Journal of Consciousness Studies
www.imprint-academic.com/jcs

Journal of Consciousness Studies, **5**, No. 4, 1998, pp. 464–82

EMERGENCE & THE MIND–BODY PROBLEM*

Michael Silberstein, Department of Philosophy, Elizabethtown College, Elizabethtown, PA 17022, USA. Email: Silbermd@acad.etown.edu

Abstract: In the first part of the paper I argue that neither *physicalism* (whether its reductive or non-reductive form) nor standard forms of *dualism* (as well as Chalmers' *fundamentalism*) can provide an explanatory framework for consciousness or cognition — neither account can existence of conscious experience nor its relationship to cognition and the brain. Physicalism and fundamentalism fail to provide an explanatory framework for consciousness because they both share, at least with respect to the physical universe, the same misguided commitment to part/whole reductionism and microreductive accounts of explanation. In addition to their lack of explanatory power, both physicalism and fundamentalism have well known absurd and troubling metaphysical consequences such as eliminativism and epiphenomenalism. In the second section of the paper I advocate a position I call *radical emergence*, arguing that microphysics (especially quantum mechanics) provides strong empirical evidence for emergence. I show that emergence provides a viable alternative for explaining consciousness and cognition — an alternative that has none of the awkward metaphysical consequences of either physicalism or fundamentalism.

I: Physicalism and Fundamentalism

Those of us who study consciousness (experience, awareness, etc.) are ultimately interested in answering the following questions:

(1) What exactly explains the existence of consciousness?
(2) What exactly is the relationship between consciousness and 'physical' systems such as the brain?
(3) What exactly is the relationship between consciousness and cognition?

Of course, answering these three questions will mean answering a host of more specific questions. For example, there are those who argue that question number one must be further divided into a number of more specific and distinct questions based on a scientifically sophisticated taxonomy of the phenomena of consciousness (see Lycan, 1996, for example). Answering the second and third questions thoroughly will probably require a deep understanding of neuroscience and cognition, and acquiring such an understanding will mean answering many, many different questions. The assumption that philosophers make however is that even after all the rich empirical complexity has been factored in, it is still reasonable to ask the wholesale question: *how/why did consciousness evolve, and what is it for?* It is widely held that this question is not merely an empirical question, but, as Lycan puts it:

> Both in philosophy and in psychology 'the problem of consciousness' is supposed to be very special. It is not just the mind–body problem; few theorists question the eventual truth of materialism in some form, but many see a deep principled difficulty for the materialist in giving a plausible account of 'consciousness'. Nor is it just the problem of *intentionality*, or mental aboutness, in particular, since intentional states need not be conscious in any sense at all. It has to do with the internal or subjective character of expe-

* I would like to give a special thanks to Paul Humphreys, James Hawthorne, Peter J. Lewis, Teed Rockwell and an anonymous referee for their invaluable comments on earlier drafts of this paper. I would also like to thank the National Endowment for the Humanities for financial support.

> rience, paradigmatically, sensory experience, and how such a thing can be accommodated in, or even tolerated by, a materialist theory of the mind. And it is a conceptual problem, not merely an empirical one — it is a 'How could . . . possibly . . .' question, not merely a 'How does . . .' question. Scientifically, there is felt to be a systematic, if not insuperable, obstacle to psychological research, and philosophically there is felt to be a conceptual tension between materialism and the phenomenal or subjective character of experience (Lycan, 1996, p. 1).

None of the more popular explanations of consciousness seem to explain much of anything (especially the purely philosophical explanatory frameworks). Perhaps the two most common reactions to the nest of worries about consciousness that are characterized in the Lycan passage, are *physicalism* and *fundamentalism* respectively. Physicalism is motivated by both methodological and ontological reductionism. Physicalism holds that consciousness is either reducible to or determined by the physical. But what physicalism does not do is tell us *how* or *why* this is the case. Non-reductive physicalism (NRP) does not explain how or why brain-states produce (non-causally determine) consciousness — or if you prefer, NRP does not explain how or why consciousness 'supervenes' on brain-states (see Kim, 1993). Is the explanation simply that such production is a brute fact? If so, it would seem that NRP is no better off than emergence, which is frequently accused of promoting mystery as explanation. Indeed, NRP is much worse off than emergence, since its commitment to the causal closure of the physical domain means that it cannot begin to explain how conscious states arise from non-conscious physical states. And what is more, given that NRP holds property dualism in addition to causal closure, it is forced to embrace epiphenomenalism (the lack of efficacy) with respect to those conscious states (see Kim, 1996).

You might think that reductive physicalism is better off, since presumably reduction/identity relations do not require further explanation. For example, it would be silly to ask the question, 'but why is water H_2O?' However, when one thinks of the reduction/identity relation realistically, as being between entities or properties (rather than between theories as neo-positivists and eliminativists tend to do), then the reduction of the qualitative to the non-qualitative seems more than puzzling. For example, it is one thing to claim that qualitative states of awareness are properties of brain-states or neurochemical processes, but it is quite another to claim that qualitative states are *nothing but* non-qualitative neurochemical processes. The latter kind of reductionism would certainly solve the problem; unfortunately, it is demonstrably false, since each of us knows that conscious states are not merely theoretical posits, but rather they are observable phenomena that require explanation. The former kind of reductionism might be true, but even so it does not explain *how* or *why* brain-states are conscious. Again, if this is just a brute fact, then reductive physicalism is no better off than emergence. In other words, the only difference between *this* characterization of reductive physicalism and NRP is that reductive physicalism maintains property monism, while NRP admits to being property-dualist.

Notice that this is merely a pseudo-scientific squabble about whether properties of the world come in essentially one *type* (flavour) or two *types* (flavours). The hard problem is not whether conscious states are essentially *physical* as opposed to essentially *mental* (I regard this as a pseudo-problem), the hard problem is *how* and *why* do brain-states give rise to conscious states, or if you prefer, *how* and *why* are brain-states conscious? Simply claiming that qualia are *physical* kinds as opposed to *mental* kinds, does not explain anything. The question of types is orthogonal to the question of reduction. The question is: what relation(s) do the qualitative *features* of 'brain' processes bear to the non-qualitative *features* of 'brain' processes? One thing is for sure: the answer to that question cannot be 'reduction/identity'.

On the other end of the explanatory spectrum we have accounts of consciousness that attempt to explain its existence by, in one way or another, positing it as a *fundamental* entity, property, or process (see for example Chalmers [1996] and Penrose [in Penrose *et al.*, 1997] just to name two). The idea is that along with the four forces, super-strings, or whatever turns out to be basic, we must posit consciousness as a fundamental ingredient of the universe. The argument behind this move (let us call it fundamentalism) seems to be that if consciousness cannot be reduced to or identified with the non-conscious, then it must be a fundamental ingredient. Notice that this explanatory strategy is also motivated by methodological and ontological reductionism, for the suggestion is that a property/entity is either fundamental or it is reducible to one that is fundamental — there are no other explanatory options. The problem with fundamentalism (aside from the fact that it flies in the face of everything we know about cosmological/biological evolution and that it, like physicalism, is generally forced into epiphenomenalism) is that it does not *explain* anything. Indeed, fundamentalism is really the ultimate way of saying that consciousness is just a brute fact. However, even if you accept fundamentalism, it still leaves much to be explained, such as: why is there such an intimate relationship between consciousness and the brain? Fundamentalism does not seem to answer *any* of our questions.

Neither physicalism nor fundamentalism seems a very promising explanation of consciousness; in fact, in many ways, they are just two-sides of the same coin. Both views presuppose that a phenomenon is either fundamental or reducible to something fundamental or at the very least, determined by something that is fundamental. Neither view can conceive of another possibility. For this reason, both physicalism and fundamentalism share the same conception of the mind/body problem, namely that:

> There are persuasive reasons to believe that the world we live in is a fundamentally material world, a world made up of material particles and their aggregates, all of which behave strictly in accordance with physical law. How can we accommodate minds and mentality in such a world? (Kim, 1996, p. 9).

According to both physicalism and fundamentalism, the mind/body problem is how to account for mind in a world that appears to be *nothing but* an aggregate of its most fundamental parts (part/whole reductionism) and therefore a world that is *essentially* physical (non-mental). Given this assumption, it would appear that either physicalism or fundamentalism is the only solution possible to the mind/body problem. It is also worth re-emphasizing that both these accounts lead to epiphenomenalism when combined with the assumption that the physical domain is causally closed (an assumption they both share as a rule). If the mental is epiphenomenal with respect to both the mental and the physical — either because the physical diachronically determines the physical and synchronically determines the mental as is the case with NRP (see Kim, 1996), or because the mental is a fundamentally distinct ingredient from the physical that cannot interact with a causally closed physical world as is the case with fundamentalism (see Chalmers, 1996) — then the mental is eliminated in any way that matters for us and for explanation (scientific, folk or otherwise).

My question is this: why do both physicalism and fundamentalism hold that the physical universe is nothing but an aggregate of its parts? For it is this assumption that leads to both these troublesome accounts of the mind. After all, we would all rather have a non-reductive account of the mental that in no way flirts with epiphenomenalism or eliminativism. Boghossian raises a similar question when he says:

> It is worth noting — if only because it so seldom is nowadays — that this rationale for the naturalistic conviction begs a question that doesn't obviously deserve to be begged. Why, indeed, must we think that no property can be real unless it is identical with, or supervenient upon, the properties that appear in the catalogues provided by physics? There is, I think, no *obvious* answer. . . . And what the naturalist needs is an argument why, in general, it is a condition of a property's being real that it supervene on [or reduce to] the properties recognized by physics (Boghossian, 1994, p. 65).

The view that the physical world is nothing but an aggregate of its parts, is known as part/whole reductionism: all wholes (objects with parts) are completely reducible to their most basic parts (entities without further parts). There are two ways of cashing this out, and the first is as follows: 'a whole is explained by being shown to be nothing but the parts, interrelated in a certain manner' (Scharf, 1989, p. 602). Scharf calls this doctrine microreduction and he goes on to say that, 'microreduction requires that compound elements (objects composed of parts) and their properties be explainable in terms of their parts and their interrelations' (*ibid.*). Kim, on the other hand, puts the point this way: 'wholes are completely determined, causally and ontologically, by their parts . . .' (1978, p. 154). Scharf is giving the *nothing but* version of part/whole reductionism and Kim is giving the weaker determination version favoured by NRP. Kim calls this doctrine mereological supervenience or microdeterminism and claims that it is the primary motivation behind most forms of physicalism — of both the reductive and the non-reductive variety (see Kim, 1993, pp. 77, 96–7, 148, 168 and 338). Both Kim and Scharf agree that, 'the program for the unity of science is a program for universal microreduction' (Scharf, 1989, p. 608). If the properties of wholes are fundamentally reducible to or determined by the properties of their most basics parts (as decided by microphysics) then it would seem that physicalism or fundamentalism must be right about the status of the mental. As Kim puts it:

> A particularly important and promising approach to consider, I believe, is to explicate mind–body supervenience as an instance of mereological supervenience. That is, we try to view mental properties as macroproperties of persons, or whole organisms, which are determined by, and dependent on, the character and organization of the appropriate parts, or subsystems, of organisms (Kim, 1993, p. 168).

The truth of part/whole reductionism would also fully justify the widely held assumption that physics, or more properly, microphysics, is causally closed and complete. To put it more strongly, if either version of part/whole reductionism is true then the widespread belief in the priority of microphysics is absolutely justified. But again, why do so many people accept the truth of part/whole reductionism, and how do they justify this belief? The answer comes to this:

> Mereological supervenience. . . . It is this metaphysical doctrine of atomism that seems to underlie and support the enormously productive research strategy of micro-reduction in modern theoretical science. And, conversely, the success of this research strategy reinforces our belief in mereological supervenience (Kim, 1993, p. 77).

In what follows I will show that the widely accepted inference from the success of *methodological* reductionism to the truth of universal *metaphysical* reductionism is not necessarily a good inference.

II: Emergence — The Probable Connection Between Quantum Mechanics and Consciousness and the Refutation of Universal Part/Whole Reductionism

Kim argues that the success of microreduction in the natural sciences is a good reason to believe in universal part/whole reductionism. The claim is that while universal part/whole reductionism flies in the face of everyday experience and common sense, it is something we must take seriously if we also take science seriously. In all honesty we must admit that science has frequently shown us that everyday experience and common sense cannot always be trusted. However, I will argue that science, including and especially microphysics, provides us with strong evidence that part/whole reductionism is false.

We can negatively define emergence as the failure of part/whole reductionism. What follows is a more positive definition of emergence or more specifically, emergent property instances are: qualitatively new properties of systems or wholes that possess causal capacities that are not reducible to *any* of the causal capacities of the parts; and such properties are potentially not even reducible to any of the relations between the parts. Emergent properties are properties of a system taken as a whole; such properties exert causal influence on the parts of the system consistent with but distinct from the causal capacities of the parts themselves. Emergence, therefore, is an explicit rejection of part/whole reductionism in both the '*nothing but*' form and the determination form.

This type of emergence is predicted by the formalism of quantum mechanics (in the form of superposition and nonseparable states) and has only recently had strong experimental confirmation (in the form of Bell-EPR experiments). I will use the term *holism* to designate such emergence within quantum mechanics (QM).

While there are several examples of holism from within QM, I will focus on EPR-Bohm systems. A number of quantum mechanical systems fit the description of EPR-Bohm systems. A simple example of such a system is as follows. Suppose that two spin $^1/_2$ particles (an electron and a positron) are produced by the decay of a single spin-zero particle at some central point, and that the two move directly outwards in opposite directions. By conservation of angular momentum, the spins of the electron and positron must add up to zero, since that was the angular momentum of the initial central particle. Therefore, when we measure the spin of the electron in some direction, whatever direction we choose, the positron always has the opposite spin value. This is true regardless of how far apart the particles are (this is true for both correlated and anti-correlated particles). It is clear that such anti-correlations are not mere coincidence, for the 'change' in the electron's spin (or what have you) will occur every time! Constant-correlation guarantees that if the two measurement events on separate parts of a EPR-Bohm system should happen to employ the *same types* of measurements, their outcomes will be opposite. The odds against the chance occurrence of such 'agreements' for millions of EPR-Bohm systems are astronomical.

There appear to be only three plausible explanations of the agreements. *Either* (i) the parts of the system *agreed in advance* on which outcomes they would yield for each kind of measurement, *or* (ii) the parts *nonlocally* influence one another so as to come to agreement when measured the same way, *or* (iii) the parts exhibit a *holistic* correlation property possessed by the system that is not locally carried by the separate parts. But agreement in advance would entail the existence of either some kind of passive-measurement-definite-values at the moment of measurement or the existence of deterministic *type*-dependent definite values. I have argued elsewhere (Hawthorne & Silberstein, 1995) that passive measurement definite values are ruled out if the theoretical predictions of QM are approximately correct. And all deterministic *type*-dependent definite values are essentially dependent on the *type* of *distant* measurement performed, which clearly must be a non-local or holistic connection, given the measurement setup. So, it seems that EPR-Bohm systems must be either nonlocally or holistically connected after all. Taking option (ii) — nonlocality — has a high price, namely, special relativity would be in serious trouble. Because the special theory of relativity weighs against the possibility of superluminal influences, this seems an unlikely possibility. There is no absolute temporal order between a pair of locally isolated (i.e. space-like separated) events; so there is no absolute sense in which one such event occurs first and then influences the development of the other. Option (iii), holism (Bell correlation properties), is the most reasonable explanation for the outcomes. If it were not for the Bell correlation properties, outcomes of measurements would be random or unconnected. No matter how one interprets quantum mechanics, if the probabilities it specifies for outcomes of measurements on EPR-Bohm systems are approximately correct, then the properties exhibited by the measurements on parts must essentially depend on the *type* of measurement performed *and* must either interact nonlocally or be holistically connected with properties of distant parts of the same system. These systems exhibit *correlation properties* that cannot be accounted for by local properties and dispositions possessed individually by their parts.

The previous argument has the following form: (1) the properties of EPR-Bohm systems can only be explained by nonlocal influences or holistic correlation properties and, (2) the existence of nonlocal influences would mean that the theory of special relativity would be in serious trouble. Therefore, since there is every reason to think that special relativity is true and we do not want to give it up, holism must be the correct explanation for the properties of EPR-Bohm systems. Let me pre-empt what I take to be likely counter-responses to both premises. The first response to premise (1) that I wish to consider is that it poses a false dichotomy. One might argue that any holistic account of EPR-Bohm systems is bound also to be nonlocal. If we get into the business of talking about irreducible properties of whole systems, then we have to countenance nonlocal properties — properties which attach to more than one point in space–time. If we refer to such a property at a given instant in time — and to tell causal stories involving these properties it seems likely we are going to have to do that — then we are appealing to an absolute standard of simultaneity between the distant parts of the system, in violation of the spirit of special relativity.

My response is that holism fits much better with special relativity then the more robustly nonlocal alternatives. I agree, of course, that all quantum correlations are 'nonlocal' in the sense that no causal influence moving at a finite speed could be transmitted from one wing of an EPR set-up to the other wing in time to account for the correlation of outcomes on measurements. So in this sense, any story about the properties of EPR-Bohm systems is going to violate the 'spirit of special relativity'. The question is, which account of EPR-Bohm systems is the least offender. My distinction for such accounts is 'nonlocally causal' versus 'nonlocally connected' in some 'noncausal way'. Of course, ultimately such a distinction depends in part on being able to make out the difference between 'causal processes or connections' on the one hand, and 'noncausal connections' on the other. For example, Bohm's account looks to be on the 'causally nonlocal' side of this distinction because on his account the wave function is first affected by a measurement on one wing, which then instantaneously changes the part of the wave function guiding the other wing, which then affects its outcome.

A more Copenhagen-like view is on the 'noncausal connection' side of the distinction. On this view, nothing on either wing has a definite property until a measurement gets made. The whole system is in a 'correlation state' until a measurement gets made. When a measurement gets made on one wing, that does introduce a new state for the other wing, but these states are not taken to 'exist in space', but merely describe the propensities of the other wing to have various outcomes when measured. No *physical causal influence* is appealed to in the theory, but only a property of the whole system (which is what I take the wave function to represent on this view).

Holism (Copenhagen or otherwise) is in general, a version of the 'noncausal connection' account. There is certainly something very different about these two accounts, where one has some effect travelling instantaneously through space to affect a particle on the other wing, and the other account has only a disposition of the whole system to correlate, with no intervening mechanism at all. On the 'nonlocally causal' account, quantum systems remain 'separable': all physical properties of objects are intrinsic (with the exception of spatio-temporal relations), and the physical state of any system is completely determined by the intrinsic properties of its constituents.

The reason that holism fits better with special relativity than the 'nonlocally causal' alternatives, is because in light of special relativity, the notion of superluminal causes is incoherent. There are many reasons why one might find superluminal causes suspect. For example, causal connections are typically understood to be mediated by some sort of physical influence (such as a causal mechanism, transference of mass-energy, etc.), but given the speed limit of light, no such physical influence is possible for superluminal causes. I personally do not believe this particular objection to superluminal causes is a good one. As far as I can tell, there are no principled reasons for ruling out the existence of 'immaterial' or non-mediated causes — superluminal or otherwise. The proper objection to superluminal causes is this: it is widely agreed that an essential feature of what it is to be the *cause* of some effect, is that causes precede their effects temporally and ontologically, causes give rise to their effects — effects arise from, come out of and are derived from their causes. Given this, the properties of EPR-Bohm systems cannot be explained by superluminal causes. On the scale of Bell experiments space-time has a Minkowski structure according to special relativity. An established feature of Minkowski space–time is that two spacelike separated events A and B have an indeterminate temporal order. Thus, in some inertial reference frames A occurs before B occurs, while in other physically equivalent reference frames B occurs before A. Given all of this, it cannot *be* the case that Bell correlations are produced by superluminal causes and *that* Bell correlations are compatible with special relativity. You can have one or the other but not both. So as I originally argued, holism is the only way that quantum mechanics and special relativity can peacefully co-exist.[1] Taken together, Bell-type arguments and their experimental confirmation attest to the existence of emergence within QM. This kind of connectedness has *appeared* to be inherent in the formalism of QM all along, reflected in system states that are superpositions of possible states of the parts. In some such quantum states the parts are locally isolated, but in many cases they are not. Prior to Bell arguments (and the various experiments that have confirmed them, see Aspect *et al.* 1982) regarding measurements on isolated parts, one might have viewed superposition states as merely an artifact of the formalism of quantum theory. The Bell results suggest that the formalism presages the existence of genuine emergent properties. Stairs makes exactly the same point when he says:

> The singlet state is irreducibly a state of the *pair* of electrons. The information it contains simply cannot be thought of as a summing up of the facts about the individual electrons. Thus, even in this extremely simple case, the facts about the whole are not determined by the facts about the parts (Stairs, 1990, p. 468).

A second objection to my original argument worth considering is aimed against both premises (1) and (2). This objection holds that there are interpretations of quantum mechanics that involve *subtle* nonlocal influences that are neither clearly causal nor holistic and which are compatible with special relativity — or at least permit an instrumentalist stance toward special relativity. It might be alleged for example, that Bohm-type hidden variable theories and GRW-type spontaneous collapse theories are both interpretations of this sort. Whether a system exhibits any emergent properties will depend on what the proper parts of the system are. For example, according to Bohm-type hidden variable theories, the complete description of a system is given by specifying a state ψ, plus a bunch of hidden variables. The hidden variables describe some local properties of the particles in the system (the positions of the particles according to Bohm's original theory). But how should we interpret the state ψ? We could interpret ψ as encoding holistic properties of the system, but Bohm interprets ψ as descriptive of a 'physical' wavefunction. On this view, the wavefunction is part of

[1] In Maudlin's excellent book *Quantum Non-Locality & Relativity* (1994), he argues that EPR-Bohm systems do entail the existence of superluminal causes. Maudlin says that the existence of EPR-Bohm systems does not entail superluminal matter–energy transport or the possibility of superluminal signalling. What EPR-Bohm systems do entail, says Maudlin, is a superluminal causal connection and superluminal information transmission. He calls such causal relations 'causal connections between events which don't support any distinction between cause and effect. The events are counterfactually connected in the right way for causation but we simply forgo any further specification of the relation' (p. 155). Maudlin concurs that the inability to send a signal between the space-like separated events rules out 'controllability causation' for EPR-Bohm systems. However, he says that this does not mean the relation between the measurement events is not a causal relation. According to Maudlin, the fact that different reference frames will disagree about which event came first should not disturb us, because there is no reason to believe that cause and effect always occur in a determinate time order with cause preceding effect. As he puts it, 'unequivocal causal connections do not require unequivocal identification of the cause and the effect' (p. 156). Given what I said above, it should be clear that almost everyone (myself included) holds that Maudlin's claim is simply doublespeak. The relation he is describing between space-like separated events, whatever else it may be, is not the *causal* relation. Maudlin is apparently driven to his radically counterintuitive account of causation because he believes that holism is too high a price for peaceful co-existence. Yet as we have seen, holism is at the very heart of quantum mechanics.

the system, just as much as each particle is, but the wavefunction is not spatially located — it is located in a multi-dimensional configuration space. The wavefunction itself has parts, the parts being regions of configuration space. If one takes a 'realist' view of the wavefunction in this way, then there are no emergent properties of the system — every property is either a property of some particle, or it is a property of some part of the wavefunction. And this includes Bell correlation properties (see Bohm and Hiley, 1993). Another example, according to GRW-type spontaneous collapse theories, ψ itself constitutes the complete description of a system, and ψ is usually taken here as describing something 'real', namely the wavefunction (see Ghirardi *et al.*, 1986, for example). If one takes this view seriously, then there are no particles, the apparent properties of particles are really local properties of the wavefunction. Once again, the wavefunction lives in a multi-dimensional configuration space, and again the proper parts of the system are regions of configuration space. Every property of the system, including Bell correlation properties, is a property of one of these parts, and there are no emergent properties. Both Bohm-type theories and GRW-type theories can explain EPR correlations without recourse to emergent properties and both allegedly solve the measurement problem. Therefore, even if holistic accounts of quantum mechanics also involve only subtle nonlocal influences, if there are subtly nonlocal alternatives without holism, then there is no reason to prefer the holistic solution.

My response to this objection is two-fold. First, the claim made by both Bohm-type theories and GRW-type theories is that the wavefunction has parts that are not spatially located, but located in configuration space. And that since it has parts as regions of configuration space, there is no place for holism here — every property is either a property of a particle or a region of configuration space. My question is, 'where in the world is configuration space?' The answer, I think, is 'nowhere in the world'. Configuration space is merely an instrumental way of keeping track of how things are likely to turn out in real space-time when measurements are made. The fact that regions of some phase-space have parts with 'phase-space properties' does not phase me unless these are parts in physical space — unless they are really real. Given the non-reality of configuration space, Bohm-type theories and GRW-type theories are in no position to avoid the dilemma between superluminal causal connections on the one hand and holism on the other. In short, such theories cannot escape holism if they wish to preserve special relativity.

This brings me to my second point. Independently of anything said here, it is well known that neither Bohm-type theories nor GRW-type theories have had any luck in coming up with Lorentz-invariant versions of quantum mechanics that can account for EPR correlations in a Lorentz-invariant manner (see Maudlin, 1994). Both the Bohm-type theories and the GRW-type theories must make use of concepts which depend on the notion of absolute simultaneity, which of course special relativity forbids for space-like separated events. Certainly these theories are not unique in this failure. Though I cannot argue the point now, I believe that holism can come to the rescue here as well. There already exists one attempt at a Lorentz-invariant version of quantum mechanics which explicitly invokes a form of holism as the solution. Gordon Fleming (1992) for example argues that the physical properties of quantum systems *depend on the choice of hyperplane*. The hyperplane is a flat, space-like, three-dimensional collection of space–time points which are all simultaneous in some reference frame. The crucial point is that hyperplanes are Lorentz-invariant surfaces. Interestingly, Maudlin argues that if the EPR-Bohm results are instances of holistic Bell correlation properties, then those properties must also be hyperplane dependent (1994, p. 212). As Maudlin puts it, 'the photon has a polarization state only as a component of a larger whole, a complete hyperplane' (1994, p. 211). Given that we already have good reason to believe in holism, the suggestion of hyperplane dependence as a way of reconciling quantum mechanics and special relativity deserves to be further developed.

Fundamental physics also provides us with other plausible claims for the existence of emergence. One such case is very surprising to many people, as it is widely regarded as the central jewel in the crown of reductionism. I have in mind the (mistaken) perception that physics has largely succeeded with the ultimate reductionist project of Grand Unification Theories (GUTs) or in producing a Theory Of Everything (TOEs). TOEs seek to combine all physical laws and principles into a single, simple formula. From this would flow a correct description of all the forces of nature, all the basic particles and fields from which the universe is composed, an account of the various masses, coupling constants and other parameters that describe how these particles interact, and a theory of space and time, including their dimensionality. Such a theory would allegedly show that there is only one fundamental force in the universe that has come to display itself as if it were four different forces; or, at least, that at some point in the very distant past the four forces were one force, but became successively distinct from one another as the universe cooled off after the Big Bang. For example, according to GUT theory (gravity is excluded in such theories), the energy scale at which the unification of all three particle forces takes place is enormously large — just below the Planck energy. Near the instant of the Big Bang, where such energies were found, the theory predicts that all three particle forces were unified by one GUT symmetry. As the universe rapidly cooled down, the original GUT symmetry was 'spontaneously' broken down successively into the present-day symmetries of the 'standard model' (see Kaku, 1993, for more details).

Let us suppose that physics succeeds in finding a GUT or a TOE; would such a discovery be inconsistent with emergence? The answer, for a number of reasons, is 'No!' Indeed, a TOE or GUT — the standard model — might actually be telling us how and when major large-scale events of emergence occurred in the history of the universe! Such theories may actually be telling us how 'the one becomes the many'. In the following passage Maudlin gives us a clue as to how this particular version of emergence might go:

> Obviously, manifest symmetry of the forces is precluded by the spontaneous symmetry breaking. Unification is to be sought in spite, rather than because, of the immediately observable properties of the forces. The mechanism of symmetry breaking allows the research program to continue in the face of the apparent dissimilarity of the forces, but it also denies us direct empirical grounds for believing that there is any hidden symmetry at all (Maudlin, 1996, p. 141).

In this passage Maudlin angles for the conclusion that we should be sceptical of GUTs and TOEs due to the absence of empirical evidence. But he also raises another possibility, namely, that spontaneous symmetry breaking explains how the universe came to be the way it is through successive acts of emergence. In short, spontaneous symmetry breaking *is* emergence. Physics has no explanation for why 'spontaneous' symmetry breaking occurs, it is simply a brute fact. And once it does occur, the emerging forces are distinct from one another and in possession of unique causal capacities.

I am not the first person to make the suggestion that the phenomenon of spontaneous symmetry breaking provides evidence for emergence. To the best of my knowledge, this claim was first made in 1972 by P.W. Anderson in his famous article 'More Is Different' (see Anderson, 1994, chapter 1). As you can see from what follows, Anderson has continued to develop this theme throughout his career:

> The role of this type of broken symmetry in the properties of inert but macroscopic material bodies is now understood, at least in principle. In this case we can see how the whole becomes not only more than but very different from the sum of its parts. Examples are legion: for one, superconductivity cannot be properly understood simply as a phenomenology without understanding electrons and their interactions. Broken Symmetry is encountered in several other contexts. One important one is in the theory of the 'big bang' during which, it is proposed, one or more broken symmetry transitions took place in the state of the vacuum, changing the nature and number of elementary particles available at each one, greatly modifying the energetics of these primeval events, and leaving behind one or more forms of debris (Anderson, 1994, p. 588).

As Anderson points out, there are several phenomena, both microscopic and macroscopic, that are best explained by our current physics in terms of spontaneous symmetry breaking. This makes it abundantly clear that not only is emergence consistent with fundamental physics such as TOEs, but fundamental physics actually provides us with an account of emergence.

It would seem that the world is much more complex and intertwined than the current cartography of scientific disciplines leads us to believe. All ontologically emergent properties/entities have unique causal capacities that constrain or supersede the behaviour of the parts in question. But some cases of emergence present us with even more radical violations of mereological supervenience or part/whole reductionism. These are cases of what Humphreys (1997) calls 'fusion' and what Healey (1997) calls 'nonseparability'. Fusion is characterized by the original or subvenient property instances going out of existence or being 'used up' in producing the emergent property or 'fused instance' (see Humphreys, 1997, p. 10). In such cases, the subvenient property instances literally no longer exist at the same time as the emergent property instance. Superposition or entangled states in quantum mechanics (such as Bell correlation properties) are paradigm cases of fusion, as are the cases discussed by Anderson, such as superconductivity and other cases of spontaneous symmetry breaking. The lesson is that at least within quantum mechanics, quantum field theory and condensed matter theory, part/whole reduction and mereological supervenience sometimes fail. Notice that fusion is a more severe violation of mereological supervenience than what Teller calls 'relational holism'. In the latter case the relational properties in question are not reducible to the intrinsic properties of the relata (see Teller, 1986). Fusion entails relational holism, but it goes even further in that the parts *cease to exist* in such cases.

One important feature exemplified by all the cases of emergence we have discussed so far is the entangled nature of parts and wholes. The kind of emergence found in quantum mechanics and quantum field theory completely explodes the ontological picture of reality as divided into a 'discrete hierarchy of levels' (a position often associated with emergence, see Kim, 1996); but rather it is 'more likely that even if the ordering on the complexity of structures ranging from those of elementary physics to those of astrophysics and neurophysiology is discrete, the interactions between such structures will be so entangled that any separation into levels will be quite arbitrary' (Humphreys, 1997, p. 15). Many have claimed that emergence entails downward causation and is therefore off limits, because downward causation would violate causal closure of the physical domain. Once we rid ourselves of the bankrupt metaphor of causally autonomous levels, we should also be able to dispense with talk about 'upward' and 'downward' causation as well. It should also be clear that given my definition of emergence and given the entangled nature of reality, the failure of causal closure is a trivial consequence. The belief in universal causal closure of the physical domain is little more than an article of faith; no one should mourn its loss given that its rejection allows for a much more unified picture of reality.

Until now there have been troubling issues about whether emergence could even make sense in a world such as ours — a world of parts and wholes — but quantum mechanics clearly gives us a model on which at least some kind of emergence makes sense. This provides us with some precedence to apply a similar view elsewhere, such as the mind. If quantum mechanics shows that emergence within physics is a coherent position, then it is also coherent to postulate the existence of other emergent properties, such as those useful for explaining consciousness and cognition. Barring physicalism or fundamentalism, the mind cries out for such an explanation.

We have seen that part/whole reductionism fails within QM. This fact does not entail that the mental is not reducible to or determined by the physical, but it does undercut the assumption of universal mereological supervenience, the primary reason for believing in physicalism in the first place. The failure of part/whole reductionism within QM also radically constrains the degree of reduction possible for the mental. The holistic emergence found within QM actually leaves open two possibilities. The first is that consciousness is a quantum mechanical property (or at least determined by quantum mechanical properties); the second is that since part/whole reductionism fails *within* quantum mechanics, it is also possible that consciousness may not be a quantum mechanical property at all, nor even be determined by such properties. The former possibility is what I will call *emergence*. The latter possibility is what I will call *radical emergence*. Both possibilities might be properly called *emergentism*, but the degree of emergence is greater in the latter case. In the former case, consciousness would be a quantum mechanical property explainable from within QM, but even in this case, consciousness would not be *nothing but* other physical properties such as charge or charm (just as Bell correlation properties are not *nothing but* the intrinsic properties of particles involved in EPR-Bohm systems). Contrary to accepted wisdom, the reduction of the mental domain to the quantum domain would *not* entail part/whole reductionism. We might say, in this case, that the quantum field has 'mental' properties that are every bit as quantum mechanical as its other properties.

In the latter case, we would have not only a failure of part/whole reductionism, but also a failure of domain or type reductionism. Consciousness would not be explicable even in principle from within QM. Consciousness would be a qualitatively new type of phenomenon from that of quantum mechanical phenomena. This does not mean that there is a separate mental *substance* that is completely detached from quantum phenomena. The quantum mechanical facts would no doubt constrain the mental. But just as Bell correlation properties have consequences for the behaviour of the individual electrons, even if such holistic properties are not determined by the properties of electrons, so the mental would have consequences for the quantum mechanical in general, even if the mental is not determined by the quantum. It is also possible that consciousness is a biological property but that biological properties are not determined by quantum properties. Holistic emergence within QM opens up the possibility of new types of phenomena whether it be biological or psychological. Radical emergence is the view that I favour, for it is consistent with our best science and it best reconciles the findings of theoretical physics with commonsense and everyday experience.[2]

It must be said that even radical emergence is compatible with monism. Indeed, I make an assumption that I call 'naturalized monism': there is one fundamental ingredient (call it a force, substance, process or what have you) from which the myriad phenomena in the universe emerges over time. Consciousness is not a fundamental or manifest ingredient of the universe, it emerges much later in the history of the universe in conjunction with brains and the like. This is also true of many so-called physical (at least macrophysical) features of the universe. Yet, consciousness is not *essentially* different in kind (or of a different type) from the other features of the universe. In short, the universe is neither *essentially* physical (non-mental) nor is it *essentially* mental (nonphysical), but it is essentially one, essentially unified. One would only need that single ingredient (however you characterize it) to create a universe such as our own.

I take naturalized monism to be a purely *meta*-physical assumption. The point is that the assumption of naturalized monism is not open to empirical confirmation or disconfirmation, because science is in no position to determine the metaphysical *essence* of the universe. As I said earlier (see p. 466 above), many people hold that the mind/body problem is 'accounting for the place of mind in a world that is essentially physical (non-mental)', and they often invoke the success of the natural sciences as their reason for making this claim (see Kim, 1996 for example). But it is absurd to

[2] As I see it, there are three possible overarching relationships between quantum mechanical phenomena and consciousness: (1) QM explains consciousness; the claim here is that one must invoke quantum phenomena in order to obtain a complete explanation of consciousness and/or cognition. For example, there are those who argue for quantum brain effects. Such accounts differ as to exactly where in the brain to locate such effects and exactly how such effects help explain consciousness and/or cognition (for detailed accounts see Lockwood, 1989; Stapp, 1993; Jibu & Yasue, 1995, and Penrose, 1996). Some also argue that the deeply non-classical nature of matter as characterized by quantum mechanics helps explain consciousness because it makes it conceivable that 'mere matter' could give rise to consciousness (see for example Lockwood, 1989, and Hodgson, 1991). (2) Consciousness explains QM; the general idea here is that the conscious observer (in one form or another, such as God or the human observer) is an ineliminable and essential ingredient in the evolution of the physical world. The supposition is that consciousness as an independent phenomena is needed to solve the measurement problem (proponents of this view include Von Neumann, 1955; Wigner, 1983, and Squires, 1994). (3) QM and consciousness are each needed to fully explain the other; this view is basically a modified conjunction of the first two. For example, Eccles (1994) argues that while consciousness exists independently of the physical world, there are quantum effects in the brain that help explain how consciousness can causally interact with the brain. It should be clear that the relationship I am positing between quantum phenomena and consciousness, namely radical emergence, is different from all three of the preceding positions. Although, of course, I do agree that the non-classical nature of quantum phenomena makes it much more plausible that consciousness emerges from 'mere matter' than if a more Cartesian view of matter were true.

think that science can tell us anything about the *essence* of the universe; a consistent naturalist of the Quinean variety will immediately concede this. Science is completely neutral with respect to transcendental claims in general. Furthermore, starting out with the purely *a priori* and unmotivated assumption that the universe is essentially physical (non-mental) makes the mind/body problem impossible to resolve in any satisfactory way — hence the problems with physicalism and fundamentalism.

There is a widely held disjunction that QM is either completely (in principle) irrelevant for explaining consciousness *or* consciousness is completely explained by QM. Churchland (Grush & Churchland,1995), Crick (1994) and many others argue that quantum mechanics is not needed to explain, nor is it able to explain consciousness. For consciousness is a neurochemical phenomena, and therefore the appropriate explanation will be found at the neurochemical or neurocomputational level. Chalmers expresses this sentiment perfectly when he accuses those who advocate quantum mechanics for explaining consciousness of engaging in the faulty heuristic that 'consciousness is mysterious and quantum mechanics is mysterious, so maybe the two mysteries have a common source' (1995, p. 207). On the other hand, Stapp (1993), Penrose (1994) and others argue that neuroscience *à la* classical physics can never explain consciousness, because in classical physics determinism and part/whole reductionism are inescapable, and no system of this sort could possibly yield consciousness or cognition. It is only by reconceptualizing matter *à la* quantum mechanics that a naturalistic explanation of consciousness is possible. Who is right? Based on what I have said about emergence, it seems that both sides are to some extent probably correct. The neurocomputational side is right that we should seek an explanation of consciousness at some level other than the quantum mechanical because consciousness is probably not a quantum phenomenon; and the quantum side is dead right in claiming that matter conceived *à la* classical physics cannot possibly yield a naturalistic explanation of consciousness.

It seems that both sides of the debate are engaging in a bit of bad faith. The neurocomputational side is primarily motivated by old fashioned reductionism (both metaphysical and methodological); yet they balk at quantum mechanical explanations of consciousness which is where their reductionist credo logically demands that they go. Hard-core reductionists like Churchland and Crick have no principled reason for stopping the buck at neuroscience; yet they treat appeals to quantum mechanics in explaining consciousness as if they were inherently absurd. It is for this reason that I am always amazed at the hostility that neurocomputationalists and other special science people have toward emergence. After all, if emergence is true and universal part/whole reductionism false, then the special sciences might actually be more than epistemic placeholders for QM. For example, functionalism might be more than just a methodological approach to studying the mind, more than just a higher-level description of physical processes as Clark indicates in the following passage:

> Ultimately, I suppose there is only physics. But our scientific world-view posits a set of overlaid structures, the chemical, the social, even the computational. Reality functionally described is not a legitimate level of scientific ontology, but only reflects our explanatory interests' (Clark, 1990b, p. 601).

If on the other hand, the brand of reductionism that so many middling reductionists such as the Churchlands put their faith in is true, then consciousness must ultimately be best explained by QM — at least in principle.

However, even if neurocomputationalists do not want to accept emergence, it seems to me that, being good naturalists, they are duty bound to do so. This is so, because naturalists believe our best science should be the final authority when it comes to metaphysics and our best science tells us that emergence is true. This puts neurocomputational reductionists in an awkward position; where once microphysics was their friend, it is now their enemy.[3] If everything reduces to the domain of microphysics then emergent properties are ubiquitous and neurocomputationalism cannot justify its scepticism about QM explanations of consciousness. If everything does not reduce to the domain of physics then why believe the mental reduces to the neurochemical in the first place? The neurocomputational reductionist might try to get around all this by claiming that quantum effects are screened off at the macro-level and this includes such effects as quantum holism. The first problem with this response is that QM provides the only explanations for certain macroscopic phenomena, such as superconductivity, ferromagnetism, crystals, lasers and solid state TVs, just to name a few. These are all cases of quantum effects manifested at the macroscopic level. The second problem is this, if despite our best evidence, quantum effects such as holism really are screened off then this fact itself is either explained by QM or it is not. If it is explained by QM then the quantum reductionist wins and neurocomputationalism is merely an epistemic placeholder for QM. If on the other hand, QM cannot explain why quantum effects are screened off, then it looks as if some kind of emergence is going to be the best explanation for this fact. In fact, even if quantum effects are screened off, given the failure of part/whole reductionism within physics, there maybe *macroscopic* emergent properties that are not explained by QM. And again, in such cases why believe the mental is reducible to the neurochemical? It seems to me that the middling reductionist, who wants to stop the reductionist buck at say the neurochemical level, is in the worst position of all.

Interestingly, the quantum accounts of consciousness are also largely motivated by reductionism.They seek that ultimate reduction, that grand unification that has been the holy grail of theoretical physics for some time now. The quantum side often appeals to this grand reductionist vision when justifying their modus operandi with respect to consciousness. All quantum reductionists (such as Penrose) are willing to concede that within quantum mechanics wholes have causal properties not reducible to the causal properties of their parts (emergence). However, the quantum reductionist wants to maintain the truth of domain or type reduction (the denial of radical emergence), claiming that consciousness is ultimately explicable from within quantum mechanics. Here is the rub: given what quantum mechanics has discovered about matter, the quantum folk should not be so quick to stick to their reductionist guns. The

[3] Many people have claimed that the success of fundamental physics has been the primary bane of emergence (see for example McLaughlin, 1992, p. 89). There is certainly some *historical* truth to this claim, for example, QM made it possible to explain various chemical phenomena such as molecular bonding, that had never been explained before. However, the fact is that QM often explains various types of bonding relations by invoking such quantum effects as superposition states or more specifically, quantum tunneling. If chemistry reduces to QM, then holism is everywhere. Has chemistry been successfully reduced to QM? The answer is ‘no’, as Cartwright puts it: ‘Notoriously, we have nothing like a real reduction of the relevant bits of physical chemistry to physics — whether quantum or classical. Quantum mechanics is important for explaining aspects of chemical phenomena but always quantum concepts are used alongside of *sui generis* — that is, unreduced — concepts from other fields. They don’t explain the phenomena on their own’ (1997, p. 163).

quantum reductionists are quick to appeal to the failure of part/whole reductionism *within* quantum mechanics when they are arguing their claim that consciousness is a quantum property. What the quantum reductionists do not see is that this very failure of reductionism from within quantum mechanics may undercut the only good empirical reason they ever had for believing in the grand unification in the first place.

I am in complete agreement with those who claim that matter *à la* classical physics cannot possibly give rise to nor have as one of its properties consciousness. As Stapp puts it; 'nothing in classical physics is more than the sum of its parts, no consciousness in, no consciousness out. Nothing in classical physics can create something that is essentially more than an aggregation of its parts. Classical physics entails reductive materialism' (1993, p. 154). Stapp's point is that from the perspective of classical physics, part/whole reductionism is an inescapable feature of the physical world. If that is the case, then there can be no naturalistic explanation of consciousness forthcoming from classical physics. The inability of the Churchlands, Crick, Dennett and the other neurocomputationalists to see the problems with classical physics and consciousness is a serious blind spot.

Penrose (Penrose & Hameroff, 1995), Stapp (1993) and others claim that quantum brain effects can help explain the existence of consciousness where classical physics fails. However even *if* there are large scale quantum brain effects, their explanatory value would be limited. Quantum reductionism is no better off than functionalism in being able to explain consciousness — even if quantum reductionism were true, it could not solve the 'hard problem' of consciousness. Penrose admits as much in the following passage: 'Although I had not explicitly asserted, in either *Emperor* or *Shadows,* the need for mentality to be "ontologically fundamental in the universe", I think that something of this nature is indeed necessary' (Penrose *et al.*, 1997, p. 175–6). So Penrose's theory about consciousness comes to this: given large scale quantum brain effects *and* given *consciousness as a fundamental ingredient of the universe*, he can explain consciousness. Somehow, I find this explanation less than satisfying — perhaps because it is no explanation at all. For Penrose and others like him, there is only the entities of (a finished) microphysics and consciousness. Everything else is reducible to one of these. What drives him to such a radically neutered and counter-intuitive picture of the world? As the following passage suggests, the answer, I think, is his inability to see how emergence could be possible; 'If someone can similarly give me an idea of what a "biology" that does not supervene on its corresponding "physics" could be like, then I might begin to take such an idea seriously' (*ibid.*, p. 182). And this is exactly the burden I have tried to meet in this paper. Most of what we know about neuroscience and evolution suggests that it is the unique macroscopic properties of the brain that are essential for consciousness. Where the quantum reductionist goes wrong is in the assumption that since classical physics cannot explain consciousness, the only other viable option is quantum mechanics. But again, quantum mechanics itself actually provides us with a third alternative. What is so crucial about quantum holism/emergence is that it undercuts the knee-jerk reductionism that has plagued science (and especially physics) for so long. In short, it opens up the possibility of a totally non-reductive explanation of consciousness and cognition.

Quantum reductionists who wish to reject radical emergence might respond that since *we* are made up of large and complex arrangements of quantum systems, then it must be the case that consciousness is explainable from within quantum mechanics.

Such an argument would of course be fallacious, since wholes often possess (or lack) properties that their parts do not. We have learned that this is especially true in quantum mechanics in a very nontrivial way. However there is a better argument for quantum reductionism. Without quantum mechanics we could not explain the behavior of solids such as the growth of crystals, the action of lasers or the properties of superfluids, the colour of the stars, the structure and function of DNA, the functioning of solid state TV sets, and more. Given the success of quantum mechanics in particular and the success of methodological reductionism in general, it is not unreasonable to believe that quantum mechanics can also explain consciousness. 'Fair enough', it is an empirical question and must be settled as such. It must be said though that while quantum mechanics has been very successful at explaining certain 'macroscopic' effects, overall the relationship between the microrealm of quantum mechanics and the everyday macroworld is a complete mystery. To put a finer point on it, 'the most serious deficiency of pioneer quantum mechanics is its inability to deal in a candid way with classical systems'(Primas, 1981, p. 16). Quantum mechanics is currently in no position to explain the relatively classical nature of the macroscopic physical universe, let alone explain consciousness.

Conclusion

Emergence, far from being mysterious or mystical, is highly confirmed by physics and has a high degree of explanatory value across the board. It would be strange indeed if the universe only exhibits emergence at the 'level' of quantum phenomena and consciousness respectively, while the rest of the universe remains pretty much as conceptualized by classical physics. It is more likely that emergence is ubiquitous and that the universe is far more intertwined and complex than the standard division of the sciences would lead us to believe. It is for this reason that I believe a complete explanation of consciousness in all its multifarious aspects, will require the kind of interdisciplinary effort we are seeing now. Given emergence, 'reflective equilibrium' (see Flanagan, 1992) is not merely a methodological consideration, it's an explanatory necessity.[4]

Emergence is often accused of promoting mystery as explanation. Kim characterizes emergence as holding that there are brute laws that necessitate the emergence of the mental under certain conditions. Beyond that, says Kim, 'why pain emerges when c-fibers are excited will forever remain a mystery; we have no choice but to accept it as an unexplainable brute fact' (Kim, 1996, p. 229). This makes it sound as if emer-

[4] It must be said again, that radical emergence offers hope to the special sciences that they are more than 'epistemic placeholders' for quantum mechanics (fundamental physics). In a world with emergent properties, in a stochastic world with probabilistic causation (see Hawthorne & Silberstein, 1995; Humphreys, 1989, for details on this) to ask what was *the* cause of a particular phenomenon will be to ask an unanswerable question in general. In such a world there will rarely be a unique cause for any given effect, but rather, the question should be which of the causal factors was most important. In such a world, there is no robust distinction between real causes and background conditions to be had. The complete true causal explanation for most events will involve multiple causal factors and therefore will be too complex to produce and too unwieldy to use.

There will often be more than one true (partial) causal explanation of the same event, and this is exactly what the various special sciences (and the natural sciences) often provide. In such a world, there is just no reason to believe that the special sciences only exist as a function of ignorance or that the natural sciences must have some universal explanatory priority.

gence is explanatorily empty — nothing but proud and bold hand-waving. This is no doubt an accurate criticism of NRP, but not of emergence. Emergence provides the possibility of giving a *causal* explanation for consciousness, and this is as much explanation as we can demand from science for any phenomena. This is because emergent processes/relations are diachronic and not synchronic. This means that emergent processes/relations unlike determination relations, might be a subset of the causation relation or vice-versa. Notice that once science produces what we take to be the *ultimate* causal explanation for a particular phenomena, there is no point in asking, 'but *why* is that particular effect produced by that particular cause?' All scientific explanations must eventually run into brute facts; what emergence teaches us is that brute facts arise much more quickly in the game than we had thought. That should come as no surprise, unless you already had your heart set on quantum field theory or string theory explaining consciousness.

Far from promoting mystery, it is emergence alone that can provide an *explanation* for consciousness. Emergence is the only explanatory framework for consciousness that allows us to avoid the absurdity of eliminativism that reductive physicalism brings with it, without being forced into fundamentalism. Given the truth of emergence, the so-called philosophical or conceptual mind/body problem (the 'hard problem' of consciousness) has been solved. We have learned that Chalmers is just wrong when he says: 'if it turns out that it [consciousness] cannot be explained in terms of more basic entities, then it must be taken as irreducible, just as happens with such categories as space and time' (Chalmers, 1997, p. 8). The existence of emergent properties tells us that Chalmers is operating with a potentially false dichotomy, because emergent properties are neither reducible nor fundamental! Given emergence, there is only the *empirical mind/body problem(s)* left to be resolved, and that is primarily a job for science not philosophy. Or perhaps I should say a philosophically sophisticated science. In my view, philosophy can best help the new multidisciplinary Science of Mind by learning its results and theories, and then use that data as the primary basis upon which to draw its metaphysical conclusions about mind and matter. In the 21st century, scientifically ignorant practioners of philosophy of mind have about as much chance of making a real contribution, as 21st century natural philosophers who know nothing of physics. If emergence is the proper account of consciousness, then once we explain *how* consciousness came into being, *how* it relates to the brain and what role, if any, it plays in cognition, then we will have also explained *why* consciousness exists.[5]

---

[5] As Lycan points out: even *within* the realm of the hard problem of consciousness, it may very well be necessary to divide the question when it comes to seeking scientific explanations (see 1996, chp. 1). For example, Lycan makes a distinction between 'qualia', 'introspective consciousness' and 'self-consciousness', as well as several types of *conscious* phenomena (1996, pp. 2–7). Given the current state of consciousness studies, there is no reason to think that science will produce a 'unified theory of consciousness' that will single-handedly answer all the hard sub-problems of consciousness. However, even if the hard problem must be divided into several distinct hard problems, that would not deflate the hard problem of consciousness, but only multiply it.

References

Anderson, P.W. (1994), *A Career in Theoretical Physics* (World Scientific Publishing).

Aspect, A., Grangier, P. and Roger, G. (1982), 'Experimental realization of Einstein-Podolsky-Rosen-Bohm *Gedankenexperiment*: A new violation of Bell's inequalities', *Physical Review Letters*, **49**, pp. 91–4.

Boghossian, H. (1994), 'Fodor's naturalism', in *Meaning in Mind: Fodor and his Critics*, ed. B. Loewer and G. Rey (Oxford: Blackwell).

Bohm, D and Hiley, B.J.(1993), *The Undivided Universe* (London: Routledge).

Cartwright, N. (1997), 'Why physics?', in Penrose *et al.* (1997).

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

Chalmers, D.J. (1996), *The Conscious Mind* (New York: Oxford University Press).

Chalmers, D.J. (1997), 'Moving forward on the problem of consciousness', *JCS*, **4** (1), pp. 3–46.

Clark, A. (1990a), *Microcognition* (Cambridge, MA: MIT Press).

Clark, A. (1990b), 'Aspects and algorithms', *Behavioral and Brain Sciences*, **13** (4), pp. 601–2.

Crick, F. (1994), 'The astonishing hypothesis',in discussion with Jane Clark, *Journal of Consciousness Studies*, **1** (1), pp. 10–16.

Eccles, J.C. (1994), *How the Self Controls its Brain* (Springer-Verlag).

Flanagan, O. (1992), *Consciousness Reconsidered* (Cambridge, MA: MIT Press).

Fleming, G. (1992), 'The objectivity and invariance of quantum predictions', in *PSA 1992*, ed. D. Hull, M. Forbes and K. Okruhlik, Vol. 1, pp. 104–13.

Ghirardi, G.C., Rimini, A. and Weber, T. (1986), 'Unified dynamics for micro and macro systems', *Phys. Rev.*, D34, pp. 470–91.

Grush, R. and Churchland, P.S. (1995), 'Gaps in Penrose's toiling', *JCS*, **2** (1), pp. 10–29.

Hawthorne, J. and Silberstein, M. (1995), 'For whom the Bell arguments toll', *Synthese*, **102**, pp. 99–138.

Healey, R.(1997), 'Nonlocality and the Aharonov-Bohm Effect', *Philosophy of Science*, **64**, pp. 18–41.

Hodgson, D.H. (1991), *The Mind Matters* (Oxford: Oxford University Press).

Humphreys, P. (1989), *The Chances of Explanation* (Princeton University Press).

Humphreys, P. (1997), 'How properties emerge', *Philosophy of Science*, **64**, pp. 1–17.

Jibu, M. and Kunio, Y. (1995), *Quantum Brain Dynamics and Consciousness* (Amsterdam: John Benjamins Publishing Co.).

Kaku, M. (1993), *Quantum Field Theory: A Modern Introduction* (Oxford: Oxford University Press).

Kim, J. (1978), 'Supervenience and nomological incommensurables', *American Philosophical Quarterly*, **15**, pp. 149–56.

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

Kim, J. (1996), *Philosophy of Mind* (Westview Press).

Lockwood, M. (1989), *Mind, Brain and the Quantum* (Oxford: Blackwell).

Lycan, W. (1996), *Consciousness and Experience* (Cambridge, MA: MIT Press).

McLaughlin, B. (1992) in *Emergence or Reduction?: Essays on the Prospects of Nonreductive Physicalism*, ed. A. Beckermann *et al.* (New York: Walter De Gruyter).

Maudlin, T. (1994), *Quantum Non-Locality And Relativity* (Oxford: Blackwell).

Maudlin, T. (1996), 'On the unification of physics', *The Journal of Philosophy*, 9303, pp. 129–44.

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

Penrose, R. and Hameroff, S. (1995), 'What "Gaps"?', *Journal of Consciousness Studies*, **2** (2), pp. 98-–111.

Penrose, R. with Shimony, A., Cartwright, N. and Hawking, S. (1997), *The Large, the Small and the Human Mind* (Cambridge: Cambridge University Press).

Primas, H. (1981), *The Philosophical Issues of Theoretical Chemistry* (Springer-Verlag).

Scharf, A. (1989), 'Quantum measurement and the program for unity of science', *Philosophy of Science*, **60** (1), pp. 601–23.

Squires, E. (1994), 'Quantum theory and the need for consciousness', *Journal of Consciousness Studies*, **1** (2), pp. 201–4.

Stairs, A. (1990), 'Quantum mechanics, mind and self', in *Essays on Personal Identity*, ed. Charles Taylor (Harvard University Press).

Stapp, H.P. (1993), *Mind, Matter and Quantum Mechanics* (Springer-Verlag).

Teller, P. (1986), 'Relational holism and quantum mechanics', *British Journal for the Philosophy of Science*, **37**, pp. 71–81.

Von Neumann, J. (1955), *Mathematical Foundations of Quantum Mechanics* (Princeton University Press).

Wigner, E.P. (1983), 'Remarks on the mind–body question', in *Quantum Theory and Measurement*, ed. Wheeler and Zurek (Princeton University Press).