1 July 1996

Turning 'The Hard Problem' Upside Down & Sideways

Piet Hut

School of Natural Sciences
Institute for Advanced Study
Princeton, NJ 08540-0631

email: piet@sns.ias.edu

Roger N. Shepard

Department of Psychology
Stanford University
Stanford, CA 94305-2130

email: roger@psych.stanford.edu

Abstract

Instead of speaking of conscious experience as arising in a brain, we prefer to speak of a brain as arising in conscious experience. From an epistemological standpoint, starting from direct experiences strikes us as more justified. As a first option, we reconsider the 'hard problem' of the relation between conscious experience and the physical world by thus turning that problem upside down.

We also consider a second option: turning the hard problem sideways. Rather than starting with the third-person approach used in physics, or the first-person approach of starting with individual conscious experience, we consider starting from an I-and-you basis, centered around the second-person.

Finally, we present a candidate for what could be considered to underlie conscious experience: 'sense'. We consider this to be a shot in the dark, but at least a shot in the right direction: somewhere between upside down and sideways.

Our notion of sense can be seen as an alternative to panpsychism. To give an analogy, using the notions of space and time is more convenient than trying to analyze the phenomenon of motion in terms of a space-based 'pandynamism'. Similarly, when approaching the phenomenon of consciousness, we prefer the triad of space, time, and sense over a space-time-based form of panpsychism.

Table of Contents

1.Introduction

The two of us — an astrophysicist and a cognitive psychologist — struggling with questions concerning the relation between the mental and the physical, but starting from the two contrasting disciplines of the cognitive and the physical sciences, discovered that we had come to essentially the same conclusions about what is wrong with prevailing views of what are variously referred to as 'the mind-body problem,' 'the problem of consciousness,' or what Chalmers (1995) has termed 'the hard problem.' This paper represents our initial attempt to articulate the core of the view we have come to share.

1.1. The Hard Problem Concerning Consciousness

The 'hard problem' is not the 'third-person' problem of providing a scientific account for how a physical system, such as a human brain, can come to carry out the information processing necessary for intelligent behavior. The reason that this is not the 'hard problem' is that no physical limitation has so far been identified concerning what a sufficiently complex physical system might be capable of in the way of adaptive information processing (and, surely, the human brain is as complex a physical system as any so far known to science).

The 'hard problem' is, instead, the 'first-person' problem of understanding how the subjective quality of experience (including, the seemingly nonphysical 'qualia' of pains, colors, odors, etc.) can be explained or understood as arising from any physical system as described in the objective terms of present day physics — whether at the level of electrochemical processes in neurons and their synapses, at the much smaller scales of interactions of elementary particles, relativistic quantum fields, or even more exotic objects postulated to underlie those fields. The reason we say this is the 'hard problem' is that (despite the contrary statements of many commentators on the problem of consciousness) no so far imagined advance in understanding the physical processes going on in a nervous system (at any level) seems capable of moving us one iota closer to bridging the chasm between physical description and subjective quality.

To set the stage for our discussion, we will list a handful of common assumptions, and then point out some of their defects.

1.2. Some Widely Shared Notions

The standard scientific approach to the 'hard' problem of consciousness (among those who recognize that there is a hard problem), appears to be based (whether explicitly or implicitly) on the following widely shared notions:

1. Conscious experience is not itself adequately described or explained in the purely physical terms of electro-chemical processes, particles, waves, or the like.

2. Conscious experience is nevertheless supposed to arise as a manifestation of (and perhaps only of) physical systems of enormous complexity — most indubitably in that most complex of known systems, the human brain.

3. Yet, even in such a system, only some processes in some spatial regions and in some (e.g., waking) states are supposed to be accompanied by conscious experience; other processes within the same system (though indistinguishable in their local, physical properties and processes) remain unconscious.

4. Consciousness is, accordingly, posited to be an emergent, but evidently nonphysical accompaniment of only certain, as yet uncharacterized, parts or phases of certain processes in certain highly complex physical systems.

5. Whether these nonphysical accompaniments of those particular physical processes have causal effects back in the physical system itself (psychophysical interactionism) or not (epiphenomenalism) remains a matter of dispute.

1.3. Some Serious Problems

Though seemingly taken for granted by many researchers, these notions underlying standard scientific approaches to the problem of consciousness face several difficulties:

1. There are no generally accepted criteria for deciding whether any externally observed physical process is or is not accompanied by conscious experience. The sciences, from fundamental physics up through neurophysiology, proceed on the assumption that verbal reports, being themselves physical processes, must be fully explainable as the physical effects of other, preceding purely physical processes. Hence, verbal reports even of conscious experiences cannot serve as guarantors of the occurrence of conscious experience. Conversely, when such reports fail to emerge from a mute physical system (whether the right cerebral hemisphere of a 'split-brain' patient, or from a chimpanzee, a rat, or a single atom), the absence of such a report does not preclude that physical events in the system under consideration have subjective accompaniments or 'qualia.'

2. Indeed, no fundamental physical property has been identified that might distinguish those physical events that are – from those that are not – accompanied by conscious experience. Speaking in the first person, I presume that measurement and analysis of the electrochemical processes in one or more neurons in my own brain whose firing is perfectly correlated with my own experience of a flash of red, say, would reveal no differences from the electrochemical processes in other neurons in my own brain whose firing has no conscious correlate. Even if I should discover that the firing of particular neurons in a particular location of my brain was necessary and sufficient for my own conscious experience of a flash of red, this would tell me nothing about why just this externally observed physical event had such a phenomenal accompaniment while another event that, at the level of fundamental physics, was indistinguishable from the first did not.

3. If nonphysical conscious experience is taken to have a causal influence back on the physical process from which it arose (psychophysical interactionism), how is this to be reconciled with the fundamental assumption of science that every physical state of a system is strictly determined by a preceding physical state of the system and its environment (determinism) — except for possible quantum mechanical influences that are purely random (indeterminism)?

4. If conscious experience does not causally affect the course of those physical processes (epiphenomenalism), then: (a) Why does it seem that I can control my own actions (free will)? (b) What function does consciousness serve; and why would it have evolved? (c) What causes some physical bodies (namely, other persons) to make those physical acts (of

speech, writing, or typing) that express the (hard) problem of consciousness (including the problems of 'solipsism,' of the existence of 'other minds,' of whether robots could feel pain, of whether your experiences of red and green are the same as mine or just the reverse, etc.).

5. Finally, of course, there is the difficulty that the standard reductionistic approach of contemporary science simply makes no provision for a nonphysical, phenomenal experience that is not ultimately composed of physical constituents (such as atoms and molecules, or particles and waves).

2. Turning the Hard Problem Upside Down

In a nut shell, we can summarize the situation as follows. There is clearly room for physical objects within experience; it is not at all clear whether there is room for experience within physical objects. To try to see how what we understand to be physical objects produce experience is in fact the 'Hard Problem,' a problem that may simply be wrongly posed. Let us try as an alternative to turn the hard problem upside down, starting from experience, in order to see how what we understand to be physical objects may arise out of it.

2.1. Motivation

Some of the difficulties reviewed in the previous section appear to us to be consequences of starting from the reductionistic and materialistic presuppositions of physical science. The standard approach builds upon an epistemologically weak foundation: What it takes for granted is a physical world containing physical brains composed of atoms, molecules, ions, electric fields, and so on. But what are directly given to any scientist are only the consciously experienced appearances (filled with 'qualia' and their relationships) that (on the basis of certain regularities and correlations) are interpreted as independently existing physical objects. The weakness of this starting point is evident to those of us who experience vivid dreams populated with what we take to be independently existing physical objects until we awake.

Indeed, the atoms, molecules, and fields that (on the standard scientific view) constitute the material basis of any object, including a brain, are in fact known only as abstractions, not themselves directly experienced. The hypothesized invisible constituents of material objects can only be referred to by words, diagrams, or equations that are themselves (from this objective standpoint) but meaningless arrangements of molecules or (from the subjective standpoint) but constellations of qualia in the scientist's own conscious experience that lead the scientist to expect other experiences to ensue upon the performance of particular operations.

2.2. Starting from Experience

Suppose, then, that we set aside the presuppositions that have been foisted upon us by the standard scientific view and build, instead, upon the foundation of what is indubitably given in our own experience. This is not to abandon what has been gained by science; only to find a more certain basis for science.

The macroscopic objects of common sense and the microscopic objects posited by science then become, alike, hypotheses whose meaning cashes out in an individual's expectations about future experiences. Phrased, most naturally, in the first person, some of these expectations concern the behavior of those 'objects' of the kind I call 'other persons.' Thus, from the experience of a colleague emitting the utterance "There is a package for you in the mail room," I infer that if I were to enter the mail room, I would have particular kinds of visual and tactile experiences. And, from the experience of reading of Galileo's discovery of the moons of Jupiter, I infer that if I build and look through a telescope in a particular direction in the night sky, I will have other particular kinds of visual experiences. This type of analysis of intersubjectivity in science was long ago advocated by the Nobel laureate physicist Percy Bridgeman (1940) and, before him, by the philosopher Rudolf Carnap (1928). It has sometimes been referred to as 'methodological solipsism' (*cf.* Shepard 1981).

It may seem that the 'hard problem,' though inverted, is still with us. Before, we started with a physical world and found it difficult to understand:

- how nonphysical conscious experience could arise in this physical world,

• why it would arise only from some particular physical processes and not from others that from the outside appear to be entirely equivalent, or

• how it could act back on — or play a functional role in — the physical world.

Now, each of us starts only with what is directly given in our own conscious experience and must explain what we mean by, and why we believe in:

• the external world of common sense as well as that of modern physics, and

• the existence of other conscious experiences (or 'minds') beyond our own.

2.3. The Hard Problem Softened

So far, we do not pretend to have given a solution to the 'Hard Problem,' or anything like a detailed alternative to the usual picture of reality that physics presents. As we will stress at the end of this paper, our strategy is not aimed at adding yet another model of consciousness to the literature. Rather, we question the underlying methodology of almost all attempts currently pursued in consciousness research. Once our questioning is taken seriously, we can explore the new terrain that opens up with a shift in the way we pose the whole problem of consciousness. But already at this early stage, simply by turning the Hard Problem upside down, we feel that this hard problem has already been softened.

When we turn the hard problem upside down, everything is grounded in our indubitable immediate experience, not in the hypothesized 'noumenal' world of unexperienced atoms, particles, or waves. Far from making our knowledge more insecure, this turn shifts us to firmer ground, away from attempts to reduce our felt reality to ever-more chimerical constructs in mathematical physics, that are posited to underly our physical reality. Fascinating as explorations of the properties of matter on ever-smaller scales are, the grounding they so far afford for the whole of reality strikes us as undesirably precarious. For one thing, several ingredients that physicists use to describe the world are intrinsically unobservable, which makes their ontological status unclear. In addition, quantum mechanics seems to preclude any meaningful interpretation in terms of a single objective reality that exists prior to any attempt at measuring it.

The problem of the existence of other minds is softened in that by starting with subjective experience (my own) instead of with an 'objective reality,' I begin with something closer to other subjective experiences (such as yours). Intersubjectivity might be viewed as expressing properties that are inherent in subjective conscious experience, but in addition are mutually agreed upon by different subjects. Taken this way, intersubjectivity provides an antidote against solipsism that is not more mysterious or artificial than any other form of knowledge, based on experience, including the more abstract varieties.

An analogy with Euclidean geometry may be helpful: once we specify the lengths of two sides of a triangle, and the magnitude of the enclosed angle, the length of the third side is fixed, and so are the magnitudes of the remaining two angles. Why is this? Wherein resides this magical power of space? How can space enforce these 'laws' of geometry, laws that physical objects obey as well, to very high accuracy? Compared to material objects, space seems like a very flimsy something, or really a nothingness, or at least a no-thing-ness. How could anyone imagine space enforcing all these 'rules'?

The reason we don't worry about such questions is that we simply 'see' that triangles are fully specified with two sides and an angle. But if that is the case, why can't we start with consciousness, and say that we simply 'see' that the different consciousnesses of different observers are obeying various rules that leave room for differences in, say, the size of an imagined chair, but not for differences in making conclusions about the size of a chair that is actually present?

In brief, as long as we have no ultimate foundation for any of our forms of knowledge, be they geometry or physics, we will have to live our lives layered on top of a sea of unanswered questions. Invoking a lack of grounding therefore is not a valid argument against one philosophical attitude, in favor of another. Rather, we should be more empirical, accepting from the outset what we indubitably experience, without becoming entangled with anxious worries about foundations.

Thus the biggest mystery is no longer consciousness but the objective physical world, which is never directly experienced but is only inferred on the basis of order and correlations within subjective experience. It seems to us more natural and epistemologically more justifiable to leave as inference what is inferred and to take as given only what is given, than the other way around. As for regularities and correlations, that are often interpreted as pointing to the brain as the seat of consciousness, we can think of many other explanatory avenues. After all, effectively jamming a printer does not 'prove' that computer processing takes place in the printer heads; nor does switching channels on a TV 'prove' that the information in the channels is produced inside the TV set. This is not to suggest that the 'source' of experience is outside the brain, in some other physical location; rather, the whole notion of a source that is reducible to spatio-temporal-physical terms, is something we question.

2.4. Radical Consequences

Inverting the usual view in this way does, however, call for some rather radical changes in the way we usually think and talk about mind and matter.

For one thing, we should no longer speak of conscious experience as taking place 'in the head' or 'in the brain.' (What, after all, would be different in our experience if it took place somewhere else?) Rather, we should speak of the head or the brain, alike, as something that appears in or is inferred from conscious experience. Nor should we point to our surrounding environment to indicate the objective physical world and to our head to indicate our subjective experience. Everything we experience (whether 'out here' or 'in here') is, alike, a part of our experience.

Spatial extension, too, is no exception. We should not follow Descartes in distinguishing the physical from the mental on the basis of whether it is spatially extended or not (i.e., contrasting 'res extensa' and 'res cogitans'). We directly experience the world and the things in it as spatially extended. Spatial extension, as known to us, is thus, by virtue of being known, a mental phenomenon. If the physical world, independent of our conscious experience possesses spatial extension, it can only be known in a more abstract, mathematical (as opposed to experiential) sense.

Temporal extension, too, takes on a whole different character, once we start from experience, rather than from an objective world view grounded in a physical description of reality. One of the most glaring aspects of time is the distinction between past, present, and future. Only in the 'now' is anything directly given in experience. Past and future are also given in the now, as present memories and present anticipations. While their *contents* point to the past and the future, *as* experienced memories and anticipation they take place in the experienced now. In contrast to this plain fact of every-day (or, better, every-moment) experience, physics has never provided an explanation for the special status of the present moment.

Whereas a global 'moving now' could be introduced by fiat (although rather artificially) in a Newtonian picture of space and time, we do not even have that luxury any more. In special relativity, in which each enduring object is represented by a four-dimensional 'world line' and in which simultaneity is relative to the motion of the observer (*i.e.,* to the orientation of that observer's world line), there is no single preferred basis for an objectively moving common 'now' within which all observers are simultaneously conscious. In general relativity, things are far worse: an observer falling into a black hole will reach the hole in a finite time according to his or her own clocks, but will be seen to 'hover' just outside the black hole's horizon for all eternity, as far as far-away observers are concerned.

Thus, in the third-person view of physics, there simply is no privileged position along the world line of any observer. The moving now has been filtered out, reduced to an arbitrary number, and as a result all times have acquired equal status. What could be farther from our experience, and what could point more blatantly to the process of reduction underlying the whole physicalistic approach.

Finally, we should resist the temptation to invoke the complexity of the brain as somehow providing an explanation of the quality of conscious experience. There is, after all, nothing complex about a momentary flash of red or twinge of pain. (Complexity may be a component of intelligent reasoning or thought, but even those who believe that all mental processes are concomitants of neural ones might not wish to exclude the possibility that the firing of a single neuron in a brain could be the sufficient condition for an experience of pain.)

2.5. Philosophical Company

Even those (if any) who find some cogency in our proposed inversion of the hard problem may conclude, at this point, that we have merely advocated a return to something very like the 1710 idealism of Bishop

George Berkeley (*esse est percipi*: "to be is to be perceived") — hardly much of an advance from the perspective of present day science! There are, however, a number of respects in which our line of thinking departs from that of traditional idealism and (as we shall try to argue in a subsequent section) can provide insights into the problem of consciousness.

Certainly, in proposing to start from what is given in experience, we do not propose to take the given to be exclusively, or even primarily, concrete pointillistic 'sense data.' In contrast with the British empiricists, and more in line with the more phenomenological approaches of Husserl (in Germany; 1913) and Nishida (in Japan; 1911) as well as with the radical empiricism of William James (in the U.S.; 1912), we find that what is given in experience is largely of a different character: Rather than a two-dimensional array of colored spots or patches (in the visual case), what we find to be given is a three-dimensional arrangement of objects that evoke expectations about what further experiences will follow upon various actions that (in the terminology of James Gibson 1979) they appear to 'afford.' (From the experience of dreams, however, we also recognize that there is no guarantee that such expectations will be confirmed on any given occasion.) Likewise, what is given is not confined to the concrete colors, shapes, sounds, tastes, odors, feels, etc. presented by any particular sensory modality. Rather, we are directly aware of relations, affordances, meanings (with, again, the caveat that the associated expectations carry no guarantees) as well as the 'abstract ideas' that Berkeley was wont to reject.

Moreover, we do not deny (as Berkeley did) the possible existence of a world behind the phenomena we directly experience — such as Kant's world of noumena. But, rather than taking the physical world presumed to be known by common sense or (in a very different form) by modern physics, we treat any notions about such a world as hypotheses that are useful to the extent that they predict and explain the regularities in our experience.

Finally, as we shall note, our line of reasoning leads to some rather novel 'panpsychist' speculations about mind that the British empiricists (as well as many present day scientists) would probably find quite counterintuitive. In contrast, twentieth-century continental philosophers, most notably Husserl, would have been much more sympathetic to our views. For Husserl, an empirical approach to reality meant dealing with reality as it presents itself to us, in the form of conscious experience. Interpretations, in terms of atoms and molecules, are exactly that: interpretations.

As Husserl stressed, seemingly simple and fundamental notions, such as that of an electron, in fact carry an enormous amount of baggage with it, including the whole methodology of science, as well as the cultural setting that leads one into a scientific attitude in the first place. An interestingly parallel and much more modern attempt to show this has recently been made by Brian Smith (1996).

All this is not to deny the significance of the *products* of science, its deep insights as well as its powerful technological applications, for better or worse. Rather, we want to stress that the *interpretation* of modern science is far from cut-and-dried, even though methods, results, and interpretations are usually presented as a package deal, the connections between them unquestioned.

3. Turning the Hard Problem Sideways

Having tried our hands at turning the 'hard problem' upside down, let us now consider alternatives. Rather than making a switch from physical reality to experience as providing a grounding for reality, perhaps we can consider both on an equal footing, without trying to make the one into a foundation for the other.

3.1. A Second-Person Point of View

Another way to approach the 'hard problem' is thus to acknowledge the third-person character (in the grammatical sense) of knowledge based on a physical description of reality. In the objective approach, there simply is no room for the 'moving now' as experienced by me, as an individual human being. In contrast, turning the hard problem upside down suggests that all knowledge starts with the subject, the first person, the 'I' who looks at the world, standing on the ground of 'my' experience. If this way of turning the problem 180 degrees around seems to be too much of a good thing, how about a more modest turn, by only 90 degrees?

Turning the hard problem sideways brings us to the remaining grammatical choice: that of the second person. Whereas an I-based attitude raises the specter of solipsism, and an it-based attitude offers only a cold objectivism, an I-and-you based orientation may combine the best of both alternatives, while avoiding the unpalatable extremes. In other words, the notion of intersubjectivity cannot be seen as a simple superposition of subjective and objective properties. Rather, acknowledging consciousness in others as being on a par with our own, we see a world around us, filled with physical objects as well as conscious experience of humans and other animals. The fact that we can and do interact with others is an aspect of conscious experience that is at least as important to us as the possibility that we humans have of reflecting upon our own existence.

3.2. Self-Reference

It may seem that the original 'hard problem' was not eliminated by simply turning it upside down. If the proposed inversion of the problem turned its daunting frown into a smile, that smile may now appear to be one that is more ironic than comforting. Again speaking in the first-person, each of us may well regard brains as things that arise within our experience (rather than regarding experience as something that arises within brains). We may even adopt the 'methodological solipsism' approach to science.

Nevertheless, one of the brains for which my own experience already provides good evidence is the brain I call my own. Moreover, everything I have experienced makes me confident that I could obtain even more detailed evidence about my own brain by seeking certain further experiences — such as those we would call functional magnetic resonance imaging (fMRI) or positron emission tomography (PET) of my own brain. In a manner reminiscent of the 'autocerebroscope' envisioned by Herbert Feigl (Feigl 1958, Meehl 1966, *cf.* Shepard 1978), then, I might have a 'first-order' experience (whether of a flash of red, a twinge of pain, a beloved face, or a scientific insight) and, simultaneously, a 'second-order' experience of an (fMRI or PET) image of my own brain in which locations corresponding to the first-order experience were 'lighting up.'

Quotation marks surround 'first-order' and 'second-order,' here, in recognition that, in a sense, all experiences are 'first-order' experiences; it is only one's interpretations of experience of the two kinds that might be said to be of 'second' order — and these interpretations, too, are parts of one's experience. Ultimately, we encounter this curious circle: part of the 'lighting up' in the brain image I experience may represent the very neuronal activity that corresponds to my experiencing the brain image of that same activation. This may lead to Gödelian paradoxes.

After all, mathematics and consciousness have in common that they can be described self-reflexively. This in contrast to, say, physics, where mathematics as the language of physics is different from physics, or chemistry, which is dependent on physics for providing its basic building blocks. Math, however, can be directly modeled by math, and consciousness can be directly studied by consciousness.

In other words, if we view (somewhat naively) the sciences as providing a tower, starting with math on the ground floor, physics on top of that, then chemistry, biology, etc., and psychology on top, then we have to conclude that only the bottom and top layer are truly self-reflexive, and therefore provide the conditions of possibility for unique forms of paradox — as exemplified by Gödel's incompleteness theorems.

3.3. Other-Reference

Even within one's own experience, then, several puzzles remain:

- How is one to understand the relation between 'first-order' experiences and 'second-order' experiences of their corresponding brain activities?
- Why is it that only some activities of my own brain (as I might experience them through brain imaging) correspond to my 'first-order' experiences while other activities (apart from the just noted 'curious circularity') evidently have no conscious manifestation?
- Where, indeed, am I to draw the line between those things in my experience that (in the 'second-order' sense) are accompanied by some 'first-order' experiences — whether experienced by what I call 'me' or by some other, independent experiencing agent or entity?

Just as we may take certain kinds of experienced regularities — and also surprises — as manifestations of an invisible and intangible (i.e., a 'noumenal') world behind the phenomenal world of experience, so too we may take certain other kinds of experienced regularities — and also surprises — as manifestations of other minds.

For example, I may have the experience of another person presenting an extended argument that leads up to a particular, unexpected conclusion that I see to be valid only after I subsequently think through the argument (or, perhaps, after I verify the conclusion by performing an actual experiment or calculation). Such an experience seems to provide compelling evidence for the occurrence of mental processes independent of my own. Granted, given the distinction (mentioned at the outset) between adaptive behavior and subjective experience, this does not in itself provide a definitive answer for the 'hard problem' of other minds experiencing the same qualia I do or, indeed, experiencing any qualia at all (*cf.* Shepard 1995).

It would, however, seem a strange and inexplicable violation of symmetry if the other bodies that I experience as being so much like my own in structure and behavior differed so radically as to have no conscious experience. Nor does there seem to be any basis for drawing a sharp dividing line between humans and other animals — including such diverse species as apes, dogs, dolphins, and birds, which have more or less similarly structured brains and which exhibit behaviors similar to those that in ourselves are the outward manifestations of felt joy, lust, love, affection, caring, anger, fear, or pain (Shepard 1993). It would seem more reasonable to grant reality the necessary structure to take care of the occurrence of consciousness, just as space seems to somehow take care of providing the correct angles and sides for a triangle, as we saw in §2.3. This aspect of reality, then, must be a different aspect from the more familiar notions of space and time. We will consider this possibility further in §4.2.

4. Upside Down and Sideways may be Compatible

Having described both of the turns we have given to the hard problem of consciousness, we have reached a point where we can begin to speculate about the relationship between the 90 degree and 180 degree rotations. We think both may turn out to be valid.

For example, we could consider matter and consciousness both as emergent properties of underlying and more fundamental aspects of reality. Just as matter might ultimately be explained as a property of space and time (as is already the case for the mass and energy of a black hole), so consciousness might be a property of another aspect of reality, 'X' for short.

We realize that such a proposal immediately faces two serious problems. If we stop at this point, merely mentioning the notion of an unspecified aspect 'X', the reader may consider us glib and/or superficial, deriding us for not providing more detail. However, if we follow the temptation to construct an *ad hoc* model, we might be able to make our point more clearly, but with the penalty attached of almost certainly being wrong, in more than just the details of the model.

In this paper, we will try to avoid both horns of this dilemma. We will resolutely avoid any premature form of model building. Our goal here is not to provide a solution, let alone 'the' solution, to the hard problem of consciousness. At present, most specific models of consciousness strike us as naive. Instead, we see a greater need for new questions, rather than new attempts at answers, and we hope that the present paper will provide a step in that direction. At the same time, we do want to illustrate our thinking as clearly as possible. To that end, we start with an analogy from physics, in §4.1, and apply that in §4.2 to our view of the role of consciousness as a derived aspect of a more fundamental property of reality, X, on a par with space and time.

4.1. Space and Time

How reasonable is it, to view matter and mind on the same level, as complementary? Starting with matter, let us imagine that a future form of physics will have succeeded in describing matter and energy as forms of excitations of spacetime (in analogy with the case of a black hole, say, where spacetime curvature directly provides a definite mass, with no need for any specific 'matter' ingredient). Is it, then, reasonable to turn the hard problem sideways, viewing conscious experience to be complementary to spacetime, neither of the two being reducible to the other?

We will address that question in the next section. As a warm-up exercise, and in order to provide a helpful analogy (Hut 1996, p. 149), we first remain on the physics side, in order to have a closer look at space and time. They, too, cannot be reduced, one to the other, even though they can be *partly* transformed into each other, according to relativity theory. But let us keep things simple, and start with everyday experience.

What is space? If we should meet somebody from another culture in which there were no word for space, how would we be able to describe our concept of space? Each specific description would even strike ourselves as being too crude. We could try to point to space as something that is present everywhere, as what remains behind after taking away all objects. But we would immediately realize how such an attempt would be almost certainly misunderstood. It would invite a view of space as a type of all-pervading substance, like air or ether. And while both can be used to some extent as metaphors, neither captures the notion of space.

What is time? This is an even more difficult question. Let us imagine that we meet someone who has a working knowledge of space, and shares the way we talk about space, but for some reason is not familiar with the notion of time. Perhaps that person has had a stroke or car accident, resulting in a form of selective amnesia. How might we begin to teach such a person what it means for us, to live in a world of space and time, rather than just in a world of space?

As in the case of space, we would be hard put to capture the notion of time in a purely verbal description. It would seem to make more sense to try to use a more action-based approach. We could take a series of snapshots of a street scene, say. We could put those pictures on a table, and point out that each picture shows the same space, but at different times.

Our 'space man,' who had somehow lost the notion of time, might nevertheless recognize a house in each of the snapshots. We then tell him that it is the same house in each picture, that he has to identify all these houses with each other. And what about that cloud, which has slightly different positions in each picture? These, too, all have to be identified with each other, as all pertaining to one and the same cloud.

In short, each object is really a summary notion for a whole series of objects, as seen through the stack of pictures. And where does time come in? Can our space man get a clue from the fact that the cloud is occupying slightly different positions in each photograph? "Yes indeed," we tell him, "there is a significant difference between the house and the cloud. The house does not move; it has no motion. The cloud has some motion."

The conversation with the space man could then continue according to the following dialogue. Puzzled by the notion of motion, he would ask:

So time is the same as motion?

No. Time is what makes motion possible.

But clearly, the cloud has more 'time' than the house. The house is the same in all pictures. I can understand that from a purely space-based picture. No need to introduce this mysterious notion of 'time.'

No, there is as much time in the house as in the cloud. In fact, time is not located anywhere. It is equally present everywhere.

Like space! So, after all, space and time are exactly the same.

No, not at all. I understand that it is hard to imagine, and indeed space and time could both be said to be everywhere, in some sense. Still, they are completely different.

Hmm. Hard to imagine indeed. And what about that middle picture? It contains a car, one that is not present in any of the other pictures. Surely, there must be an error of some sort.

No, it simply means that the car went by so fast that it did not register in the other 'nearby' snapshots. It had a greater amount of motion, but that does not make it more or less real.

Clearly, even with a stack of snapshots, it would not be easy to get the idea of time across. And of course, this whole process of explanation would unroll in time. It could never happen in the first place if we were dealing with a purely 'spacey' being, one that did not partake in time at all. It is this observation that connects our example to the hard problem of consciousness.

4.2. Space, Time, and X

In the world described by physics, there is room for space and time (or at least for space-time), but there is no room for conscious experience (*cf.* McGinn 1995). Of course, in any laboratory experiment, and in any derivation of a piece of theoretical physics, consciousness of the physicists is necessary, if nothing else, to comprehend the final results. But let us imagine that someone did not understand the notion of consciousness, just as in our previous example someone did not understand the notions of time and motion.

The analogy here would be between consciousness and motion. A car seems to have more motion than a house (in the rest frame of the house at least), and a human brain seems to have more consciousness than a rock (from a human perspective at least). In the previous section, the spaceman had to learn two things, in turn: motion, which was visible, and time, which was inferred, invisible, but considered a more fundamental aspect of reality. So in this section, we find our work cut out for us: the space-time-person (the physicist) has to be shown, first, the presence of conscious experiences, and then the more fundamental aspect of reality that provides the condition of possibility for experience. Let us call the latter X, for lack of a better designation at this point. X then stands to consciousness as time stands to motion.

Since the notion of some aspect of reality, X, is rather abstract, we can try to use a more familiar label, one that at least points in the direction we are contemplating. One possibility would be to use the label 'sense' to stand for X (Hut 1996; Hut and van Fraassen 1996; *cf.* Rota 1989). Like space, like time, sense[1] is for us what water is for a fish. Our lives are embedded in it, given by it, irremovably linked to and through it. If we would lack any understanding of the world around us, in other words if nothing would make sense, we would not have any understanding of either space or time. But since the world does make sense to us, we can explore what it means, this notion of sense. Like space and time, sense it not something that can be directly experienced as such. We can form an impression of space through separations between particular objects. Similarly we can form an impression of time through observing specific changes taking place in time. And we can look at the same object or situation with a different 'depth' of sense, possibly looking at it 'from a different angle' (figuratively speaking).

Sure, we can interpret our world as a world of things. But what is a thing? When we look carefully, then we find that what we considered to be an object appears in our consciousness as a bundle of meanings, draped around sense impressions that are far, far less complete and filled in and filled up than the 'real thing' we feel to be present, three-dimensionally, continuous in time. What then remains of the solidity of the object? It is recognized in its givenness for us through the *sense* of solidity we have. Its continuity? This follows from our *sense* of continuity and identity. Its reality? Nothing but a *sense* of reality. The indubitability of its reality?

[1] The word 'sense' is used here in its aspect of 'meaning,' *not* in connection with 'sense experience' – the word sense seems to convey more directly a grasp of something than the more abstract word meaning.

The only thing we have a real handle on is our *sense* of indubitability of its reality. And this brings us back to sense as X, not so much a form of 'background field,' but rather even more fundamentally, a primordial aspect of reality. Sense is seen then as a dimension of reality, on a par with space and time (*cf.* Tarthang Tulku 1977).

No more can we walk out of sense than we can walk out of space or out of time. Still, we have to learn to see this, to see what such a statement may mean, experientially. In that respect we are initially in a situation very much like the space man trying to find time. Using this analogy, we could conjecture that sense is 'everywhere,' just as space and time are.

More accurately, space is 'everywhere,' time is 'everywhen,' and sense is 'in every which sense.' Space is also there where there are no objects present. Time is also there where there are no specific events to be located. And sense, then, could be postulated to exist also there where no specific information would be at hand.

This idea, that we live in a world of sense, and that we can move around in sense, may sound strange. But it is clear that sense pervades our lived world. It seems hard to escape the conclusion that everything we know, as we know it, is what we know it to be through the way it makes sense to us. In this light, even nonsense is yet another form of sense.

Physics, then, describes a simplification of the real world, by projecting it down along the X axis. There is an analogy within physics itself. Projecting physics down along the time axis, we are left with space only, and we can then study statics. Of course, statics is nothing more than a limiting case of kinematics, in which all motions have been reduced to zero, or ignored. Similarly, physics studies only a limiting case of our reality in which the presence of the X 'dimension' of reality has been ignored, with the consequence that those aspects we call value and beauty have been reduced to zero as well. This is not to say that a physicist cannot find beauty or sense in physics; after all, we need to apply motion in order to set up an equilibrium situation that can be studied with statics alone. It is just that beauty and value *as such* do not appear as part of the formalism of physics.

4.3. Beyond Pandynamism and Panpsychism

A graphic way to illustrate the role of consciousness in our view of reality as woven out of space, time, and X is presented in figure 1. Confronted with the hard problem of the relation between conscious experience and the physical world, our first move is to postulate the presence of a third aspect of reality, besides the dimensions of space and time that underlie our description of the world in terms of physics. We have used the term 'X' here, for lack of a specific model, or even a specific set of notions as to what might be the structure of this extra 'dimension' of reality (Hut and van Fraassen 1996).

Our second move is to make an analogy between motion and physics, on the one hand, and consciousness and reality including X, on the other. As we argued in §4.1, if we would have limited a study of physics to the field of statics, we would by definition have had no way of dealing with motion. Any attempt to introduce the notion of motion in purely spatial terms would then have led to confusion and puzzlement. A first attempt to broaden a description of reality, beyond statics, could have postulated a new and as yet unknown 'field of potential motion', pervading all of space — since, after all, motion is possible in any position in space. Such a mysterious field could then be considered to somehow 'carry' motion, and such a theory could properly be called a form of 'pandynamism.'

Similarly, attempts to postulate the pervasive presence of (proto-) consciousness under the banner of panpsychism, may seem puzzling, if not downright obscure, when starting with physics on its natural stage of space and time. Our conclusion is that attempts to *embed* consciousness in space and time are doomed to failure, just as equivalent attempts to *embed* motion in space only. Yes, motion does *take place in* space, but it also partakes in time. Similarly, consciousness certainly *takes place in* space and time, but in addition seems to require an additional aspect of reality, namely X, in order for us to give a proper description of its relation with the world as described in physics.

At this point, we can only guess what the third term will be, following the succession of statics and dynamics. In our figure, we have used the term 'radical empiricism', coined by William James (1912), since the ideas expressed by James are sufficiently simple and general as well as close enough to our views to serve as a place holder in the figure.

4.4. Taking Turns

The 180^o turn we started out with, when we turned the hard problem upside-down, is natural when we analyze the world from within our own experience, individually and collectively. In such a move, the first thing to notice is that our own sense of self and individuality is given with (and as part of) experience. Experience has a self, not the other way around. However, once we have accepted conscious experience as fundamental, the question arises of the ontological status of experience. Can experience itself provide the foundation of all of reality, as a basic sort of substratum (and what would that even mean?), or would it be more accurate to view experience as something that points to a more basic aspect of reality, an aspect X that provides the condition of possibility for experience? The latter seems to us more plausible, in light of the intersubjectivity we encounter in our world (*cf.* §2.3).

When starting with space and time as the conditions of possibility for our physical reality, and some form of X as the condition of possibility for our conscious experience, we naturally find ourselves having made something that might be more aptly described as a 90^o turn. The hard problem has been turned sideways, as we saw with the use of a few analogies. A bullet has more motion than a mountain, but not more time[2]; the whole world is drenched in time; it may be drenched in a form of pre-experience (X) as well. In fact, you can't say which is drenched in which: space in time, time in space, time and space in pre-experience, pre-experience and space in time, *etc.* We suggest that all three are equiprimordial, co-eval aspects of reality, intimately interwoven, and only together making up reality, 'as we know it'.

This change of perspective has implications that go beyond questions of epistemology or ontology. In the usual interpretation, we grant that different bodies, and their respective brains, share the same space and time, while we assume their associated experiences to be wholly separated. In our view, however, the experiences of different individuals are more intimately connected, in sharing the pre-experience aspect of reality, as well as the space and time aspects. It is here that intersubjectivity, dealing with different subjects in the second person, on an I-and-you basis, acquires its 'inter'.

[2] slightly less, in fact, according to special relativity.

Differences between turning the hard problem upside down and turning it sideways may lose their meaning. How we may describe such a reinterpretation may reflect more our predeliction for interpretation. We can illustrate this with another example from physics. Do we consider matter and energy on the same level? Is it equally reasonable to call energy a rarefied form of matter, as it is to call matter a solidified form of energy? Intuitively, the latter may seem more accurate. Once matter is seen to be no longer bound to a fixed mass or form of representation, matter partakes in the greater degree of mobility and fluidity of energy, at least potentially. Considering both forms, matter and energy, to have more 'energy-like' properties might seem like the natural conclusion. But ultimately, it is more a matter of convention, of defining what one means exactly with the word 'energy,' which determines whether matter is seen as a form of energy, or whether matter and energy are considered to be on a par.

Somewhat similarly, we could allow the hard problem of consciousness also to be turned side-ways *as well as* upside-down, depending on the view we would take. On a more practical level of describing how human minds interact with material objects, for example, an effective dualism may carry us quite far, in a side-ways approach. But as soon as we reflect on the lack of consistency of a dualistic view, and the unsatisfactory nature of splitting up the world into parallel and seemingly incommensurable aspects, we can take the more fundamental view in which the hard problem is turned upside down. The question then is: what is the nature of the 'conscious mind' that is seen as more fundamental than 'matter'? And how does consciousness 'arise' in a mind, or vice versa? As long as all these terms, consciousness and mind and arising, are used as vaguely and variously as they are today, the above question is not even well-posed. If the word 'mind' is used for what we have designated as X, the hard problem can be considered to be turned upside down. If the word 'mind' is associated with conscious experience, a turn side-ways may form a better description.

In either case, however, we may question whether it still makes sense to use the labels consciousness or experience. Maybe a label such as 'appearance' is more appropriate. Something appears. That's a given. From moment to moment we find ourselves in an ongoing flux of appearing. Everything else, events, experience, consciousness, let alone material objects, are late-comers, results from taking interpretive stances.

Even though we may have been trained in accepting such interpretive stances, to the exclusion of everything else, at pre-kindergarten age, still this by itself is no reason not to question such stances. When approaching the hard problem of consciousness, let us try to remain as open-minded as we possibly can.

Aren't these statements rather rough and qualitative forms of speculation? Yes indeed, and that is all we feel ready for, now. Is there no hope to flesh these ideas out, make them perhaps more quantitative? Yes, there is. Physics and math were not created overnight, so it is unreasonable to expect — let alone demand — miracles the day we abandon the overly rigid 'straight-up' view of nature that has been with us for the last few centuries. Instead, let us try to flesh out some of our speculations a bit further, using yet other analogies. One appropriate analogy is found in the fact that we all are accomplished experimenters in dealing with reality transformations, in the process called dreaming.

Within a dream, we identify ourselves with our own bodily presence, and consider other people and things to be independent of us, and to reside 'outside' ourselves. But after we wake up, we switch perspective: we view each and every element of the dream as being part of us, something constructed in our consciousness. Not only can we then consider ourselves as having scripted the role of each person we have met in the dream, but the role of each animal, plant, or inert object as well.

What is more, we must conclude that we have provided the supporting background notions of the dream-time and the dream-space that have formed the stage for all (seemingly) material objects to appear. Yet, somehow our usual identification with our body seems to prevent us from more than rarely dreaming ourselves to be an animal or a plant — or a rock, a piece of trash, or a patch of empty space that generously allows each and any object to pass through without being modified by it in the least.

Granted, in dreams we may also believe in the independent existence both of the physical world we seem to perceive around us and of the minds of the other persons with whom we interact in the dream. Yet, on awakening, that 'physical world' and the 'other minds' that expressed themselves in it vanish. Their evanescence does not, however, preclude a dependence of such directly experienced phantasms on a (noumenal) something beyond themselves. The prevailing scientific view is, in fact, that both the order and the surprises within the dream arose from something external to the experienced dream — namely, the restless activity of our own physical brain.

In short, there may be some justification — in waking and dreaming consciousness alike — for hypothesizing the existence of something behind what we experience as an explanation for both its predictable or conservative aspect and its unanticipated or creative aspect. But whatever that something may be, it would be a category mistake to take particular elements from *within* experience, whether based on dreaming or waking, as fundamental. However abstract our notions of atoms, quantum fields, or more exotic constructs may be, all of these notions are ultimately grounded *in* experience. As such, they cannot even be considered as candidates for whatever it might be, if anything, that could be considered to underlie conscious experience. Such a candidate had better be far more pervasive. We view our attempt to label such a candidate with 'sense' or better 'X' as no more than a shot in the dark — but, we hope, a shot in the right direction.

Acknowledgments. We thank Sean Murphy and Jonathan Shear, as well as an anonymous referee[3] for comments on the manuscript. This work was supported in part by a grant to the authors from the Alfred P. Sloan Foundation, for research on limits to scientific knowledge. We thank John Casti and Joseph Traub, organizers of the workshop on 'Limits to Scientific Knowledge,' held at the Santa Fe Institute in May 1994, for providing a natural opportunity for the two of us to meet each other. We also thank SFI for their hospitality, and the Sloan Organization for sponsoring that workshop.

References

[3] This referee also alerted us to some parallels between our notion of 'turning the hard problem upside down' and several papers by Velmans, most recently Velmans, M. (1996), 'What and where are conscious experiences?', in *The Science of Consciousness: Psychological, Neuropsychological and Clinical Reviews*, ed. M. Velmans (Routledge).

Bridgeman, P.W. (1940), 'Science: Public or private?,' *Philosophy of Science*, **7**, 36-48.

Carnap, R. (1928), *Der Logische Aufbau der Welt.* English translation, by Rolf A. George (1967): *The Logical Structure of the World* (Berkeley: Univ. of Cal. Pr.).

Chalmers, D.J. (1995), 'Facing up to the problem of consciousness,' *Journal of Consciousness Studies*, **2** (3), pp. 200-19.

Feigl, H. (1958), 'The "mental" and the "physical"', in *Minnesota studies in the Philosophy of Science*, Vol. II, eds. G. Maxwell and M. Scriven (Minneapolis: Univ. of Minnesota Pr.), pp. 370-497.

Gibson, J.J. (1979), *The Ecological Approach to Visual Perception* (Boston: Houghton-Mifflin).

Husserl, E. (1913), *Ideen zu einer reinen Phänomenologie and phänomenologischen Philosophie I, Erstes Buch*, translated as '*Ideas Pertaining to a Pure Phenomenology and to a Phenomenological Philosophy, first book*, 1982 (Dordrecht: Kluwer)

Hut, P. (1996), 'Structuring reality: the role of limits,' in *Boundaries and Barriers*, eds. J. Casti and A. Karlqvist (Reading, MA: Addison-Wesley), pp. 148-187.

Hut, P. and van Fraassen, B. (1996), 'Elements of reality: A dialogue', submitted to *Journal of Consciousness Studies.*

James, W. (1912), *Essays in Radical Empiricism*, reprinted in *Essays in Radical Empiricism & A Pluralistic Universe*, 1967, by W. James (Peter Smith).

McGinn, C., 1995, 'Consciousness and Space', *Journal of Consciousness Studies*, **2** (3), pp. 220-30.

Meehl, P. E. (1966), 'The compleat autocerebroscopist,' in *Mind, matter, method: Essays in Philosophy and Science in Honor of Herbert Feigl*, eds. P.K. Feyerabend and G. Maxwell (Minneapolis: Univ. of Minnesota Pr.), pp. 103-81.

Nishida, K. (1911), *Zen no Kenkyuu*, translated as '*An Inquiry into the Good*, 1990 (Yale Univ. Pr.).

Rota, G.-C. (1989), '*Fundierung* as a logical concept,' *The Monist*, **72**, pp. 70-77.

Shepard, R.N. (1978), 'The mental image,' *American Psychologist*, **33**, Figs. 1, 2 & 4.

Shepard, R.N. (1981), 'Psychophysical complementarity,' in *Perceptual organization*, eds. M. Kubovy and J. Pomerantz (Hillsdale, NJ: Erlbaum), p. 280.

Shepard, R.N. (1993), 'On the physical basis, linguistic representation, and conscious experience of colors,' in *Conceptions of the mind*, ed. G. Harman (Hillsdale, NJ: Erlbaum), pp. 217-45.

Shepard, R.N. (1995), 'What is an agent that it experiences P-consciousness? And what is P-consciousness that it moves an agent?', *Behavioral and Brain Science*, **18**, pp. 267-8.

Smith, B.C. (1996), *On the Origin of Objects* (Cambridge: M.I.T. Press).

Tarthang Tulku (1977), *Time, Space, and Knowledge* (Berkeley: Dharma Publ.).

Figure 1: A schematic summary of our view of the role of consciousness, in analogy with that of motion. Motion is a phenomenon that cannot be captured in the study of statics. Rather, it points to the presence of time as an aspect of reality equally fundamental as space. Similarly, it seems that consciousness cannot be captured in a study of purely physical phenomena, unless we extend our notion of physics to include at least one new element, on a par with space and time, indicated here by 'X'. We offer this suggestion as an sharpening of the notion of panpsychism, just like the notion of time would be a sharpening of the concept of 'pandynamism' that someone could introduce as a condition of possibility for motion.