Existential Phenomenology and the Brave New World of *The Matrix*

By Hubert Dreyfus

HE MATRIX RAISES SEVERAL FAMILIAR PHILOSOPHICAL PROBLEMS IN SUCH FAS-cinating new ways that students all over the country are assigning it to their philosophy professors. In so doing, they have offered us a great opportunity to illustrate some of the basic insights of existen-tial phenomenology. *The Matrix* might seem to renew Descartes's worry that, since all we ever experience are 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. But there is a way of understanding *The Matrix* that denies the mediation of mental states and shows those living in the Matrix to be in direct touch with Matrix reality. The Matrix world is public and objective, not a private subjective dream. Still, there is clearly a sense in which the Matrix world, while not merely mental, is not real either. There *is* after all a demon—the AI intelligences and their computer—that has in some sense fooled all those who accept the reality of the Matrix world. Thus, the film's account of our situation is even more disturbing than Descartes's claim that we are each confined to our own mind. The Matrix world is a vivid illustration of Descartes's additional pre-scient claim that we could *never* be in direct touch with the real world (if there is one) because we are all what we would now call *brains in vats.*[1]

But then, in choosing to leave the Matrix world for the "desert of the real," Neo and his friends are simply choosing one of two sets of systematic appearances. And, although we tend to disapprove of Cypher's choosing to return from the harsh "real world" to the Matrix world where he feels com-fortable, he is just choosing a different, more satisfying set of appearances, and so being quite sensible. Before we can rationally approve of Neo and condemn Cypher, we have to rethink what we mean by experience, illusion, and our contact with the real world, as well as freedom and control.

Hubert Dreyfus is Professor of Philosophy in the Graduate School at the University of California, Berkeley. His research interests bridge the analytic and Continental traditions in 20ᵗʰ century philosophy and include phenomenology, existentialism, philosophy of psychology, philosophy of literature, and philo-sophical implications of artificial intelligence. His most recent work is On the Internet, *and he is working on a second edition of* Being-in-the-World, *his com-mentary on Heidegger's* Being and Time.

1. 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 a sensible question to us now, it has not always been 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's 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 lives. Saint Augustine had to work hard to convince them otherwise. For example, he called attention to the fact that one did not have to read out loud. In his *Confessions*, he points out that Saint 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 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, Descartes 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, all that each of us can directly experience is the content of his or her own mind. Our access to the world is always *indirect*. Descartes then 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]

For all we could ever know, Descartes concluded, the objective external world, including our body, 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, Leibniz postulated that each of us is a windowless 'monad'.[6] A monad is a self-contained world of experience, which gets no input from external objects and other embodied 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 external world.

A generation after Leibniz, 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, each person's experiences were coordinated with the mental representations of all the others to produce what he called the phenomenal world.[7] Finally, 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 directly, whether the world and other minds existed or not, was the contents of each person's 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 Heidegger in Germany and Maurice Merleau-Ponty in France, in opposition to Husserl, contested the Cartesian view that our contact with the world and even our own bodies is mediated by internal mental representations.[9] 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.[10]

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 readiness (itself corrigible) to encounter a façade rather than the house I was set to deal with.

2. Brains in Vats

So it looks like the inner/outer distinction introduced by Descartes holds only for thoughts. At the basic level of involved skillful coping, one is simply what Merleau-Ponty calls an empty head turned toward the world. But this doesn't at all show that *The Matrix* is old-fashioned or mistaken. On the contrary, it shows that *The Matrix* has gone further than philosophers who hold that we can't get outside our *mind*. *The Matrix* suggests the more convincing condition—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 human beings' 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."[11] Likewise, Descartes used the phantom limb phenomenon to argue that other nerves brought information about the body to the brain and that from there the information passed to the mind.

It seemed to follow that, since we are each a brain in a cranial vat,[12] we can *never* be in direct contact with the world or even with our own bodies. So even if phenomenologists like Heidegger, Merleau-Ponty, and Taylor seem 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 our nervous system copy the complex feedback loop between the brain's out-going behavior-producing impulses and the incoming perceptual ones, we would have the experience of directly coming to grips with things in the world. Yet, in the brain in the vat case, there would be no house and no city—indeed, no real world—to interact with, and so we would be confined to our inner experiences after all. As Morpheus says to Neo in the construct:

> How do you define "real"? If you're talking about what you can feel, what you can smell, what you can taste and see, then "real" is simply electrical signals interpreted by your brain...

But this Cartesian conclusion is mistaken. The inner electrical impulses are the *causal basis* of what one can feel and taste, but we don't feel and taste *them.* Even if I have only a phantom limb, my pain is not in my brain but in my phantom hand. What the existential phenomenologist can and should claim is that, in a Matrix world, which has its causal basis in bodies in vats outside that world, the Matrix people whose brains are getting computer-generated inputs and responding with action outputs are directly coping with *perceived* reality, and that this reality isn't *inner*. Even in the Matrix world, people directly cope with chairs by sitting on them and need baseballs to bring out their batting skills. 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 construct have a lot of false beliefs: for example, that its Matrix body is its real body whereas its real body is in a vat? No. If the ordinary Matrix dweller has a pain in his damaged foot, it's in his Matrix foot, not in the foot of a body in a vat—a foot that is not damaged and about which he knows nothing at all. It's a mistake to think that each of us is experiencing a set of neural firings in a brain in a cranial vat. True, each of us has a brain in his or her skull and the brain provides the causal basis of our experience, but we aren't 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 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.

Since the only body a Matrix dweller sees, feels, and moves is the one he has in the Matrix world, the AI programmers could have given him a Matrix body radically unlike the body in the vat. After all, the brain in the vat started life as a baby brain and could have been given any content the AI programmers 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.[13]

But there is still at least one problem. The Matricians' beliefs about the properties and uses of their perceived bodies—and of chairs, cities, and the world—may be shared and reliable, and in that sense true, but what about their *causal* beliefs? They believe, as we do, that germs cause disease, that the sun causes things to get warm, that gravity causes things to fall, and so forth. Aren't all these beliefs false? That depends on their understanding of causality. People don't normally have explicit *beliefs* about the nature of causality. Rather, they simply take for granted a shared sense that they are coping with a shared world whose contents are causing their experiences. 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 that they can cope with things successfully. If, however, the philosophers in the Matrix world believe that there is a physical universe with causal powers that make things happen in the world, and so accounts for the reliability of equipment and the success of science, they are mistaken. But, if they claim that causality is merely our response to the constant conjunctions of experiences as Hume did, or that it is a necessary succession according to a rule as Kant held, then their causal beliefs would be true of the causal relations in the Matrix world.[14]

Kant claims that we experience a public, objective world and that science then relates these appearances by rules we call laws, but we can't know the ground of the phenomena we perceive. Specifically, according to Kant, we experience the world *as* in space and time, but things in themselves aren't in space and time. So Kant says we can know the *phenomenal* world of objects and their law-like relations, but we can't know the *things in themselves* that are the ground of these *appearances*.

The Matricians are in the same epistemological position that we are all in according to Kant. If there are Kantians in the Matrix world, they would understand that they are experiencing a coordinated system of appearances, and would understand too that they couldn't know the things in themselves that are the ground of these appearances; that is, they couldn't know the basis of their shared experience of the world and the 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. And, as long as Kantians, and everyone else in the Matrix, didn't claim to know about things in themselves, most of their beliefs would 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. 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 about 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 have and couldn't have had while in the Matrix. We have seen that existential phenomenologists acknowledge that we are sometimes mistaken about particular things and have to retroactively take back our set to cope with 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 real world can see that much of what they took for granted about the basis of their experience before was mistaken. They can, for example, understand that what they took to be a world that had been around for millions of years was a recently constructed computer program.

Of course, things are not so simple. Neo's current beliefs might still all be false. His experience is, after all, sustained by a brain in a skull in a vat, and the AI programmers might now be feeding that brain the experience of being outside the Matrix and 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—–boundary conditions with which we must get in sync in order to cope successfully. In this way, 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 the boundary conditions *and the causal structures discovered by science* are both the effect of an unknowable thing in itself that is the ground of appearances as postulated by Kant, or whether the cause of all appearances is a computer—this is something we could never know from inside our world. But Neo—once he is on the hovercraft—does know that, as in waking from a dream, his current understanding of reality supercedes and crosses out his former one.[15]

3. An Ethical Interlude

THE DISTINCTION BETWEEN A MATRIX PERSON AND THE BODY THAT IS THE CAUSAL BASIS of that person has serious ethical implications. In the movie, innocent people doing their job, like the police officers, are killed with casual unconcern, if not with relish.[16] But when we remember that each time a Matrician is killed an associated human body somewhere in a vat dies, it seems that the killing of a virtual person in the Matrix must be morally wrong because it causes the death of a real human being.

But this can't be the right way to think about the moral issue. The bodies in the vats are not people; they are the causal basis of the people in the Matrix. They happen to be human bodies made of protoplasm but they could just as well be computers made of silicone as long as they processed the inputs and outputs the way the human brain does. It is important to bear

in mind that a body in the vat doesn't have a human personality apart from the active, vulnerable, feeling person in the Matrix of whom it is the causal basis.

Thus, when Neo is in the Matrix world, there are two Neo-related bodies, one is an active embodied Neo coping in the Matrix world and the other is a non-coping, Neo-causing body in a vat outside the Matrix world, but there is only one Neo and he stays the same in the Matrix world and later in the Hovercraft because he has the same concerns, memories, and so forth—whatever accounts for personal identity—and there never was a Neo in the vat, any more than there is a person in your skull. It follows that when Morpheus and his followers kill the people in the Matrix world it is murder, not because the killers cause a human organism in a vat to die, but because they kill Matricians who have personalities, act freely, love, suffer, and so forth. True, the way the Matrix world is set up, if one were to kill a body in a vat, the associated person in the Matrix would die. But the point to note is that the moral priorities are the reverse of one's first intuitions. The killing of a person in the Matrix world is intrinsically wrong because killing a person is wrong, and incidentally it results in the death of a human body in a vat; while killing the human body in the vat is only derivatively wrong because that body happens to be the causal basis of a person in the Matrix world. In our world, the tight causal connection between our biological body and our personhood keeps us from noticing these moral distinctions.

4. A New *Brave New World*

WE ARE NOW IN A POSITION TO UNDERSTAND AND TRY TO 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. After all, most of the beliefs of average Matricians are true; when they sit on a chair it usually supports them, when they enter a house they see the inside, they have bodies that can be injured, and they can cope by acting in some ways and not others. Their background sense that in their actions they are coping with something independent of them, and that others are coping with it too, is justified. Likewise, living in the Matrix world does not seem to be less moral than living in our everyday world. The Matricians are dealing with real other people, and they 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 like Trinity. None of the above concerns give us a grip on what, if anything, is wrong with the Matrix world.

To understand what's wrong with living in the Matrix, we have to understand the source of the power of the Matrix world. Part of the power comes from the way the inputs and outputs from the computer are plugged directly into 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. When we lose our balance, the perceived world loses its stability. These correlations produce a powerful perceptual effect 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 seat watching a movie, or just as the moon looks bigger on the horizon even though we know it isn't.

The inputs to the perceptual system of the brain in the vat produce the perceptual world whether we believe it is real or not. [17] But, once one realizes that the causality in the Matrix world is only virtual, since, according to *The Matrix*, our experience of *causality* is not built into our perceptional system, one can violate the Matrix's causal laws.[18] By the end of the movie Neo can fly, and if he wanted to, he could bend spoons. About the causal principles governing the Matrix world, Morpheus tells Neo, "it is all in your mind."[19]

What, then, is the source of the sinister 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—powers which we learn early on in the movie are free from the control of direct sensory-motor correlations. It must be some sort of mind control.

It seems that 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 to keep doing (and not doing) what "one" does and doesn't do. (As in: one eats peas with a fork, one doesn't throw food at the dinner table, and one goes out the door rather than the window.) 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 obeying shared, social norms, and that it is almost impossible for them 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 norms 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, you know there is something lacking in this world; "it's like a splinter in your mind." 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 seeing and doing things has 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 weirdness (*Unheimlichkeit*) of the world. How fitting, then, that a barely expressible unease seems to pervade Neo's life—an anxiety that prompts him to begin the process of breaking free by subverting the system. Finally, Neo has a dramatic version of an anxiety attack. When he hears that the world he has been taking for granted is a computer-generated simulation used to turn people into energy resources, he falls to the floor and throws up.

5. A Really *Brave New World*

IT'S TEMPTING TO THINK THAT *THE MATRIX* IS A GNOSTIC, BUDDHIST, OR PLATONIC/ Christian parable, in which what we take to be reality turns out to be a dream, and we are led to wake from the world of appearances to a 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 higher disembodied life. But this association would be all wrong. True, the conformist Matrix world is a sort of tranquilizing cover-up promoted by the AI programmers who take care of those who, like Neo, get out of line. And we are led to expect that Neo will lead people 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 life. In the film, salvation means the absolute opposite of the traditional religious vision.

In the first place, those who see through the Matrix can get over some of the limitations of having a body, as exemplified by their flying. But if, in the movie, the liberated ones were free of *all* bodily constraints, we couldn't make sense of what they were doing and neither could they. They wouldn't be liberated; they would be disoriented. [22]

Moreover, it is important to realize that flying, like other violations of our ordinary bodily limitations, 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 earthbound, vulnerable bodies and suffer cold, bad food, and death. It may look at the end of the film as if Neo evades death, but his "resurrection" in the hovercraft is not into a world where death has been overcome by a miraculous divine love; rather, he has been saved by an earthly 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 presumably has gotten over is not death but the herd's fear of death, thereby overcoming what, according to Heidegger, is the most serious constraint that normally limits people's freedom.

We are thus led to expect that, in return for accepting embodied 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 admirable to confront risky reality rather than remain in the safe and tran-

quilized Matrix, whatever the quality of experience in each, we need an account of human nature so that we can understand what human beings need that the Matrix world fails to provide.

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 human nature and the natural world that changed what counted as human beings and things. What mattered in the world of Homer was to be a hero and to 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 and to read the text of God's world in order to know God's will; and with 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 trying to get the most out of their possibilities by maximizing the quality of their private experiences and thereby treating themselves as resources.

But doesn't this just show, as Sartre famously observed, that there is no human nature? Here, Heidegger makes an important meta-move. As the history of the West suggests, 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 new worlds—not just breaking the rules of the current game but inventing new games. Nietzsche says we should "become those we are—human beings who are new, unique, incomparable, who give themselves laws, who create themselves."[23] Jesus created a new world by defining us in terms of our desires rather than our actions, and Descartes invented the inner and so helped disclose the Modern World. On a less dramatic scale, Martin Luther King Jr. opened a new world for African-Americans. It is just such a freedom to open up new worlds that the Matrix World lacks. As Neo says to the AI intelligences at the end of the film, "I know you are afraid…of change."

Heidegger thinks that our freedom to disclose new worlds is our special human freedom, and he holds that this freedom implies that there is no fixed preexistent set of possible worlds. Each world 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 program such a radically open world even if they wanted to.

Thus, there is something stifling about the Matrix World. It is a closed, preprogrammed world that has no place for radical change. If changing the rules just amounts to being able to bend spoons, fly, and stop bullets, this may be fun but it doesn't seem like any kind of creativity. Being creative must mean more than just being disruptive.[24] There must be a human need

to disclose new worlds latent in the higher mental powers left free even in the computer-generated perceptual world. That would be the splinter in the mind.

If being world-disclosers is our nature, that would explain why we feel a special joy when we are opening new worlds even on a more local scale than the transformations brought about by Jesus or Descartes.[25] Once we experience world disclosing, we understand why it's better to be in the real world than in the Matrix, even if, in the world of the Matrix, one can enjoy steak and good wine. Real salvation comes from transcending the world-foreclosing limits of the Matrix program. What's ultimately important to us, then, is not whether most of our beliefs are true, or whether we are brave enough to face a risky reality, but whether we are locked into a world of routine activities or are free to transform the world and ourselves.

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 them 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 free the people in the Matrix by showing them that the order they take for granted is ungrounded and can be creatively changed. Neo says, "I'm going to show these people a world without rules and control. A world where everything is possible." But what do control and possibility mean here?

Early in the film, Morpheus says, "What is the Matrix? Control. The Matrix is a computer-generated dream world, built to keep us under control." James Pryor, at the end of his essay on *The Matrix* on the Warner Brothers web site, tries heroically to make sense of this claim by speculating on what the AI programmers might do to the Matrix dwellers. He rightly points out that the AI intelligences could sabotage the Matricians' projects or reset their world back to 1980 if they so chose. If the machines had done any such things, Pryor would have the right to say as he does:

> In the movie, humans are…not in charge of their own lives.…They have only a very limited ability to shape their own futures.…The worst thing about living in the Matrix would not be something metaphysical or epistemological. The worst thing would be something political. It would be the fact that you're a slave.[26]

But I fear that Morpheus is simply mistaken, at least concerning what has happened in the *Matrix* series thus far. If you're a slave, there must be a master who controls what you can do or, as in *Brave New World*, who even controls what you want to do—and, of course, if you knew you were in such a world, you would want your freedom. Having their causal basis used as a battery, however, doesn't interact with the Matricians' psychic lives and doesn't limit what they can decide, what they can desire, or what they can do. What Morpheus doesn't understand (and Pryor doesn't bring out) is that there is nothing in being used as a battery that is essentially enslaving. That

is, although the Matricians' causal basis is being *used* to generate electricity, the Matricians are not being *controlled*. Their "enslavement" in the Matrix is like our relation to our selfish genes.[27] No one feels there is something morally wrong with our world because our DNA is using us to propagate itself; likewise, the simple fact that the Matricians are linked to bodies that are serving some purpose outside their lives can't be what's wrong with living in the Matrix.

There is, indeed, a very subtle way that the AI computers have foreclosed the Matrix dwellers' freedom but it is not by limiting the possibilities available to them *in their world*. The limitation in question has nothing to do with being brains in vats as long as the inputs to the brains are modeled on the way things normally behave in the world, and the outputs depend on the Matrix dwellers' decisions. The problem isn't epistemological, nor is it metaphysical, nor (pace Morpheus and Pryor) is it political. The problem is what Heidegger would call *ontological*. It has to do not with freedom to choose in the current world but with freedom to change worlds. By suppressing all unconventional behavior in their fear of change, and, in any case, having no way to introduce radical freedom into their programs, the AI intelligences have suppressed the Matricians' most essential human capacity—a way of being the computers can't understand but dimly fear: our ontological capacity for opening radically new worlds.

So Neo is on the right track when he promises "to show these people …a world where everything is possible." But, by the end of the movie, Neo as the One (or the anti-one as Heidegger would see it) has begun freeing the people *in* the Matrix from their conformism by showing them that they have the freedom to bend the rules. He has not, however, freed them *from* the Matrix by showing them how to open new worlds. But, of course, there are two more movies to come. We can hope that, before number three is over, Neo will get to Zion and lead people in disclosing a really brave new world.[28] φ

Notes

[1] See Hilary Putnam, *Reason, Truth, and History* (New York: Cambridge University Press, 1981), chapter 1.

[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 I know, there is no other reference to private feelings in Homer. Rather, there are many public displays of emotions, and shared visions of gods, monsters, and future events.

[3] Saint Augustine, *Confessions*, trans. R.S. Pine-Coffin (Penguin, l961), 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 inner perceptions and appetites. As Leibniz put it: "A monad has no windows."

[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), and Maurice Merleau-Ponty, *Phenomenology of Perception*, trans. C. Smith (London: Routledge & Kegan Paul, 1962).

[10] Charles Taylor, "Overcoming Epistemology," in *Philosophical Arguments* (Cambridge, MA: Harvard University Press, 1995), 12.

[11] René Descartes, "Dioptric," *Descartes: Philosophical Writings*, ed. and trans. Norman Kemp Smith (Modern Library, l958), 150.

[12] 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." *Intentionality: An essay in the philosophy of mind* (Cambridge University Press, l983), 230.

[13] There are limits, of course. The Matrix programmers can't give a human being a dog's body. It's also unlikely they could make a brain in a female body the causal basis of a man's body in the Matrix world. The hormones of the body in the vat wouldn't match the physical attributes of the body in the Matrix world.

Still, 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 a person would quite likely go crazy trying to reconcile the body they had experienced all their life as theirs with the alien body they found themselves in on the hovercraft.

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

[15] Neo's new understanding of reality is described by him 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 a dream. 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 usually don't notice them. They are not the result of a systematic correlation between input and output to the brain's perceptual system that is designed to reproduce the consistent, intersubjective experience that we have when awake.

[16] Morpheus justifies these killings by explaining that the Matricians have been told that the intruders are terrorists and so the police and other defenders of law and order will kill Morpheus and friends if they don't strike first.

[17] 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 I'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 world but understands it is a program—but even so, he should continue to *see* it as a world.

[18] If our experience of causality originates from our (receptive) disposition to expect certain consequences, as Hume claimed, we could retrain those habits; if our experience of causality is our (spontaneous) application of rules, we should be free to change them. Only if, as the

gestaltists claim, we *see* causal relations, would there be a question how, in the special case of causality, we could override perception.

[19] If one jumps from a building believing the fall is an illusion, the computer, nonetheless, gives one the *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*, namely getting hurt, and that realization somehow blocks the visual and tactile experiences one would have had as one hit 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 believe in the existence of a spoon, when one's brain gives out the neural output for the act of bending the 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 stop bullets) and that affects what he sees—in this case, the bullets stopping. Unfortunately, how this suspension of causality is supposed to work is not explained in the film.

[20] Not to be confused with Neo as "the One" who will save people 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: MIT 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] Samuel Todes shows in detail how the fact 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, and so forth structures our spatio-temporal world. The filmmakers have successfully met the challenge of discovering which body-relative invariances can be violated if what is going on is still to make sense. Even in the Matrix, 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? [Samuel Todes, *Body and World* (Cambridge, Mass.: MIT Press, 2001)].

[23] F. Nietzsche, *The Gay Science* (Vintage Books Edition, March 1974), 335.

[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] See Charles Spinosa, Fernando Flores, and Hubert Dreyfus, *Disclosing New Worlds: Entrepreneurship, Democratic Action, and the Cultivation of Solidarity* (Cambridge, MA: MIT Press, 1997).

[26] James Pryor, "What's So Bad About Living in The Matrix?,"
<http://whatisthematrix.warnerbros.com>.

[27] See Richard Dawkins, *The Selfish Gene* (Oxford: Oxford University Press, 1989).

[28] An earlier version of this paper appeared in the Philosophy Section of the website
<http://whatisthematrix.warnerbros.com>, ed. Christopher Grau.