 Mind & Matter Vol. 6(1), pp. 87–121

Conscious Realism and the Mind-Body Problem

Donald D. Hoffman
Department of Cognitive Science
University of California at Irvine, USA

Abstract

Despite substantial efforts by many researchers, we still have no scientific theory of how brain activity can create, or be, conscious experience. This is troubling, since we have a large body of correlations between brain activity and consciousness, correlations normally assumed to entail that brain activity creates conscious experience. Here I explore a solution to the mind-body problem that starts with the converse assumption: these correlations arise because consciousness creates brain activity, and indeed creates all objects and properties of the physical world. To this end, I develop two theses. The *multimodal user interface* theory of perception states that perceptual experiences do not match or approximate properties of the objective world, but instead provide a simplified, species-specific, user interface to that world. *Conscious realism* states that the objective world consists of conscious agents and their experiences; these can be mathematically modeled and empirically explored in the normal scientific manner.

1. Introduction

What is the relationship between consciousness and biology? This question, a version of the classic mind-body problem, has in some form troubled philosophers at least since the time of Plato, and now troubles scientists. Indeed, a list of the top 125 open questions in *Science* puts the mind-body problem at number two, just behind the question (Miller 2005): What is the universe made of? The mind-body problem, as *Science* formulates it, is the question: What is the biological basis of consciousness?

One reason for this formulation is the large body of empirical correlations between consciousness and brain activity. For instance, damage to cortical area V1 is correlated with the loss of conscious visual perception (Celesia *et al.* 1991). If V1 is intact but certain extrastriate cortical regions are damaged, there is again a loss of conscious visual perception (Horton and Hoyt 1991). Damage to the lingual and fusiform gyri are correlated with achromatopsia, a loss of color sensation (Collins 1925, Critchley 1965), and magnetic stimulation of these areas is correlated with chromatophenes, conscious experiences of unusual colors (Sacks 1995, p. 28;

Zeki 1993, p. 279). Damage to area V5 is correlated with akinetopsia, a loss of motion sensation (Zihl *et al.* 1983, 1991; Rizzo *et al.* 1995); magnetic inhibition of V5 is also correlated with akinetopsia (Zeki *et al.* 1991). In many tasks in which subjects view a display inducing binocular rivalry, so that they consciously perceive the stimulus presented to one eye and then periodically switch to consciously perceive the stimulus presented to the other eye, there are changes in cortical activity precisely correlated with changes in conscious perception (Alais and Blake 2004), changes that can be measured with fMRI (Lumer *et al.* 1998, Tong *et al.* 1998), EEG (Brown and Norcia 1997), MEG (Tononi *et al.* 1998), and single unit recording (Leopold and Logothetis 1996). Such correlated activity can be found in ventral extrastriate, parietal, and prefrontal cortices (Rees *et al.* 2002).

Such correlations, and many more not mentioned here, persuade most researchers that brain activity causes, or is somehow the basis for, consciousness. As Edelman (2004, p. 5) puts it: "There is now a vast amount of empirical evidence to support the idea that consciousness emerges from the organization and operation of the brain." Similarly, Koch (2004, pp. 1–2) argues:

> The fundamental question at the heart of the mind-body problem is, *what is the relation between the conscious mind and the electro-chemical interactions in the body that give rise to it?* How do [conscious experiences] emerge from networks of neurons?

Consensus on this point shapes the current scientific statement of the mind-body problem. It is not the neutral statement that opened this section, viz.: What is the relationship between consciousness and biology? Instead, as *Science* makes clear, it is a statement that indicates the expected nature of the solution: What is the biological basis of consciousness? Given this consensus, one would expect that there are promising theories about the biological basis of consciousness, and that research is proceeding to cull and refine them. Indeed such theories are numerous, both philosophical and scientific, and the volume of empirical work, briefly highlighted above, is large and growing.

For instance, following the demise of behaviorism in the 1950s, there have been many philosophical theories. Type physicalist theories assert that mental state types are numerically identical to certain neural state types (Place 1956, Smart 1959); token physicalist theories assert instead that each mental state token is numerically identical to some neural state token (Fodor 1974). Reductive functionalist theories assert that the type identity conditions for mental states refer only to relations, typically causal relations, between inputs, outputs, and each other (Block and Fodor 1972). Non-reductive functionalist theories make the weaker claim that functional relations between inputs, outputs and internal system states give rise to mental states but are not identical with such states (Chalmers 1996). Representationalist theories (e.g., Tye 1996, 2000) identify conscious experiences with certain tracking relationships, i.e., with certain causal covariations, between brain states and states of the physical world. The "biological naturalism" theory of Searle (1992, 2004) claims that consciousness can be causally reduced to neural processes, but cannot be eliminated and replaced by neural processes.

This brief overview does not, of course, begin to explore these theories, and it omits important positions, such as the emergentism of Broad (1925), the anomalous monism of Davidson (1970), and the supervenience theory of Kim (1993). However it is adequate to make one obvious point. The philosophical theories of the mind-body problem are, as they advertise, philosophical and not scientific. They explore the conceptual possibilities where one might eventually formulate a scientific theory, but they do not themselves formulate scientific theories. The token identity theories, for instance, do not state precisely which neural state tokens are identical to which mental state tokens. The non-reductive functionalist theories do not state precisely which functional relations give rise, say, to the smell of garlic versus the smell of a rose, and do not give principled reasons why, reasons that lead to novel, quantitative predictions. These comments are not, of course, intended as criticisms of these theories, but simply as observations about their intended scope and limits.

It is from the scientists that we expect theories that go beyond statements of conceptual possibilities, theories that predict, from first principles and with quantitative precision, which neural activities or which functional relations cause which conscious experiences. Scientists have produced several theories of consciousness.

For instance, Crick and Koch (1990, cf. Crick 1994) proposed that certain 35-75 Hz neural oscillations in cerebral cortex are the biological basis of consciousness. Subsequently Crick and Koch (2005) proposed that the claustrum may be responsible for the unified nature of conscious experience. Edelman and Tononi (2000, p. 144; cf. Tononi and Sporns 2003) proposed that "a group of neurons can contribute directly to conscious experience only if it is part of a distributed functional cluster that, through reentrant interactions in the thalamocortical system, achieves high integration in hundreds of milliseconds." Baars (1988) proposed that consciousness arises from the contents of a global workspace, a sort of blackboard by which various unconscious processors communicate information to the rest of the system. Hameroff and Penrose (1996, cf. Penrose 1994) proposed that quantum coherence and quantum-gravity-induced collapses of wave functions are essential for consciousness. Stapp (1993, 1996) proposed that the brain evolves a superposition of action templates, and the collapse of this superposition gives rise to conscious experience.

Again, this brief overview does not begin to explore these theories and, for brevity, omits some. But the pattern that emerges is clear. The theories so far proposed by scientists are, at best, hints about where to look for a genuine scientific theory. None of them remotely approaches the minimal explanatory power, quantitative precision, and novel predictive capacity expected from a genuine scientific theory. We would expect, for instance, that such a theory could explain, in principle, the difference in experience between, e.g., the smell of a rose and the taste of garlic. How, precisely, is the smell of a rose generated by a 40 Hz oscillation, a reentrant thalamocortical circuit, information integration, a global-workspace entry, the quantum state of microtubules, or the collapse of evolving templates? What precise changes in these would transform experience from the smell of a rose to the taste of garlic? What quantitative principles account for such transformations? We are not asking about advanced features of consciousness, such as self-consciousness, that are perhaps available to few species. We are asking about an elementary feature available, presumably, to a rat. But no current theory has tools to answer these questions and none gives guidance to build such tools. None begins to dispel the mystery of conscious experience. As Pinker (1997, p. 564) points out, "... how a red-sensitive neuron gives rise to the subjective feel of redness is not a whit less mysterious than how the whole brain gives rise to the entire stream of consciousness."

In short, the scientific study of consciousness is in the embarrassing position of having no scientific theory of consciousness. This remarkable situation provokes several responses. The first concludes that, although consciousness arises naturalistically from brain activity, humans lack the cognitive capacities required to formulate a scientific theory. As McGinn (1989) puts it, "we know that brains are the *de facto* causal basis of consciousness, but we have, it seems, no understanding whatever of how this can be so." Pinker (1997) agrees. After asking how conscious experience arises from physical systems he answers (Pinker 1997, pp. 146–147):

> Beats the heck out of me. I have some prejudices, but no idea of how to begin to look for a defensible answer. And neither does anyone else. The computational theory of mind offers no insight; neither does any finding in neuroscience, once you clear up the usual confusion of sentience with access and self-knowledge.

A second response concludes that we must keep experimenting until we find the empirical fact that leads to a theoretical breakthrough. This is a defensible position and, indeed, the position of most researchers in the field.

A third response claims there is no mind-body problem, on at least two different grounds: There is no mind to reduce to body, or no body to which mind can be reduced. The first of these two arguments is sometimes asserted by eliminative materialists, who claim that nothing in reality corresponds to our folk psychological notions of consciousness (Churchland 1981, Churchland 1986, Dennett 1978). As neuroscience progresses we will not reduce such notions to neural activity; we will abandon them, much as we abandoned phlogiston. We will instead adopt the language of neurophysiology.

The second argument, that there is no body to which mind can be reduced, is made most notably by Chomsky (1980, 2000), who argues that there has been no coherent formulation of the mind-body problem since Newton introduced action-at-a-distance and, thereby, destroyed any principled demarcation between the physical and non-physical. Chomsky concludes that consciousness is a property of organized matter, no more reducible than rocks or electromagnetism (Chomsky 2000, p. 86). However, what counts as matter is no clearer than what counts as physical. And why should we expect, in the non-dualistic setting that Chomsky endorses, that consciousness is a property of matter rather than *vice versa*?

This is a natural point of departure for the theory developed here. The dualistic formulation of the mind-body problem, in which consciousness arises from non-conscious neurobiology or physics, has failed to produce a scientific theory. But the search space of scientific theories is large, and it is reasonable, given the failure of explorations in the dualistic region, to explore elsewhere. That is the intent here: to explore a non-dualistic, but mathematically rigorous, theory of the mind-body problem, one that does not assume consciousness is a property of organized matter. To this end, we first develop a non-dualistic theory of perception that questions a key assumption of current perceptual theories.

2. Perception as Faithful Depiction

Current scientific theories of perception fall into two main classes: direct and indirect (see, e.g., Fodor and Pylyshyn 1981, Hoffman 1998, Palmer 1999).

Indirect theories, which trace their lineage through Helmholtz (1910/1962) and Alhazen (956–1039; cf. Sabra 1978), typically claim that a goal of perception is to match, or at least approximate, useful properties of an objective physical world (Marr 1982). The physical world is taken to be objective in the sense that it does not depend on the perceiver for its existence. According to indirect theories, the information transduced at sensory receptors is not sufficiently rich, by itself, to determine a unique and correct match or approximation. Therefore the perceiver must infer properties of the world using constraining assumptions. For instance, in the perception of a three-dimensional shape from visual motion, the perceiver might use a *rigidity* assumption: If the image data could have arisen, in principle, by projection of the motion of a rigid three-dimensional body, then the visual system infers that the image data are, in fact, the projection of that rigid body (Ullman 1979). This inference might be couched in the mathematical framework of regularization theory (Poggio *et al.* 1985) or Bayesian inference (Knill and Richards 1996).

Direct theories, which trace their origin to Gibson (1950, 1966, 1979/1986), agree with indirect theories that a goal of perception is to match an objective physical world, but argue that the sensory data are sufficiently rich that perceivers can, without inference, pick up true properties of the world, especially affordances, directly from these data.

The debate between direct and indirect theories raises interesting issues (Fodor and Pylyshyn 1981, Ullman 1980). But what is pertinent here is that both agree on this: *A goal of perception is to match or approximate true properties of an objective physical environment.* We can call this the *hypothesis of faithful depiction* (HFD). This hypothesis is widespread and rarely questioned in the scientific study of perception.

For instance, Stoffregen and Bardy (2001) state:

> We analyze three hypotheses about relations between ambient arrays and physical reality: (1) that there is an ambiguous relation between ambient energy arrays and physical reality, (2) that there is a unique relation between individual energy arrays and physical reality, and (3) that there is a redundant but unambiguous relation, within or across arrays, between energy arrays and physical reality.

The first hypothesis is endorsed by indirect theories, and the second by some direct theories. They conclude in favor of the third hypothesis, viewing it as an extension of standard direct theories. Nowhere do they question the assumption of faithful depiction that is shared by all three; nor do any of the more than 30 commentaries on their article.

Yuille and Buelthoff (1996, p. 123) endorse HFD: "We define vision as perceptual inference, the estimation of scene properties from an image or sequence of images." The commitment to HFD is clear in such terms as 'estimate', 'recover', and 'reconstruct', which appear repeatedly throughout the literature of computational vision.

Lehar (2003, p. 375) endorses HFD: "The perceptual modeling approach reveals the primary function of perception as that of generating a fully spatial virtual-reality replica of the external world in an internal representation."

Searle (2004, p. 171) endorses HFD: "In visual perception, for example, if I see that the cat is on the mat, I see how things really are (and thus achieve mind-to-world direction of fit) only if the cat's being on the mat causes me to see the situation that way (world-to-mind direction of causation)."

Purves and Lotto (2003) appear, on first reading, to reject HFD. They reject, for instance, "the seemingly sensible idea that the purpose of vision is to perceive the world as it is..." (p. 5). They suggest instead that (p. 287)

> what observers actually experience in response to any visual stimulus is its accumulated statistical meaning (i.e., what the stimulus has turned out to signify in the past) rather than the structure of the stimulus in the image plane or its actual source in the present.

Thus Purves and Lotto do not, in fact, recommend rejection of HFD *tout court*. They simply recommend rejecting a version of the hypothesis that focuses exclusively on the *present* stimulus and the *present* state of the physical world. The purpose of vision is to perceive the world, not just as it is, but as it has been.

Noë and Regan (2002) also appear, on first reading, to reject HFD. They reject, for instance, the position that "... the visual system builds up a detailed internal representation of the three-dimensional environment on the basis of successive snapshot-like fixations of the scene ..." (p. 575). They propose instead that "what one sees is the aspect of the scene to which one is attending – with which one is currently interacting..." (p. 575). Thus Noë and Regan also do not reject HFD *tout court*. They claim that "perceivers are right to take themselves to have access to environmental detail and to learn that the environment is detailed" (p. 576) and that "the environmental detail is present, lodged, as it is, right there before individuals and that they therefore have access to that detail by the mere movement of their eyes or bodies" (p. 578). Thus they support a version of HFD that is careful to observe the limits of perceptual attention and the critical role of sensorimotor interactions.

HFD is so universally accepted that it appears in textbooks. For instance, Palmer (1999, p. 6, his italics) endorses HFD as follows:

> Evolutionarily speaking, visual perception is useful only if it is reasonably accurate ... Indeed, vision is useful precisely because it is so accurate. By and large, *what you see is what you get.* When this is true, we have what is called *veridical perception* ... perception that is consistent with the actual state of affairs in the environment. This is almost always the case with vision...

I, too, endorsed HFD, describing the central questions about visual perception as follows (Hoffman 1983, p. 154): "First, why does the visual system need to organize and interpret the images formed on the retinas? Second, how does it remain true to the real world in the process? Third, what rules of inference does it follow?" But I now think HFD is false. Our perceptual systems do not try to approximate properties of an objective physical world. Moreover evolutionary considerations, properly understood, do not support HFD, but require its rejection.

I propose that perception is like a multimodal user interface (Hoffman 1998, 2003). A successful user interface does not, in general, resemble what it represents. Instead it dumbs down and reformats in a manner useful to the user. Because it simplifies, rather than resembles, a user interface usefully and swiftly informs the actions of the user. The features in an interface usually differ from those in the represented domain, with no loss of effectiveness. A perceptual user interface, simplifying and reformatting for the niche of an organism, gives that organism an adaptive advantage over one encumbered with constructing a complex approximation to the objective world. The race is to the swift; a user interface makes one swift by not resembling the world.

This is not what textbooks or most perceptual experts say and therefore invites spelling out. I begin by discussing user interfaces and virtual worlds.

3. User Interfaces

Suppose you wish to delete a file on your PC. You find the icon for the file, click on it with your mouse, drag it to the recycle-bin icon, and release. Quick and easy. The file icon might be blue and square. The recycle bin might be shaped like a trash can. All for ease of use. Of course what goes on behind the icons is quite complex: A central processor containing millions of transistors executes binary commands encoded as voltages in megabytes of memory, and directs the head on a hard drive to change the magnetic structure of a disk revolving thousands of times per minute. Fortunately, to delete a file you do not need to know anything about this complexity. You just need to know how to move colorful icons.

The icons, and the entire graphical-windows interface, are designed to help the user by hiding the complexity of the computer (see, e.g., Schneiderman 1998). This is accomplished, in part, by *friendly formatting.* The windows interface and its contents are designed not to resemble the actual complexity of the computer and its inner workings, but instead to present needed information to the user in a format that is friendly, i.e., that is easy and natural to use. Although the actual file in the computer is a complex array of voltages and magnetic fields with no simple geometry, the file icon is a rectangle because this is a simple symbol easily interpreted by human users. Nothing about the shape of the file icon resembles the shape of the file itself. This is no failure of the icon, no misrepresentation of reality. It is, instead, what makes the icon useful.

Few souls delight to search the guts of a computer with voltmeter and magnetometer to find a file. We prefer to find a rectangular blue icon in a pretty display. But nothing about the file itself, the voltages and magnetic fields inside the computer, is blue. Is this a gross misrepresentation by the icon? Of course not. The color of the icon is not intended to resemble anything about the file but simply to indicate, say, what kind of file it is or how recently it was modified. The icon sits at some spot on the display, perhaps in the upper right. But this does not mean that the file itself is in the upper right of the computer. The location of an icon on the display is, in part, simply a convenient way to keep track of it. There is, in short, no resemblance between properties of the icon and properties of the file. This is no problem, no failure of veridicality. It is the intended consequence of friendly formatting.

The interface also helps the user by means of *concealed causality.* Not only is the structural complexity of the computer hidden behind icons, but also its causal complexity. When you drag the file icon to the recycle bin and release, does moving the file icon to the recycle bin icon cause deletion of the file? No. Icons have no causal powers within the computer. They are patterns of pixels on the display, and send no signals back to the computer. The complex causal chain within the computer that deletes the file is hidden, behind the interface, from the user. And nothing in the movement of the file icon to the recycle-bin icon resembles anything in this causal chain. Is this a failure or misrepresentation of the interface? To the contrary, it is the reason for the interface. Hiding causal complexity helps the user to quickly and easily delete a file, create a new one, modify an illustration, or format a disk, without distraction by a myriad of causal details.

Although the icons of the interface have no causal powers they are nonetheless useful by providing *clued conduct.* The icons effectively inform actions of the user, allowing the user to trigger the appropriate, but hidden, causal chains.[1] In the case of deleting a file, the icon of the file informs the user how to click the mouse, and the icon of the recycle bin informs the user how to release the mouse, so that appropriate causal chains are triggered inside the computer, resulting in deletion of the file. Icons inform an effective perception-action loop, without themselves having causal powers in the computer.

To the extent that a user interface succeeds in providing friendly formatting, concealed causality, and clued conduct, it will also offer *ostensible objectivity.* Usually the user can act as if the interface is the total reality of the computer. Indeed some users are fooled; we hear humorous stories of a child or grandparent who wondered why an unwieldy box was attached to the screen. Only for more sophisticated purposes, such as debugging a program or repairing hardware, does dissolution of this illusion become essential.

[1]Here, and throughout the paper, the verb *trigger* means "to initiate a sequence of actions, typically causal and complex." To say, for instance, that stress triggers cardiovascular disease means that stress initiates a complex causal sequence of biochemical interactions that eventuate in the disease.

4. Virtual Worlds

Suppose you and a friend play virtual tennis at an arcade. You don your helmet and bodysuit, and find yourself in Roland-Garros stadium, home of the French Open. After admiring the clay court and stadium, you serve to open the first set, and are soon immersed in play. The stadium, court, net, ball, and racquet that you experience are all, of course, part of a sophisticated user interface, one that exhibits the four qualities described in the last section. First, it sports friendly formatting. You see red clay, a yellow ball, a graphite tennis racquet, and a green stadium. These are much easier to interpret and use than the complex supercomputer and megabytes of software that control the game.

It conceals causality and clues conduct. When you hit a killer drop volley, it might appear that the head of the racquet caused the ball to sneak across the net. But of course the racquet and ball are just pixels in the user interface, and send no signals back to the supercomputer. The racquet and ball serve only to inform your actions and these, transmitted back via the body suit, trigger a complex but hidden causal sequence within the supercomputer, resulting in the proper updating of registers corresponding to the positions of racquet and ball. A good programmer could update these registers directly. But this would be so slow and cumbersome that even the deftest coder would lose the match to a modestly talented player who simply acted on the user interface. That is the power, and purpose, of the interface.

Finally, the commercial success of the game depends, in large part, on its ostensible objectivity. Customers want to play tennis, blissfully ignorant of the supercomputer and software hard at work in a back room. Tennis is, for them, the reality. Nothing in their tennis reality resembles the hidden supercomputer, the true causal nexus that makes the game possible. Customers can play as if the tennis ball and racquet had causal powers, even though this is merely a convenient, and entertaining, fiction.

5. Perception as a Multimodal User Interface

I reject HFD, the hypothesis that a goal of perception is to match or approximate properties of an objective physical world. Instead I propose the hypothesis of multimode user interfaces (MUI): *The conscious perceptual experiences of an agent are a multimodal user interface between that agent and an objective world.*

To say that a world is objective means that the world's existence does not depend on the agent. MUI theory claims nothing about the ontology of that objective world. It requires no resemblance between properties of the interface and the world. As virtual tennis illustrates, they can be as dissimilar as tennis balls and integrated circuits. MUI is a weaker hypothesis than HFD: Both say perception represents an objective world; but HFD claims, in addition, that perception *resembles* that objective world. MUI theory makes no such claim.

For instance, if you experience a rock or tree, HFD claims that, barring illusion, there must be a rock or tree in the objective world whose properties approximate those of your experience. MUI theory is not committed to this claim. It allows countless possibilities for what in the objective world triggered your experience. Chances are there is no match between properties of experience and the objective world. Instead perceptual experiences are, in the typical case, much less complex and differently formatted than the objective properties that trigger them. This failure to match, due to adaptive simplification and reformatting, is key to the usefulness of perceptual experiences. Concern about veridicality of perception is a category error. The proper concern is whether perception usefully informs action.

According to MUI theory, the objects of everyday experience – tables, chairs, mountains, moon – are not public. If, for instance, I hand you a glass of water, it is natural, but false, to assume that the glass I once held is the same glass you now hold. Instead, according to MUI theory, the glass I held was, when I observed it, an icon of my MUI, and the glass you now hold is, when you observe it, an icon of your MUI, and they are numerically distinct. There are two glasses of water, not one. And if a third person watches the transaction, there are three glasses.

This claim seems, to most, absurd, and straightforwardly refuted. Searle (2004, pp. 275ff), for instance, argues against the denial of public physical objects as follows: First, we all assume, quite naturally, that we sometimes communicate successfully. This requires that we have public meanings in a public language, so that we can both mean, or intend, the same thing by utterances such as "this glass of water". But this requires that we have publicly available objects of reference, e.g., a publicly available glass of water, so that when I say "this glass of water" I am referring to the same object as you do when you say "this glass of water". This implies that we share perceptual access to the same object, which makes it a public object. Thus, concludes Searle, there are public physical objects and the correct philosophy of perception is direct realism.

This argument is seen false by counterexample. Bob and Tom, playing virtual tennis, can talk meaningfully about "the tennis ball" they hit; they can agree that Tom hit "the tennis ball" out of court, thus losing a point. There is, patently, no public tennis ball. Instead, a supercomputer in the back room feeds signals to the helmet displays of Bob and Tom and each, in consequence, constructs his own tennis-ball experience. But Bob's tennis-ball experience is numerically distinct from Tom's. And there is no other tennis ball around to serve the role of public tennis ball. Thus public physical objects are not required for meaningful communication.

This counterexample is instructive, for it shows why Searle's argument fails. Bob and Tom can speak meaningfully about "the tennis ball" because their experiences are properly coordinated. Searle assumes that such coordination *requires* a public tennis ball. But this assumption is false: the coordination in the counterexample is accomplished not by a public tennis ball, but by a hidden supercomputer.

According to MUI theory, everyday objects such as tables, chairs and the moon exist only as experiences of conscious observers. The chair I experience only exists when I look, and the chair you experience only exists when you look. We never see the same chair. We only see the chair icons we each construct each time we look.

There are several arguments for the absurdity of this claim. First, that chair cannot exist only when I look at it. For I can look away and still touch it. So it still exists. Or I can look away and you can look at it, and confirm to me that it is still there. So again it still exists.

But this argument is easily refuted by the virtual-tennis counterexample. Bob can claim that the tennis ball he and Tom are hitting exists even when he does not look at it. After all, he can look away and still touch the tennis ball. Or he can look away and Tom can look at it. So, Bob can claim, the tennis ball still exists even when he does not look at it. But Bob's claim is patently false.

A second argument: If you think that this train thundering down the tracks is just an icon of your user interface, and does not exist when you do not perceive it, then why don't you step in front of it ? You will soon find out that it is more than an icon. And I will see, after you are gone, that it still exists.

This argument confuses taking something *literally* and taking it *seriously.* If your MUI functions properly, you should take its icons *seriously*, but not *literally.* The point of the icons is to inform your behavior in your niche. Creatures that do not take their well-adapted icons seriously have a pathetic habit of going extinct. The train icon usefully informs your behaviors, including such laudable behaviors as staying off of train-track icons. The MUI theorist is careful about stepping before trains for the same reason that computer users are careful about dragging file icons to the recycle bin.

A third argument: Look, if that wall is just an icon I construct, why can't I walk through it ? Shouldn't it do what I want ?

Not at all. You construct the subjective Necker cube that you see in Figure 1. But it doesn't do everything you want. For instance, sometimes you see a cube with corner A in front and sometimes a different cube with corner B in front. But try to make yourself switch, at will and instantly, between the two cubes and you will find that your cube constructions are stubborn (for a model of this, see Atmanspacher *et al.* 2004). Or try to see the edges of the cube as wiggly rather than straight. No chance. The fact that we construct our icons does not entail that they do whatever we wish. We are triggered to construct icons by our interactions with the objective world (whatever its nature might be) and, once so triggered, we construct our icons according to certain probabilistic rules (see, e.g., Hoffman 1998). The objective world and our rules for icon construction make the icons stubborn. Still, these icons exist only in our conscious perceptions.

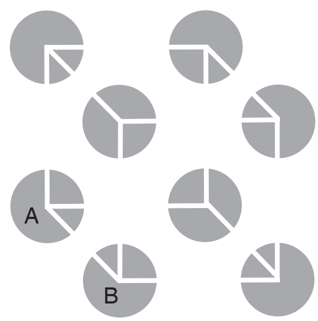

Figure 1: The subjective Necker cube (reproduced from Bradley and Petry 1977).

A fourth argument: Of course tables, chairs and the moon are just our icons, and exist only in our conscious experiences. But what's new? Physicists have long told us that the apparent solidity of a table is an illusion. It is mostly empty space with quarks and leptons darting about. Our perception of a table's surface approximates the envelope of this activity, and in this sense HFD is correct: There are no objective tables, just objective particles.

The mistake here is analogous to a computer user who admits that file icons on the display are just conventional symbols, not the actual files, but then puts a magnifying glass over an icon, sees its pixels, and concludes that these pixels are the actual file. File icons are indeed composed of pixels, but these pixels are part of the interface, not elements of the file. Similarly, tables are indeed composed of quarks and leptons, but quarks and leptons are part of the MUI, not elements of the objective world. The MUI may be hierarchically organized, but different levels of this hierarchy are part of the MUI, not of the objective world.

Placing subatomic particles in the MUI rather than in the objective world is compatible with quantum theory. Indeed, the Copenhagen interpretation of quantum theory asserts that the dynamical properties of such particles have real values only in the act of observation (see, e.g., Albert 1992, Wheeler and Zurek 1983, Zurek 1989). That is, they are part of the observer's MUI. Quantum physics does not contradict MUI theory.

A fifth argument: Ideas similar to MUI theory are found in various forms of idealism. But, as Searle (2004, p. 48) says,

> idealism had a prodigious influence in philosophy, literally for centuries, but as far as I can tell it has been as dead as a doornail among nearly all the philosophers whose opinions I respect, for many decades, so I will not say much about it.

This is a simple misunderstanding. MUI theory is not idealism. It does not claim that all that exists are conscious perceptions. It claims that our conscious perceptions need not resemble the objective world, whatever its nature is.

A sixth objection runs as follows: MUI theory implausibly claims that everything we see is not real, but created by an interface between us and the world.

This objection highlights an ambiguity of the word *real*. To say that something is real can mean either that it exists, or that it exists independent of any observers. A headache is real in the first sense, but not in the second: If I have a headache, then I am inclined to say that the headache is real, and to feel cross with anyone who says otherwise; however I would not claim that the headache exists independent of me, nor that anyone else could experience my headache, nor that I could experience the headache of anyone else. Each of us has our own private headaches, and each such headache is real, but entirely dependent for its existence on the observer who has it. I typically have little idea what causes a headache, and therefore little reason to assert that my headache resembles these unknown causes. Indeed, it almost surely does not. But the headache is not, thereby, a mystical veil between me and its unknown causes; instead, it is a simple guide to useful behavior, such as taking an aspirin, and spares me the further headache of ascertaining the complex causes of its genesis.

MUI theory does not claim that everything we see is unreal, but says instead that all sensory perceptions are real in the sense that headaches are real: They exist and are observer dependent. They exist so long as they are experienced.

This sixth objection also highlights a similar ambiguity of the word *world*: This word can refer to a sensory world or to an observer-independent world. When we speak of the visual world, we use *world* in the first sense. The visual world is observer-dependent; it disappears, for instance, when we close our eyes. Similarly, our auditory worlds are silenced if we plug our ears, and our olfactory worlds cease if we pinch the nose. The word *world* can also refer to entities hypothesized to be objective, i.e., to exist independent of any observation. HFD asserts that our sensory worlds resemble or approximate an objective world. MUI theory rejects this assertion.

MUI theory does not claim that our sensory perceptions are created by an interface between us and the world, as in the old sense datum theories. Instead, MUI theory simply acknowledges that our sensory worlds of space and time, objects, motions, colors, sounds, touches, tastes, smells and pains are observer-dependent and are not likely, on evolutionary grounds, to resemble the objective world, whatever form that world might have. This point is simple, but can be counterintuitive since we habitually assume, from early childhood, that the objective world resembles our sensory worlds.

A seventh objection is that MUI theory is logically faulty, because it is simply not true that real user interfaces do not imitate the physical world; on the contrary, they do their best to reproduce a physical-like world.

This objection is correct in noting that the user interface on a typical computer employs icons that imitate shapes and colors familiar from everyday sensory perception. However, these icons do not imitate the diodes, resistors, voltages and magnetic fields inside the computer that they represent. The icons purposely hide all this complexity, so that computer users can get on with their work.

The idea that our sensory perceptions in everyday life are useful precisely because they *do not* resemble what they represent is, for most people, counterintuitive. Fortunately, the recent introduction and widespread popularity of user interfaces on personal computers gives a ready-to-hand metaphor that most can grasp: the typical computer user understands that icons of the interface are useful precisely because they simplify, and in no way resemble, the complex world of hardware and software they represent.

An eighth objection focuses on the notion of resemblance, as follows: MUI theory recognizes that a virtual replica of the world must share some causality with its target (a virtual tennis ball must behave causally like the real one, more or less). However MUI theory does not see that this is a kind of isomorphism between the world and the user interface. It seems to consider only pictorial isomorphisms as relevant. This is not the case.

This objection is correct in noting that a tennis ball in a realistic virtual-reality game behaves much like a normal tennis ball. But the point of the virtual-reality example is not the relation between virtual tennis balls and normal tennis balls, but rather the relation between virtual tennis balls and supercomputers. The point is that the virtual tennis ball in no way resembles, pictorially or otherwise, the structural or causal properties of the supercomputer that is running the virtual tennis game. Then, by analogy, the reader is invited to envision the possibility that a *normal* tennis ball might in no way resemble, pictorially or otherwise, the structural or causal properties of whatever observer-independent entities it represents.

So the analogy offered here is as follows: Virtual tennis ball is to supercomputer as normal tennis ball is to the observer-independent world. The supercomputer is vastly more complex, structurally and causally, than a virtual tennis ball; the observer-independent world is, in all likelihood, vastly more complex, structurally and causally, than a normal tennis ball. In mathematical terms, the functions relating the supercomputer to the virtual tennis ball, or the observer-independent world to the normal tennis ball, are not isomorphisms or bijections, but are instead many-to-one maps that lose much information.

A ninth objection questions the entire metaphor of virtual reality: The whole issue of virtual reality is dependent on the creation of real stimuli (for instance, a head mounted display projects real lights and real colors to the subject's head). There is no evidence about the possibility of creating a super virtual reality world (like that in the *Matrix* movie). There is no empirical ground on which an argument can be built.

The evidence that our sensory worlds *might* be virtual worlds that in no way resemble an observer-independent world comes from quantum physics. There are many interpretations of quantum theory, and this is no place to enumerate them. Suffice it to say that proponents of the standard interpretation, the Copenhagen interpretation, often respond to the empirical evidence for quantum entanglement and violation of Bell's inequalities by rejecting local realism, and in particular by claiming that definite physical properties of a system do not exist prior to being observed; what does exist in observer-independent reality is, on their view, unknown. Which definite physical properties are instantiated at any instant depends entirely on how and what we choose to observe, i.e., on the particular observables we choose. If we choose to observe momentum, we get a value of momentum. But this value did not exist before we observed, and ceases to exist if we next choose to measure, say, position.

Thus the possibility that our sensory worlds *might* be virtual worlds, akin to a user interface, comports well with the empirical evidence of quantum physics, and is endorsed by some physicists. This is not to say, of course, that quantum theory *requires* this interpretation. Proponents of decoherence approaches, for instance, reject this interpretation. And most proponents of the Copenhagen interpretation embrace it only for the microscopic realm, not the macroscopic; but this saddles them with the unsolved problem of providing a principled distinction between microscopic and macroscopic.

6. Conscious Realism

MUI theory, we have seen, makes no claim about the nature of the objective world. In this section I propose a theory that does: *conscious realism.* One could accept MUI theory and reject conscious realism. But they fit well, and together provide a novel solution to the mind-body problem. Conscious realism is a proposed answer to the question of what the universe is made of. Conscious realism asserts that *the objective world, i.e., the world whose existence does not depend on the perceptions of a particular observer, consists entirely of conscious agents.*

Conscious realism is a non-physicalist monism. What exists in the objective world, independent of my perceptions, is a world of conscious agents, not a world of unconscious particles and fields. Those particles and fields are icons in the MUIs of conscious agents, but are not themselves fundamental denizens of the objective world. Consciousness is fundamental. It is not a latecomer in the evolutionary history of the universe, arising from complex interactions of unconscious matter and fields. Consciousness is first; matter and fields depend on it for their very existence. So the terms "matter" and "consciousness" function differently for the conscious realist than they do for the physicalist. For the physicalist, matter and other physical properties are ontologically fundamental; consciousness is derivative, arising from or identified with complex interactions of matter. For the conscious realist, consciousness is ontologically fundamental; matter is derivative, and among the symbols constructed by conscious agents.

According to conscious realism, when I see a table, I interact with a system, or systems, of conscious agents, and represent that interaction in my conscious experience as a table icon. Admittedly, the table gives me little insight into those conscious agents and their dynamics. The table is a dumbed-down icon, adapted to my needs as a member of a species in a particular niche, but not necessarily adapted to give me insight into the true nature of the objective world that triggers my construction of the table icon. When, however, I see you, I again interact with a conscious agent, or a system of conscious agents. And here my icons give deeper insight into the objective world: they convey that I am, in fact, interacting with a conscious agent, namely you.

Conscious realism is not panpsychism nor does it entail panpsychism. Panpsychism claims that all objects, from tables and chairs to the sun and moon, are themselves conscious (Hartshorne 1937/1968, Whitehead 1929/1979), or that many objects, such as trees and atoms, but perhaps not tables and chairs, are conscious (Griffin 1998). Conscious realism, together with MUI theory, claims that tables and chairs are icons in the MUIs of conscious agents, and thus that they are conscious experiences of those agents. It does not claim, nor entail, that tables and chairs are conscious or conscious agents. By comparison, to claim, in the virtual-tennis example, that a supercomputer is the objective reality behind a tennis-ball icon is not to claim that the tennis-ball icon is itself a supercomputer. The former claim is, for purposes of the example, true but the latter is clearly false.

Conscious realism is not the transcendental idealism of Kant (1781/

2003). Exegesis of Kant is notoriously difficult and controversial. The standard interpretation has him claiming, as Strawson (1966, p. 38) puts it, that "reality is supersensible and that we can have no knowledge of it". We cannot know or describe objects as they are in themselves, the noumenal objects, we can only know objects as they appear to us, the phenomenal objects (see also Prichard 1909). This interpretation of Kant precludes any science of the noumenal, for if we cannot describe the noumenal then we cannot build scientific theories of it. Conscious realism, by contrast, offers a scientific theory of the noumenal, viz., a mathematical formulation of conscious agents and their dynamical interactions. This difference between Kant and conscious realism is, for the scientist, fundamental. It is the difference between doing science and not doing science. This fundamental difference also holds for other interpretations of Kant, such as that of Allison (1983).

Many interpretations of Kant have him claiming that the sun and planets, tables and chairs, are not mind-independent, but depend for their existence on our perception. With this claim of Kant, conscious realism and MUI theory agree. Of course many current theorists disagree. For instance, Stroud (2000, p. 196), discussing Kant, says:

> It is not easy to accept, or even to understand, this philosophical theory. Accepting it presumably means believing that the sun and the planets and the mountains on earth and everything else that has been here so much longer than we have are nonetheless in some way or other dependent on the possibility of human thought and experience. What we thought was an independent world would turn out on this view not to be fully independent after all. It is difficult, to say the least, to understand a way in which that could be true.

But it is straightforward to understand a way in which that could be true. There is indeed something that has been here so much longer than we have. But that something is not the sun and the planets and the mountains on earth. It is dynamical systems of interacting conscious agents. The sun and planets and mountains are simply icons of our MUI that we are triggered to construct when we interact with these dynamical systems. The sun you see is a momentary icon, constructed on the fly each time you experience it. Your sun icon does not match or approximate the objective reality that triggers you to construct a sun icon. It is a species-specific adaptation, a quick and dirty guide, not an insight into the objective nature of the world.

One reader commented that conscious realism and MUI theory entail not just that the objects of our experience are created by subjects, but also that particles and all the rest are so created. Eventually the theory will claim that natural selection and time are a creation of the user interface. It is more noumenic than Kant.

This comment is correct, pace Kant. Space, time, particles, and therefore natural selection are all within the user interface. But this claim comports well with recent attempts in physics to construct a theory of everything – including space, time and particles – from more fundamental constituents, such as quantum information and quantum computing (e.g., Lloyd 2006), loop quantum gravity (Smolin 2006), and others (e.g., Callender and Huggett 2001). Space-time, classically conceived as a smooth manifold, appears untenable at the Planck scale. Instead there appear to be "pixels" of space and time. The intuition that space-time is a fundamental constituent of an observer-independent reality seems destined to be overturned by theories of quantum gravity.

The ontology of conscious realism proposed here rests crucially on the notion of conscious agents. This notion can be made mathematically precise and yields experimental predictions (Bennett *et al.* 1989, 1991; Bennett *et al.* 1993a,b; Bennett *et al.* 1996). Space precludes presenting the mathematics here, but a few implications of the definition of conscious agent should be made explicit. First, a conscious agent is not necessarily a person. All persons are conscious agents, or heterarchies of conscious agents, but not all conscious agents are persons. Second, the experiences of a given conscious agent might be utterly alien to us; they may constitute a modality of experience no human has imagined, much less experienced. Third, the dynamics of conscious agents does not, in general, take place in ordinary four-dimensional space-time. It takes place in state spaces of conscious observers, and for these state spaces the notion of dimension might not even be well-defined. Certain conscious agents might employ a four-dimensional space-time as part of their MUI. But again, this is not necessary.

From these comments it should be clear that the definition of a conscious agent is quite broad in scope. Indeed, it plays the same role for the field of consciousness that the notion of a Turing machine plays for the field of computation (Bennett *et al.* 1989).

7. The Mind-Body Problem

We now use MUI theory and conscious realism to sketch a solution to the mind-body problem. Exactly what that problem is depends, of course, on one's assumptions. If one adopts *physicalism*, then the central scientific problem is to describe precisely how conscious experience arises from, or is identical to, certain types of physical systems.

As we discussed before, there are no scientific theories of the physicalist mind-body problem. If one adopts *conscious realism*, then the central mind-body problem is to describe precisely how conscious agents construct physical objects and their properties.

Here there is good news. We have substantial progress on the mind-body problem under conscious realism, and there are real scientific theories. We now have mathematically precise theories about how one type of conscious agent, namely human observers, might construct the visual shapes, colors, textures, and motions of objects (see, e.g., Hoffman 1998; Knill and Richards 1996, Palmer 1999).

One example is Ullman's (1979) theory of the construction of three-dimensional objects from image motion. This theory is mathematically precise and allows one to build computer-vision systems that simulate the construction of such objects. There are many other mathematically precise theories and algorithms for how human observers could, in principle, construct three-dimensional objects from various types of image motions (e.g., Faugeras and Maybank 1990, Hoffman and Bennett 1986, Hoffman and Flinchbaugh 1982, Huang and Lee, 1989, Koenderink and van Doorn 1991, Longuet-Higgins and Prazdny 1980). We also have precise theories for constructing three-dimensional objects from stereo (Geiger *et al.* 1995, Grimson 1981, Marr and Poggio 1979), shading (Horn and Brooks 1989), and texture (Aloimonos and Swain 1988, Witkin 1981). Researchers debate the empirical adequacy of each such theory as a model of human perception, but this is just normal science.

Almost without exception the authors of these perceptual theories are physicalists who accept HFD and conceive of their theories as specifying methods by which human observers can *reconstruct* or approximate the true properties of physical objects that, they assume, exist objectively, i.e., independently of the observer (a claim about physical objects that is explicitly denied by conscious realism). But each of these perceptual theories can equally well be reinterpreted simply as specifying a method of object *construction*, not reconstruction. The mathematics is indifferent between the two interpretations. It does not require the hypothesis of independently existing physical objects. It is perfectly compatible with the hypothesis of conscious realism, and the mind-dependence of all objects. So interpreted, the large and growing literature in computational vision, and computational perception more generally, is concrete scientific progress on the mind-body problem, as this problem is posed by conscious realism. It gives mathematically precise theories about how certain conscious agents construct their physical worlds. The relationship between the conscious and the physical is thus not a mystery, but the subject of systematic scientific investigation and genuine scientific theories.

What one gives up in this framework of thinking is the belief that physical objects and their properties exist independently of the conscious agents that perceive them. Piaget claimed that children, at about nine months of age, acquire object permanence, the belief that physical objects exist even when they are not observed (Piaget 1954; but see Baillargeon 1987). Conscious realism claims that object permanence is an illusion. It is a useful fiction that substitutes for a situation which, for the child, is too subtle to grasp: Something continues to exist when the child stops observing, but that something is not the physical object that the child sees when it observes. That something is, instead, a complex dynamical system of conscious agents that triggers the child to create a physical-object icon when the child interacts with that system. For the child it is much simpler, and rarely problematic, to simply assume that the physical object it perceives is what continues to exist when it does not observe. Indeed, only when one faces the subtleties of, e.g., quantum theory or the mind-body problem, does the utility of the illusion of object permanence finally break down, and a more sophisticated, and comprehensive, ontology become necessary.

With physicalist approaches to the mind-body problem, one faces a difficult question of causality: If conscious experience arises somehow from brain activity, and if the physical world is causally closed, then how, precisely, does conscious experience cause anything? It seems, for instance, that I eat pistachio ice cream because I feel hungry and I like the taste of pistachio. Do my conscious experiences in fact cause my eating behaviors? No, say non-reductive functionalists, such as Chalmers (1996), who claim that functional properties of the brain give rise to, but are not identical with, conscious experiences. Instead they often endorse epiphenomenalism: Brain activity gives rise to conscious experiences but, since the physical realm is causally closed, conscious experiences themselves have no causal consequences. It seems like I eat pistachio because it tastes good, but this is an illusion. Moreover, I believe that I consciously experience the taste of pistachio, but I would believe this whether or not I in fact consciously experience this taste. This is a radical claim and close to an outright *reductio* of the position. Reductive functionalists, by contrast, do not endorse epiphenomenalism, since they claim that conscious experiences are *identical* to certain functional states of the brain, and conscious experiences therefore possess the causal properties of those functional states. However, reductive functionalism has recently been disproved by the "scrambling theorem" which shows that, if one grants that conscious experiences can be represented mathematically, then conscious experiences and functional relations are not numerically identical (Hoffman 2006).

Conscious realism leads to a different view of causality, a view I call *epiphysicalism*: Conscious agents are the only locus of causality, and such agents construct physical objects as elements of their MUIs; but physical objects have no causal interactions among themselves, nor any other causal powers. Physical objects, as icons of a conscious agent's MUI, can *inform*, but do not *cause*, the choices and actions of a conscious agent. When a cue ball hits an eight ball and sends it careening to the corner pocket, the cue ball does not cause the movement of the eight ball any more than the movement of a file icon to the recycle bin causes the bin to open or a file to be deleted. A useful user interface offers, as discussed above, concealed causality and ostensible objectivity. It allows one to act, in all but the most sophisticated situations, as if the icons had causal powers, and in complete ignorance of the true causal chains. The perceptual conclusions of one conscious observer might be among the premises of a second conscious observer and, thereby, inform but not cause the perceptions of the second (Bennett *et al.* 1989). Attractors in the asymptotic stochastic behavior of a system of conscious agents might be among the premises of other conscious agents and thereby inform, but not cause, their behavior (Bennett *et al.* 1989).

So, in particular, epiphysicalism entails that the brain has no causal powers. The brain does not cause conscious experience; instead, certain conscious agents, when so triggered by interactions with certain other systems of conscious agents, construct brains (and the rest of human anatomy) as complex icons of their MUIs. The neural correlates of consciousness are many and systematic not because brains cause consciousness, but because brains are useful icons in the MUIs of certain conscious agents. According to conscious realism, you are not just one conscious agent, but a complex heterarchy of interacting conscious agents, which can be called your *instantiation* (Bennett *et al.* 1989 give a mathematical treatment). One complex symbol, created when certain conscious agents within this instantiation observe the instantiation, is a brain.

Does this view entail that we should stop the scientific study of neural correlates of consciousness? No. If we wish to understand the complex heterarchy of conscious agents in human instantiations, we must use the data that our MUIs provide, and that data takes the form of brain icons. Brains do not create consciousness; consciousness creates brains as dramatically simplified icons for a realm far more complex, a realm of interacting conscious agents. When, for instance, we stimulate primary visual cortex and see phosphenes, the cortex does not cause the phosphenes. Instead, certain interactions between conscious agents cause the phosphenes, and these interactions we represent, in greatly simplified icons, as electrodes stimulating brains.

One objection to conscious realism and MUI theory runs as follows: It is completely obscure how this user interface could present its content. If the physical world is not accessible and completely out of reach, where is the user interface creating its virtual world? On which mental screen? What is the stuff its content is made of?

The key to this objection is the concept "the physical world". The objection assumes a physicalist ontology, in which the physical world is an observer-independent world comprising, *inter alia*, space-time, matter and fields. If one assumes a physicalist ontology, then it is indeed obscure how our sensory experiences, which constitute our user interface, can be understood. This is just the classic, physicalist, mind-body problem: Is there a Cartesian theater in the brain that mysteriously displays our experiences, or are there multiple drafts in multiple brain areas that can mysteriously turn into experiences ? What stuff are these experiences made of, if the fundamental constituents of the universe are mindless and physical ? This physicalist mind-body problem is still a mystery, awaiting its first genuine scientific theory.

Conscious realism, in direct contradiction to physicalism, takes our conscious experiences as ontologically fundamental. If experiences are ontologically fundamental, then the question simply does not arise of what screen they are painted on or what stuff they are made of. Compare: If space-time and leptons are taken to be ontologically fundamental, as some physicalists do, then the question simply does not arise of what screen space-time is painted on or what stuff leptons are made of. To ask the question is to miss the point that these entities are taken to be ontologically fundamental. Something fundamental does not need to be displayed on, or made of, anything else; if it did, it would not be fundamental. Every scientific theory must take something as fundamental; no theory explains everything. Conscious realism takes conscious experiences as fundamental. This might be counterintuitive to a physicalist, but it is not *ipso facto* a logical error.

A related objection is as follows: MUI theory claims that the conscious perceptual experiences of an agent are a multimodal user interface between that agent and an objective world. If the user interface is providing a completely independent world, how should it be multimodal ? Where are the different sensory modalities coming from ? Are they created internally ? Internally to what ? MUI theory claims that there is no brain or body since they are just placeholders inside the user interface.

The answer here, again, is that conscious experiences, in all their qualitative varieties, are fundamental. Because they are fundamental, they are not existentially dependent on the brain, or any other physical system. Different qualitative modalities of conscious experience are part of the basic furniture of the universe.

Is this a flight from science to mysticism ? Not if we give a mathematically precise theory of conscious experiences, conscious agents, and their dynamics, and then make empirically testable predictions. This is the reason for the previous references to mathematical models of conscious agents. Science is a methodology, not an ontology. The methodology of science is just as applicable to the ontology of conscious realism as to that of physicalism.

Another objection notes that there seems to be a difference when I meet an object and when I meet someone else. If I meet an object (or whatever it is, since by the MUI hypothesis, we cannot know), a simplified version of it is created by my super-user interface. If I meet another conscious agent, we both see each other and we both interact together. However, the other conscious agent should be equally inaccessible to me, like the noumenic object. How do we get outside of our epistemic jail, the super-user interface ?

To answer this, consider what you see when you look into a mirror. All you see is skin, hair, eyes, lips. But as you stand there, looking at yourself, you know first hand that the face you see in the mirror shows little of who you really are. It does not show your hopes, fears, beliefs, or desires. It does not show your consciousness. It does not show that you are suffering a migraine or savoring a melody. All you see, and all that the user interfaces of others can see, is literally skin deep. Other people see a face, not the conscious agent that is your deeper reality. They can, of course, infer properties of you as a conscious agent from your facial expressions and your words; a smile and a laugh suggest certain conscious states, a frown and a cry others. Such inferences are the way we avoid an epistemic jail, but all such inferences are unavoidably fallible. When we look at a rock, rather than a face, we get much less information about the conscious agents that triggered us to construct the rock. This is no surprise. The universe is complex, perhaps infinitely so. Thus our user interfaces, with their endogenous limits, necessarily give us less insight into some interactions with that universe, and more into others. When we look at ourselves in the mirror, we see first hand the limitations of our user interface and the presence, behind that interface, of a conscious agent.

8. Evolution

One major objection to conscious realism invokes evolution. We now know, the argument goes, that the universe existed for billions of years before the first forms of life, and probably many millions more before the first flickers of consciousness. Natural selection, and other evolutionary processes first described by Darwin, then shaped life and consciousness into "endless forms, most beautiful and most wonderful". This contradicts the claim of conscious realism, viz., that consciousness is fundamental and that matter is simply a property of certain icons of conscious agents. There are four responses to this objection.

First, although it is true that evolutionary theory has been interpreted, almost exclusively, within the framework of a physicalist ontology, the mathematical models of evolution do not require this ontology. They can be applied equally well to systems of conscious agents and, indeed, such an application of evolutionary game theory (Maynard-Smith 1982, Skyrms 2000) is quite natural. Systems of conscious agents can undergo stochastic evolution, and conscious agents can be synthesized or destroyed in the process (Bennett *et al.* 1989, 2002). There is simply no principled reason why evolution requires physicalism. Evolutionary changes in genes and body morphology can be modeled by evolution whether those genes and bodies are viewed as mind-dependent or mind-independent. The mathematics does not care. Nor does the fossil evidence. A dinosaur bone dated to the Jurassic can be interpreted along physicalist lines as a mind-independent object or, with equal ease, as a mind-dependent icon that we construct whenever we interact with a certain long-existing system of conscious agents.

For the conscious realist there is, no doubt, interesting and fundamental work to be done here: We want a rigorous mathematical theory of the evolution of conscious agents which has the property that, when this evolution is projected onto the relevant MUIs, it gives us back the current physicalist model of evolution. That is, we must exhibit physicalist evolutionary models as special cases, in fact projections, of a richer and more comprehensive evolutionary theory. But this is nothing special about evolution. We want the same for all branches of science. For instance we want, where possible, to exhibit current laws of physics as projections of more general laws or dynamics of conscious agents. Some current laws of physics, or of other sciences, might be superseded or discarded as the science of conscious realism advances, but those that survive should be exhibited as limiting cases or projections of the more complete laws governing conscious agents and their MUIs.

Second, according to conscious realism it simply is not true that consciousness is a latecomer in the history of the universe. Consciousness has always been fundamental, and matter derivative. The picture of an evolving unconscious universe of space-time, matter and fields that, over billions of years, fitfully gives birth first to life, then to consciousness, is false. The great psychological plausibility of this false picture derives from our penchant to commit a reification fallacy, to assume that the icons we create are in fact objects independent of us and fundamental in the universe. We embrace this fallacy because our MUI successfully informs our behavior and has ostensible objectivity, because we construct the icons of our MUI so quickly and efficiently that most of us never discover that we in fact construct them, and because we first commit the fallacy in infancy and are rarely, if ever, encouraged to challenge it. The illusion of object permanence starts by nine months, and does not go easy.

Third, standard evolutionary theory itself undercuts the reification fallacy that underlies HFD. Natural selection prunes perceptual systems that do not usefully guide behavior for survival, but natural selection does *not* prune perceptual systems because they do not approximate objective reality (see, e.g., Radnitzky and Bartley 1987). The perceptual systems of roaches, we suspect, give little insight into the complexities of objective reality. The same for lice, maggots, nematodes and an endless list of creatures that thrived long before the first hominoid appeared and will probably endure long after the last expires. Perceptual systems arise without justification from random mutations and, for 99 percent of all species that have sojourned the earth, without justification they have disappeared in extinction. The perceptual icons of a creature must quickly and successfully guide its behavior in its niche, but they need not give truth. The race is to the swift, not to the correct. As Pinker (1997, p. 561) puts it:

> We are organisms, not angels, and our minds are organs, not pipelines to the truth. Our minds evolved by natural selection to solve problems that were life-and-death matters to our ancestors, not to commune with correctness...

Shepard (2001, p. 601) hopes otherwise:

> Possibly we can aspire to a science of mind that, by virtue of the evolutionary internalization of universal regularities in the world, partakes of some of the mathematical elegance and generality of theories of that world.

It is, one must admit, *logically* possible that the perceptual icons of *Homo sapiens*, shaped by natural selection to permit survival in a niche, might also just happen to faithfully represent some true objects and properties of the objective world. But this would be a probabilistic miracle, a cosmic jackpot against odds dwarfing those of the state lottery. The smart money is on humble icons with no pretense to objectivity.

But this last response might not go far enough, for it grants that natural selection, understood within a physicalist framework, can shape conscious experience. Perhaps it cannot. Natural selection prunes *functional* propensities of an organism relevant to its reproductive success. But the scrambling theorem proves that conscious experiences are not identical with functional propensities (Hoffman 2006). Thus natural selection acting on functional propensities does not, *ipso facto*, act as well on conscious experiences. A non-reductive functionalist might counter that, although conscious experiences are not identical to functional properties, nevertheless conscious experiences are caused by functional properties, and thus are subject to shaping by natural selection. The problem with this, as we have discussed, is that no one has turned the idea of non-reductive functionalism into a genuine scientific theory, and the failure appears to be principled. Thus the burden of proof is clearly on those who wish to claim that natural selection, understood within a physicalist framework, can shape conscious experience. Understood within the framework of conscious realism, natural selection has no such obstructions to shaping conscious experiences.

A second evolutionary objection raised against MUI theory and conscious realism finds it strange that criteria of efficiency should control the user interface. Efficiency with respect to what if, as MUI theory claims, there is no way to access the real world? The logic here is a little bit like that of Descartes. Where he suggested that the mental world is similar to the physical one, MUI theory suggests that the mental world is built in such a way to be a useful schema of the physical one. Useful with respect to what? And why should we need a simplified version?

In answering this objection, we must again be careful how we use our terms. In particular, as discussed before, the phrase *real world* could mean the real worlds of our sensory perceptions, whose existence is observer-dependent. Or it could mean a world that is objective, in the sense that it is observer-independent. It is the latter interpretation that is probably intended by the objection. If so, then MUI theory does not claim there is no access to the real world, but rather that our access is *via* sensory systems that radically simplify, and probably in no way resemble, that real world. There is access, just no resemblance.

Similarly, when this objection speaks of the physical world, it presumably assumes a physicalist ontology, with physical objects and properties that are observer-independent. If so, MUI theory and conscious realism together do not claim that our sensory worlds are built to be a useful schema of the physical world, for they reject the ontology of physicalism. If there is no observer-independent physical world, then there is no reason to build schemas of it. MUI theory asserts, instead, that the physical world, the world of space-time, objects, matter and so on, is itself a sensory user interface that is observer-dependent. This might be counter-intuitive to a physicalist, but it is not logically self-contradictory. It can be made mathematically precise, and is consistent with quantum theory.

With these provisos, we can now address the main question of this objection, which is why criteria of efficiency and usefulness should control the user interface. The reason is that, according to conscious realism, there is a reality independent of any particular observer, and to interact intelligently or appropriately with that reality one's sensory perceptions must be a useful and efficient guide to that reality. Conscious realism is not solipsism. There is a reality independent of my perceptions, and my perceptions must be a useful guide to that reality. This reality consists of dynamical systems of conscious agents, not dynamical systems of unconscious matter. Moreover, this reality is quite complex. So if my sensory systems are to be efficient, they must dramatically simplify this complexity, and yet still provide a useful guide.

A third objection to MUI theory runs as follows: Inexplicably, the table I see is created by my personal user interface, but your table is created in a way that is coherent with my own. An ironic reader would ask whether they are using the same operating system.

To answer this, it is important to note that MUI theory does not require that your user interface be functionally identical to mine. Evolutionary considerations suggest that they might be functionally similar, since we are of the same species. This is the reason this paper sometimes employs the phrase "*species-specific* user interface". But evolutionary considerations also suggest that our interfaces will differ slightly in function, since random variations are essential for the operation of natural selection. Functional coherence, then, between our user interfaces is not unexpected. However, the scrambling theorem establishes that functional coherence, or even functional identity, does not logically entail identity, or even similarity, between our conscious experiences (Hoffman 2006).

9. Conclusion

Abraham Pais, describing his interactions with Einstein, wrote (Pais 1979, p. 907):

> Einstein never ceased to ponder the meaning of the quantum theory . . . We often discussed his notions on objective reality. I recall that during one walk Einstein suddenly stopped, turned to me and asked whether I really believed that the moon exists only when I look at it.

MUI theory says that the moon you see is, like any physical object you see, an icon constructed by your visual system. Perception is not objective reporting but active construction. A perceptual construction lasts only so long as you look, and then is replaced by new constructions as you look elsewhere. Thus the answer to Einstein's question, according to MUI theory, is that the moon you see only exists when you look at it. Of course the moon Jack sees might continue to exist even when the moon Jill sees ceases to exist because she closes her eyes. But the moon Jack sees is not numerically identical to the moon Jill sees. Jack sees his moon, Jill sees hers. There is no public moon.

Something does exist whether or not you look at the moon, and that something triggers your visual system to construct a moon icon. But that something that exists independent of you is not the moon. The moon is an icon of your MUI, and therefore depends on your perception for its existence. The something that exists independent of your perceptions is always, according to conscious realism, systems of conscious agents. Consciousness is fundamental in the universe, not a fitfully emerging latecomer.

The mind-body problem is, for the physicalist, the problem of getting consciousness to arise from biology. So far no one can build a scientific theory of how this might happen. This failure is so striking that it leads some to wonder if *Homo sapiens* lacks the necessary conceptual apparatus. For the conscious realist, the mind-body problem is how, precisely, conscious agents create physical objects and properties. Here we have a vast and mathematically precise scientific literature, with successful implementations in computer vision systems.

To a physicalist, the conscious-realist mind-body problem might appear to be a bait and switch that dodges hard and interesting questions: What is consciousness for ? When and how did it arise in evolution ? How does it now arise from brain activity ? Now, the switch from the ontology of physicalism to the ontology of conscious realism changes the relevant questions. Consciousness is fundamental. So to ask what consciousness is for is to ask why something exists rather than nothing. To ask how consciousness arose in a physicalist evolution is mistaken. Instead we ask how the dynamics of conscious agents, when projected onto appropriate MUIs, yields current evolutionary theory as a special case. To ask how consciousness arises from brain activity is also mistaken. Brains are complex icons representing heterarchies of interacting conscious agents. So instead we ask how neurobiology serves as a user interface to such heterarchies. Conscious realism, it is true, dodges some tough mysteries posed by physicalism, but it replaces them with new, and equally engaging, scientific problems.

Nobody explains everything. If you want to solve the mind-body problem you can take the physical as given and explain the genesis of conscious experience, or take conscious experience as given and explain the genesis of the physical. Explaining the genesis of conscious experience from the physical has proved, so far, intractable. Explaining the genesis of the physical from conscious experience has proved quite feasible. This is good news: We do not need a mutation that endows a new conceptual apparatus to transform the mind-body problem from a mystery to a routine scientific subject, we just need a change in the direction in which we seek an explanation. We can start with a mathematically precise theory of conscious agents and their interactions. We can, according to the norms of methodological naturalism, devise and test theories of how conscious agents construct physical objects and their properties, even space and time themselves. In the process we need relinquish no method or result of physicalist science, but instead we aim to exhibit each such result as a special case in a more comprehensive, conscious realist, framework.

Acknowledgments

For helpful comments on previous drafts I thank Mike Braunstein, Mike D'Zmura, Michael Gokhale, Perry Hoberman, David Hoffman, Julie Kwak, John Pyles, Barron Ramos, Whitman Richards, Kim Romney, Jay Saul, Ian Scofield, Carol Skrenes, and Ken Vaden. I also thank two anonymous reviewers whose comments helped greatly to clarify the paper. The mathematical formulation of conscious realism was developed in collaboration with Bruce Bennett and Chetan Prakash. To both I am most grateful. Sole responsibility for any errors here is, of course, mine. This research was supported in part by a grant from the US National Science Foundation.

References

Alais D. and Blake R., eds. (2004): *Binocular Rivalry*, MIT Press, Cambridge.

Albert D. (1992): *Quantum Mechanics and Experience*, Harvard University Press, Cambridge.

Allison H.E. (1983): *Kant's Transcendental Idealism: An Interpretation and Defense*, Yale University Press, New Haven.

Aloimonos Y. and Swain M.J. (1988): Shape from texture. *Biological Cybernetics* **58**, 345–360.

Atmanspacher H., Filk T., and Römer H. (2004): Quantum Zeno features of bistable perception. *Biological Cybernetics* **90**, 33–40.

Baars B. (1988): *A Cognitive Theory of Consciousness*, Cambridge University Press, Cambridge.

Baillargeon R. (1987): Object permanence in 3 and 4 month-old infants. *Developmental Psychology* **23**, 655–664.

Bennett B.M., Hoffman D.D., and Kakarala R. (1993a): Modeling performance in observer theory. *Journal of Mathematical Psychology* **37**, 220–240.

Bennett B.M., Hoffman D.D., and Murthy P. (1993b): Lebesgue logic for probabilistic reasoning, and some applications to perception. *Journal of Mathematical Psychology* **37**, 63–103.

Bennett B.M., Hoffman D.D., and Prakash C. (1989): *Observer Mechanics: A Formal Theory of Perception*, Academic Press, San Diego.

Bennett B.M., Hoffman D.D., and Prakash C. (1991): Unity of perception. *Cognition* **38**, 295–334.

Bennett B.M., Hoffman D.D., and Prakash C. (2002): Perception and evolution. In *Perception and the Physical World: Psychological and Philosophical Issues in Perception*, ed. by D. Heyer and R. Mausfeld, Wiley, New York, pp. 229–245.

Bennett B.M., Hoffman D.D., Prakash C., and Richman S. (1996): Observer theory, Bayes theory, and psychophysics. In *Perception as Bayesian Inference*, ed. by D. Knill and W. Richards, Cambridge University Press, Cambridge, pp. 163–212.

Block N.J. and Fodor J.A. (1972): What psychological states are not. *Philosophical Review* **81**, 159–181.

Bradley M.M. and Petry M.C. (1977): Organizational determinants of subjective contour: The subjective Necker cube. *American Journal of Psychology* **90**, 253-262.

Broad C.D. (1925): *Mind and Its Place in Nature*, Routledge and Kegan Paul, London.

Brown R.J. and Norcia A.M. (1997): A method for investigating binocular rivalry in real-time with the steady-state VEP. *Vision Research* **37**, 2401–2408.

Callender C. and Huggett N. (2001): *Physics Meets Philosophy at the Planck Scale*, Cambridge University Press, Cambridge.

Celesia G.G., Bushnell D., Cone-Toleikis S., and Brigell M.G. (1991): Cortical blindness and residual vision: Is the second visual system in humans capable of more than rudimentary visual perception? *Neurology* **41**, 862–869.

Chalmers D.J. (1996): *The Conscious Mind: In Search of a Fundamental Theory*, Oxford University Press, Oxford.

Chomsky N. (1980): *Rules and Representations*, Columbia University Press, New York.

Chomsky N. (2000): *New Horizons in the Study of Language and Mind*, Cambridge University Press, Cambridge.

Churchland P.M. (1981): Eliminative materialism and the propositional attitudes. *Journal of Philosophy* **78**, 67–90.

Churchland P.S. (1986): *Neurophilosophy: Toward a Unified Science of the Mind/Brain*, MIT Press, Cambridge.

Collins M. (1925): *Colour-Blindness*, Harcourt, Brace & Co, New York.

Crick F. (1994): *The Astonishing Hypothesis: The Scientific Search for the Soul*, Scribners, New York.

Crick F. and Koch C. (1990): Toward a neurobiological theory of consciousness. *Seminars in the Neurosciences* **2**, 263–275.

Crick F.C and Koch C. (2005): What is the function of the claustrum? *Philosophical Transactions of the Royal Society B* **360**, 1271–1279.

Critchley M. (1965): Acquired anomalies of colour perception of central origin. *Brain* **88**, 711–724.

Davidson D. (1970): Mental events. In *Experience and Theory*, ed. by L. Foster and J. Swanson, Humanities Press, New York, pp. 79–101.

Dennett D. (1978): Why you can't make a computer that feels pain. In *Brainstorms*, MIT Press, Cambridge, pp. 190–229.

Edelman G.M. (2004): *Wider Than the Sky: The Phenomenal Gift of Consciousness*, Yale University Press, New Haven.

Edelman G.M. and Tononi G. (2000): *A Universe of Consciousness: How Matter Becomes Imagination*, Basic Books, New York.

Faugeras O.D. and Maybank S. (1990): Motion from point matches: Multiplicity of solutions. *International Journal of Computer Vision* **4**, 225–246.

Fodor J.A. (1974): Special sciences: Or, the disunity of science as a working hypothesis. *Synthese* **28**, 97–115.

Fodor J.A. and Pylyshyn Z. (1981): How direct is visual perception? Some reflections on Gibson's 'Ecological Approach'. *Cognition* **9**, 139–196.

Geiger D., Ladendorf B., and Yuille A. (1995): Occlusions and binocular stereo. *International Journal of Computer Vision* **14**, 211–226.

Gibson J.J. (1950): *The Perception of the Visual World*, Houghton-Mifflin, New York.

Gibson J.J. (1966): *The Senses Considered as Perceptual Systems*, Houghton-Mifflin, New York.

Gibson J.J. (1979/1986): *The Ecological Approach to Visual Perception*, Lawrence Erlbaum, Hillsdale.

Griffin D.R. (1998): *Unsnarling the World Knot: Consciousness, Freedom, and the Mind-Body Problem*, University of California Press, Berkeley.

Grimson W.E.L. (1981): *From Images to Surfaces*, MIT Press, Cambridge.

Hameroff S.R. and Penrose R. (1996): Conscious events as orchestrated space-time selections. *Journal of Consciousness Studies* **3**(1), 36–53.

Hartshorne C. (1937/1968): *Beyond Humanism: Essays in the Philosophy of Nature*, University of Nebraska Press, Lincoln.

Helmholtz H.L.F. von (1910/1962): *Handbook of Physiological Optics, Volume III*, Dover, New York.

Hoffman D.D. (1983): The interpretation of visual illusions. *Scientific American* **249**, 154–162.

Hoffman D.D. (1998): *Visual Intelligence: How We Create What We See*, W.W. Norton, New York.

Hoffman D.D. (2003): Does perception replicate the external world ? *Behavioral and Brain Sciences* **26**, 415–416.

Hoffman D.D. (2006): The scrambling theorem: A simple proof of the logical possibility of spectrum inversion. *Consciousness and Cognition* **15**, 31–45.

Hoffman D.D. and Bennett B.M. (1986): The computation of structure from fixed-axis motion: rigid structures. *Biological Cybernetics* **54**, 71–83.

Hoffman D.D. and Flinchbaugh B.E. (1982):. The interpretation of biological motion. *Biological Cybernetics* **42**, 197–204.

Horn B.K.P. and Brooks M., eds. (1989): *Shape from Shading*, MIT Press, Cambridge.

Horton J.C. and Hoyt W.F. (1991): Quadratic visual field defects: A hallmark of lesions in extrastriate (V2/V3) cortex. *Brain* **114**, 1703–1718.

Huang T. and Lee C. (1989): Motion and structure from orthographic projections. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **11**, 536–540.

Kant I. (1781/2003): *Critique of Pure Reason*, Dover, New York.

Kim J. (1993): *Supervenience and Mind*, Cambridge University Press, Cambridge.

Knill D. and Richards W., eds. (1996) *Perception as Bayesian Inference*, Cambridge University Press, Cambridge.

Koch C. (2004): *The Quest for Consciousness: A Neurobiological Approach*, Roberts & Company, Englewood.

Koenderink J.J. and Doorn A.J. van (1991): Affine structure from motion. *Journal of the Optical Society of America A* **8**, 377–385.

Lehar S. (2003): Gestalt isomorphism and the primacy of subjective conscious experience: A Gestalt bubble model. *Behavioral and Brain Sciences* **26**, 375–444.

Leopold D.A. and Logothetis N.K. (1996): Activity changes in early visual cortex reflect monkeys' percepts during binocular rivalry. *Nature* **379**, 549–553.

Lloyd S. (2006): *Programming the Universe*, Knopf, New York.

Longuet-Higgins H.C. and Prazdny K. (1980): The interpretation of a moving retinal image. *Proceedings of the Royal Society of London B*, **208**, 385–397.

Lumer E.D., Friston K.J. and Rees G. (1998). Neural correlates of perceptual rivalry in the human brain. *Science* **280**, 1930–1934.

Marr D. (1982): *Vision*, Freeman, San Francisco.

Marr D. and Poggio T. (1979): A computational theory of human stereo vision. *Proceedings of the Royal Society of London B* **204**, 301–328.

Maynard-Smith J. (1982): *Evolution and the Theory of Games*, Cambridge University Press, Cambridge.

McGinn C. (1989): Can we solve the mind-body problem? *Mind* **98**, 349–366.

Miller G. (2005): What is the biological basis of consciousness? *Science* **309**, 79.

Noë A. and Regan J.K. (2002): On the brain-basis of visual consciousness: A sensorimotor account. In *Vision and Mind: Selected Readings in the Philosophy of Perception*, ed. by A. Noë and E. Thompson, MIT Press, Cambridge, pp. 567–598.

Pais A. (1979): Einstein and the quantum theory. *Reviews of Modern Physics* **51**, 863–914.

Palmer S.E. (1999): *Vision Science: Photons to Phenomenology*, MIT Press, Cambridge.

Place U.T. (1956): Is consciousness a brain process? *British Journal of Psychology* **45**, 243–255.

Penrose R. (1994): *Shadows of the Mind*, Oxford University Press, Oxford.

Piaget J. (1954): *The Construction of Reality in the Child*, Basic, New York.

Pinker S. (1997): *How the Mind Works*, W.W. Norton & Co, New York.

Poggio T., Torre V. and Koch C. (1985): Computational vision and regularization theory. *Nature* **317**, 314–319.

Prichard H.A. (1909): *Kant's Theory of Knowledge*, W.W. Norton & Co, Oxford.

Purves D. and Lotto R.B. (2003): *Why We See What We Do: An Empirical Theory of Vision*, Sinauer, Sunderland.

Radnitsky G. and Bartley W.W., eds. (1987): *Evolutionary Epistemology, Rationality, and the Sociology of Knowledge*, Open Court, La Salle.

Rees G., Kreiman G., and Koch C. (2002): Neural correlates of consciousness in humans. *Nature Reviews Neuroscience* **3**, 261–270.

Rizzo M., Nawrot M., and Zihl J. (1995): Motion and shape perception in cerebral akinetopsia. *Brain* **118**, 1105–1127.

Sabra A.I. (1978): Sensation and inference in Alhazen's theory of visual perception. In *Studies in Perception: Interrelations in the History of Philosophy and Science*, ed. by P.K. Machamer and R.G. Turnbull, Ohio State University Press, Columbia, pp. 160–185.

Sacks O. (1995): *An Anthropologist on Mars*, Vintage Books, New York.

Searle J.R. (1992): *The Rediscovery of the Mind*, MIT Press, Cambridge.

Searle J.R. (2004): *Mind: A Brief Introduction*, Oxford University Press, Oxford.

Schneiderman B. (1998): *Designing the User Interface: Strategies for Effective Human-Computer Interaction*, Addison-Wesley, Reading.

Shepard R. (2001): Perceptual-cognitive universals as reflections of the world. *Behavioral and Brain Sciences* **24**, 581–601.

Skyrms B. (2000): Game theory, rationality, and evolution of the social contract. *Journal of Consciousness Studies* **7**(1/2), 269–284.

Smart J.J.C. (1959): Sensations and brain processes. *Philosophical Review* **68**, 141–156.

Smolin L. (2006): *The Trouble with Physics*, Houghton Mifflin, Boston.

Stapp H.P. (1993): *Mind, Matter, and Quantum Mechanics*, Springer, Berlin.

Stapp H.P. (1996). The hard problem: A quantum approach. *Journal of Consciousness Studies* **3**(3), 194–210.

Stoffregen T.A. and Bardy B.G. (2001): On specification and the senses. *Behavioral and Brain Sciences* **24**, 195–261.

Strawson P.F. (1966): *The Bounds of Sense, an Essay on Kant's Critique of Pure Reason*, Methuen, London.

Stroud B. (2000): *The Quest for Reality: Subjectivism and the Metaphysics of Colour*, Oxford University Press, Oxford.

Tong F., Nakayama K., Vaughan J.T., and Kanwisher N. (1998): Binocular rivalry and visual awareness in human extrastriate cortex. *Neuron* **21**, 753–759.

Tononi G. and Sporns O. (2003): Measuring information integration. *BMC Neuroscience* **4**, 31–50.

Tononi G., Srinivasan R., Russell D.P., and Edelman G.M. (1998): Investigating neural correlates of conscious perception by frequency-tagged neuromagnetic responses. *Proceedings of the National Academy of Sciences of the USA* **95**, 3198–3203.

Tye M. (1996): T *Ten Problems of Consciousness*, MIT Press, Cambridge.

Tye M. (2000): *Consciousness, Color, and Content*, MIT Press, Cambridge.

Ullman S. (1979): *The Interpretation of Visual Motion*, MIT Press, Cambridge.

Ullman S. (1980): Against direct perception. *Behavioral and Brain Sciences* **3**, 373–415.

Wheeler J.A. and Zurek W.H., eds. (1983): *Quantum Theory and Measurement*, Princeton University Press, Princeton.

Whitehead A.N. (1929/1979): *Process and Reality: An Essay in Cosmology*, Free Press, New York.

Witkin A.P. (1981): Recovering surface shape and orientation from texture. *Artificial Intelligence* **17**, 17–45.

Yuille A. and Buelthoff H. (1996): Bayesian decision theory and psychophysics. In *Perception as Bayesian Inference*, ed. by D. Knill and W. Richards, Cambridge University Press, Cambridge, pp. 123–161.

Zeki S. (1993): *A Vision of the Brain*, Blackwell, Oxford.

Zeki S., Watson J.D.G., Lueck C.J., Friston K.J., Kennard C., and Frackowiak R.S.J. (1991): A direct demonstration of functional specialization in human visual cortex. *Journal of Neuroscience* **11**, 641–649.

Zihl J., Cramon D. von, and Mai N. (1983): Selective disturbance of movement vision after bilateral brain damage. *Brain* **106**, 313–340.

Zihl J., Cramon D. von, Mai N., and Schmid C.H. (1991): Disturbance of movement vision after bilateral posterior brain damage: Further evidence and follow up observations. *Brain* **114**, 2235–2252.

Zurek W.H., ed. (1989): *Complexity, Entropy and the Physics of Information*, Addison-Wesley, Reading.

Received: 08 January 2007
Revised: 15 May 2007
Accepted: 27 July 2007

Reviewed by two anonymous referees.