

How Not To Be A Reductivist*

William Hasker

Some current positions in the philosophy of mind, while ostensibly non-reductive, are in fact reductivist in ways that are seriously problematic. An example is found in the “naturalistic dualism” of David Chalmers: by maintaining the causal closure of the physical domain, Chalmers makes the rationality of conscious experience inexplicable. This can only be remedied by abandoning causal closure and acknowledging that micro processes in the brain go differently in the presence of conscious experience than they would without it. But this move has startling consequences: once it has been made, major objections to mind-body dualism disappear, and determinism is seen to be a theory that is completely lacking in empirical support. Thomas Nagel and John Searle are cited as examples of philosophers who make a serious effort to face up to the consequences of not being reductivists.

Introduction

No sensible person should want to avoid reductionism altogether. Many of the greatest successes of the sciences have come through reductive explanations, and it would be quixotic (at best) to deprive oneself of all those explanations. The kind of reductionism to be avoided is the kind that gives reductive explanations which deny or ignore the reality and importance of that which is in fact real and important. One example would be the reduction of mental states and actions given by logical behaviorism; it is now widely recognized that such explanations, by ignoring phenomenal experience, fail to do justice to the reality of one’s mental life. Another example, perhaps more controversial in some circles, is the equating of moral goodness with fitness defined in evolutionary terms. To many of us it seems evident that moral goodness possesses a significance quite different from that which attaches to evolutionary success, and if that is so a reduction of the former to the latter must fundamentally distort the nature of morality. The way not to be a reductivist in the philosophy of mind is to avoid giving this sort of explanation of mental facts – one that denies either their existence or the significance we rightly take them to possess.

The main thesis of this paper is that not being a reductivist in this sense is harder than is often supposed. “Non-reductive” accounts of mentality are relatively common these days; it is widely recognized that certain kinds of reductivism should be avoided. However, on closer examination it often turns out that these accounts are reductive after all in a way that undermines the significance of mentality – and, it will be argued, results in a position that is in the end simply untenable. Not being a reductivist requires going further in the other direction than a great many philosophers are willing to go.

The notion that avoiding reductionism is difficult is not a novel one. Jaegwon Kim, at least since the time of his American Philosophical Association presidential address, “The Myth of Nonreductive Materialism” (1991a), has repeatedly emphasized the difficulty of giving a nonreductive

* This article will be appearing in the forthcoming book, entitled *Mind: Its Place in the World: Non-reductionist approaches to the Ontology of Consciousness*, ed. Alexander Batthyany, Dimitri Constant, and Avshalom Elitzur.

account of the mind. However, the direction in which Kim proceeds from this recognition is quite different from the one taken here. His response is to look for room to maneuver in the narrow space left by physicalist commitments he is unwilling to abandon, even while he admits that the remaining options are intuitively somewhat unattractive. My argument will be that these options are not merely unattractive but indefensible, and that some of the core commitments of physicalism must be surrendered.

A Case Study: Chalmers' *Conscious Mind*

An interesting example of a “non-reductive” view that turns out to be reductive after all is found in David Chalmers’ book, *The Conscious Mind: In Search of a Fundamental Theory* (1996). Chalmers has become known for his distinction between the “hard problem” and the “easy problem” of consciousness. The easy problem (which is “easy” only in contrast to the hard problem!) deals with such questions as “How can a human subject discriminate sensory stimuli and react to them appropriately? How does the brain integrate information from many different sources and use this information to control behaviour? How is it that subjects can verbalize their internal states?” (2002, p. 92) The hard problem, on the other hand, is “the question of how physical processes in the brain give rise to subjective experience” (p. 92). This distinction is correlated with another distinction Chalmers makes, between “phenomenal” and “psychological” aspects of mind. The “phenomenal” concept of mind characterizes mind by the way it *feels*; it is “the concept of mind as conscious experience, and of a mental state as a consciously experienced mental state” (1996, p.11). The psychological¹ concept of mind, on the other hand, characterizes mind by what it *does*; this is “the concept of mind as the causal or explanatory basis for behavior,” and “A state is mental in this sense if it plays an appropriate role in the explanation of behavior.” He adds, “According to the psychological concept, it matters little whether a mental state has a conscious quality or not. What matters is the role it plays in a cognitive economy” (p. 11). Many of our ordinary concepts concerning the mind, Chalmers acknowledges, have both a phenomenal and a psychological aspect. However, the distinction is significant in that the psychological aspect of mind, Chalmers thinks, is well on the way to being understood through the standard scientific means of neurology, cognitive science, and so on. In contrast, he argues that the phenomenal aspect of mind not only *has not* been explained, but in principle *cannot* be explained by the existing methods of the sciences: this is the “hard problem.” Phenomenal consciousness is *irreducible* in that its existence is unexplainable, not just by our present physical theories, but by any successor theory that is sufficiently similar to them to merit the name, “physical.”² Conscious experience fails to be *logically supervenient* on the physical state of the organism or the world. Borrowing an image from Kripke, Chalmers states, “When God

¹Chalmers acknowledges an element of stipulation here; this is not necessarily the standard use of ‘psychological’.

²Chalmers does not provide an explicit list of requirements for physical properties; he characterizes them as “the fundamental properties that are invoked by a completed theory of physics. Perhaps these will include mass, charge, spatio-temporal position; properties characterizing the distribution of various spatio-temporal fields, the exertion of various forces, and the form of various waves, and so on” (1996, p. 33).

created the world, after ensuring that the physical facts held, *he had more work to do*. He had to ensure that the facts about consciousness held” (p. 124).³ Given the insufficiency of physics to explain consciousness, doing this will require the new “fundamental theory” of the book’s subtitle. Such a theory will include new fundamental properties, and also new fundamental laws; the latter “will be *psychophysical* laws, specifying how phenomenal (or protophenomenal) properties depend on physical properties” (p. 124). Chalmers is inclined to call his view dualism rather than a version of materialism, albeit a “naturalistic” dualism and one which involves a duality of properties, not of substances. “One might say: You can’t have your materialist cake and eat your consciousness too” (p. 168).

All this goes to show that Chalmers does present a theory of the mind that is seriously nonreductive. As one might expect, he has taken considerable flak for his deviations from materialist orthodoxy. Nevertheless, our focus here will be on the sense in which he remains, even after all this, a reductivist. In order to bring this out we need to keep in mind Chalmers’ commitment to one of the central theses of materialism or physicalism: *the causal closure of the physical domain*. He writes,

The best evidence of contemporary science tells us that the physical world is more or less causally closed: for every physical event, there is a physical sufficient cause. . . . A small loophole may be opened by the existence of quantum indeterminacy, but . . . this probably cannot be exploited to yield a causal role for a nonphysical mind. . . . [I]t remains plausible that *physical* events can be explained in physical terms (p. 125).

As he also says, on his view “the physical domain remains autonomous,” and “the view makes experience *explanatorily irrelevant*” (pp. 126, 156).

Now, let us pose the question: What precisely is it that can’t be explained by physics? As we have seen, Chalmers’ answer is, “how physical processes in the brain give rise to subjective experience.” This answer is correct so far as it goes, but it is seriously incomplete. A more complete answer would be, “how physical processes in the brain give rise to subjective experience *in such a way that subjective experience reliably corresponds to the way things really are in the world.*” This, of course, is a special form of the problem of accounting for the *rationality* of the functioning of the conscious mind. Our minds aren’t anything like perfectly rational, to be sure, but they are rational to a significant degree, and it’s to this rationality that we owe whatever ability we have to grasp the truth about things – including, of course, the truth about the mind itself. A theory of mind that can’t account for this rationality has little to be said for it.

So how will naturalistic dualism account for rationality? Chalmers himself doesn’t address this question, but the answer is not far to seek, and the key to the puzzle is provided by Darwinian evolution. The central idea of “Darwinist epistemology” is simply that an organism’s conscious states confer a benefit in the struggle to survive and reproduce. Such responses as discomfort in the presence of a chemical irritant, or the awareness of light or warmth or food, enhance the

³Chalmers is not religious and the reference to God is merely part of a heuristically useful “creation myth.”

organism's ability to respond in optimal fashion. For more complex animals there is the awareness of the presence of predator or of prey, and the ability to devise simple strategies so as to increase the chances of successful predation or of escape therefrom. As the organisms and their brains become more complex, we see the emergence of systems of beliefs and of strategies for acquiring beliefs, and the strategies that lead to the acquisition of true rather than false beliefs confer an adaptive advantage. Natural selection guarantees a high level of fitness, including cognitive fitness.⁴

It seems clear that this is the *only* way of accounting for rationality that is available to naturalistic dualism and similar accounts of the mind. It is out of the question to suppose any sort of pre-planning or design in the original state of things; that sort of explanation leads to immediate expulsion from the naturalistic camp. So there must be some sort of "filtering mechanism" by which the conscious states of organisms are sorted into those which are rationally substandard and those which to some degree approximate the rational ideal. And the only such mechanism that anyone has proposed is Darwinian evolution with its survival of the fittest; we have to place our faith in the "blind watchmaker."

Now let's compare this survey of Darwinist epistemology with a concise summary of evolutionary theory, as follows: Certain complex assemblages of organic chemicals develop a kind of dynamic stability in their interactions with the environment, together with a capacity for self-replication, which leads us to say they are alive. A variety of random physical forces leads to variations in the self-replicating assemblages, and some of the assemblages are more successful than others in maintaining and reproducing themselves. Over time, some of these assemblages become more complex than the earliest forms by many orders of magnitude, and their behaviors and interactions with the surrounding environment also become more complex. Nevertheless, the entire process is governed by, and explicable in terms of, the ordinary laws of physics and chemistry. Put differently, it is never necessary to go outside of the physical configurations and the physical laws in order to predict the future behavior of these assemblages; this is the "causal closure of the physical domain" (Hasker, 1997, pp. 40-41).

The incongruity between these accounts is not difficult to spot; in fact, Chalmers himself has stated it very clearly: "The process of natural selection cannot distinguish between me and my zombie twin" (p. 120). The problem is not merely that the evolutionary theory doesn't *mention* the adaptive benefits of awareness and cognition. If that were the only problem, it could easily be maintained that awareness and cognition are among the necessary preconditions for the more successful behaviors and interactions with the environment that are featured in the account. (That, in fact, is exactly what evolutionary epistemology affirms to be the case.) The problem, rather, is that the account of evolution *precludes* the kind of role for awareness and cognition that is posited in the epistemological account. It does this by its last two sentences, which affirm the causal closure of the physical domain. Those sentences guarantee that the conscious state of the organism, as such, can have *no influence whatever* on the organism's behavior and thus on its propensity to survive. Conscious experience is *invisible* to the forces of natural selection, and

⁴See Stephen Stich (1990), pp. 38-39, for a summary of the argument that Darwinian evolution guarantees rationality. Stich is a sceptic about this argument, but he provides this formulation as a basis for discussion.

the central contention of evolutionary epistemology has been decisively undermined.

It's important to realize how stark are the choices imposed by this argument. Some might suppose that we can retreat to understanding rationality in functional terms, and ignore the issue of the rationality of conscious experience. But this is a non-starter, for the simple reason that our *only* evidence for the truth of anything we believe derives ultimately from just that conscious experience. Another response (which is at least suggested by Chalmers' account⁵) is to assert that it is simply a fact that the psychophysical laws correlate conscious experiences with physical states in a way that is appropriate and generally truth-conducive, and leave the matter there. But where do these marvelously appropriate correlations come from? A philosopher can't simply present them to himself as a birthday present! And as we've seen, any sort of pre-arrangement or pre-established harmony in these laws places one immediately outside the naturalist camp.

We have to affirm the rationality of conscious experience, and we aren't allowed simply to assume such rationality without offering any explanation. It seems fairly clear what needs to be done here. We have to affirm, contrary to Chalmers and most contemporary philosophers of mind, that *conscious experience is explanatorily relevant*. If this is so, then consciousness is no longer invisible to evolutionary selection; conscious experiences that correspond appropriately to the external environment can indeed be selected for, and brain structure and consciousness can co-evolve in response to environmental pressures. This may or may not be *sufficient* to enable an explanation of human rationality (I suspect it is not), but it is certainly *necessary* for any such explanation.

This apparently simple and obvious step, however, has profound consequences. To begin with, it breaches the causal closure of the physical domain, which is axiomatic for Chalmers as for most contemporary physicalists. Processes in the brain will go differently, in the presence of conscious experience, than they would if governed merely by the ordinary laws of physics and chemistry. We will have a species of *emergent causation*, in which there are causal principles at work in situations involving consciousness which are different from those operative in simpler physical situations. The universal sway of the fundamental physical laws no longer obtains.

It should not be supposed that this result can be avoided by such standard devices as "supervenient causation," or even by mind-body identity theory. These views may salvage a kind of causal role for mental states, either as supervenient upon physical brain-states or as identical with those states. But on these views, *the subjective content of mental states is irrelevant to their causal efficacy*.⁶ Once again, conscious experience becomes invisible to evolutionary selection, and there is no explanation for the rationality of such experience.

⁵According to Chalmers, on his view the evolution of consciousness is "not a problem. Like the fundamental laws of physics, psychophysical laws are eternal, having existed since the beginning of time. . . . [A]s the universe developed, it came about that certain physical systems evolved that satisfied the relevant conditions" (p. 171). However, he doesn't consider the problem of explaining human rationality, so the remarks quoted shouldn't properly be considered an answer to that problem.

⁶Kim makes this point with regard to Davidson, in Kim, 1991b, p. 106. However, the point applies equally to Kim's own supervenience view.

Chalmers briefly considers the notion of emergent causation, and it is instructive to consider his reasons for rejecting it. He asks us to “note that nothing in the story about emergent causation requires us to invoke *phenomenal* properties anywhere. The entire causal story can be told in terms of links between configurations of physical properties. There will still be a possible world that is physically identical but that lacks consciousness entirely. It follows that at best phenomenal properties *correlate* with causally efficacious configurations” (1996, p. 379 n. 41). His point is this: the phenomenal properties will be (on his view) strictly correlated with a particular physical configuration – so why can’t the emergent causal powers be attributed to the physical configuration alone, with the result that the causal powers are purely physical and phenomenal properties remain explanatorily irrelevant as before?

The right answer to this begins by pointing out a fundamental characteristic of the physical laws we are familiar with: all such laws are *mechanistic* in that they are *non-intentional* and *non-teleological*. To put the point graphically, the stones in an avalanche do not go where they go because it would be a good idea for them to go there.⁷ Conscious experience, on the other hand, is riddled through and through with teleology. It’s just a fact that we often decide to do things because we think the consequences of doing them will accord with our desires. So if we attribute the emergent causality, which in many instances is clearly teleological, to the physical configuration alone, we are attributing to it behavior quite different from that produced by all the other physical configurations we know about. At the same time, we are ignoring the teleological character of conscious experience, which seems on the face of it to provide by far the most plausible explanation for goal-directed behavior. Chalmers’ claim amounts to the stipulation that however we act in order to achieve our goals, *conscious experience is never in any degree the explanation for our so acting*. But why should we accept this?

Chalmers’ other reason is that “there is no evidence for such emergent principles of causation. As far as we can tell, all causation is a consequence of low-level physical causation, and ‘downward causation’ never interferes with low-level affairs” (p. 378). This bears looking at closely. Is Chalmers meaning to say that we have sufficient knowledge of the causal processes in the brain that we can rule out those processes ever going any differently than they would if governed only by the standard laws of physics? Presumably not – at least, a claim of that sort would amount to little more than an enormous bluff. But, he may respond, we have no evidence that this is *not* the case. I reply that we do indeed have such evidence, as shown by the following argument:

1. Human beings are rational to a significant (though highly imperfect) degree.
2. If human beings are rational, there is an explanation for the fact that human beings are rational.
3. There is an explanation for the fact that human beings are rational.
4. If conscious experience is explanatorily irrelevant, there is no explanation for the fact that human beings are rational. (Argued for above.)
5. Conscious experience is explanatorily relevant.
6. If the physical realm is causally closed, conscious experience is explanatorily irrelevant.

⁷I owe this example to Victor Reppert.

(Also argued for above, and conceded by Chalmers.)

7. The physical realm is not causally closed.

If Chalmers, and other naturalistic dualists or physicalists, reject the conclusion of this argument, they owe us an accounting of which premise they reject, and why. So far as I can see, the only premises Chalmers might conceivably reject are 2 and 4. To reject 2, however, amounts to giving up; it acknowledges a huge explanatory void in one's worldview that can never be filled. As regards 4, I have argued that (a) Chalmers' naturalistic assumptions leave him with no possible way of explaining rationality apart from an evolutionary explanation, and (b) if the physical realm is closed, an evolutionary explanation fails completely.⁸

It should by now be clear in what sense Chalmers is, and in what sense he is not, a reductivist. He emphatically rejects the *logical reduction* of phenomenal properties to physical properties. Instead, he insists on a dualism of properties, and makes strenuous efforts to deal with the difficulties this creates for his position. He is, all the same, an *explanatory reductivist*, in that all the explanatory work is in the end done by the configurations of physical particles. We've argued in this section that this position is indefensible – but avoiding it requires us to abandon a cornerstone of physicalist (and naturalistic dualist) doctrine, namely the causal closure of the physical domain.

Some Consequences of Non-Reductivism

Even if the abandonment of causal closure is forced, as I've argued that it is, this represents a major cost for many philosophers, one they will be extremely reluctant to incur. Interestingly, however, some very important additional consequences emerge free of charge, once causal closure is gone. Three such consequences are detailed below.

1. *Mind-body supervenience does not obtain.* This is an important result, and requires careful examination. Consider Jaegwon Kim's definition of "weak supervenience":

A weakly supervenes on B if and only if necessarily (for any property F in A, if an object x has F, there exists a property G in B such that x has G), and if any y has G it has F (1991c, p. 64; parentheses added for clarity)

⁸As has been stated, Chalmers never addresses this issue, but he comes close enough that it is hard to see how he manages to overlook it. In the chapter on "The Paradox of Phenomenal Judgment" he develops an implication of his position, namely that *the content of our conscious experience plays no role in explaining the judgments we make about that content*. How then can he avoid the question, "What explains the fact, *if indeed it is a fact*, that those judgments about content correspond to the actual content?"

Weak supervenience is thus a sort of uniform co-variance, within a possible world, of two classes of properties – in particular, of physical and mental properties. The argument against causal closure does not, so far, preclude such co-variance. There is, however, an additional ingredient in the notion of supervenience that is not captured by definitions such as the one given; this is the idea that supervenient properties *depend*, unilaterally, on the corresponding subvenient properties. (An action is wrong *because* it possesses certain non-moral properties; it doesn't possess those properties because it is wrong.) (See Kim 1991d and 1991e.) It is this notion of *unilateral dependence of the mental upon the physical* that is given up when causal closure is abandoned. On the contrary, it will now be the case that the physical properties of an organism develop differently, when accompanied by conscious experience, than they would in the absence of such experience.⁹ And this, as we shall see, is an extremely important result.

2. *Major objections to mind-body dualism no longer apply.* To see this, consider the following objection against dualism by Richard Taylor:

What we must conceive . . . is a physical change in the brain, this change being wrought, not by some other physical change in the brain or elsewhere, but by an *idea*. . . . We can suppose . . . just to get some sort of picture before us, that it is a change consisting of the diffusion of sodium ions into certain of the brain's cells. Conceive, then, if possible, how an *idea* can effect such a change as this, how an *idea* can render more permeable the membranes of certain brain cells . . . Try, I say, to form a conception of this, and then confess that, as soon as the smallest attempt at any description is made, the description becomes unintelligible and the conception an impossible one (Taylor, 1974, p. 25).

Anyone possessing even distant awareness of such developments as relativity, quantum mechanics, and string theory might wonder how Taylor's, or our, inability to imagine something in an underexplored area of science has any bearing on what is physically the case. But however that may be, the situation Taylor declares "unintelligible and impossible" is precisely what we have now seen to be required by the rationality of human thought – namely, that *a conscious state, the entertaining of an idea, has effects on the physical state of the brain*. This very old objection to dualism has gone by the boards.

A different but closely related objection has been stated recently by William G. Lycan: "In any case it does not seem that immaterial entities could cause motion consistently with any of the conservation laws of physics, such as that regarding matter-energy; physical energy would have to vanish and reappear inside human brains" (1996, p. 168). The core of this objection is simply that, given mind-body interaction, *the fundamental laws of physics will not always be satisfied by events in the brain*. And this is assuredly true; it's just what interaction amounts to. Some persons, for whatever reason, find failure of the matter-energy conservation laws especially unpalatable, but this response is not universal. One philosopher who does not find it

⁹A caveat: If we adopted Chalmers' proposal, and attributed the novel causal powers entirely to the physical configuration rather than to the conscious state, then a kind of supervenience might still obtain. I judge that, once causal closure has been given up, Chalmers' interpretation is extremely implausible.

compelling is W. D. Hart, who hypothesizes an exchange of energy between mind and brain: physical energy is converted into “psychic energy,” and back again (Hart 1988; see Hasker 1999, p. 301). If this is viewed as undesirable, it is certainly conceivable that the mind should influence the brain without changing the total amount of energy in the brain; for instance, by directing a nerve impulse into a different channel than it would otherwise follow. But *something* has to be different in the brain as a result of conscious thought, and some sort of conservation law will undoubtedly be violated. Or perhaps “violated” is not the right word; what happens is that *because of the impingement of new fundamental forces* – in this case, *mental* forces – things go differently in the physical brain than they would otherwise have gone. Some will find this shocking, but it is a straightforward implication once causal closure has been abandoned. The minimal form of “interference” would be for the mind to control indeterministic quantum events that are then “amplified” to produce large-scale results. This would still mean, however, that physical principles are violated; in this case, the principle that states that quantum events are truly random and are not controlled by a hidden causality. And it requires that there be in the brain some mechanism by which, not on rare occasions but as a regular occurrence, random or quasi-random quantum events are amplified so as to produce macroscopic results. It is my understanding that so far there is no evidence for such a mechanism, though the possibility can’t be ruled out.

I fully understand that many philosophers will find all this shocking and entirely unacceptable. But the only way to escape from it is to reaffirm causal closure and revert to a position according to which rationality is an unexplained, and forever unexplainable, miracle. On the other hand, I by no means wish to claim that the truth of mind-body dualism has been established by the arguments given here. There is much more to be said on this subject, but this is not the place to say it.

3. *There is no empirical evidence for determinism.* Such evidence for determinism as we have had consists almost entirely of the extremely accurate predictions that are routinely made in the physical sciences. Contrary to what is sometimes supposed, merely statistical predictions (such as are common in the social sciences) constitute no evidence whatever for determinism, since such predictions are equally well explained on the assumption of indeterminism. This is clearly shown by calculations of gambling odds, made on the assumption that the fall of the cards or dice is completely random. Further evidence is provided by quantum mechanics, which makes remarkably accurate statistical predictions on the basis of an indeterministic theory. The indeterminist must, to be sure, acknowledge the existence of propensities which make things more likely to happen in one way rather than another, but few indeterminists will find this problematic.

The incredibly precise and accurate predictions made in the physical sciences, on the other hand, provide a strong indication that genuinely deterministic processes are at work. Even the evidence for quantum indeterminacy disturbs this picture only slightly. Such indeterminacy usually is important only on the sub-atomic scale, while statistical effects render most macroscopic phenomena effectively deterministic. But abandoning the causal closure of the physical domain *completely eliminates* the significance of physical determinism as evidence for a general thesis of determinism. Even if all purely physical processes are strictly deterministic (as many indeterminists have been willing to concede), abandoning causal closure opens the way for

indeterminism precisely where it is of most concern – in the thoughts, deliberations, and decisions of conscious beings.

Nor is there an escape from this result in a purely psychological determinism, in which an agent's actions are caused by the "strongest motive." As has often been observed, this formula tends to be empirically vacuous, since the strongest motive can only be identified retrospectively, as the motive which did in fact lead to action. If on the other hand we were to try to identify the "strongest motive" *before* the decision has been made, there is no reason to think the results would support determinism. According to John Searle, "Insofar as psychological determinism is an empirical hypothesis like any other, then the evidence we presently have available to us suggests it is false." At the time when he wrote this, Searle held firmly to physical determinism, and suggested a "modified form of compatibilism," in which "psychological libertarianism is compatible with physical determinism." (1984), p. 97). (As we shall see, Searle's commitment to physical determinism is at present very much in question.)

So to repeat: there is no empirical evidence for determinism. This does not, of course, establish that determinism is false or that indeterminism is true. The possibility of undetectable causal factors, ranging anywhere from unconscious motivation to secret divine decrees, can never be ruled out completely. Nevertheless, the picture of determinism as purely a metaphysical hypothesis, one which entirely lacks empirical evidence and has no prospects of acquiring any in the foreseeable future, is very different from the way we have been accustomed to regard these topics. For one thing, it removes the pressure to accept a compatibilist view of free will on the ground that "determinism might turn out to be true." There simply isn't any credible scenario which would vindicate such fears.

I must confess that I find it remarkable – and I expect many others will find it so as well – that such important results follow immediately, and without further questionable assumptions, from the failure of causal closure. And causal closure fails because we have assumed, first, that human beings are to a significant degree rational, and second, that this fact requires an explanation. I submit that for either of these assumptions to be false would be far *more* remarkable (not to mention intellectually stultifying) than are the results reached in this section.

Fleshing Out Non-Reductivism: Searle and Nagel

The non-reductive views developed so far present a formidable challenge to the reigning reductivist orthodoxy, but the picture they present remains seriously incomplete. Completing the picture is no doubt the task of an entire philosophy. But short of that, we shall in this section explore the progress made by two philosophers in fleshing out non-reductivism in the direction of a more complete understanding of things. It's especially significant that neither of these philosophers has expressed any sympathy for a "philosophy of transcendent reality," such as absolute idealism or theism. Their non-reductivism isn't motivated by such lofty metaphysical aspirations; instead, it has been developed "from below," as it were, as a response to the perceived inadequacies of reductivist views of particular regions of the intellectual landscape.

Throughout the course of his philosophical work, John R. Searle has argued against reductivism in many different ways. Perhaps most famous is his "Chinese room argument," which aims to

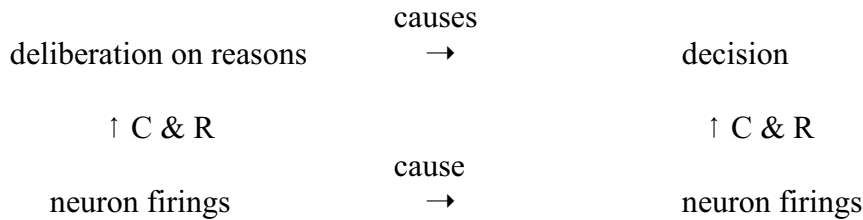
establish that understanding, meaning, and intentionality cannot be understood purely in causal or computational terms (1980). In the story a human being, who knows nothing of the Chinese language, is ensconced in a room with a large number of pieces of paper bearing Chinese characters, as well as sets of instructions for manipulating these pieces of paper. He is presented with messages consisting of Chinese characters and, by following the instructions, produces appropriate sequences of characters in response, but he knows nothing about the meaning either of the messages presented to him or the ones he has himself assembled. The “Chinese room” fulfills the causal and computational requirements for understanding, since it produces appropriate outputs in response to Chinese-language inputs, but in fact there is no one in the story who understands the messages.

The anti-reductivism of most interest to us here is found in Searle’s recent book, *Rationality in Action*. Much of the argument of that book revolves around the notion of “the gap” that exists in all cases of rational action. According to Searle, “‘The gap’ is the general name that I have introduced for the phenomenon that we do *not* normally experience the stages of our deliberations and voluntary actions as having causally sufficient conditions or as setting causally sufficient conditions for the next stage” (2001, p. 50). He also says, “The operation of rationality presupposes that there is a gap between the set of intentional states on the basis of which I make my decision, and the actual making of the decision” (p. 13). This gap is crossed by the activity of a “non-Humean self.” (“You cannot account for the rational self just in terms of a Humean bundle of disconnected perceptions” (p. 289).) The gap, according to Searle, is an obvious feature of everyday experience. Suppose you have gone to a restaurant, and the waiter asks for your order. You can’t just say, “Look, I am a determinist, che sarà, sarà. I will just wait and see what I order! I will wait and see what my beliefs and desires cause!” Searle continues, “This refusal to exercise my freedom is itself only intelligible to you as an exercise of freedom” (p. 14). Searle’s point here serves to underscore the failure of psychological determinism, as noted in the previous section of this essay.¹⁰

Important questions arise, however, when we attempt to combine this psychological account of human action with the neurobiological account. Searle queries, “There is no doubt that the gap is psychologically real, but is it otherwise empirically real? Is it neurobiologically real? If human freedom really exists, it must be a feature of brain function” (p. 269). Searle’s analysis of this question can be illustrated by a pair of diagrams.¹¹ The first diagram illustrates the situation in which it is assumed that the agent’s reasons, consisting of her desires and beliefs, constitute a sufficient cause of the decision:

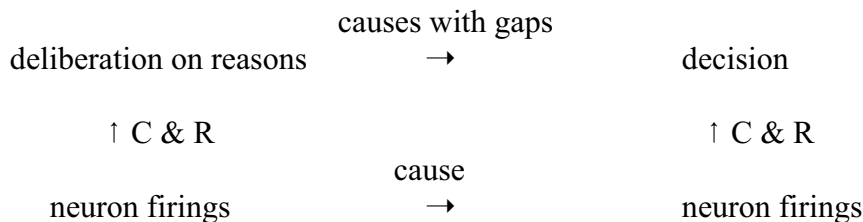
¹⁰It should be noted, however, that Searle is not claiming that the experience of the gap demonstrates that determinism is false. His point is, first, that desires and beliefs are not in themselves sufficient to bring about action; a decision is required. Furthermore, “the reasons preceding the decisions and the actions are not experienced by the agent as setting causally sufficient conditions for the decisions and actions” (p. 62).

¹¹The second of these diagrams is Searle’s, on p. 283; the first is modified from one given by Searle.



One set of neuron firings “causes and realizes” (C & R) the deliberation, another set causes and realizes the decision, and the whole process is regarded as entirely deterministic. This way of conceptualizing the situation is quite similar to Kim’s supervenient causation, though Searle does not employ the language of supervenience.

When we add the gap, however, the diagram must be altered:



Now the assumption is that “the indeterminacy at the psychological level is matched by a completely deterministic system at the neurobiological level” (p. 283). In Searle’s view, some distinctly uncomfortable conclusions emerge from this way of understanding action:

This result . . . is intellectually very unsatisfying because, in a word, it is a modified form of epiphenomenalism. It says that the psychological processes of rational decision making do not really matter. The entire system is deterministic at the bottom level, and the idea that the top level has an element of freedom is simply a systematic illusion. It seems to me at t_1 that I have a choice between the Burgundy and the Bordeaux and that the causes operating on me are not sufficient to determine the choice. But I am mistaken. The total state of my brain at t_1 is entirely sufficient to determine every bodily movement as well as every thought process from t_1 to t_2 to t_3 . . . and the only thing we can say about psychological indeterminism at the higher level is that it gives us a systematic illusion of free will” (p. 285).

In addition to undermining our lived experience of free will, Searle states that

the hypothesis seems to me to run against everything we know about evolution. It would have the consequence that the incredibly elaborate, complex, sensitive, and – above all – biologically expensive system of human and animal conscious rational decision making would actually make no difference whatever to the life and survival of the organisms. (p. 286).¹²

¹²Interestingly, Searle does not raise the question pursued in the first part of this essay, namely, *How is it even possible* for the system of conscious rational decision making to evolve, given that it makes no physical difference and is thus invisible to evolutionary selection?

Given these unwelcome consequences of the hypothesis of “psychological libertarianism with neurobiological determinism,” the situation is ripe for the emergence of a competing hypothesis, which takes the form of “system causation with consciousness and indeterminacy.” On this view, “the absence of causally sufficient conditions at the psychological level is matched by a parallel lack of causally sufficient conditions at the neurobiological level” (p. 286). But this in turn forces us to

examine critically the assumptions built into our diagrammatic representation with its metaphors of ‘bottom-up,’ ‘top-down,’ ‘levels of description,’ etc. . . . The problem is this: the idea that consciousness is a higher-level or surface feature of the brain gives us a picture of consciousness as like the paint on the surface of the table. . . . All of that is wrong. Consciousness is no more on the surface of the brain than liquidity is on the surface of the water. Rather the idea we are trying to express is that consciousness is a *system feature*. . . . What we have to suppose, if we believe that our conscious experience of freedom is not a complete illusion, is that the whole system moves forward toward the decision making, and toward the implementing of the decision in actual actions; that the conscious rationality at the top level is realized all the way down, and that means that the whole system moves in a way that is causal, but not based on causally sufficient conditions. (p. 287)

Clearly, this second hypothesis involves extremely difficult metaphysical questions, which cannot be pursued further here. What is remarkable is that Searle has even proposed such a hypothesis, one which challenges the foundations of most contemporary thinking about the mind and the brain. Nevertheless, Searle himself is troubled by its implications, and states, “Frankly, I do not find either hypothesis intellectually attractive” (p. 296); he is unwilling to decide definitely between them. The problem he finds with the second hypothesis “is to see how the consciousness of the system could give it a causal efficacy that is not deterministic. And it is not enough help to be told that we could accept the randomness of quantum mechanical accounts that are not deterministic. Conscious rationality is not supposed to inherit the randomness of quantum mechanics.¹³ Rather, conscious rationality is supposed to be a causal mechanism that proceeds causally, though not on the basis of antecedently sufficient causal conditions” (pp. 297-98). Searle himself would be the first to admit that his non-reductivism is a work still in progress.

Thomas Nagel’s *The Last Word* is a sustained attack on a number of different varieties of reductivism. With respect to logic, science, history, and ethics he argues that none of these intellectual practices can be understood by taking up a viewpoint external to the practice in question and “explaining” the practice from that viewpoint. Such reductive explanations invariably issue in judgments which conflict with the first-order judgments made within the disciplines themselves – and in a fair competition, the disciplines win out over the reductive explanations.

¹³It should be noted that if reliance on quantum indeterminacy were sufficient, the second diagram could be modified by simply adding indeterminism at the neural level paralleling the psychological indeterminacy. I judge that Searle is right to reject this solution.

Especially interesting are the conclusions Nagel reaches in his final chapter, significantly entitled, “Evolutionary Naturalism and the Fear of Religion.” Here Nagel admits quite candidly, “I hope there is no God! I don’t want there to be a God; I don’t want the universe to be like that” (p. 130). Nagel goes on to speculate,

My guess is that this cosmic authority problem is not a rare condition and that it is responsible for much of the scientism and reductionism of our time. One of the tendencies it supports is the ludicrous overuse of evolutionary biology to explain everything about life, including everything about the human mind. Darwin enabled modern secular culture to heave a great collective sigh of relief, by apparently providing a way to eliminate purpose, meaning, and design as fundamental features of the world. Instead they become epiphenomena, generated incidentally by a process that can be entirely explained by the operation of the nonteleological laws of physics on the material of which we and our environments are all composed. (p. 131)

Nagel himself, even though he shares in the “cosmic authority problem,” strenuously resists this facile appeal to Darwinism. At this point he brings into the discussion some views of C. S. Peirce concerning the nature of science. Peirce is well known for his pragmatic theory of belief, but in fact he strongly contrasts belief, or willingness to act on the truth of a proposition, with science: “We *believe* the proposition we are ready to act upon. . . . But pure science has nothing at all to do with *action*” (Peirce 1992 p. 112; cited in Nagel 1997 p. 127). Nagel cites with approval Peirce’s assertion that

The only end of science, as such, is to learn the lesson that the universe has to teach it. In Induction it simply surrenders itself to the force of facts. But it finds . . . that this is not enough. It is driven in desperation to call upon its inward sympathy with nature, its instinct for aid, just as we find Galileo at the dawn of modern science making his appeal to *il lume naturale* (Peirce 1992 p. 176, cited in Nagel 1997 pp. 128-29).

In these sentiments Peirce’s view of science is revealed as being less pragmatic than Platonic. Nagel states that he finds this view of Peirce’s “not only eloquent but entirely congenial” (p. 129). But he has a worry:

The reason I call this view alarming is that it is hard to know what world picture to associate it with, and difficult to avoid the suspicion that the picture will be religious, or quasi-religious. . . . Even without God, the idea of a natural sympathy between the deepest truths of nature and the deepest layers of the human mind, which can be exploited to allow gradual development of a truer and truer conception of reality, makes us more *at home* in the universe than is secularly comfortable. (pp. 129-30)

Nagel, however, wishes to resist any move in the direction of a religious interpretation. He suggests that “the capacity of the universe to generate organisms with minds capable of understanding the universe is itself somehow a fundamental feature of the universe,” and while admitting that this “has a quasi-religious ‘ring’ to it, something vaguely Spinozistic,” maintains that “one can admit such an enrichment of the fundamental elements of the natural order without going over to anything that should count literally as a religious belief” (p. 132). Still, he is

forced to admit (with an acknowledgment to Mark Johnson) that “if one asks, ‘*Why* is the natural order such as to make the appearance of rational beings likely?’ it is very difficult to imagine any answer to the question that is not teleological” (p. 138 n.). I think it is fair to say that for Nagel, not being a reductivist *while remaining naturalistic* has proved very difficult indeed.

Conclusion

I have argued that some ostensibly “non-reductive” approaches to the mind are in fact still reductive in a way that leads to insoluble problems, in particular the problem of explaining the rationality of the conscious mind. In order to avoid this kind of reductionism, one must abandon a cornerstone of contemporary naturalism and materialism, namely the causal closure of the physical domain. This having been done, certain other results follows as a matter of course: mind-body supervenience is abrogated; major objections to mind-body dualism must be abandoned, and determinism is seem to be devoid of empirical support. These consequences have the further result of opening up difficult metaphysical questions which non-reductivists will have to cope with. Not being a reductivist is anything but an easy task – but there is no going back.

References

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

Chalmers, David (2002). The Puzzle of Conscious Experience. *Scientific American*, Special Edition, *The Hidden Mind*, pp. 90-98.

Hart, W. D. (1988). *Engines of the Soul*. Cambridge: Cambridge University Press.

Hasker, William (1997). Theism and Evolutionary Biology. In *A Companion to the Philosophy of Religion*, ed. Philip Quinn and Charles Taliaferro, Oxford: Blackwell, pp. 426-32.

Hasker, William (1999). *The Emergent Self*. Ithaca: Cornell University Press.

Kim, Jaegwon (1991a). The Myth of Nonreductive Materialism. In *Supervenience and Mind: Selected Philosophical Essays*. Cambridge: Cambridge University Press, pp. 265-84.

Kim, Jaegwon (1991b). Epiphenomenalism and Supervenient Causation. In *Supervenience and Mind*, pp. 92-108.

Kim, Jaegwon (1991c). Concepts of Supervenience. In *Supervenience and Mind*, pp. 53-78.

Kim, Jaegwon (1991d). Supervenience as a Philosophical Concept. In *Supervenience and Mind*, pp. 131-60.

Kim, Jaegwon (1991e). Postscripts on Supervenience. In *Supervenience and Mind*, pp. 161-71.

Lycan, William G. (1996). Philosophy of Mind. In *The Blackwell Companion to Philosophy*, ed. N. Bunnin and E. P. Tsui-James. Oxford: Blackwell.

Nagel, Thomas (1997). *The Last Word*. New York: Oxford University Press.

Peirce, C. S. (1992). *Reasoning and the Logic of Things*, ed. Kenneth Laine Ketner. Cambridge: Harvard University Press.

Searle, John R. (1980). Minds, Brains, and Programs. *Behavioral and Brain Sciences* 3, pp. 417-24.

Searle, John R. (1984). *Minds, Brains, and Science*. Cambridge: Harvard University Press.

Searle, John R. (2001). *Rationality in Action*. Cambridge, Mass.: MIT Press.

Stich, Stephen (1990). Evolution and Rationality. In *The Fragmentation of Reason: Preface to a Pragmatic Theory of Cognitive Evaluation*. Cambridge: MIT Press.

Taylor, Richard (1974). *Metaphysics*, 2nd edition. Englewood Cliffs: Prentice-Hall.