still so terrible. This is because there's nothing *wrong* with the Matrix *per se*; indeed, I've argued that our reality might just as well *be* the Matrix. What we want, now as always, one way or another, is to have control over it ourselves. What we would do with such power is a question, I suppose, for psychologists; but, looking at what people have done so far, I at any rate hope we remain enslaved and deceived by *something* for a long time to come.¹⁸

Iakovos Vasiliou

Footnotes

1. Another topic raised by the film, which I will not discuss beyond this note, would be to assess the moral background of the plot. Are the humans clearly in the right? After all, it was they who blotted out the sun in an attempt to exterminate the machines. Particularly in light of the machines' claim

that they are simply the next evolutionary step, we ought to think about whether there is some objectionable "speciesism" at work in the humans' assessment of the situation. For my purposes I'll assume the humans are morally justified in the fight for liberation, which, I might add, is certainly a defensible position. For even if machines are the next evolutionary step, and some human beings are guilty of having acted wrongly towards them, that would hardly justify the involuntary enslavement of the entire human race in perpetuity. Moreover, the existence of a "more advanced" species than our own (however that is to be determined) surely should not deprive us of our human rights.

2. And in part because we too would like to control reality; see below.

3. The Stoics thought of the natural world, of the universe as a whole, as itself a rational creature with an overall goal or purpose.

4. Although this theme is present throughout, see especially *Meditations I* and *IV*.

5. The idea that reality is tricky and tries to hide its nature from us is very old, even without, as in Homer's case, any gods acting as agents of deception. For example, the Presocratic philosopher Heraclitus (c.540-c.480 BC) writes (fr. 53) "an unapparent connection is stronger than an apparent one" and (fr. 123) "nature/the real constitution [of things] (*physis*) loves to hide itself."

6. Morpheus and company are an interesting amalgam of technological sophistication and religious symbolism.

7. Or, more precisely, those of us who accomplished this.

8. Of course for Morpheus and his crew, and for the machines if they were sufficiently reflective, the same questions could be raised about what makes the reality *outside* of the Matrix the way *it* is – who is responsible for *that*? And then we can imagine them responding in the sorts of ways I have described, pinning the blame on God, the laws of nature, etc.

9. This detail is meant simply to avoid the possibility of unease over the issue of whether human beings are being used as batteries, voluntarily or not.

10. I shall assume that my reader has read that essay, where Grau clearly explains Nozick's example. (The essay can be found [here](#).)

11. See Grau, "Dream Skepticism". The threat of solipsism seems to me to be the same in the Matrix or in the ordinary world; and that is not my concern here. I am simply taking the truth of the "God's eye" perspective offered the viewer of the film for granted. *The Matrix* tells us and shows us that we are all hooked up to the same Matrix.

12. I think that perhaps Colin McGinn's essay too quickly assimilates the Matrix to dreaming, and Neo's control over it to "lucid" dreaming. Although McGinn may be right that the Matrix must be dealing with "images" rather than "percepts", there are important disanalogies between Matrix-experience and dream-experience. First, in a dream, there is only your own mind involved. The Matrix must be, at a minimum, a group dream. I am arguing above that the fact that one mind is really interacting with other minds is critical to assessing the value of the Matrix reality. This complicates the apparently clear idea of controlling one's dream, since it is not simply one mind at work that can "alter" the images one is conscious of. I am not sure of the coherence of the hypothesis here. For example, when the young boy bends the spoon, Neo "sees" this. So the boy's control of his environment is perceivable *both* by the boy's mind *and* by Neo's. So he must be changing something that is, "in reality", in Neo's mind – namely, Neo's image of the spoon. But what if Neo straightens the spoon at the same time the boy bends it? Whose lucid dream will win out, and be perceived by the other minds? The one with the stronger will? Second, the "images" that are in your mind in the Matrix can, and regularly do, *really* kill people; that is, kill their bodies *outside* of the Matrix. Except in some bad horror movies, dream images cannot really kill you, or make you bleed. The difficulty of understanding how something which is a mere "image" is supposed to have this sort of effect seems therefore to cause some problems for calling the state of ordinary people in the Matrix "dreaming". See also next note.

13. We could certainly, if we wish, *call* the experience of the Matrix "a dream", as the movie does.

But we should remember that Neo, while in the Matrix and before he has met Morpheus, has a dream while he is "asleep". So we need some distinction between that sort of "dream" and Neo's "waking" "group-dream" within the Matrix.

14. This sentence implies that contemporary physics represents humans' best understanding of the true nature of reality, which is certainly a contentious claim.

15. The question of whether I know something is in fact real or an illusion remains as legitimate or illegitimate as always. As throughout this essay, I am simply bypassing any sceptical questions, since it is part of my argument that being in the Matrix does not affect them.

16. All human being might be considered "scientists" insofar as we are curious about and have a conception of what the reality of things are: what causes them, how they come into being, how they are destroyed, etc. But we are also interested in other people, objects, and activities because of their inherent value, a value they retain regardless of the correct explanation of their reality.

17. Given a true account of what it is to have a mind, I would surely care if what appeared to be a person did not fulfill those criteria, for then he would not be a person after all. For example, if someone somehow showed that a machine could not have a "free mind", then I would care whether my friend was a machine or not, but only secondarily, given that *ex hypothesi* as a machine he would not have a free mind. My point is only that it is "having a free mind" or "being a person" that is the source of value, not the correct theory about what makes someone a person. I am claiming that ignorance of or deception about the right physical or metaphysical account of mind does not thereby cast doubt on the value of having a mind. Scepticism about other minds – the questions of whether there really are other minds and how we could tell whether there are – is not addressed at all by what I am saying. I am taking for granted the truth of what the film tells us: there are other minds. The problem of other minds, like solipsism mentioned above, is equally a problem in or out of the Matrix.

18. I am grateful to Chris Grau and Bill Vasiliou for comments on and discussion about an earlier version of this essay.

THE MATRIX – OUR FUTURE?

- KEVIN WARWICK -

Is *The Matrix* merely a science fiction scenario, or is it, rather, a philosophical exercise? Alternatively, is it a realistic possible future world? The number of respected scientists predicting the advent of intelligent machines is growing exponentially. Steven Hawking, perhaps the most highly regarded theoretical scientist in the world and the holder of the Cambridge University chair that once belonged to Isaac Newton, said recently, "In contrast with our intellect, computers double their performance every 18 months. So the danger is real that they could develop intelligence and take over the world." He added, "We must develop as quickly as possible technologies that make possible a direct connection between brain and computer, so that artificial brains contribute to human intelligence rather than opposing it."¹ The important message to take from this is that the danger—that we will see machines with an intellect that outperforms that of humans—is real.

I. The Facts

But is it just a danger—a potential threat—or, if things continue to progress as they are doing, is it an inevitability? Is the Matrix going to happen whether we like it or not? One flaw in the present-day thinking of some philosophers lies in their assumption that the ultimate goal of research into Artificial Intelligence is to create a robot machine with intellectual capabilities approaching those of a human. This may be the aim in a limited number of cases, but the goal for most AI developers is to make use of the ways in which robots can outperform humans—rather than those in which they can only potentially become our match.

Robots can sense the world in ways that humans cannot—ultraviolet, X-ray, infrared, and ultrasonic perception are some obvious examples—and they can intellectually outperform humans in many aspects of memory and logical mathematical processing. And robots have no trouble thinking of the world around them in multiple dimensions, whereas human brains are still restricted to conceiving the same entity in an extremely limited three dimensional way. But perhaps the biggest advantage robots have over us is their means of communication—generally an electronic form, as opposed to the human's embarrassingly slow mechanical technique called speech, with its highly restricted coding schemes called languages.

It appears to be inevitable that at some stage a sentient robot will appear, its production having been initiated by humans, and begin to produce other, even more capable and powerful robots. One thing overlooked by many is that humans do not reproduce, other than in cloning; rather, humans *produce* other humans. Robots are far superior at producing other robots and can spawn robots that are far more intelligent than themselves.

Once a race of intellectually superior robots has been set into action, major problems will appear for humans. The morals, ethics, and values of these robots will almost surely be drastically different from those of humans. How would humans be able to reason or bargain with such robots? Why indeed should such robots want to take any notice at all of the silly little noises humans would be making? It would be rather like humans today obeying the instructions of cows.

So a war of some kind would be inevitable, in the form of a last gasp from humans. Even having created intelligent, sentient robots in the first place, robots that can out-think them, the humans' last hope would be to find a weak spot in the robot armoury, a chink in their life-support mechanism. Naturally, their food source would be an ideal target. For the machines, obtaining energy from the sun—a constant source—would let them bypass humans, excluding them from the loop. But as we know, humans have already had much success in polluting the atmosphere and wrecking the ozone layer, so blocking out the sun's rays – scorching the sky, in effect – would seem to be a perfectly natural line of attack in an attempt to deprive machines of energy.

In my own book, *In the Mind of the Machine²*, I had put forth the idea that the machines would, perhaps in retaliation, use humans as slave labourers, to supply robots with their necessary energy. Indeed, we must consider this as one possible scenario. However, actually using humans as a source of energy—batteries, if you like—is a much sweeter solution, and more complete. Humans could be made to lie in individual pod-like wombs, acting rather like a collection of battery cells, to feed the machine-led world with power.

Probably in this world of machine dominance there would be a few renegade humans causing trouble, snapping at the heels of the machine authorities in

an attempt to wrestle back power for humans, an attempt to go back to the good old times. So it is with the Matrix. It is a strange dichotomy of human existence that as a species we are driven by progress—it is central to our being—yet at the same time, for many there is a fruitless desire to step back into a world gone by, a dream world.

Yet it is in human dreams that the Matrix machines have brought about a happy balance. Simply treating humans as slaves would always bring about problems of resistance. But by providing a port directly into each human brain, each individual can be fed a reality with which he or she is happy, creating for each one a contented existence in a sort of dream world. Even now we know that scientifically it would be quite possible to measure, in a variety of ways, the level of contentment experienced by each person. The only technical problem is how one would go about feeding a storyline directly into a brain.

So what about the practical realities of the brain port? I myself have, as reported in 'I, Cyborg',³ had a 100-pin port that allowed for both signal input and output connected into my central nervous system. In one experiment conducted while I was in New York City, signals from my brain, transmitted via the Internet, operated a robot hand in the UK. Meanwhile, signals transmitted onto my nervous system were clearly recognisable in my brain. A brain port, along the lines of that in the Matrix, is not only a scientific best guess for the future; I am working on such a port now, and it will be with us within a decade at most.

II. Human or Machine

With the port connected into my nervous system, my brain was directly connected to a computer and thence on to the network. I considered myself to be a Cyborg: part human, part machine. In *The Matrix*, the story revolves around the battle between humans and intelligent robots. Yet Neo, and most of the other humans, each have their own brain port. When out of the Matrix, they are undoubtedly human; but while they are in the Matrix, there can be no question that they are no longer human, but rather are Cyborgs. The real battle then becomes not one of humans versus intelligent robots but of Cyborgs versus intelligent robots.

The status of an individual whilst within the Matrix raises several key issues.

For example, when they are connected are Neo, Morpheus, and Trinity individuals within the Matrix? Or do they have brains which are part human, part machine? Are they themselves effectively a node on the Matrix, sharing common brain elements with others? It must be remembered that ordinarily human brains operate in a stand-alone mode, whereas computer-brained robots are invariably networked. When connected into a network, as in the Matrix, and as in my own case as a Cyborg, individuality takes on a different form. There is a unique, usually human element, and then a common, networked machine element.

Using the common element, 'reality' can be downloaded into each brain. Morpheus describes this (as do others throughout the film) as 'having a dream.' He raises questions as to what is real. He asks how it is possible to know the difference between the dream world and the real world. This line of questioning follows on from many philosophical discussions, perhaps the most prominent being that of Descartes, who appeared to want to make distinctions between dream states and 'reality', immediately leading to problems in defining what was real and what was not. As a result he faced further problems in defining absolute truths.

Perhaps a more pertinent approach can be drawn from Berkeley, who denied the existence of a physical world, and Nietzsche, who scorned the idea of objective truth. By making the basic assumption that there is no God, my own conclusion is that *there can be no absolute reality, there can be no absolute truth* — whether we be human, Cyborg, or robot. Each individual brain draws its conclusions and makes assumptions as to the reality it faces at an instant, dependant on the input it receives. If only limited sensory input is forthcoming, then brain memory banks (or injected feelings) need to be tapped for a brain to conceive of a storyline. At any instant, a brain links its state with its common-sense memory banks, often coming to unlikely conclusions.

As a brain ages, or as a result of an accident, the brain's workings can change; this often appears to the individual to be a change in what is perceived rather than a change in that which is perceiving. In other words, the individual thinks it must be the world that has changed, not his or her brain. Where a brain is part of a network, however, there is a possibility for alternative viewpoints to be proposed by different nodes on the network. This is not something that individual humans are used to. An individual brain

tends to draw only one conclusion at a time. In some types of schizophrenia this conclusion can be confused and can change over time; it is more usually the case, though, that such an individual will draw a conclusion about what is perceived that is very much at variance with the conclusion of other individuals. For the most part, what is deemed by society to be 'reality' at any point, far from being an absolute, is merely a commonly agreed set of values based on the perceptions of a group of individuals.

The temptation to see a religious undertone in *The Matrix* is interesting — with Morpheus cast as the prophet John the Baptist, Trinity perhaps as God or the holy spirit, Neo clearly as the messiah, and Cypher as Judas Iscariot, the traitor. But, far from a Gandhi-like, turn the other cheek, approach, Neo's is closer to one that perhaps was actually expected by many of the messiah himself, taking on his role as victor over the evil Matrix: a holy war against a seemingly invincible, all-powerful machine network.

But what of the machine network, the Matrix, itself? With an intellect well above that of collective humanity, surely its creativity, its artistic sense, its value for aesthetics would be a treat to behold. But the film keeps this aspect from us — perhaps to be revealed in a sequel. Humans released from the Matrix grip, merely regard it as an evil, perhaps Cypher excluded here. Meanwhile the Agents are seen almost as faceless automatons, ruthless killers, strictly obeying the will of their Matrix overlord. Possibly humans would see both the Matrix and Agents as the enemy, just as the Matrix and Agents would so regard humans — but once inside the Matrix the picture is not so clear. As a Cyborg, who are your friends and who are your enemies? It is no longer black and white when you are part machine, part human.

III. In and Out of Control

Morpheus tells Neo that the Matrix is control. This in itself is an important revelation. As humans, we are used to one powerful individual being the main instigator, the brains behind everything. It is almost as though we cannot even conceive of a group or collection running amuck, but believe, rather, that there is an individual behind it all. In the second world war, it was not the Germans or Germany who the allies were fighting but Adolf Hitler; meanwhile in Afghanistan, it is Bin-Laden who is behind it all. Yet in the Matrix we are faced with a much more realistic scenario, in that it is not some crazed individual up to no good, but the Matrix — a network.

When I find myself in a discussion of the possibility of intelligent machines taking over things, nine times out of ten I am told—following a little chuckle to signify that I have overlooked a blindingly obvious point—that "If a machine causes a problem you can always switch it off." What a fool I was not to have thought of it!! How could I have missed that little snippet?

Of course it is not only the Matrix but even today's common Internet that gives us the answer, and cuts the chuckle short. Even now, how is it practically possible to switch off the Internet? We're not talking theory here, we're talking practice. Okay, it is of course possible to unplug one computer, or even a small subsection intranet, but to bring down the whole Internet? Of course we can't. Too many entities, both humans and machines, rely on its operation for their everyday existence. It is not a Matrix of the future that we will not be able to switch off, it is a Matrix of today that we cannot switch off, over which we cannot have ultimate control.

Neo learns that the Matrix is a computer-generated dream world aimed at keeping humans under control. Humans are happy to act as an energy source for the Matrix as long as they themselves believe that the reality of their existence is to their liking; indeed, how are the human nodes in a position to know what is computer-generated reality and what is reality generated in some other way?

A stand-alone human brain operates electrochemically, powered partly by electrical signals and partly by chemicals. In the western world we are more used to chemicals being used to change our brain and body state, either for medicinal purposes or through narcotics, including chemically instigated hallucinations. But now we are entering the world of e-medicine. Utilising the electronic element of the electrochemical signals on which the human brain and nervous system operate, counterbalancing signals can be sent to key nerve fibre groups to overcome a medical problem. Conversely, electronics signals can be injected to stimulate movement or pleasure. Ultimately, electronic signals will be able to replace the chemicals that release memories and "download" memories not previously held. Why live in a world that is not to your liking if a Matrix state is able to keep your bodily functions operating whilst you live out a life in a world in which you are happy with yourself? The world of the Matrix would appear to be one that lies in the direction humanity

is now heading—a direction in which it would seem, as we defer more and more to machines to make up our minds for us, that we wish to head.

IV. Ignorance and Bliss

In a sense, *The Matrix* is nothing more than a modern day "Big Brother," taking on a machine form rather than the Orwellian vision of a powerful individual using machines to assist and bring about an all-powerful status. But *1984*, the novel in which the story of Big Brother was presented, was published in 1948. *The Matrix* comes fifty years later. In the meantime, we have witnessed the likes of radar, television for all, space travel, computers, mobile phones, and the Internet. What would Orwell's Big Brother have been like if he had had those technologies at his disposal – would Big Brother have been far from the Matrix?

With the first implant I received, in 1998, for which I had no medical reason (merely scientific curiosity), a computer network was able to monitor my movements. It knew what time I entered a room and when I left. In return it opened doors for me, switched on lights, and even gave me a welcoming "Hello" as I arrived. I experienced no negatives at all. In fact, I felt very positive about the whole thing. I gained something as a result of being monitored and tracked. I was happy with having Big Brother watching me because, although I gave up some of my individual humanity, I benefited from the system doing things for me. Would the same not be true of the Matrix? Why would anyone want to experience the relatively tough and dangerous life of being an individual human when he or she could be part of the Matrix?

So here we come on to the case of Cypher. As he eats his steak he says, "I know that this steak doesn't exist. I know when I put it in my mouth, the Matrix is telling my brain that it is juicy and delicious!" He goes on to conclude that "Ignorance is bliss." But is it ignorance? His brain is telling him, by whatever means, that he is eating a nice juicy steak. How many times do we nowadays enter a fast-food burger bar in order to partake of a burger that, through advertising, our brains have been conditioned into believing is the tastiest burger imaginable. When we enter we know, because we've seen the scientific papers, that the burger contains a high percentage of water, is mainly fat, and is devoid of vitamins. Yet we still buy such burgers by the billion. When we eat one, our conditioned brain is somehow telling us that it

is juicy and delicious, yet we know it doesn't quite exist in the form our brain is imagining.

We can thus understand Cypher's choice. Why be out of the Matrix, living the dangerous, poor, tired, starving life of a disenfranchised human, when you can exist in a blissfully happy life, with all the nourishment you need? Due to the deal he made with Agent Smith, once Cypher is back inside he will have no knowledge of having made any deal in the first place. He appears to have nothing at all to lose. The only negative aspect is that before he is reinserted he may experience some inner moral human pangs of good or bad. Remember that being reinserted is actually good for the Matrix, although it is not so good for the renegade humans who are fighting the system.

Robert Nozick's thought experiment puts us all to the test, and serves as an immediate exhibition of Cypher's dilemma. Nozick asks, if our brains can be connected, by electrodes, to a machine which gives us any experiences we desire, would we plug into it for life? The question is, what else could matter other than how we feel our lives are going, from the inside? Nozick himself argued that other things do matter to us, for example that we value being a certain type of person, we want to be decent, we actually wish to do certain things rather than just have the experience of doing them. I disagree completely with Nozick.

Research involving a variety of creatures, principally chimpanzees and rats, has allowed them to directly stimulate pleasure zones in their own brain, simply by pressing a button. When given the choice of pushing a button for pleasure or a button for food, it is the pleasure button that has been pressed over and over again, even leading to starvation (although individuals were quite happy even about that). Importantly, the individual creatures still had a role to play, albeit merely that of pressing a button. This ties in directly with the Matrix, which also allows for each individual mentally experiencing a world in which he or she is active and has a role to play.

It is, however, an important question whether or not an individual, as part of the Matrix, experiences free will or not. It could be said that Cypher, in deciding to re-enter the Matrix, is exercising his free will. But once inside, will he still be able to exhibit free will then? Isn't it essentially a similar situation to that proposed by Nozick? Certainly, within the mental reality projected on an individual by the Matrix, it is assumed that a certain amount of mental

free will is allowed for; but it must be remembered, at the same time, that each individual is lying in a pod with all his or her life-sustaining mechanisms taken care of and an interactive storyline being played down into his or her brain. Is that free will? What is free will anyway, when the state of a human brain is merely partly due to a genetic program and partly due to life's experience? Indeed, exactly the same thing is true for a robot.

In the Matrix, no human fuel cells are killed, not even the unborn—there is no abortion. Yet, naturally dying humans are allowed to die naturally and are used as food for the living. Importantly, they are not kept alive by chemicals merely for the sake of keeping them alive. The Matrix would appear to be more morally responsible to its human subjects than are human subjects to themselves. Who therefore wouldn't want to support and belong to the Matrix, especially when it is making life easier for its subjects?

Neo is kidnapped by Luddites, dinosaurs from the past when humans ruled the earth. It's not the future. We are in reality heading towards a world run by machines with an intelligence far superior to that of an individual human. But by linking into the network and becoming a Cyborg, life can appear to be even better than it is now. We really need to clamp down on the party-pooper Neos of this world and get into the future as soon as we can—a future in which we can be part of a Matrix system, which is morally far superior to our Neolithic morals of today.

Kevin Warwick

www.kevinwarwick.com

Footnotes

1. Hawking, S., "Hawking's plan to offset computer threat to humans", Ananova, www.ananova/news, 1 September 2001
2. *In the Mind of the Machine*, Arrow, 1998. Available on www.amazon.co.uk
3. *I, Cyborg*, Century, 2002. Available on www.amazon.co.uk

WAKE UP!

GNOSTICISM & BUDDHISM IN *THE MATRIX*

- FRANCES FLANNERY-DAILEY & RACHEL WAGNER -

At the beginning of *The Matrix*, a black-clad computer hacker known as Neo falls asleep in front of his computer. A mysterious message appears on the screen: "Wake up, Neo."¹ This succinct phrase encapsulates the plot of the film, as Neo struggles with the problem of being imprisoned in a "material" world that is actually a computer simulation program created in the distant future by Artificial Intelligence (AI) as a means of enslaving humanity, by perpetuating ignorance in the form of an illusory perception called "the Matrix." In part, the film crafts its ultimate view of reality by alluding to numerous religious traditions that advance the idea that the fundamental problem which humanity faces is ignorance and the solution is knowledge or awakening. Two religious traditions on which the film draws heavily are Gnostic Christianity and Buddhism.² Although these traditions differ in important ways, they agree in maintaining that the problem of ignorance can be solved through an individual's reorientation of perspective concerning the material realm.³ Gnostic Christianity and Buddhism also both envision a guide who helps those still trapped in the limiting world of illusion, a Gnostic redeemer figure or a bodhisattva, who willingly enters that world in order to share liberating knowledge, facilitating escape for anyone able to understand. In the film, this figure is Neo, whose name is also an anagram for the "One."

Although as a "modern myth"⁴ the film purposefully draws on numerous traditions,⁵ we propose that an examination of Gnostic Christianity and Buddhism well illuminates the overarching paradigm of *The Matrix*, namely, the problem of sleeping in ignorance in a dreamworld, solved by waking to knowledge or enlightenment. By drawing syncretistically on these two ancient traditions and fusing them with a technological vision of the future, the film constructs a new teaching that challenges its audience to question "reality."

I. Christian Elements in *The Matrix*

The majority of the film's audience probably easily recognizes the presence of some Christian elements, such as the name Trinity⁶ or Neo's death and Christ-like resurrection and ascension near the end of the film. In fact, Christian and biblical allusions abound, particularly with respect to nomenclature:⁷ Apoc (Apocalypse), Neo's given name of Mr. Ander/son (from the Greek andras for man, thus producing "Son of

Man"), the ship named the Nebuchadnezzar (the Babylonian king who, in the Book of Daniel, has puzzling symbolic dreams that must be interpreted),⁸ and the last remaining human city, Zion, synonymous in Judaism and Christianity with (the heavenly) Jerusalem.⁹ Neo is overtly constructed as a Jesus figure: he is "the One" who was prophesied to return again to the Matrix, who has the power to change the Matrix from within (i.e., to work miracles), who battles the representatives of evil and who is killed but comes to life again.

This construction of Neo as Jesus is reinforced in numerous ways. Within minutes of the commencement of the movie, another hacker says to Neo, "You're my savior, man, my own personal Jesus Christ."¹⁰ This identification is also suggested by the Nebuchadnezzar's crew, who nervously wonder if he is "the One" who was foretold, and who repeatedly swear in Neo's presence by saying "Jesus" or "Jesus Christ."¹¹ In still another example, Neo enters the Nebuchadnezzar for the first time and the camera pans across the interior of the ship, resting on the make: "Mark III no. 11." This seems to be another messianic reference, since the Gospel of Mark 3:11 reads: "Whenever the unclean spirits saw him, they fell down before him and shouted, ' You are the Son of God!'"

II. Gnosticism in *The Matrix*

Although the presence of individual Christian elements within the film is clear, the overall system of Christianity that is presented is not the traditional, orthodox one. Rather, the Christian elements of the film make the most sense when viewed within a context of Gnostic Christianity.¹² Gnosticism was a religious system that flourished for centuries at the beginning of the Common Era, and in many regions of the ancient Mediterranean world it competed strongly with "orthodox" Christianity, while in other areas it represented the only interpretation of Christianity that was known.¹³ The Gnostics possessed their own Scriptures, accessible to us in the form of the Nag Hammadi Library, from which a general sketch of Gnostic beliefs may be drawn.¹⁴ Although Gnostic Christianity comprises many varieties, Gnosticism as a whole seems to have embraced an orienting cosmogonic myth that explains the true nature of the universe and humankind's proper place in it.¹⁵ A brief retelling of this myth illuminates numerous parallels with *The Matrix*.

In the Gnostic myth, the supreme god is completely perfect and therefore alien and mysterious, "ineffable," "unnamable," "immeasurable light which is pure, holy and immaculate" (Apocryphon of John). In addition to this god there are other, lesser divine beings in the pleroma (akin to heaven, a division of the universe that is not

Earth), who possess some metaphorical gender of male or female.¹⁶ Pairs of these beings are able to produce offspring that are themselves divine emanations, perfect in their own ways.¹⁷ A problem arises when one "aeon" or being named Sophia (Greek for wisdom), a female, decides "to bring forth a likeness out of herself without the consent of the Spirit," that is, to produce an offspring without her consort (Apocry. of John). The ancient view was that females contribute the matter in reproduction, and males the form; thus, Sophia's action produces an offspring that is imperfect or even malformed, and she casts it away from the other divine beings in the pleroma into a separate region of the cosmos. This malformed, ignorant deity, sometimes named Yaldabaoth, mistakenly believes himself to be the only god.

Gnostics identify Yaldabaoth as the Creator God of the Old Testament, who himself decides to create archons (angels), the material world (Earth) and human beings. Although traditions vary, Yaldabaoth is usually tricked into breathing the divine spark or spirit of his mother Sophia that formerly resided in him into the human being (especially Apocry. of John; echoes of Genesis 2-3). Therein lies the human dilemma. We are pearls in the mud, a divine spirit (good) trapped in a material body (bad) and a material realm (bad). Heaven is our true home, but we are in exile from the pleroma.

Luckily for the Gnostic, salvation is available in the form of gnosis or knowledge imparted by a Gnostic redeemer, who is Christ, a figure sent from the higher God to free humankind from the Creator God Yaldabaoth. The gnosis involves an understanding of our true nature and origin, the metaphysical reality hitherto unknown to us, resulting in the Gnostic's escape (at death) from the enslaving material prison of the world and the body, into the upper regions of spirit. However, in order to make this ascent, the Gnostic must pass by the archons, who are jealous of his/her luminosity, spirit or intelligence, and who thus try to hinder the Gnostic's upward journey.

To a significant degree, the basic Gnostic myth parallels the plot of *The Matrix*, with respect to both the problem that humans face as well as the solution. Like Sophia, we conceived an offspring out of our own pride, as Morpheus explains: "Early in the 21st century, all of mankind was united in celebration. We marveled at our own magnificence as we gave birth to AI."¹⁸ This offspring of ours, however, like Yaldabaoth is malformed (matter without spirit?). Morpheus describes AI as "a singular consciousness that spawned an entire race of machines," a fitting parallel for the Gnostic Creator God of the archons (angels) and the illusory material world. AI creates the Matrix, a computer simulation that is "a prison for your mind." Thus,

Yaldabaoth/ AI traps humankind in a material prison that does not represent ultimate reality, as Morpheus explains to Neo: "As long as the Matrix exists, the human race will never be free."

The film also echoes the metaphorical language employed by Gnostics. The Nag Hammadi texts describe the fundamental human problem in metaphorical terms of blindness, sleep, ignorance, dreams and darkness / night, while the solution is stated in terms of seeing, waking, knowledge (gnosis), waking from dreams and light / day.¹⁹

Similarly, in the film Morpheus, whose name is taken from the Greek god of sleep and dreams, reveals to Neo that the Matrix is "a computer generated dreamworld." When Neo is unplugged and awakens for the first time on the Nebuchadnezzar in a brightly lit white space (a cinematic code for heaven), his eyes hurt, as Morpheus explains, because he has never used them. Everything Neo has "seen" up to that point was seen with the mind's eye, as in a dream, created through software simulation. Like an ancient Gnostic, Morpheus explains that the blows he deals Neo in the martial arts training program have nothing to do with his body or speed or strength, which are illusory. Rather, they depend only on his mind, which is real.

The parallels between Neo and Christ sketched earlier are further illuminated by a Gnostic context, since Neo is "saved" through gnosis or secret knowledge, which he passes on to others. Neo learns about the true structure of reality and about his own true identity, which allows him to break the rules of the material world he now perceives to be an illusion. That is, he learns that "the mind makes it [the Matrix, the material world] real," but it is not ultimately real. In the final scene of the film, it is this gnosis that Neo passes on to others in order to free them from the prison of their minds, the Matrix. He functions as a Gnostic Redeemer, a figure from another realm who enters the material world in order to impart saving knowledge about humankind's true identity and the true structure of reality, thereby setting free anyone able to understand the message.

In fact, Neo's given name is not only Mr. Anderson / the Son of Man, it is Thomas Anderson, which reverberates with the most famous Gnostic gospel, the Gospel of Thomas. Also, before he is actualized as Neo (the one who will initiate something "New," since he is indeed "the One"), he is doubting Thomas, who does not believe in his role as the redeemer figure.²⁰ In fact, the name Thomas means "the Twin," and in ancient Christian legend he is Jesus' twin brother. In a sense, the role played by Keanu Reeves has a twin character, since he is constructed as both a doubting Thomas and as a Gnostic Christ figure.²¹

Not only does Neo learn and pass on secret knowledge that saves, in good Gnostic fashion, but the way in which he learns also evokes some elements of Gnosticism. Imbued with images from eastern traditions, the training programs teach Neo the concept of "stillness," of freeing the mind and overcoming fear, cinematically captured in "Bullet Time" (digitally mastered montages of freeze frames / slow motion frames using multiple cameras).²² Interestingly enough, this concept of "stillness" is also present in Gnosticism, in that the higher aeons are equated with "stillness" and "rest" and can only be apprehended in such a centered and meditative manner, as is apparent in these instructions to a certain Allogenes: "And although it is impossible for you to stand, fear nothing; but if you wish to stand, withdraw to the Existence, and you will find it standing and at rest after the likeness of the One who is truly at rest...And when you becomes perfect in that place, still yourself..." (Allogenes) The Gnostic then reveals, "There was within me a stillness of silence, and I heard the Blessedness whereby I knew my proper self" (Allogenes).²³ When Neo realizes the full extent of his "saving gnosis," that the Matrix is only a dreamworld, a reflective Keanu Reeves silently and calmly contemplates the bullets that he has stopped in mid-air, filmed in "Bullet Time."

Yet another parallel with Gnosticism occurs in the portrayal of the Agents such as Agent Smith, and their opposition to the equivalent of the Gnostics - that is, Neo and anyone else attempting to leave the Matrix. AI created these artificial programs to be "the gatekeepers - they are guarding all the doors, they are holding all the keys." These Agents are akin to the jealous archons created by Yaldabaoth who block the ascent of the Gnostic as he/she tries to leave the material realm and guard the gates of the successive levels of heaven (e.g., Apocalypse of Paul).²⁴

However, as Morpheus predicts, Neo is eventually able to defeat the Agents because while they must adhere to the rules of the Matrix, his human mind allows him to bend or break these rules.²⁵ Mind, though, is not equated in the film merely with rational intelligence, otherwise Artificial Intelligence would win every time. Rather, the concept of "mind" in the film appears to point to a uniquely human capacity for imagination, for intuition, or, as the phrase goes, for "thinking outside the box." Both the film and the Gnostics assert that the "divine spark" within humans allows a perception of gnosis greater than that achievable by even the chief archon / agent of Yaldabaoth:

And the power of the mother [Sophia, in our analogy, humankind] went out of Yaldabaoth [AI] into the natural body which they had fashioned [the humans grown on farms by AI]... And in that moment the rest of the powers [archons / Agents] became jealous, because he had come into being through all of them and they had given their power to the man, and his intelligence ["mind"] was

greater than that of those who had made him, and greater than that of the chief archon [Agent Smith?]. And when they recognized that he was luminous, and that he could think better than they... they took him and threw him into the lowest region of all matter [simulated by the Matrix]. (Apocry. of John 19-20)

It is striking that Neo overcomes Agent Smith in the final showdown of the film precisely by realizing fully the illusion of the Matrix, something the Agent apparently cannot do, since Neo is subsequently able to break rules that the Agent cannot. His final defeat of Smith entails entering Smith's body and splitting him in pieces by means of pure luminosity, portrayed through special effects as light shattering Smith from the inside out.

Overall, then, the system portrayed in *The Matrix* parallels Gnostic Christianity in numerous respects, especially the delineation of humanity's fundamental problem of existing in a dreamworld that simulates reality and the solution of waking up from illusion. The central mythic figures of Sophia, Yaldabaoth, the archons and the Gnostic Christ redeemer also each find parallels with key figures in the film and function in similar ways. The language of Gnosticism and the film are even similar: dreaming vs. waking; blindness vs. seeing;²⁶ light vs. dark.²⁷

However, given that Gnosticism presumes an entire unseen realm of divine beings, where is God in the film? In other words, when Neo becomes sheer light, is this a symbol for divinity, or for human potential? The question becomes even more pertinent with the identification of humankind with Sophia - a divine being in Gnosticism. On one level, there appears to be no God in the film. Although there are apocalyptic motifs, Conrad Ostwalt rightly argues that unlike conventional Christian apocalypses, in *The Matrix* both the catastrophe and its solution are of human making - that is, the divine is not apparent.²⁸ However, on another level, the film does open up the possibility of a God through the figure of the Oracle, who dwells inside the Matrix and yet has access to information about the future that even those free from the Matrix do not possess. This suggestion is even stronger in the original screenplay, in which the Oracle's apartment is the Holy of Holies nested within the "Temple of Zion."²⁹ Divinity may also play a role in Neo's past incarnation and his coming again as the One. If, however, there is some implied divinity in the film,³⁰ it remains transcendent, like the divinity of the ineffable, invisible supreme god in Gnosticism, except where it is immanent in the form of the divine spark active in humans.³¹

III. Buddhism in *The Matrix*

When asked by a fan if Buddhist ideas influenced them in the production of the movie, the Wachowski brothers offered an unqualified "Yes."³² Indeed, Buddhist ideas pervade the film and appear in close proximity with the equally strong Christian imagery. Almost immediately after Neo is identified as "my own personal Jesus Christ," this appellation is given a distinctively Buddhist twist. The same hacker says: "This never happened. You don't exist." From the stupa-like³³ pods which encase humans in the horrific mechanistic fields to Cypher's selfish desire for the sensations and pleasures of the Matrix, Buddhist teachings form a foundation for much of the film's plot and imagery.³⁴

The Problem of Samsara. Even the title of the film evokes the Buddhist worldview. The Matrix is described by Morpheus as "a prison for your mind." It is a dependent "construct" made up of the interlocking digital projections of billions of human beings who are unaware of the illusory nature of the reality in which they live and are completely dependent on the hardware attached to their real bodies and the elaborate software programs created by AI. This "construct" resembles the Buddhist idea of samsara, which teaches that the world in which we live our daily lives is constructed only from the sensory projections formulated from our own desires. When Morpheus takes Neo into the "construct" to teach him about the Matrix, Neo learns that the way in which he had perceived himself in the Matrix was nothing more than "the mental projection of your digital self." The "real" world, which we associate with what we feel, smell, taste, and see, "is simply electrical signals interpreted by your brain." The world, Morpheus explains, exists "now only as part of a neural interactive simulation that we call the Matrix." In Buddhist terms, we could say that "because it is empty of self or of what belongs to self, it is therefore said: 'The world is empty.' And what is empty of self and what belongs to self? The eye, material shapes, visual consciousness, impression on the eye -- all these are empty of self and of what belongs to self."³⁵ According to Buddhism and according to *The Matrix*, the conviction of reality based upon sensory experience, ignorance, and desire keeps humans locked in illusion until they are able to recognize the false nature of reality and relinquish their mistaken sense of identity.

Drawing upon the Buddhist doctrine of Dependent Co-Origination, the film presents reality within the Matrix as a conglomerate of the illusions of all humans caught within its snare. Similarly, Buddhism teaches that the suffering of human beings is dependent upon a cycle of ignorance and desire which locks humans into a repetitive cycle of birth, death, and rebirth. The principle is stated in a short formula in the Samyutta-nikaya:

If this is that comes to be;
from the arising of this that arises;
if this is not that does not come to be;
from the stopping of this that is stopped.³⁶

The idea of Dependent Co-Origination is illustrated in the context of the film through the illusion of the Matrix. The viability of the Matrix's illusion depends upon the belief by those enmeshed in it that the Matrix itself is reality. AI's software program is, in and of itself, no illusion at all. Only when humans interact with its programs do they become enmeshed in a corporately-created illusion, the Matrix, or samsara, which reinforces itself through the interactions of those beings involved within it. Thus the Matrix's reality only exists when actual human minds subjectively experience its programs.³⁷

The problem, then, can be seen in Buddhist terms. Humans are trapped in a cycle of illusion, and their ignorance of this cycle keeps them locked in it, fully dependent upon their own interactions with the program and the illusions of sensory experience which these provide, and the sensory projections of others. These projections are strengthened by humans' enormous desire to believe that what they perceive to be real is in fact real. This desire is so strong that it overcomes Cypher, who can no longer tolerate the "desert of the real" and asks to be reinserted into the Matrix. As he sits with Agent Smith in an upscale restaurant smoking a cigar with a large glass of brandy, Cypher explains his motives:

"You know, I know this steak doesn't exist. I know that when I put it in my mouth, the Matrix is telling my brain that it is juicy and delicious. After nine years, you know what I realize? Ignorance is bliss."³⁸

Cypher knows that the Matrix is not real and that any pleasures he experiences there are illusory. Yet for him, the "ignorance" of samsara is preferable to enlightenment. Denying the reality that he now experiences beyond the Matrix, he uses the double negative: "I don't want to remember nothing. Nothing. And I want to be rich. Someone important. Like an actor." Not only does Cypher want to forget the "nothing" of true reality, but he also wants to be an "actor," to add another level of illusion to the illusion of the Matrix that he is choosing to re-enter.³⁹ The draw of samsara is so strong that not only does Cypher give in to his cravings, but Mouse also may be said to have been overwhelmed by the lures of samsara, since his death is at least in part due to distractions brought on by his sexual fantasies about the "woman in the red dress" which occupy him when he is supposed to be standing alert.

Whereas Cypher and Mouse represent what happens when one gives in to samsara, the rest of the crew epitomize the restraint and composure praised by the Buddha. The scene shifts abruptly from the restaurant to the mess hall of the Nebuchadnezzar,

where instead of being offered brandy, cigars and steak, Neo is given the "bowl of snot" which is to be his regular meal from that point forward. In contrast to the pleasures which for Cypher can only be fulfilled in the Matrix, Neo and the crew must be content with the "single-celled protein combined with synthetic aminos, vitamins, and minerals" which Dozer claims is "everything the body needs." Clad in threadbare clothes, subsisting on gruel, and sleeping in bare cells, the crew is depicted enacting the Middle Way taught by the Buddha, allowing neither absolute asceticism nor indulgence to distract them from their work.⁴⁰

The Solution of Knowledge/Enlightenment. This duality between the Matrix and the reality beyond it sets up the ultimate goal of the rebels, which is to free all minds from the Matrix and allow humans to live out their lives in the real world beyond. In making this point, the film-makers draw on both Theravada and Mahayana Buddhist ideas.⁴¹ Alluding to the Theravada ideal of the arhat, the film suggests that enlightenment is achieved through individual effort.⁴² As his initial guide, Morpheus makes it clear that Neo cannot depend upon him for enlightenment. Morpheus explains, "No one can be told what the Matrix is. You have to see it for yourself." Morpheus tells Neo he must make the final shift in perception entirely on his own. He says: "I'm trying to free your mind, Neo. But I can only show you the door. You're the one that has to walk through it." For Theravada Buddhists, "man's emancipation depends on his own realization of the Truth, and not on the benevolent grace of a god or any external power as a reward for his obedient good behavior."⁴³ The Dhammapada urges the one seeking enlightenment to "Free thyself from the past, free thyself from the future, free thyself from the present. Crossing to the farther shore of existence, with mind released everywhere, no more shalt thou come to birth and decay."⁴⁴ As Morpheus says to Neo, "There's a difference between knowing the path and walking the path." And as the Buddha taught his followers, "You yourselves should make the effort; the Awakened Ones are only teachers."⁴⁵ As one already on the path to enlightenment, Morpheus is only a guide; ultimately Neo must recognize the truth for himself.

Yet *The Matrix* also embraces ideas found in Mahayana Buddhism, especially in its particular concern for liberation for all people through the guidance of those who remain in samsara and postpone their own final enlightenment in order to help others as bodhisattvas.⁴⁶ The crew members of the Nebuchadnezzar epitomize this compassion. Rather than remain outside of the Matrix where they are safer, they choose to re-enter it repeatedly as ambassadors of knowledge with the ultimate goal of freeing the minds and eventually also the bodies of those who are trapped within the Matrix's digital web. The film attempts to blend the Theravada ideal of the arhat

with the Mahayana ideal of the bodhisattva, presenting the crew as concerned for those still stuck in the Matrix and willing to re-enter the Matrix to help them, while simultaneously arguing that final realization is an individual process.

Neo as the Buddha. Although the entire crew embodies the ideals of the bodhisattva, the filmmakers set Neo apart as unique, suggesting that while the crew may be looked at as arhats and bodhisattvas, Neo can be seen as a Buddha. Neo's identity as the Buddha is reinforced not only through the anagram of his name but also through the myth that surrounds him. The Oracle has foretold the return of one who has the ability to manipulate the Matrix. As Morpheus explains, the return of this man "would hail the destruction of the Matrix, end the war, bring freedom to our people. That is why there are those of us who have spent our entire lives searching the Matrix, looking for him." Neo, Morpheus believes, is a reincarnation of that man and like the Buddha, he will be endowed with extraordinary powers to aid in the enlightenment of all humanity.

The idea that Neo can be seen as a reincarnation of the Buddha is reinforced by the prevalence of birth imagery in the film directly related to him. At least four incarnations are perceptible in the film. The first birth took place in the pre-history of the film, in the life and death of the first enlightened one who was able to control the Matrix from within. The second consists of Neo's life as Thomas Anderson. The third begins when Neo emerges, gasping, from the gel of the eerily stupa-like pod in which he has been encased, and is unplugged and dropped through a large black tube which can easily be seen as a birth canal.⁴⁷ He emerges at the bottom bald, naked, and confused, with eyes that Morpheus tells him have "never been used" before. Having "died" to the world of the Matrix, Neo has been "reborn" into the world beyond it. Neo's fourth life begins after he dies and is "reborn" again in the closing scenes of the film, as Trinity resuscitates him with a kiss.⁴⁸ At this point, Neo perceives not only the limitations of the Matrix, but also the limitations of the world of the Nebuchadnezzar, since he overcomes death in both realms. Like the Buddha, his enlightenment grants him omniscience and he is no longer under the power of the Matrix, nor is he subject to birth, death, and rebirth within AI's mechanical construct.⁴⁹

Neo, like the Buddha, seeks to be free from the Matrix and to teach others how to free themselves from it as well, and any use of superhuman powers are engaged to that end. As the only human being since the first enlightened one who is able to freely manipulate the software of the Matrix from within its confines, Neo represents the actualization of the Buddha-nature, one who can not only recognize the "origin of pain in the world of living beings," but who can also envision "the stopping of the pain,"

enacting "that course which leads to its stopping."⁵⁰ In this sense, he is more than his bodhisattva companions, and offers the hope of awakening and freedom for all humans from the ignorance that binds them.

The Problem of *Nirvana*. But what happens when the Matrix's version of reality is dissolved? Buddhism teaches that when samsara is transcended, nirvana is attained. The notion of self is completely lost, so that conditional reality fades away, and what remains, if anything, defies the ability of language to describe. In his re-entry into the Matrix, however, Neo retains the "residual self-image" and the "mental projection of [a] digital self." Upon "enlightenment," he finds himself not in nirvana, or no-where, but in a different place with an intact, if somewhat confused, sense of self which strongly resembles his "self" within the Matrix. Trinity may be right that the Matrix "cannot tell you who you are," but who you are seems to be at least in some sense related to who you think you are in the Matrix. In other words, there is enough continuity in self-identity between the world of the Matrix and "the desert of the real" that it seems probable that the authors are implying that full "enlightenment" has not yet been reached and must lie beyond the reality of the Nebuchadnezzar and the world it inhabits. If the Buddhist paradigm is followed to its logical conclusions, then we have to expect at least one more layer of "reality" beyond the world of the crew, since even freed from the Matrix they are still subject to suffering and death and still exhibit individual egos.

This idea is reinforced by what may be the most problematic alteration which *The Matrix* makes to traditional Buddhist teachings. The Buddhist doctrine of ahimsa, or non-injury to all living beings, is overtly contradicted in the film.⁵¹ It appears as if the filmmakers deliberately chose to link violence with salvific knowledge, since there seems to be no way that the crew could succeed without the help of weaponry. When Tank asks Neo and Trinity what they need for their rescue of Morpheus "besides a miracle," their reply is instantaneous: "Guns -- lots of guns." The writers could easily have presented the "deaths" of the Agents as nothing more than the ending of that particular part of the software program. Instead, the Wachowski brothers have purposefully chosen to portray humans as innocent victims of the violent deaths of the Agents.⁵² This outright violation of ahimsa stands at direct odds with the Buddhist ideal of compassion.

But why link knowledge so directly with violence? The filmmakers portray violence as redemptive,⁵³ and as absolutely essential to the success of the rebels. *The Matrix* steers sharply away at this point from the shared paradigms of Buddhism and Gnostic Christianity. The "reality" of the Matrix which requires that some humans must die as

victims of salvific violence is not the ultimate reality to which Buddhism or Gnostic Christianity points. Neither the "stillness" of the *pleroma* nor the unchanging "nothingness" of *nirvana* are characterized by the dependence on technology and the use of force which so characterizes both of the worlds of the rebels in *The Matrix*.

The film's explicit association of knowledge with violence strongly implies that Neo and his comrades have not yet realized the ultimate reality. According to the worldviews of both Gnostic Christianity and Buddhism that the film evokes, the realization of ultimate reality involves a complete freedom from the material realm and offers peace of mind. The Wachowskis themselves acknowledge that it is "ironic that Morpheus and his crew are completely dependent upon technology and computers, the very evils against which they are fighting."⁵⁴ Indeed, the film's very existence depends upon both technology's capabilities and Hollywood's hunger for violence. Negating itself, *The Matrix* teaches that *nirvana* is still beyond our reach.

IV. Concluding Remarks

Whether we view the film from a Gnostic Christian or Buddhist perspective, the overwhelming message seems to be, "Wake up!" The point is made explicit in the final song of the film, *Wake Up!*, by, appropriately, *Rage Against the Machine*. Gnosticism, Buddhism and the film all agree that ignorance enslaves us in an illusory material world and that liberation comes through enlightenment with the aid of a teacher or guide figure. However, when we ask the question, "To what do we awaken?", the film appears to diverge sharply from Gnosticism and Buddhism. Both of these traditions maintain that when humans awaken, they leave behind the material world. The Gnostic ascends at death to the *pleroma*, the divine plane of spiritual, non-material existence, and the enlightened one in Buddhism achieves *nirvana*, a state which cannot be described in language, but which is utterly non-material. By contrast, the "desert of the real," is a wholly material, technological world, in which robots grow humans for energy, Neo can learn martial arts in seconds through a socket inserted into the back of his brain, and technology battles technology (Nebuchadnezzar vs. AI, electromagnetic pulse vs. Sentinels). Moreover, the battle against the Matrix is itself made possible through technology - cell phones, computers, software training programs. "Waking up" in the film is leaving behind the Matrix and awakening to a dismal cyber-world, which is the real material world.

Or perhaps not. There are several cinematic clues in the scene of the construct loading program (represented by white space) that suggest that the "desert of the real" Morpheus shows Neo may not be the ultimate reality. After all, Morpheus, whose

name is taken from the god of dreams, shows the "real" world to Neo, who never directly views the surface world himself. Rather, he sees it on a television bearing the logo "Deep Image." Throughout the film, reflections in mirrors and Morpheus's glasses, as well as images on television monitors point the viewer toward consideration of multiple levels of illusion.⁵⁵ As the camera zooms in to the picture on this particular television and the viewer "enters" the image, it "morphs" the way the surveillance screens do early in the film, indicating its unreality. In addition, the entire episode takes place while they stand in a construct loading program in which Neo is warned not to be tricked by appearances. Although sense perception is clearly not a reliable source for establishing reality, Morpheus himself admits that, "For a long time I wouldn't believe it, and then I saw the fields [of humans grown for energy] with my own eyes... And standing there, I came to realize the obviousness of the truth." We will have to await the sequels to find out whether "the desert of the real" is itself real.⁵⁶

Even if the film series does not ultimately establish a complete rejection of the material realm, *The Matrix* as it stands still asserts the superiority of the human capacity for imagination and realization over the limited "intelligence" of technology. Whether stated in terms of matter/ spirit, body/ mind, hardware/ software or illusion/ truth, the ultimate message of *The Matrix* seems to be that there may be levels of metaphysical reality beyond what we can ordinarily perceive, and the film urges us to open ourselves to the possibility of awakening to them.

Frances Flannery-Dailey & Rachel Wagner

[Note: This essay originally appeared in *The Journal of Religion and Film*]

Endnotes

1. All unidentified quotes are from *The Matrix* (Warner Bros. release, 1999).
2. In an online chat with viewers of the DVD, the Wachowskis acknowledged that the Buddhist references in the film are purposeful. However, when asked "Have you ever been told that the Matrix has Gnostic overtones?", they gave a tantalizingly ambiguous reply: "Do you consider that to be a good thing?" From the Nov. 6, 1999 "Matrix Virtual Theatre," at "Wachowski chat"
3. Elaine Pagels notes that the similarities between Gnosticism and Buddhism have prompted some scholars to question their interdependence and to wonder whether "...if the names were changed, the 'living Buddha' appropriately could say what the *Gospel of Thomas* attributes to the living Jesus." Although intriguing, she rightly maintains that the evidence is inconclusive, since parallel traditions may emerge in different cultures without direct influence. Elaine Pagels, *The Gnostic Gospels*, (New York: Random House, 1979, repr. 1989), xx-xxi
4. James Ford recently explored other Buddhist elements in *The Matrix*, which he rightly calls a "modern myth," in his article "Buddhism, Christianity and *The Matrix*: The Dialectic of Myth-Making in Contemporary Cinema," for the *Journal of Religion and Film*, vol.4 no. 2. See also Conrad Ostwalt's focus on apocalyptic elements of the film in "Armageddon at the Millennial Dawn," JRF vol.

4, no. 1

5. A viewer asked the Wachowski brothers, "Your movie has many and varied connections to myths and philosophies, Judeo-Christian, Egyptian, Arthurian, and Platonic, just to name those I've noticed. How much of that was intentional?" They replied, "All of it" (Wachowski chat).

6. Feminists critics can rejoice when Trinity first reveals her name to Neo, as he pointedly responds, "The Trinity?... Jesus, I thought you were a man." Her quick reply: "Most men do."

7. The Wachowski brothers indicate that the names were "all chosen carefully, and all of them have multiple meanings," and also note this applies to the numbers as well (Wachowski chat).

8. In a recent interview in *Time*, the Wachowskis refer to Nebuchadnezzar in this Danielic context, (www.time.com/time/magazine/article/0,9171,22971,00.html , "Popular Metaphysics," by Richard Corliss, *Time*, April 19, 1999 Vol. 153, no. 15). Nebuchadnezzar is also the Babylonian king who destroyed the Jerusalem Temple in 586 B.C.E., and who exiled the elite of Judean society to Babylon. Did the Wachowski brothers also intend the reference to point to the crew's "exile" from Zion or from the surface world?

9. The film also suggests Zion is heaven, such as when Tank says, "If the war was over tomorrow, Zion is where the party would be," evoking the traditional Christian schema of an apocalypse followed by life in heaven or paradise. Ironically, the film locates Zion "underground, near the Earth's core, where it is still warm," which would seem to be a cinematic code for hell. Is this a clue that Zion is not the "heaven" we are led to believe it is?

10. Neo's apartment number is 101, symbolizing both computer code (written in 1s and 0s) and his role as "the One." Near the end of the film, 303 is the number of the apartment that he enters and exits in his death / resurrection scene, evoking the Trinity. This in turn raises questions about the character of Trinity's relationship to Neo in terms of her cinematic construction as divinity.

11. The traitor Cypher, who represents Judas Iscariot, among other figures, ironically says to Neo, "Man, you scared the B'Jesus outta me."

12. We would like to thank Donna Bowman, with whom we initially explored the Gnostic elements of *The Matrix* during a public lecture on film at Hendrix College in 2000.

13. Gnosticism may have had its origins in Judaism, despite its denigration of the Israelite God, but the issue is complex and still debated within scholarly circles. It is clear, however, that Gnostic Christianity flourished from at least the 2nd - 5th c. C.E., with its own scriptures, and most likely also its own distinctive rituals, entrance requirements and a creation story. See Gershom Scholem, *Jewish Gnosticism, Merkabah Mysticism, and Talmudic Tradition* (New York: Jewish Theological Seminary of America, 1960), Elaine Pagels, *The Gnostic Gospels* (New York: Vintage Books, 1979, repr. 1989), Bentley Layton, *The Gnostic Scriptures* (New York: Doubleday, 1995), Kurt Rudolph, *Gnosis: The Nature and History of Gnosticism* (San Francisco: HarperSanFrancisco, 1987).

14. This corpus lay dormant for nearly 2000 years until its discovery in 1945 in Nag Hammadi, Egypt. The complete collection of texts may be found in James M. Robinson, ed. *The Nag Hammadi Library*, revised edition, (New York: HarperCollins, 1990; reprint of original Brill edition, 1978). These documents are also available on-line at The Nag Hammadi Library Section of The Gnostic Society Library.

15. Gnostic texts are cryptic, and no single text clearly explains this myth from beginning to end. The literature presupposes familiarity with the myth, which must be reconstructed by modern readers. The version of the myth presented here relies on such texts as *Gospel of Truth*, *Apocryphon of John*, *On the Origin of the World* and *Gospel of Thomas*. See *The Nag Hammadi Library*, pp. 38-51, 104-123, 124-138, 170-189.

16. Since the divine beings are composed only of spiritual substances and not matter, there are no physical gender differences among the beings.

17. Depending on the text, a plethora of divine beings populate the pleroma, many with Jewish, Christian or philosophical names, e.g. the Spirit, forethought, thought, foreknowledge,

indestructibility, truth, Christ, Autogenes, understanding, grace, perception, Piger-Adamas (*Apocryphon of John*).

18. Humanity's characterization also resonates with the Tower of Babel story in Genesis 11:1-9; in both we admire the work of our own hands.

19. The bulk of the following excerpt from the Gnostic "Gospel of Truth" might just as well be taken from the scenes in *The Matrix* in which Morpheus explains the nature of reality to Neo: Thus they [humans] were *ignorant* of the Father, he being the one whom they did not *see*... there were many *illusions* at work... and (there were) empty fictions, *as if they were sunk in sleep and found themselves in disturbing dreams*. Either (there is) a place to which they are fleeing, or without strength they come (from) having chased after others, or they are involved in striking blows, or they are receiving blows themselves, or they have fallen from high places, *or they take off into the air though they do not even have wings*. Again, sometimes (it is as) if people were murdering them, though there is no one even pursuing them, or they themselves are killing their neighbors... (but) *When those who are going through all these things wake up, they see nothing, they who were in the midst of all these disturbances, for they are nothing*. Such is the way of those who have cast ignorance aside from them like sleep, not esteeming it as anything, nor do they esteem its works as solid things either, but *they leave them behind like a dream in the night*... This is the way each one has acted, as though asleep at the time when he was ignorant. And this is the way he has [come to knowledge], as if he had awakened. (*Gospel of Truth*, 29-30)

20. This is perhaps most evident in the subway fight between Neo and Agent Smith. At a point in the film when Morpheus says of Neo, "He is just beginning to believe," Agent Smith calls him "Mr. Anderson," and while fighting he replies, "My name is Neo." The Wachowskis confirm this interpretation when they state "Neo is Thomas Anderson's potential self" (Wachowski chat).

21. This twin tradition was especially popular in Syrian Christianity. See also Pagels, p. xxi, where she wonders if the tradition that Thomas, Jesus' twin, went to India points to any historical connection between Buddhism and Hinduism on the one hand and with Gnosticism on the other.

22. See the online chat with the special effects creators in the "Matrix Virtual Theater" from March 23, 2000.

23. *Nag Hammadi Library*, pp. 490-500. Compare the Gnostic idea of stillness with these Buddhist sayings from the *Dhammapada*: "The bhikku [monk], who abides in loving-kindness, who is delighted in the Teaching of the Buddha, attains the State of Calm, the happiness of stilling the conditioned things" and "Calm is the thought, calm the word and deed of him who, rightly knowing, is wholly freed, perfectly peaceful and equipoised." Quoted in Walpola Sri Rahula, *What the Buddha Taught* (New York: Grove Weidenfeld, 1974) p.128, 136.

24. See *Nag Hammadi Library*, pp. 256-59. We are grateful to Brock Bakke for the initial equation of agents with archons.

25. In Gnosticism "Mind" or the Greek "nous" is a deity, such as in the text "Thunder, Perfect Mind," *Nag Hammadi Library*, 295-303.

26. Note that as Morpheus and Neo enter the elevator of the apartment building of the Oracle, images of "seeing" symbolize prophecy and knowledge: a blind man (evoking blind prophets such as Tiresias) sits in the lobby beneath some graffiti depicting a pair of eyes. Interestingly, the Oracle - a sibyl / seer - wears glasses to look at Neo's palm.

27. Note too the metonymic use of color to convey this dualism: black and white clothing, floors, furniture, etc.

28. Ostwalt, "Armageddon" in JRF Vol. 4, no. 1. The parallel with apocalypticism does not work quite as well as one with Gnosticism because like Gnosticism, the film understands salvation to be individual (rather than collective and occurring all at once), to be attained through knowledge, and most importantly to entail leaving behind the material Earth (that is, not resulting in a kingdom of God made manifest on the Earth).

29. In its description in the original screenplay, the Temple of Zion evokes both the Oracle of Delphi (three legged stool, priestesses) and the Jerusalem Temple (polished marble, empty throne which is the mercy seat or throne of the invisible God).

30. A viewer asked the Wachowski brothers, "What is the role or {sic} faith in the movie? Faith in oneself first and foremost – or in something else?" They answered, "Hmmm...that is a tough question! Faith in one's self, how's that for an answer?" This reply hardly settles the issue (Wachowski chat).

31. Specifically, these humans are Neo (the Gnostic Redeemer / Messiah) and Morpheus and Trinity, both of whom are named for gods. As a godhead, this trio does not quite make sense in terms of traditional Christianity. However, the trio is quite interesting in the context of Gnosticism, which portrays God as Father, Mother and Son, a trinity in which the Holy Spirit is identified as female, e.g. *Apocryphon of John* 2:9-14. For further reading on female divinities in Gnosticism, see Pagels, pp. 48-69.

32. The brothers explain, "There's something uniquely interesting about Buddhism and mathematics, particularly about quantum physics, and where they meet. That has fascinated us for a long time" (Wachowski chat). In the Time interview with Richard Corliss (see note 8), Larry Wachowski adds that they became fascinated "by the idea that math and theology are almost the same. They begin with a supposition you can derive a whole host of laws or rules from. And when you take all of them to the infinity point, you wind up at the same place: these unanswerable mysteries really become about personal perception. Neo's journey is affected by all these rules, all these people trying to tell him what the truth is. He doesn't accept anything until he gets to his own end point, his own rebirth." The film's presentation of the Matrix as a corporate network of human conceptions (or samsara) which are translated into software codes that reinforce one another illustrates this close relationship.

33. Stupa: a hemispherical or cylindrical mound or tower serving as a Buddhist shrine.

34. Of course, the most transparent reference to Buddhist ideas occurs in the waiting room at the Oracle's apartment, where Neo is introduced to the "Potentials." The screenplay describes the waiting room as "at once like a Buddhist temple and a kindergarten class." One of the children, clad in the garb of a Buddhist monk, explains to Neo the nature of ultimate reality: "There is no spoon." One cannot help wondering if this dictum only holds within the Matrix or if there is in fact "no spoon" even in the real world beyond it.

35. *Samyutta-nikaya* IV, 54. In Edward Conze, ed. *Buddhist Texts Through the Ages* (New York: Philosophical Library, 1954), p. 91.

36. *Samyutta-nikaya* II, 64-65. Ibid.

37. The entire process depends upon human ignorance, so that almost all who are born into the Matrix are doomed to be born, to die, and to re-enter the cycle again. When asked about the film's depiction of the liquefaction of humans, the Wachowskis reply that this black ooze is "what they feed the people in the pods, the dead people are liquefied and fed to the living people in the pods." Tongue in Buddhist cheek, the brothers explain this re-embodiment: "Always recycle! It's a statement on recycling." (Wachowski Chat) Even in the "real world" beyond the Matrix, the human plight is depicted as a relative and inter-dependent cycle of birth, death, and "recycling."

38. (Ed. Note: This clip can be viewed [here](#). (Hit your back button to return to this essay.))

39. This dialogue also points to the "reality" (or the "Matrix") which we ourselves inhabit. In our world, and in the world of Joe Pantoliano, he is an actor. Therefore, the world of which both the actor Joe Pantoliano and we are now a part may be seen as the "Matrix" into which he has been successfully re-inserted, and thus the film itself may be seen as a part of the software program of our own "Matrix." The argument, of course, is seductively circular.

40. Take, for example, this quote from the *Sabbasava-sutta*: "A bhikku [monk], considering wisely, lives with his eyes restrained . . . Considering wisely, he lives with his ears restrained . . . with his nose restrained . . . with his tongue . . . with his body . . . with his mind restrained . . . a bhikku, considering wisely, makes use of his robes -- only to keep off cold, to keep off heat . . . and to cover

himself decently. Considering wisely, he makes use of food – neither for pleasure nor for excess . . . but only to support and sustain this body . . ." (Quoted in Rahula 103).

41. James Ford has argued that the film embodies in particular the Yogacara school of Buddhism. Instead of pointing to that which is absolutely different than the world as *nirvana*, Yogacarins point to the world itself, and through the processes enacted in meditation, come to the realization that "all things and thought are but Mind-only. The basis of all our illusions consists in that we regard the objectifications of our own mind as a world independent of that mind, which is really its source and substance" (Edward Conze, *Buddhism*. New York: Philosophical Library, 1959), p. 167. The Matrix exists only in the minds of the human beings which inhabit it, so that in *The Matrix*, as in Yogacara, "The external world is really Mind itself" (p. 168). Yet a problem arises when one realizes that for the Yogacara school, the Mind is the ultimate reality, and therefore *samsara* and *nirvana* become identified. By contrast, the film insists on a *distinction* between *samsara* (the Matrix) and *nirvana* (that which lies beyond it). Because *The Matrix* maintains a duality between the Matrix and the realm beyond it, Yogacara is of limited help in making sense of the Buddhist elements in the film, nor is it helpful in supporting the idea that beyond the Matrix and beyond the Nebuchadnezzar there is an ultimate reality not yet realized by humans (see note 4).

42. According to Theravada teachings, *arhat* ("Worthy One") is a title applied to those who achieve enlightenment. Because, according to Theravada beliefs, enlightenment can only be achieved through individual effort, an *arhat* is of limited aid in helping those not yet enlightened and so would not necessarily choose to re-enter *samsara* to aid others still enmeshed within it.

43. Rahula, p. 2.

44. Quoted in Rahula, 135.

45. Quoted in Rahula, 133.

46. A bodhisattva is one who postpones final entry into *nirvana* and willingly re-enters or remains in *samsara* in order to guide others along the path to enlightenment. The Buddha's compassion serves as their primary model for Mahayana Buddhists, since they point out that he too remained in *samsara* in order to help others achieve enlightenment through his teachings and example.

47. The screenplay describes Neo as "floating in a womb-red amnion" in the power plant.

48. In the screenplay, Trinity does not kiss him but instead "pounds on his chest," precipitating his resuscitation. The screenplay states directly: "It is a miracle." This fourth "life" can be viewed as the one to which the Oracle refers in her predictions that Neo was "waiting for something" and that he might be ready in his "next life, maybe." This certainly appears to be the case, since Neo rises from the dead and defeats the Agents.

49. These four "lives" suggest that Neo is nothing other than "the One" foretold by the oracle, the reincarnation of the first "enlightened one," or Buddha, who "had the ability to change whatever he wanted, to remake the Matrix as he saw fit." Buddhist teaching allows that those who have been enlightened are endowed with magical powers, since they recognize the world as illusory and so can manipulate it at will. Yet supernatural powers are incidental to the primary goal, which is explained in the very first sermon spoken by the Buddha: "The Noble Truth of the cessation of suffering is this: It is the complete cessation of that very thirst, giving it up, renouncing it, emancipating oneself from it, detaching oneself from it" (*Dhammacakkappavattana-sutta*. Quoted in Rahula, 93.)

50. *Buddhacarita* 1:65. E. B. Cowell, trans., *Buddhist Mahayana Texts, Sacred Books of the East*, vol. 49 (Oxford: Oxford University Press, 1894).

51. See, for example, in the *Dhammapada*: "Of death are all afraid. Having made oneself the example, one should neither slay nor cause to slay" (Verse 129) (*Dhammapada*, trans. John Ross Carter and Mahinda Paliawadana. New York: Oxford University Press, 1987), p. 35.

52. The idea that violence as salvific is made explicit by the writers. Whereas they *could* have chosen to present the "deaths" of the Agents as of the same illusory quality as other elements within the software program, instead, they choose to depict *actual* humans *really* dying through the inhabitation of their "bodies" by the Agents. This addition is completely unnecessary to the overall

plot line; indeed, the "violence" which takes place in the Hotel could still be portrayed, with the reassuring belief that any "deaths" which occur there are simply computer blips. The fact that the writers so purposefully insist that actual human beings die (i.e. die also within the power plant) while serving as involuntary "vessels" for the Agents strongly argues for *The Matrix's* direct association of violence with the knowledge required for salvation.

53. See the article by Bryan P. Stone, "Religion and Violence in Popular Film," JRF Vol. 3, no. 1.

54. When asked whether this irony was intentional, the Wachowskis reply abruptly but enthusiastically "Yes!" (Wachowski chat).

55. This is especially true in the "red pill / blue pill" scene where Neo first meets Morpheus, and Neo is reflected differently in each lens of Morpheus's glasses. The Wachowskis note that one reflection represents Thomas Anderson, and one represents Neo (Wachowski chat).

56. A viewer asked the pertinent question of the Wachowskis: "Do you believe that our world is in some way similar to *The Matrix*, that there is a larger world outside of this existence?" They replied: "That is a larger question than you actually might think. We think the most important sort of fiction attempts to answer some of the big questions. One of the things that we had talked about when we first had the idea of *The Matrix* was an idea that I believe philosophy and religion and mathematics all try to answer. Which is, a reconciling between a natural world and another world that is perceived by our intellect" (Wachowski chat).

CONTRIBUTORS

Hubert Dreyfus was educated at Harvard and teaches philosophy at the University of California, Berkeley. His research interests bridge the Analytic and Continental traditions in 20th-century philosophy. He has written books on Heidegger (*Being-in-the-World*, MIT Press), and on Artificial Intelligence. (*What Computers (Still) Can't Do*, MIT Press). Dreyfus recently published *On the Internet* (Routledge), and is working on a book with Charles Taylor tentatively entitled, *Retrieving Realism*.

Stephen Dreyfus is a graduate of Video Symphony, just beginning his professional career as a digital film editor. He has worked as an Assistant Editor on several independent films. A long time amateur philosopher, he is always looking for new and interesting ways to bring his surrealist stories and ideas to the entertainment world. His email address is lorde_red@lycos.com.

Frances Flannery-Dailey received her Ph.D. from the University of Iowa, and is Assistant Professor of Religion at Hendrix College in Conway, Arkansas. She teaches courses in Bible, Religion and Culture and Judaism. Her main area of research is apocalypticism in early Judaism (300 B.C.E.-200 C.E.), and she is currently writing a book for Brill Publishers entitled *Dreamers, Mystics and Heavenly Priests: Dreams in Second Temple Judaism*.

Christopher Grau was educated at Johns Hopkins University and New York University. In addition to editing the "Philosophy & *The Matrix*" section of *The Matrix* website, Chris teaches philosophy at Brooklyn College. He has previously taught at Dartmouth College, Johns Hopkins University, and the University of Maryland, Baltimore County. His current research involves the ethical ramifications of theories of personal identity, though he has also published on the free will problem.

Richard Hanley was educated at Sydney University and the University of Maryland, College Park. He is the author of *The Metaphysics of Star Trek* (reprinted in paperback as *Is Data Human?*), and is co-editor of the forthcoming *Blackwell Guide to the Philosophy of Language*. He works on metaphysics, philosophy of language, and ethics, and dabbles in time travel fiction; and is gainfully employed in the Philosophy department at the University of Delaware.

Colin McGinn was educated at Oxford University. He has written widely on philosophy and philosophers in such publications as the *New York Review of Books*, the *London Review of Books*, the *New Republic*, and the *New York Times Book Review*. McGinn has written fourteen books, including *The Making of a Philosopher*; *The Mysterious Flame*; *The Character of Mind*; *Ethics, Evil and Fiction*; and the novel *The*

Space Trap. He is currently a Professor of Philosophy at Rutgers University.

Iakovos Vasiliou is Associate Professor of Philosophy at Brooklyn College, City University of New York. He has previously taught at Cornell, Johns Hopkins, and Georgia State University. He has published articles on Plato, Aristotle, and Wittgenstein and is currently working on a book on moral epistemology in Plato and Aristotle.

Rachel Wagner was educated at Wake Forest University and the University of Iowa. She is currently a Visiting Instructor of Religion at Southwestern University in Georgetown, Texas. At Southwestern, she teaches comparative courses in Islam, the Problem of Evil, Religion and Film, and Religion and Literature, incorporating materials from many different religious traditions. Her primary training is in Western Religions and biblical studies, and her dissertation is based on William Blake's epic poem *Jerusalem*.

Kevin Warwick is a Professor of Cybernetics at the University of Reading, UK. His book *In the Mind of the Machine* gives a warning of a future in which machines are more intelligent than humans. In 1998 and 2002 he received surgical implants which shocked the scientific community. The second of these linked his nervous system to the internet. His experiments are reported on in his autobiography *I, Cyborg*.

Further Reading:

Dancy, Jonathan. *Introduction to Contemporary Epistemology*, Blackwell, 1985.

Descartes. *The Philosophical Writings of Descartes*, tr: John Cottingham, Robert Stoothoff, Dugald Murdoch. Cambridge University Press, 1984

Stroud, Barry. *The Significance of Philosophical Scepticism*, Oxford, 1984.

Dancy, Jonathan. *Introduction to Contemporary Epistemology*, Blackwell, 1985.

Descartes. *The Philosophical Writings of Descartes*, tr: John Cottingham, Robert Stoothoff, Dugald Murdoch. Cambridge University Press, 1984.

Nagel, Thomas. *The View from Nowhere*, Oxford, 1986.

Putnam, Hilary. *Reason, Truth, and History*, Cambridge University Press, 1981.

Strawson, P.F. *Skepticism and Naturalism: Some Varieties*, Columbia University Press, 1983.

Johnston, Mark. "Reasons and Reductionism," *Philosophical Review*, 1992.

Nagel, Thomas. "Death," *Nous*, 1970.

Nozick, Robert. *Anarchy, State, and Utopia*, Basic Books, 1971.

Unger, Peter. *Identity, Consciousness, and Value*, Oxford, 1990.