

The *Natural-Artificial* Distinction: A Limit Concept for Truth and Reality and its Application Across Matrices

Avi M. Miller
Princeton University

The aim of this paper is to attempt to reconcile two seemingly conflicting intuitions regarding the nature of reality. The first is that the experience of 'a brain in a vat' cannot represent reality. The second is that every person experiences the world differently; no two people share the same reality or conception of the world. The former implies that there is an objective reality while the latter puts that belief into question. The brain-in-a-vat intuition is further challenged when we consider its theological analogue. Would our brain-in-a-vat experience still be unreal if God were the creator of the matrix rather than an evil genius or complex machine? Many religious and theological perspectives cohere with such a vat model and it would be strange to think that holding such religious views implicates a disbelief in the reality of the world. In hopes of answering these questions, I will present an argument for the unreality of vat experience which, unlike other arguments previously offered, respects a broad and liberal notion of reality, one that allows for the possibility of intersubjectivity within human experience. My purpose is not to define reality or prove the reality of any particular experience or matrix, but rather to define a limit concept of reality and prove strictly that vat experiences cannot be real. I believe that such a limit concept can be captured by introducing the *natural-artificial* distinction, although not without its own set of complications and worries.

The aim of this paper is to attempt to reconcile two seemingly conflicting intuitions regarding the nature of reality. The first is that the experience of ‘a brain in a vat’ cannot represent reality. The second is that every person experiences the world differently; no two people share the same reality or conception of the world. The former implies that there is an objective reality whilst the latter puts that belief into question. The brain-in-a-vat intuition is further challenged when we consider its theological analogue. Would our ‘brain in a vat’ experience still be unreal if God were the creator of the matrix rather than an evil genius or complex machine? Many religious and theological perspectives cohere with such a vat model and it would be strange to think that holding such religious views implicates one to disbelieve in the reality of the world.

In this essay, I will begin by briefly contextualizing the brain-in-a-vat hypothesis. I will distinguish what I shall call basic matrices (“worlds”) from vat matrices and show how from such a distinction we can present a more formal argument for the unreality of vat experience. I will then critique the argument by appealing to the plausibility of intersubjectivity of human experience. In light of such a critique, I shall present the strongest case I can make *against* the unreality of vat experience, that is, for the reality of vat experience despite our intuitions and arguments otherwise. I hope to successfully challenge that argument and present a new one for the unreality of vat experience which, unlike the first, respects a broad and liberal notion of reality and that allows for the possibility of intersubjectivity within human experience. In such a manner, I hope to reconcile our two competing intuitions and, in so doing, also our theological worry. My purpose, to be clear, is not to define reality or prove the reality of any particular experience or matrix, but rather to define a limit concept of reality and prove strictly that vat experiences cannot be real. I believe that such a limit concept can be captured by introducing the *natural-artificial* distinction, although not without its own set of complications and worries.

Let us begin, then, with the premise accepted by most philosophers (with the notable exception of direct realists like Thomas Reid) that human experience is mediated by our senses. If this is the case, there arises one of the oldest and most contentious questions in philosophy: What is the *source* of our sense data? What lies on the other side of our senses? Descartes argued for the existence of an external material world. Berkeley, in contrast, argued against an external material world and

The Natural-Artificial Distinction

instead argued that God placed our sense data in our minds. One famous skeptical response is the possibility that we are merely brains in a vat. Our sensations and experiences are just brain stimulations administered by machines or computers while, in truth, we are just sitting in a vat.

The brain-in-a-vat scenario can be easily misunderstood. For unlike the theories of Descartes or Berkeley, the brain in a vat 'theory,' if we can call it that, presupposes a human world outside the vat. The classical 'vat' story goes something like this: man lives in the world and then one day gets plugged into a computer simulation. For the purposes of this essay, it does not matter whether or not a person ever actually experienced the world outside the simulation or if it is even within his power to do so. The fact remains that in the case of the brain in a vat there is an original, *human* world denied to *humans*. This is not the case with either Descartes or Berkeley. Although *God* may, for either thinker, exist in a 'world' outside of human sensation, such a world cannot conceivably be one occupied by *humans*. Thus we must be careful to note that the brain-in-a-vat hypothesis analyzes the source and status of only the simulated sensations, not the sensations of a person had he lived in the actual world presupposed. The brain-in-a-vat hypothesis does not offer an explanation of the original source of human sense data (i.e. *before* they were plugged into the machines) but rather offers a line of skeptical reasoning that puts into question whether the sense data we perceive accurately reflects the world we live in.

Notwithstanding this important distinction, the classical brain-in-a-vat theory still describes a possible world and explains the source of at least our immediate sensations if not our original ones. For the purposes of this essay, any possible network of inputs and outputs that accounts for human sensation and which ultimately produces the 'world' which we can immediately perceive through our senses I shall call a matrix. Matrices can either include or exclude a material external world. As we said above, Descartes argued that God created an external world which, in turn, is the source of our sense data. Berkeley argued that God directly places those sensations in our minds with no corresponding material external world. These are two alternative matrices, two possible worlds with different accounts of what the world is and how we perceive it. Yet just as matrices need not have an external world, they also need not have God as its creator. In other matrices, the Big Bang may explain the existence of an external world without a creator at all and evolutionary theory may explain how man adapted senses to perceive it. Alternatively,

if there was no Big Bang, another evolutionary theory may argue that humans evolved senses that originate in their biochemistry rather than in any correspondence with an external world. In the latter matrix, the world we perceive does not correspond to anything outside of us, but rather results from our neurochemistry—a Berkeleyan matrix, so to speak, without a God. The table below summarizes the four basic matrices I just described.

TABLE OF BASIC MATRICES

External World	No External World
Creator of External World - God (Descartes)	God directly implants sensations into man's mind (Berkeley)
No creator - Evolution/Big Bang (Brian Ellis)	Human evolution/ neurochemistry alone accounts for sensation

Brain-in-a-vat matrices can be said to be a special, complicated form of any one of these basic matrices. Some being – whether an evil or beneficent genius, a computer, machine, or man himself – manipulated our senses (however they were naturally or divinely construed) so that they are now entirely dependent on the program. The easiest example to use is an external world matrix. Regardless of who created the external world or if there even be a creator, some being reconfigured our senses so that they no longer correspond to that external world. The same idea is true of matrices that do not involve external worlds. Some being reconfigured our senses so that they no longer correspond to their original source—perhaps God, perhaps their neurochemistry. This distinction between basic matrices and the special case of brains in a vat is important to keep in mind as this paper progresses.

Thus far all we have said concerned possible sources of sense experience. But now we encounter a new claim that, if it turns out that the immediate *source* of our sense experiences is some brain-in-a-vat computer simulation, such experiences are not *real*. The intuition for such a claim nicely follows from what was said above. Brain-in-a-vat matrices presuppose a natural, human matrix outside the vat. This predisposes one to say that the original natural matrix world represents

The Natural-Artificial Distinction

reality, whereas the vat-world represents some fictional, imaginary, or deceptive non-reality. The vat-world is ontologically distinct from the presupposed *real* world, and thus *unreal*, no matter how closely it may correspond to it.

To clarify the point above, compare vat matrices with Berkeley's idealist matrix. In some ways, Berkeley's matrix shares quite a lot in common with vat matrices. Both matrices have a creator; machines, computers, etc. or God. In both matrices, the material facts which correspond to human experience are not the facts we believe them to be. Inside the matrix, we believe that it is our eating a steak that causes us to taste its fine grilled flavor. In truth, however, we were sitting in a vat doing no such thing and either God or some machine stimulated our neurons in order to make us believe that we had a steak on our plate, we were eating it, and tasting it.

The crucial distinction however between the Berkeleyan matrix and the vat matrix is that the former does not presuppose some more natural human state. In Berkeley's matrix, God created man and placed him within this world *from the beginning of his creation*, and thus there was never a time 'before he was plugged in' to the divine matrix. Moreover, man in his present form cannot exist outside this world; he was born within it and cannot escape it. The reason why this distinction is so significant is that it denies the possibility of having the same intuition for Berkeley's matrix as we had earlier regarding vat matrices. For there is no pre-vat, original real world for the Berkeleyan matrix not to correspond to, that would lead us to conclude those experiences are unreal.

This is certainly not to say that from this we can conclude that human experience for Berkeley (and Descartes) attains the status of the *real*. On the contrary, all I mean to argue is that their respective matrices cannot be deemed *unreal* for the same reason as brains in a vat. In summary, brain-in-a-vat matrices are inherently different from the other basic matrices in that there is an original human world that has since ceased being the source of human sensation. This difference seems to motivate our intuition that vat experiences are not real. Whether we can affirm in the positive that our sensations in basic matrices are real is still up for further scrutiny.

Let us now move beyond our intuitions and clearly articulate and examine the standard argument for the unreality of vat experiences (or vat-worlds). The argument may go as follows: when we are plugged into a vat world, we sever the connection our senses had to the *objective* real

Avi M. Miller

world that we, as a species, had always lived in. Inside the vat-world, we may go to school, have friends, get married, vacation at a tropical beach, but in reality, all we are doing is sitting in a vat. This kind of ‘objectivity of facts’ argument is the one James Pryor mentions in his article ‘What is so bad about living in the Matrix?’ No vat experience can be said to be real, because reality is not defined by the neuron messages in our brain, but rather by the objective facts outside it. We are sitting in vat, no matter what our brain thinks.

To further understand what is going on in this argument, I think it will be helpful to reconstruct it using Berkeley’s crucial distinction between human understanding and human will. Let me just warn the reader in advance that I do not intend to reconstruct Berkeley as much as merely use his conceptual tools. I believe that, in so doing, our analysis and eventual critique of the argument will be much richer. Berkeley writes, “A spirit is one simple, undivided, active being: as it perceives ideas it is called the *understanding*, and as it produces or otherwise operates about them it is called the *will*” (On the Principles of Human Understanding, §26). “Will” narrowly defined is the imagination. One could be sitting in a lecture hall and imagine himself sunbathing on a tropical beach. One can imagine snow in august or a moon made out of blue cheese. Berkeley writes,

I find I can excite ideas in my mind at pleasure, and vary and shift the scene as oft as I think fit. It is no more than willing, and straightway this or that idea arises in my fancy; and by the same power it is obliterated and makes way for another. This making and unmaking of ideas doth very properly denominate the mind active. This much is certain and is grounded on experience. (*Principles*, §28)

“Understanding” refers to perceiving the sense data that presents itself to us, and not that which we actively conjure up. We close our eyes, open them, and if we are sitting in a lecture hall those certain sensations of the room impose themselves on us: the chairs, the desks, the podium, the chalkboard. No one wills these sensations, they just appear to us. In essence, the understanding/will distinction can be said to be the sense/imagination distinction. In Berkeley’s words,

Whatever power I may have over my own thoughts, I find the ideas actually perceived by sense have not a like dependence on my will. When in broad daylight I open my eyes, it is not in my power to

The Natural-Artificial Distinction

choose whether I shall see or not, or to determine what particular objects shall present themselves to my view; and so likewise as to the hearing and other senses, the ideas imprinted on them are not creatures of my will. There is therefore some other will or spirit that produces them. (*Principles*, §29)

Using this distinction, Berkeley claims that our understanding is more real than the will and therefore our sense data is more real than our imagination. He argues this by stating that our understanding is much more orderly than our will; that is to say, our passive sensations much more structured than our active imaginations. Though I believe this type of ‘argument by order or design’ fails in the end, I will not defend myself here since it is only tangentially relevant. More importantly, though, we can apply Berkeley’s terminology to what we have said above and say that the understanding is more real because it represents *objective facts*, whereas the will does not. In such a spirit of reconstructing the ‘objectivity of facts’ argument, I believe I can extend this to mean that any vat-world is less real or unreal compared to the presupposed natural world. The vat-world is, by definition, a manifestation of will, whereas the presupposed world outside the vat is, in a way, the world of understanding or the world of objective facts. Some may say that the vat-world is not the product of our own will but rather the will of some other being like a machine. Still it cannot be regarded as real as the objective facts of the presupposed natural world. For if we say that our own will or imagination is less real than our understanding, certainly someone else’s imagination imposed upon us is less real than our own understanding.

I have thus presented ‘the objectivity of facts’ argument for the unreality of vat experiences. The argument, however, is vulnerable to a serious challenge. By accepting its premises, we are forced to dismiss all too quickly the possibility of intersubjectivity of human experience. What I would like to do now is not so much prove definitively such a possibility as much as make it seem very plausible. By doing so, I will have shown the need for a new argument that will prove the unreality of vat-worlds while still respecting human intersubjectivity, or alternatively a strong reason to give up the claim that vat-worlds are unreal altogether.

The crux of my objection to the ‘objectivity of facts’ argument is the fine line drawn between understanding and will. It seems that upon closer scrutiny, human beings cannot strictly separate understanding and will. Perhaps, in the narrowest of senses we can separate our wild

Avi M. Miller

imaginations from the unwilling sensations of the understanding. We can distinguish our imagined fantasy of sunbathing on a beach from the brute sense of sitting in a boring lecture. But in a much broader (and I believe correct) understanding of will, stands all our beliefs, opinions, upbringing, schooling, and past experiences—all of which shape how we see and experience the world. If we can appreciate the ease with which we can distinguish the content of our imagination as we daydream in class from what is actually going on in the lecture hall, we should also appreciate the difficulty of the grayer, more common and complex cases when it is not so easy to tell the difference between how the world *really is* and how the world *seems to us* at a given moment. To a certain degree, we all see the world the way we want to see the world. Where one person sees the hand of God, another person sees chance. Moreover, we are all very selective in what we consciously perceive. As individuals, we cannot help but experience the world uniquely. I see the world, experience the world, and interpret the world quite differently than my Rwandan classmate or my best friend since childhood, or anyone else for that matter. Even if we can distinguish our imagination from that which is not imagined, it seems an almost impossible task to distinguish the sensations ‘given to us’ and our will that selects, processes and interprets them. I am not asserting here that objective brute facts are a myth, although perhaps they are. I am merely asserting that we cannot know them (or most of them), if they do exist.

To put this idea in Berkeleyan terms, in every act of the ‘understanding,’ the ‘will’ asserts itself. We can never purely ‘understand,’ that is to say we are never passively accepting sensations that are given to us without some activity of our own, whether it is selecting which sense data to perceive, interpreting that data uniquely, or some other operation during the complex process of perception. We may be able to purely ‘will’ or imagine something out of the thin blue that has absolutely no correspondence with our sensory inputs. However it is difficult to claim that we can perceive something purely without the interference of any active human processing (which inevitably alters the content). Again, the claim cannot be defended here in full, but I hope to have at least demonstrated its viability.

There is an interesting discussion found in Van Fraassen’s *The Scientific Image* that questions at least some of what has been just asserted. Van Fraassen argues that there is an important distinction between *observing* and *observing that*. He claims that *observing* means

The Natural-Artificial Distinction

pure physical perception whereas *observing that* requires a conceptual awareness of what that perception contextually signifies. He uses the example of a Stone Age person observing a tennis match. Certainly, they perceive the same physical things we do, but perceiving *that* they signify a tennis match would require much more information than was available to the Stone Age person. Thus we can all *observe* objective facts, even if our comprehension or interpretations of such facts are different. In Berkeleyan terms, Van Fraassen seems to be implying that indeed we can distinguish will from understanding, even if both are simultaneously present. We *understand* (or *observe*) objective facts and *will* (or *observe that*) its interpretation. As such, we can clearly identify and separate out the objective facts, i.e. the physical perception of the tennis match, from the confines of our will, i.e. the comprehension of the tennis match.

My reply to this is that differences in *observing that* may actually cause differences in simple *observing* in a way that will make it quite difficult, if not impossible, to identify the objective facts within our perceptions. Certain psychological theories of perception, for example, argue that theoretical, cognitive constructs often help guide and determine what we actually decide to selectively perceive. One study showed that linguistic differences between German and English actually create differences in ordinary perception. English has a progressive that is not available in German. In English, we can say 'he is crossing the street,' but in German we must say 'he crosses the street.' It has been shown that when a native English speaker perceives a person walking from A to B, their eyes tend to track the person's movement. When a native German speaker perceives that very same person, his eyes fixate on A and then B, without following his movement. This study has shown that linguistic differences (*observing that*) actually cause differences in observing.

It seems quite likely then that, by extension, our beliefs, experiences, assumptions, and cultural and religious heritage may also help determine simple *observing*. Psychological research supports this hypothesis as well. It seems, therefore, not so unreasonable to say that the Stone Age person might actually physically perceive a very different tennis game than the 21st century person sitting beside him. Perhaps the Stone Age person is concentrating on the fans clapping in the second row rather than the players and the ball. Or perhaps he fails to see the competitive or athletic aspects of the game or the stakes involved or the rules that govern the game. In summary, it still seems that we cannot abstract away from our embedded selves and distinguish *observing* from *observing that*.

Avi M. Miller

What I have tried to do is illustrate cases where it isn't clear that there is an objective fact of the matter, or that if there is, subjective experience makes it impossible to determine what that might be. I would like to suggest that human experience really is intersubjective more generally. Some may find the treatment of the matter way too hasty and they may argue that the examples I bring are not illustrative at all. Perhaps they are right. In my defense, I cannot possibly give an account here that generalizes this idea sufficiently to imply that most of our experiences fail to represent objective fact. That requires too much work that goes well beyond the confines of this essay. In the end, I leave it up to the reader to decide whether he believes human experience is intersubjective. For those who are convinced it is not, then perhaps the work for you is done. Yet for those who believe that a definition of reality should at least leave conceptual space open for the possibility of incorporating elements of human intersubjectivity, I ask that we press on. I only remind the reader that even if an account for human intersubjectivity is now elusive, it is far from impossible.

To complete this discussion and bring the point home, let us return now to the 'objectivity of facts' argument. The argument claims that vat experiences are not real because they do not represent the objective facts outside the vat. But if we accept the possibility of intersubjectivity within human experience, then even our experiences outside the vat do not represent the objective facts, if objective facts even exist at all. This is so because, as we said above, our experiences even outside the vat are inescapably the products of both our understanding and will. How then can we explain why vat experiences are any less real than experiences outside the vat?

We are thus left with three options. One, we can claim that despite everything we have said to the contrary, man experiences and perceives objective facts. In that case, then there would be no problem in continuing to use the objectivity of facts argument to prove the unreality of brain-in-a-vat matrices. Two, we can continue to entertain the possibility that human experience is intersubjective and then, in order to remain consistent, claim that vat experiences indeed also represent reality. Three, we can continue to entertain the possibility that human experience is intersubjective and then search for a new argument to show that vat experiences cannot represent reality. The first option will not be further considered in this paper. I will now turn to consider the second option in more detail before dedicating the remainder of this essay to the

The Natural-Artificial Distinction

third option.

The second option embraces the idea that vat experiences are real just as our subjective experiences outside the vat are real. This essay thus far seems to have taken the majority opinion for granted that vat experiences should not be considered real. But this might be too quick, for perhaps what is important or real is merely the sum of my experiences alone. Pryor partially makes this point (although does not subscribe to it) by using the example of eating steak. Why should it matter at all that a brain in a vat is not actually eating a steak in his vat, but merely having the experience of eating one due to the stimulation of certain neurons in his brain? Why can't the simulated experience of eating a steak be just as real as when there is an actual steak? More generally, why should reality be contingent upon the source of the experience? This is an absolutely crucial question that I will return to. But what I am suggesting now with the steak example is that, from this perspective, vat experiences may turn out to be real, contrary to what we have said all along.

To make the strongest case I can for the reality of vat experiences, I wish to consider Brian Ellis' remarks in his essay *What Science Aims To Do*. Although Ellis does not directly address brains in vats, what he does say about the nature of reality and truth will ultimately become very relevant to our purposes. In short, Ellis claims that reality and truth are dependent on the types of beings we are; they can only be understood within the framework of our limited human epistemic abilities. For us, reality is human reality and truth is human truth. What I hope to suggest is that if this is so, then perhaps brains in vats live in their own vat reality, different from the reality of machines or whoever is operating the program from outside, just as humans experience reality differently from bats. Likewise, perhaps brains in vats can attain vat truths different from the machine truths outside the vat, just as human truth may be very distinct from some alien truth. If Ellis is right that "it is enough if there is a way that the world is *for us*," then there may be a strong case for the reality and truth of vat experiences (p. 73).

Let me now attempt to develop this line of thinking in greater detail. Ellis presents an account of the subjectivity of human experience very different from the one I offered above. Ellis believes, like most realists, that reality is not just some construction of ideas and sense impressions but exists independently of anyone's knowledge or beliefs about it. "One cannot change the basic structure of the world by changing one's beliefs about it" (p. 70). For him there exists an external material world that is

Avi M. Miller

the source of our sense experiences. So far it seems fair to conclude that such claims endorse the view that objective facts exist.

However in this case, unlike most (metaphysical) realists, Ellis further argues that “the way the world is is relative to the sorts of beings we are” (Ellis, p. 71). What he means is that *our understanding and experience* of reality is necessarily dependent on the sorts of beings we are – our biology, evolutionary history, epistemic values, etc. This fact does not make it wrong to believe what we do about the world just as it would not make it wrong to believe something different if we were different beings. Ellis writes,

I am not denying the existence of a [common] world... What I am denying is that there is any *way* that the world is independently of how it is for various kinds of beings. Different beings may have different perspectives on the world, but there is no reason to think that there is any perspective which can claim priority... [Thus] no objective stance which would define the way the world is independently of our epistemic values seems to be possible. (p. 72)

This position does not deny the existence of objective facts, only the possibility that we can ever know them. As an illustration, Ellis writes that if we ask “‘Why do things behave as if the theory T were true?’ the perfectly adequate answer is that it is because the world is, for us, a T world” (p. 73). In essence, Ellis makes a similar point I had made above concerning the subjectivity of human experience; the difference is that I argued more ambitiously for *interpersonal* subjectivity of human experience, while he is arguing for *interspecies* subjectivity of experience.

Ellis’ concept of reality, then, seems to be integrally tied to the resources available to a particular species – he allows something to be real for one species and not for another. He maintains his realist objective thesis by claiming that while reality is dependent on our experience of it, it is not causally dependent. To illustrate this point, he gives the example of moving the position of our little finger. By moving our finger, we do not *causally* change Sirius in any way, but we do change its relationship to our little finger. In the same way, if our epistemic values change, our beliefs about the world change; therefore, we also change the way the world is *for us*. But we do not causally change the nature of reality in the sense that reality would change *for other beings*. In this manner, he thus synthesizes his realism and subjectivism (or, as he calls it, relativism).

To return now to our original question concerning vat experiences,

The Natural-Artificial Distinction

it seems there is a compelling reason to argue that vat experiences are real. If reality is integrally tied to the resources available to a particular species, then brains in vats could be said to exist in their own reality according to the resources available to them. Moreover, as we said above, there is no absolute reality, only various realities according to the various limits of the species. If this is so, and the human species was brains in vats, then our vat worlds should represent a reality *for us*.

Moreover, we may say something analogous about truth. Thus far, we have not mentioned truth in much detail. It will suffice to say that, as a general rule, the same arguments that have been presented above concerning the reality and unreality of experience can likewise be applied to the truth and untruth of facts in that world. For example, just as above we originally held the basic intuition that vat experiences are not real, here too the immediate intuition is that we cannot attain truth inside the vat unless it corresponds to the truth outside the vat. For example, if the machines programmed the same laws of physics into the computer simulation as were found outside the vat, then there would exist truth inside the vat. But it is true solely on account of it being true outside the vat. Yet if a vat 'truth' contradicts a truth outside the vat, the immediate intuition is that the vat 'truth' cannot be considered true.

However, if we take Ellis' pragmatic theory of truth seriously, it seems a strong case can be made for the possibility of vat truth that may or may not cohere with other understandings of truth outside the vat. Ellis introduces his pragmatism by contrasting it to the correspondence theory of truth which examines truth as a "relationship between what is true and what it is true of" (p. 68). Truth, in the correspondence sense, is some metaphysical entity, independent of rational evaluation and our epistemic values. Ellis rejects the correspondence theory by the following skeptical challenge: if truth corresponds to some unknowable metaphysical truth, how do we know if we have ever attained truth or will ever attain truth, or more fundamentally, whether these metaphysical entities even exist?

Instead, Ellis supports the pragmatic theory of truth which posits no such metaphysical entities. Truth, for pragmatic theorists, is that which is ultimately justifiable, the culmination of human science guided by its epistemic values and limits. We cannot escape our own limits, nor can we ever observe the world from a nonhuman perspective. Thus we cannot hold ourselves to some standard other than human. "Truth is just the culmination of the process of investigating and reasoning about nature

Avi M. Miller

in accordance with these [i.e. our human epistemic] values” (Ellis, p. 69). “Human truth is all we can aspire to,” says Ellis, while metaphysical truth seems surprisingly unattainable and incomprehensible. “If there is another kind of truth for beings of another kind, then that’s their problem. There is not and cannot be any absolute truth, and therefore there cannot be any way that the world is independently of how we, or some other kind of creature, would evaluate its beliefs about it” (p. 72).

Does the pragmatic theory imply then that truth can be attained within a vat world, and not just accidentally so? In a correspondence theory, it seems that, on the contrary, there cannot be vat truth unless it *corresponds* to some metaphysical absolute truth outside the vat. Yet in the pragmatic theory, it would seem that if all we meant by truth was that which is ultimately justifiable to us, then vat worlds can contain their own vat truths distinct from truths outside the vat. The final science of brains in a vat should be deemed truth, regardless of any *correspondence* outside the vat.

To summarize what has been said thus far, I first presented the ‘objectivity of facts’ argument for the unreality of vat experiences. After raising the objection that this argument fails to account for the possibility of human intersubjectivity, I have just put forward the strongest case I can make for the reality of vat experiences, as well as the possibility of truth within a vat world. I will now argue that this cannot be so. The third option I suggested above was to find a new argument that demonstrates why vat experiences are not real and the possibility of truth is only accidental, while still maintaining the intersubjectivity thesis. I believe that the argument I will present does just that while adequately responding to the claims made earlier in favor of the reality of vat experiences.

To clarify, the purpose here is not to define reality. On the contrary, my hope is to show the unreality of vat experiences while still allowing for the broadest possible understanding of human reality, one that includes the intertwining of human understanding and will, imagination and sense. I want to maintain that everyone experiences and sees the world somewhat differently, but that there is still an out of bounds, so to speak, for reality. In a way then, I am trying to define a limit concept, unreality. My argument only aspires to establish the negative thesis that vat experiences are unreal; it does not assert the reality of any experience. My aim is to locate one boundary that delineates reality from nonreality that will hopefully key in on exactly what makes the experiences of a

The Natural-Artificial Distinction

brain in a vat unreal.

The argument I shall make is that there exists a *natural* condition to reality. The argument begins with the Berkeleyan distinction between understanding and will that we used above. Recall that I questioned such a strict distinction on the grounds that one cannot perceive the world (i.e. *understand*) without asserting consciously or unconsciously one's unique makeup (i.e. *will*). Nonetheless, human understanding and will are distinct even if they overlap; one does not supervene on the other. Moreover, even though one cannot understand without willing, one can still will without understanding. If one imagines sunbathing on a tropical beach during a philosophy lecture, he incorporates none of the given sense data of the lecture hall. In contrast, watching a person cross the street incorporates both our given sense data of the scene as well as certain aspects of our will—our linguistic constructs, among other things. Thus there is still such a thing as pure will, even if there is no pure understanding. Let us reshape Berkeley's distinction to look like a spectrum, with will on the one side and understanding on the other. Most of the spectrum is a certain combination of both. But on one extreme, there exists pure will which is humanly possible. On the other extreme is pure understanding which is not humanly possible.

To return now to the difference between vat matrices and basic matrices, we now see that a brain in a vat, by definition, never perceives the sense data given to him by his understanding. All the sense data he receives, whether he knows it or not, is the product of the creator's will. Even what he *understands*, is really just the creator's *will*. And much of what he wills – his beliefs, assumptions, past experiences – are the manipulations of a higher will. Even his dreams, hopes and desires are affected by his perception of the world around him. Vat matrices are most akin to the case of imagining oneself on a tropical beach, in that in neither case are we perceiving the sense data that *would have* been given to us had it not been for the computer's will or our own. And indeed, as we said above, the latter case is one of the clearest examples of an experience we usually intuit is not real. In contrast, experiences outside a vat in any of the basic matrices usually represent some intertwining of understanding and will, passively receiving the given sense data and actively interpreting it.

With this distinction now in place, the question becomes: why must experiences that originate from the pure will necessarily be unreal? This is essentially the question we asked above and have put off until now:

Avi M. Miller

Why should reality, or rather, non-reality be contingent upon the source of the experience? The answer, I believe, is that experiences that lie solely within the will do not satisfy the *natural* condition of reality. Beyond the distinction of will and understanding lies the distinction between natural and artificial. *Natural* perception refers to *actively* perceiving (for human perception, contra Berkeley, can never be passive in the sense that one always brings his own beliefs, assumptions, etc. with every act of perception) *the given sense data* in such an honest manner as is appropriate according to oneself or one's species. This is purposefully vague. For example, perhaps a person's reality need not always be guided by the most *literal* or most *probable* construal of sense data. What is important however is that the natural condition of reality limits the will to *some* construal of sense data. It is clear how vat matrices fail this condition. The *natural* condition respects both the given sensations of our understanding as well as the range of appropriate human interpretations of them. Questions like: 'What exactly does this range look like?' and 'When does too much will deem something unreal?' will be questions largely disputed.

Again, to clarify, I am not equating the *natural* with the *real*, but equating the *artificial* with the *unreal*. Reality is constrained by the *natural* but is not defined by it. I think this constraint can be supported by a second look at Ellis. As we said above, Ellis seems to connect reality to the conditions of a particular species. The suggestion we found earlier was that if we are brains in a vat, then our vat worlds should represent the unique reality *for us*. But indeed, those who argue that way entirely miss the point. The point is that the creator of the vat world is interfering with the given, *natural* conditions of the human race—its biology, psychology, philosophy, and rationality among other things. True, Ellis claims that reality varies according to the varying conditions of species. But nowhere does Ellis claim that a species is at liberty to change the given conditions of their species, because there is no absolute reality independent of epistemic values. And if the change is forced upon them, worse off for them. From Ellis' perspective and mine as well, vat worlds hide the true reality of a particular species by interfering and manipulating their *natural* senses.

We began this essay by noting that any brain in a vat matrix presupposes that humans originally existed outside the vat world. I would now add that any vat world implies that humans *naturally* exist outside the vat while the vat world is merely *artificial*. This is so regardless of the

The Natural-Artificial Distinction

creator of the vat. Moreover, this is so regardless of whether the brain in a vat ever experienced the world outside the vat or if it is even possible to escape. This is true for two reasons. Firstly, should he theoretically escape the vat, say by an act of mercy from the creator, he would be able to exist as before being plugged in. If he cannot still exist outside the vat, then as we said above, a new species evolved with a new reality. Secondly, the creator of the vat is assumed to be someone who did not create man himself, but who manipulated our sense data artificially. I conclude that in any *artificially* created world, experience is not real and the possibility of truth is absent or, at most, discovered accidentally.

In contrast, if we turn now to an analysis of the basic matrices, these matrices all represent *natural* worlds. Suppose, for example, we live in an external world that came into existence as a result of the Big Bang and evolution. The world is a *natural* one. The sensations which present themselves to a person's mind are naturally given to him due to evolutionary forces. Contrast this with the vat worlds discussed above. In those cases, our sensations are imposed upon us not by the whim of nature, but by the whim of some conscious being *over and above nature*. Some may be tempted to say that the vat creators who manipulate our senses are acting according to nature, not over and above it. Yet such a claim misses the point entirely. For just as we would say that a tiger living in the San Diego Zoo is not living in its *natural* habitat, despite its being there through nature's course of events, so too would we say that a machine that puts a human in a vat and re-circuits his brain so that his experiences do not correspond to the material facts they once did, is an action *over and above Nature*.

The Berkeleyan matrix poses an interesting question. As we noted above, it is quite similar to a brain-in-a-vat scenario. What's different is that God is the source of our sense data, not some simulation, and that man has historically never known any other world outside of God's control. How can we explain what is *natural* about such a world? A distinction is in order here, between artificial creation on the one hand and divine creation on the other. I have alluded to this above. God creates the sense data like the creators of vats, but unlike them he also created man. This is crucial. For in the Berkeleyan case, man cannot exist outside the matrix. His being is tied to the matrix world by his very own *creation*, in a story similar to the natural evolutionary one. In a Berkeleyan matrix, then, I argue that experience is real and there is a possibility of truth. The possibility exists because man is *naturally* constituted by those sensations

Avi M. Miller

given to him by God.

To summarize, the *natural-artificial* distinction is crucial for understanding reality and the possibility of truth (or pragmatic truth). For we can agree with Ellis above that truth is human truth. But it is the *natural* abilities of man that are necessary and sufficient for attaining truth. An artificially created vat world denies man of some of his capabilities, interfering with his capacity to attain truth. In an artificially created matrix, there is only an accidental possibility of truth because he is removed from his natural, real world outside it. The artificial matrix hinders his epistemic abilities. In any of the basic matrices, however, either divinely created or not created by any being at all, the possibility of truth is present, for the matrix is indeed the natural world for man.

I will conclude by acknowledging further complications that arise from my suggestion. Where exactly do we draw the line between *natural* and *artificial*? It is natural to dream, so perhaps are my experiences while I am dreaming real? Perhaps man is so *naturally* constituted that he could be said to live in almost two worlds, the world of wakeful states and the world of dream states. Moreover, can drug-induced hallucinations or the hallucinations of schizophrenics be deemed *real* even though they are *natural*? Off hand, I would argue that the dreams and schizoid hallucinations are more *natural* than drug-induced hallucinations, but I acknowledge the problem is still there. Indeed, it may well turn out that the *natural-artificial* distinction is only helpful in demonstrating why vat matrices are unreal, while leaving the question of reality still as wide open as ever.

The Natural-Artificial Distinction

References

Berkeley, G. (1710). *A Treatise Concerning the Principles of Human Understanding*.

Ellis, B. (1985). *What Science Aims to do*. Images of Science: Essays on Realism and Empiricism with a reply from Bas C. Van Fraassen. Edited by Paul M. Churchland and Clifford A. Hooker. University of Chicago Press

Pryor, J. "What's So Bad About Living in the Matrix?" Retrieved on January 26, 2008 from http://whatisthematrix.warnerbros.com/rl_cmp/new_phil_fr_pryor.html

Van Fraassen, B. (1980). *The Scientific Image*. Oxford. Clarendon Press.