

TABLE OF CONTENTS

Introduction	2
The Argument from Consciousness	3
“Consciousness”	4
Alternative Positions	6
The Argument against Materialism	8
Supervenience and Explanation	8
Against Materialism	11
The Argument for Scientific Inexplicability	17
Establish Correlations	17
Producing a Causal Account	18
Positing Fundamental Laws	19
The Argument for Substance Dualism	26
Where do <i>I</i> go?	26
The Substantiality of Persons	29
Chalmers’ Theory of Consciousness	33
Underlying assumptions	33
The Resultant Theory	36
Assessing the Theory	38
Conclusion	43
Reconsidering the Alternatives	43
Naturalism and Non-materialism?	44
The Old Problem	45
The Theological Alternative	47
References	49

Acknowledgements

Thanks first my tutors – Peter Byrne for draft reading and formative input and Paul Helm for initial assistance. Unreferenced but influential books include *Consciousness Reconsidered* by Owen Flanagan and *In Search of Divine Reality* by Lothar Schäfer. The JCS Online and Quantum-Mind email lists have been an excellent source of ideas and debate. Finally, I am grateful to David Chalmers, Anthony Freeman, Lothar Schäfer, Jonathan Shear, Vidyasankar Sundaresan and Charles Tart for helpful and stimulating discussion.

INTRODUCTION

Of the areas in which philosophy is still forging new territory, the philosophy of mind is one of the most exciting and innovative. Since the publication of Gilbert Ryle's *The Concept of Mind* in 1949, many philosophers have taken up the challenge of attempting to understand how human and animal¹ minds fit into, interact with and get a hold on the world outside.

One branch of philosophy that stands to be significantly affected by any consensus that emerges is the philosophy of religion. Mainstream philosophy of religion remains strongly grounded in classical theism, of which the Cartesian dualist-interactionist theory of the mind², built on folk-psychological intuitions, is a central component. Just as the physical, finite world's existence is explained by the actions of an infinite, spiritual god, so too are the physical actions of man partially explained by his spiritual nature – the posited non-materialism of our souls provides a crucial analogy with which we can understand God. Furthermore, theistic beliefs in life after death, resurrection or even reincarnation entail certain views on the metaphysical nature of the self. If people are nothing but their bodies or brains, it is difficult to conceive of how one might survive one's death as an individual, be reincarnated in another body, or be resurrected in one's own.

Many concepts in *eastern* religion are also tied up with these issues. In Advaita, the equation of *Atman* with *Brahman* can be understood as an ontological statement regarding the relationship of an individual's pure consciousness to some sort of primal formless consciousness underlying the universe as a whole. In Buddhism, *Nirvana* not only entails a selfless attitude towards the world but also a change in conscious phenomenology. Some recent studies such as Varela, Thompson and Rosch (1991) and Austin (1998) have begun to examine this fascinating common ground.

¹ Most philosophers agree that animals have similar minds to humans, albeit less complex.

² For determinists, substitute epiphenomenalist for interactionist, but Cartesian dualism remains.

The Argument from Consciousness

In this paper, I will focus on the argument in natural theology that consciousness provides significant grounds for belief in classical theism. More precisely, *that the undeniable fact of subjective experience beckons an extra-scientific, personal explanation.*

The argument has its origins in Locke (1959, IV, III, 28):

It is evident that the bulk, figure, and motion of several bodies about us produce in us several sensations, as of colours, sounds, tastes, smells, pleasure, and pain, &c. These mechanical affections of bodies having no affinity at all with those ideas they produce in us, (there being no conceivable connexion between any impulse of any sort of body and any perception of a colour or smell which we find in our minds,) we can have no distinct knowledge of such operations beyond our experience; and can reason no otherwise about them, than as effects produced by the appointment of an infinitely Wise Agent, which perfectly surpass our comprehensions.

Locke's appeal to inconceivability and his invocation of a 'god of the gaps' principle make his argument weaker than we might like. Nonetheless, the themes he raises are similar to many of those in modern approaches to the issue.

The best known contemporary argument from consciousness to theism is to be found in Swinburne (1979, 1986, 1996a) and it is on this which I will concentrate. Other significant treatments are given in Adams (1987) and Taliaferro (1994). A critique of the argument is found in Mackie (1982) while it is also discussed by Olding (1991).

As well as considering literature within the philosophy of religion, I will also broaden this study into mainstream philosophy of mind by taking a modern theory of consciousness and seeing how well it fares against the gauntlet thrown down by Swinburne *et al.* While there are dozens of candidate naturalistic theories, I will focus on that of David Chalmers as outlined in *The Conscious Mind* (1996). The main reason for this choice is that Chalmers is

wholly in agreement with the theologians' arguments against materialism³. To pit a theory such as that of Dennett (1991) against Swinburne would require too much debate over whether the functionalist analysis of consciousness succeeds (and I think it manifestly does not). Instead, I will actually *use* Chalmers' superb explication of the argument against materialism as a basis from which natural theology can proceed.

This paper will develop as follows: by way of introduction, I will explain what I mean by the term *consciousness*, outline some of the issues it raises and list some of the positions taken up in response. Having done so, I will use Chalmers' book to argue that materialism is refuted by the existence of consciousness. Next, I will develop and criticise the two main themes of the argument for theism from consciousness: (a) that consciousness is scientifically inexplicable and (b) that only substance dualism can explain our 'selves'. Then, I will return to Chalmers to see how well the naturalistic theory he outlines in *The Conscious Mind* can respond to the theistic argument, and what difficulties his view would leave us with. Lastly, I will consider in what further directions the issue may be pursued.

“Consciousness”

Distinctions are made between consciousness *understood*⁴ as a functional process and consciousness *understood* as an intrinsic, qualitative phenomenon. In all cases, when I use the word *consciousness*, I will be referring to the latter – to the type of consciousness which none of us can directly observe in others and which we have direct epistemic access to in ourselves. This corresponds to Block's *p-consciousness* as opposed to *a-consciousness* (1995) and the phenomenon to which Chalmers' *hard* as opposed to *easy* problem (1995) pertains.

The distinction is best explicated by example. Consider a robot constructed out of silicon

³ The term *materialism* will require more precise formulation – Chalmers argues against *reductive* or *eliminative* materialism.

⁴ This distinction will be justified later – for now, it can be considered only as *prima facie*.

chips which was capable of behaving in all ways exactly like a human being. When asked “how do you feel this morning” it might reply “not bad, but I have a strange sense of being out of place in the cosmos”. And yet, it may also be true that this robot would not actually have any feelings at all, in other words, *there is nothing it would be like to be the robot* (Nagel 1974). Such a robot would qualify as being conscious in a *functional* sense, yet there would be something amiss – it would entirely lack *phenomenology*.

The corollary is even more unpleasant. One day, you wake up in bed to discover that you are incapable of moving a single muscle or limb in your body. Eventually you see your family crowd around you, trying to communicate. While you are perfectly conscious of their attempts, you are incapable of responding in any manner. Not knowing that you are, in fact, perfectly sane but simply incapable of expressing it, they package you off to a mental home where you spend the rest of your life totally imprisoned by your body. In this case, the functional side of consciousness is gone, yet the phenomenal side remains intact.

The consciousness that this paper discusses is of the phenomenal variety. By definition, if there is no answer to the question “What is it like to be X?” for a certain X, that X lacks phenomenal consciousness. And if there is an answer, the *content* of phenomenal consciousness is defined as whatever that answer may be. While there is likely to be a panoply of processes taking place with the mind at any one time, it is only those of which a subject is phenomenally conscious with which we are concerned.

An immediate fact to take note of is that phenomenal consciousness is a thoroughly surprising phenomenon, from a classical scientific point of view. If a race of non-*phenomenally* conscious aliens swooped down to take over the earth, what reason would they have to think⁵ that any of the organisms on earth had phenomenology? Indeed, one may doubt whether they could even understand what it could be – our objections of “no, you cannot do this, for we are sentient subjects with moral concerns” would be likely to meet with as much

⁵ That is, *think* in a functional sense.

understanding as would colours to a person who could only see in shades of grey. There seems to be no reason why outsiders need *necessarily* realise we are more conscious than viruses, bacteria or plants, if indeed we are.

Alternative Positions

What are the body and the mind made out of? Five standard positions may be enumerated – materialism, idealism, neutral monism, property dualism and substance dualism. Each of materialism, idealism and neutral monism is a variation on the notion that there is only one type of substance and property. In the case of materialism, the nature of this one substance and its properties are material in the sense in which that is currently understood – describable in terms of mass, energy, space, time and fields – any mental processes take place only within this realm. Idealism adopts the opposite view in claiming that the single substance is made out of thoughts or concepts – any apparent material processes are a result of the interaction of these ideas. Lying between materialism and idealism, neutral monism holds that there is only one type of substance and property, yet the precise nature of this is either epistemically inaccessible or at least not yet understood.

By contrast, property and substance dualism hold that there are two types of properties in the cosmos. While property dualism attributes these both to a single substance (and is thus close to but not identical with neutral monism), substance dualism attributes them to two different types of substance, the mental/spiritual and the material/physical. While substance dualism remains the mainstay of theistic metaphysics, property dualism is often seen as a way to hold onto physicalism while taking the phenomenon of consciousness seriously. Chalmers sees himself largely in the property-dualist vein.

Under either property or substance dualism, there are further issues concerning the interaction between the two properties or substances. Do mental properties or substances have any effect on material ones? If not, we have a case of *strong epiphenomenalism*, where nothing about the physical world is affected by what takes place in conscious minds. And if so, can

the decisions made within minds exert the intended influence on physical properties or substance? If not, we end up with *weak epiphenomenalism* where, despite mentality making some difference to physicality, it cannot get what it wants, so to speak.

Many other questions are also debated within the philosophy of mind, such as the claims of panpsychism or panprotopsychism (e.g. Nagel, 1979), the possibilities of weak or strong artificial intelligence (e.g. Searle, 1980), the nature of selfhood and personal identity (e.g. Shoemaker and Swinburne, 1984) and doubts over whether the entire metaphysical debate is a worthwhile endeavour (e.g. Johnston, 1993). Many of these issues will be touched upon in the subsequent discussion.

THE ARGUMENT AGAINST MATERIALISM

The argument of David Chalmers rests upon what seems to many people clear upon reflection – that no matter how detailed our physical description of a particular entity, we have no reason to infer that the entity harbours phenomenology. But whereas this common-sense assertion may be easily undermined by questioning our intuitions, Chalmers proffers a philosophically rigorous argument against the dominant materialism of the day.

Supervenience and Explanation

To begin to address an issue as complex and perplexing as consciousness, it is essential to construct a framework within which argument may proceed. Chalmers provide this in *The Conscious Mind* by way of *supervenience*, defined as follows (33):

Supervenience is a relation between two sets of properties: *B*-properties—intuitively, the *high-level* properties—and *A*-properties, which are the more basic *low-level* properties.

B-properties *supervene* on *A*-properties if no two possible situations are identical with respect to their *A*-properties while differing in their *B*-properties.

Thus, for example, if we say that the *A*-properties of an object refer to its entire atomic and molecular configuration in space at a particular time, and the *B*-properties to whether or not that object is lumpy, we can safely say that it is not possible for two objects, of identical physical configuration, to be of different lumpiness. Lumpiness thus *supervenes* on the physical.

Logical Supervenience

However, a distinction must be made between *logical* (or *conceptual*) supervenience and *natural* (or *nomic* supervenience). Logical supervenience is defined as follows (35):

B-properties *supervene logically* on *A*-properties if no two *logically possible* situations are identical with respect to their *A*-properties but distinct with respect to their *B*-properties... It is use-

ful to think of a logically possible world as a world that it would have been in God's power (hypothetically!) to create, had he so chosen... In determining whether it is logically possible that some statement is true, the constraints are largely *conceptual*.

So we see that, if B-properties logically supervene on A-properties, the A-properties in question *entail* the B-properties, since it is not possible for those particular A-properties to hold without the B-properties holding too.

Our prior example of lumpiness qualifies for such supervenience. One could rigorously define lumpiness in terms of some statistical property of the entity concerned, perhaps in terms of the variance of density over its volume. It is not possible for a substance to be uniformly even and yet be 'lumpy', just as a substance with significantly varying density cannot be denied the term. That lumpiness is not all-or-nothing is beside the point – to the extent that a substance can be attributed or denied lumpiness, this can be done solely on the basis of its density distribution.

An important subtlety must be addressed, however. For any property such as lumpiness, attributed by sentient beings on the basis of some phenomenal assessment, it may be retorted that it *could not* supervene on the physical, since two agents may differ over how 'lumpy' they consider a substance to be – perhaps one always had lumpy porridge for breakfast and so became more tolerant of lumpiness in general. However, this would be to confuse ontology with epistemology. If lumpiness is a *genuine* property of substances, the inter-subjective differences do not reflect on the property itself. In the same way, we might say that, notwithstanding possible differences in the way that individuals may perceive colours, the *actual* colour of an object depends only on the wavelengths of electromagnetic radiation that it reflects.

Natural Supervenience

The second type of supervenience, *natural* supervenience is described thus (37):

Natural supervenience holds when, among all naturally possible situations, those with the same distribution of A-properties have the same distribution of B-properties... This happens

when the same clusters of A-properties in our world are always accompanied by the same B-properties, and when this correlation is not just coincidental but *lawful*: that is, when instantiating the A-properties will always bring about the B-properties, wherever and whenever this happens.

A good example of natural supervenience is gravitational attraction, as understood within pre-relativistic physics. It is in no way a *conceptual* or *logical* truth that two masses will be attracted to each other in proportion to their total mass divided by the distance between their centres of gravity squared – it is also logically possible that masses would lie about in space undergoing no attraction to each other whatsoever, or even repulsion. But it so happens that, in our world as empirically observed, masses are indeed attracted to each other in this way. So gravitational attraction is *naturally supervenient* on mass-space configuration.

However, this example brings out an important caveat. It is possible that one could gain a new understanding of what the world is made out of, in terms of say C-properties, such that two sets of properties A and B, which previously only had a relationship of natural supervenience between them, may *each* be logically supervenient on the newly-discovered C-properties. Indeed, this is precisely what happened when Einstein came along with his new picture of the relationship between matter-energy and space-time. Before his revolution, it had been understood that, by positing a single property called ‘mass’, there was a relationship of natural supervenience between the amount of force required to accelerate an object a certain amount and the amount of gravitational attractive force it exerted on other objects. Afterwards, these two phenomena became *logically* supervenient on his theory of general relativity, assuming of course that it accurately describes the world. They just ‘drop out’ of the equations, so to speak.

Any attempt at reductive explanation via logical supervenience is sensitive to the concepts under which the phenomena to be explained are understood. Chalmers says (43):

If someone objected to a cellular explanation of reproduction, “This explains how a cellular process can lead to the production of a complex physical entity that is similar to the original

entity, but it doesn't explain *reproduction*," we would have little patience—for that is all that "reproduction" *means*. In general, a reductive explanation of a phenomenon is accompanied by some rough-and-ready *analysis* of the phenomenon in question, whether implicit or explicit.

Therefore, in considering whether consciousness can be reductively explained by any particular cognitive or neuroscientific theory, we must constantly return to the question of what exactly we *mean* by 'consciousness' – the *what it is like to be*.

Against Materialism

Different meanings are attached to the term *materialism* and we must clarify which form of materialism Chalmers' arguments are aimed at. He has no objection to materialism construed as a monism of substance. Rather, the materialism he attacks is that which claims that "all the positive facts about the world are... logically supervenient on the physical facts" (41). Chalmers defines physical properties⁶ as (33)

the fundamental properties that are invoked by a completed theory of physics. Perhaps these will include mass, charge, spatio-temporal position; properties characterising the distribution of various spatio-temporal fields, the exertion of various forces, and the forms of various waves; and so on. The precise nature of these properties is not important. If physics changes radically, the relevant class of properties may be quite different from those I mention, but the arguments will go through all the same.

In producing an argument for physical properties *in general* (as opposed only to those of which we are currently aware), Chalmers is advancing a strong claim. What all such properties (present and future) have in common is that they are concerned only with *structure* and *function* (1996, 106–7), not with any intrinsic quality which cannot be seen or detected from the outside.

⁶ A *fact* for Chalmers is simply the instantiation of a particular property.

Conceivability and The Zombie

Much of the work in Chalmers' book is done by examples which claim to demonstrate how these properties could never *entail* a particular variety (or indeed, any) conscious phenomenology. The simplest example is that of the *zombie*, described as follows (94–5):

This creature is molecule for molecule identical to me, and identical in all the low-level properties postulated by a completed physics, but he lacks conscious experience entirely... We can imagine that right now I am gazing out of the window, experiencing some nice green sensations from seeing the trees outside, having pleasant taste experiences through munching on a chocolate bar, and feeling a dull aching sensation in my right shoulder. What is going on in my zombie twin? He is physically identical to me, and we may as well suppose that he is embedded in an identical environment. He will certainly be identical to me *functionally*: he will be processing the same sort of information, reacting in a similar way to inputs, with his internal configurations being modified appropriately and with indistinguishable behaviour resulting... It is just that none of this functioning will be accompanied by any real experience. There will be no phenomenal feel. There is nothing it is like to be a zombie.

Although, as Chalmers admits, "it is unlikely that zombies are naturally possible" (96), such a being seems eminently *conceivable*. However, care must be taken before making the leap from *conceivability* to logical *possibility*. Chalmers states (66–7):

Let us say that a statement is conceivable (or conceivably true) if it is true in some conceivable world. This should not be confused with other senses of "conceivable." For example, there is a sense according to which a statement is conceivable if for all we know it is true, or if we do not know that it is impossible. In this sense, both Goldbach's conjecture and its negation are conceivable. But the false member of the pair will not qualify as conceivable in the sense I am using, as there is no conceivable world in which it is true (it is false in every world).

So what is conceivable on first glance may turn out to be logically impossible. Chalmers argues that the zombie will not be disqualified in a manner similar to the negation of Goldbach's conjecture. But some, such as Dennett (1995) and Cottrell (1999), counter that if he were to reflect on his zombie a little more, he would discover some incoherence in the de-

scription. As Cottrell says, “one may fool oneself into thinking one has imagined something when one has not really confronted its detailed implications” (12).

However, it is also the case that what is inconceivable to even the greatest of philosophers may turn out not only to be *logically* possible, but perhaps even *naturally* possible. In the *Discourse*, Descartes supplies an example relevant to our topic (1985, 140):

We can certainly conceive of a machine so constructed that it utters words, and even utters words corresponding to bodily actions causing a change in its organs... But it is not conceivable that such a machine should produce different arrangements of words so as to give an appropriately meaningful answer to whatever is said in its presence, as the dullest of men can do.

No one would claim that artificial intelligence has yet achieved this lofty goal, but it is within the range of some laboratories’ aspirations. Certainly what Descartes could not imagine is easily *conceivable* in an age of computer-simulated intelligence.

Despite these pitfalls, it seems that in many issues we have no alternative but to use conceivability as a guide to logical possibility. Although a proposition’s logical *impossibility* can be conclusively demonstrated by deriving a contradiction from it, how can one ever be certain that something is logically *possible*? Outside of certain narrow fields, the best we can do is describe the proposition in full and carefully consider and work through its implications, checking for any hidden incoherence.

Doing so for Chalmers’ zombie, I can find no flaws in its conception. The inability of others to conceive of it remains somewhat mysterious to me – as Chalmers says, “it is difficult to argue across this divide, and discussions are often reduced to table pounding. We may simply have to learn to live with this basic division” (xiii). Perhaps the problem lies with the metaphysical beliefs of Chalmers’ opponents – for the *committed* materialist, the zombie is as logically impossible as solipsism to the man on the street.

In any event, Chalmers supplements his argument from the zombie with many other well-known thought experiments in the field, such as the inverted spectrum (originating in Locke, 1959, XI, XXXII, 28) and Jackson’s famous knowledge argument (1982). Both of these

reiterate the point that knowing all there is to know about a being's structure and function will not tell us *what it is like to be* that being, in the absence of further yet-to-be-discovered psychophysical laws. Nonetheless, the committed materialist may continue to assert that these thought experiments harbour logical contradictions which the non-materialist cannot see. The reader must pause to reflect and decide for himself.

If these example *are* coherent, they lead to what Levine coined the 'explanatory gap' (1983). The notion is as follows: for any *explanation* we attempt to provide as to why a particular phenomenal property is instantiated, there will be a gap that cannot be bridged between the physical and phenomenal description of a particular entity. No matter how complex or sophisticated our physical description of an entity becomes, one cannot make the explanatory 'leap' over to phenomenal properties. There is always the further question: "Why is the physical process accompanied by conscious experience? And why by *this* conscious experience?"

That there is this gap means that materialism must be false, where materialism is construed in the reductive manner outlined. However, it does not mean that the weaker claim of *physicalism* is false, since this insists only that there is no extra-physical *substance*. Chalmers says (124–5):

The arguments do not lead us to a dualism such as that of Descartes, with a separate realm of mental substance that exerts its own influence on physical processes... a move to a Cartesian dualism would be a stronger reaction than is warranted.

Later, we will hear the arguments of some who disagree.

The Conceptual Fallacy?

We may note that nothing in Chalmers' argument precludes the possibility of a scientific revolution occurring, such that both physical and mental properties as currently understood *logically* supervene on a new posited fundamental property. Many criticisms of his book have been along the lines that just because our physical and mental *concepts* radically differ,

nothing is proven about whatever *actual property* is referred to by those concepts. In a review of *The Conscious Mind*, Levine says (1998, 878):

The problem... lies with a crucial assumption that is not adequately defended; namely, that competence with a concept automatically affords us a priori access to a description of its primary intension (or, better, the properties that determine its primary intension).

However, the possibility *is* considered by Chalmers. He only demands that, should we embrace it, we face the metaphysical consequences (129):

I should also note that although I call the view a variety of dualism, it is possible that it could turn out to be a kind of monism. Perhaps the physical and the phenomenal will turn out to be two different aspects of a single encompassing kind, in something like the way that matter and energy turn out to be two aspects of a single kind. Nothing that I have said rules this out, and in fact I have some sympathy with the idea. But it remains the case that if a variety of monism is true, it cannot be a *materialist* monism. It must be something broader.

In an ever stronger objection, Searle insists that the entire framework for the debate is outdated (1997, 50):

Where the mind is concerned, we have inherited a Cartesian vocabulary and with it a set of categories that include "dualism," "monism," "materialism," and all the rest of it. If you take these categories seriously, if you think our questions have to be asked and answered in those terms, and if you also accept modern science (is there a choice?), I believe you will eventually be forced to some version of materialism. But materialism in its traditional forms is more or less obviously false, above all in its failure to account for consciousness.

One could summarise our metaphysical vocabulary's aetiology as follows: in the beginning, there was animism and the Aristotelian 'dualism' of matter and form. When the search for law-like objective scientific principles took off, we formed a view of the *world* that excluded subjective entities. But the *self* was clearly left out so a dualism of the mental and the physical ensued. Then, a failure to make scientific sense of this dualism led naturalists to monism. And since the science of the physical world is so well understood, a materialist monism was adopted. But now that we face the problem of consciousness, Searle suggests

that we go back and reconsider whether historical accident has led us to a vocabulary that accurately reflects reality.

A similar point is raised by Wittgenstein, who believed that (Johnston, 1993, 237)

the attempt to reach profound truths on the basis of purely conceptual investigations was futile – according to him, grammar is autonomous and conceptual analysis can teach nothing about reality nor resolve the puzzle of how we should understand the world... In his own terms, therefore, Wittgenstein believed metaphysics was doomed to failure.

While I resist the abandonment of metaphysics *tout court*, these points do add some crucial perspective to the investigation. Nonetheless, in granting the possibility of some form of radical monism, Chalmers seems to take the bite out of Searle's criticism. When he says, "perhaps the physical and the phenomenal will turn out to be two different aspects of a single encompassing kind" (129), he is explicitly allowing that today's dualistic metaphysical vocabulary may be replaced by only a secondary distinction between the third- and first-person perspectives.

In any event, the theological arguments to which I now turn rely heavily on the Cartesian categories – if the categories reflect reality, and consciousness exists, then Chalmers' argument succeeds. Let us now see how a refutation of materialism can be developed into an argument for a personal god.

THE ARGUMENT FOR SCIENTIFIC INEXPLICABILITY

The main thrust of Swinburne's argument from consciousness in *The Existence of God* (1979, 160–75) is that it is unlikely that science will ever be able to explain the existence and contents of consciousness. This is broadly similar to the argument pursued by Adams in his essay 'Flavours, Colours and God'. But, considered in isolation, the argument will turn out to be rather indirect, so I will supplement it with Swinburne's argument for substance dualism in *The Evolution of the Soul*. How substance dualism provides support for theism will be discussed later.

We pick up Swinburne's argument in *The Existence of God* from where we left Chalmers, having established that "some kind of dualism of entities or properties or states is inevitable" (1979, 166). Swinburne goes on to outline three steps which science will have to take in order to establish a naturalistic explanation of consciousness – for the first two, he cites difficulties, but he believes that the most significant problems lie in the third.

Establish Correlations

The first stage is to (167)

establish for all mental events and states a one-one or perhaps one-many correlation between the occurrence of mental events of specifiable kinds, and the simultaneous occurrence of brain-events of specifiable kinds...

For example, we might imagine amassing enough evidence to suggest that whenever a certain subset of neurons in a brain fires in a certain pattern, a conscious experience of feeling a headache is undergone. We could further imagine discovering such correlations for every type of experience subjects undergo, enabling us to predict *what it is like to be* them given the relevant neurological data. We might even be able to induce experiences by manually stimulating these sets of neurons.

Swinburne argues that any such discoveries will be "dubious" since "the lack of public

observability of one half of the correlations means that there is a certain doubt about the data.” (167). Why he does not see it reasonable to apply his *principle of credulity* used later for developing his argument from religious experience (254–260) is left unanswered. But in any event, if such correlations could be established across many *different* subjects and subjects were able to see *for themselves* the predictive use of the discoveries, there would be little reason to doubt that Swinburne’s first condition had been fulfilled.

Producing a Causal Account

The second stage, according to Swinburne, is to show “what causes what” (167):

To show that the brain-events are the ultimate determinant of what goes on, the materialist will need to show that the occurrence of all mental events is predictable from knowledge of brain-events alone... whereas the occurrence of all brain-events is not predictable from knowledge of mental events alone.

Swinburne assumes that the naturalist will believe that causation between brain events and mental events flows only in one direction. While true of someone who insists that the domain of physical events is causally closed, two other possibilities are available to the broad-minded scientist: (a) a dualist-interactionism where mental events also lie within the scope of purely scientific (i.e. non-personal) explanation and (b) a dualist-interactionism which includes some raw notion of agency. While the latter option would entail an admission of personal causation into the explanatory picture (and thus take a step in Swinburne’s direction), it has no *necessary* connection with theism.

For this stage, Swinburne cites two difficulties. The first is “man’s experience of choice” (168) which, according to his principle of credulity, he believes should not be doubted without good reason. But as I have already argued, there is no reason why the scientist cannot also invoke agency as part of her picture of the mind. In any case, as Swinburne admits, “freedom of choice may be an illusion” (167).

The second is that, under quantum theory, (168)

the basic physical laws are statistical and probabilistic. They only allow us to infer from one brain-event B_1 that it was (physically) very probable that a subsequent brain-event B_2 would occur... But this would leave open the possibility that the explanation of the occurrence of brain-event B_2 correlated with intention I_2 was to be explained fully by the joint action of brain-event B_1 and intention I_2 .

In other words, quantum theory's stochastic nature means we cannot rule out the possibility that mental events are having some effect on the causative connections between brain events. But this difficulty applies to *any* statistical scientific theory – perhaps our understanding of thermodynamics, based on the probabilistic motions of molecules, is missing the influence of unspecified hidden factors. In any event, as already mentioned, a non-theistic explanation need not rule out mental causation – in concluding this paper, I will briefly explain how quantum theory has led some philosophical scientists to incorporate agent-based causation into a naturalistic world-view.

Positing Fundamental Laws

Swinburne puts most store by the problems he says science will experience in the third stage. He describes how this would have to proceed (169):

We could then have a dictionary in which, observing a man's brain-events, we could look up and see which mental events he now had. Would all this mean that we had got a scientific explanation of the existence of mental events, intentions, beliefs, and indeed persons? I think not. For to explain the existence of mental events we need to cite not merely the cause, the brain-event, which apparently brings about the mental event, but also the scientific law in virtue of which the brain-event brings about the mental event.

After citing the same epistemic limitations as before, he proceeds to elucidate the core problem, which I quote selectively at length (170–1):

Infinitely many totally different theories about unobservables can be constructed which predict the same events as each other... The evidence that one of them is the true theory lies in simplicity, the naturalness of the connections in the laws which it postulates...

Suppose that the materialist's programme of establishing correlations so far has been moderately successful. What will he have? Lots and lots of correlations of the following kind: brain-event B_1 correlated with a red after-image, B_2 with a blue one,... B_4 with an intention to move an arm, B_5 with an intention to sign a cheque... What he needs is a neat set of laws showing a natural connection between redness and this kind of brain-event, blueness and that kind, the intention to move an arm with the kind of brain-event, to sign a cheque with that kind; laws which fit together with each other in a theory from which we can deduce new correlations hitherto unobserved...

Although it is theoretically possible that a scientific theory of this kind should be created, still the creation of such a theory does not look a very likely prospect. Brain-states are such different things qualitatively from experiences, intentions, beliefs, etc. that a *natural* connection between them seems almost impossible...

This is a good explication of the well-known problem of phenomenological representation, which will have to be overcome by any comprehensive science of consciousness. In order to be able to find detailed correlations between subjectively experienced mental events and objectively accessible physical events, we need to represent the two types of event in a manner allowing isomorphisms to be found between them. But the modes of representation are so entirely different (like 'ethics' and 'rhubarb' in Colin McGinn's terms) that it is hard to imagine how such an enterprise could even get started.

Adams' paper makes a broadly similar point (1987, 245):

It is difficult to see how science would even try to explain the correlation between phenomenal qualia and brain states (or whatever other physical states the qualia are most directly correlated with). For what science is geared up to do is to find laws governing physical states, described in terms of properties that are geometrical or electrical or at any rate quite different from phenomenal qualia.

This problem has also often been raised outside natural theology, most notably by McGinn (1989). It may be separated into two strands – the *practical* problem of complexity and a further *in principle* objection to the entire enterprise.

The Problem of Complexity

Consider how it was that science was able to explain warmth, coolness, tepidity and intense heat with reference to a single kinetic property of molecules. The feat was achieved by placing the former descriptions on a continuous linear scale ('temperature') and demonstrating how position on this scale is directly related to another continuous value – the mean speed squared at which molecules within a body are moving. All phenomena⁷ related to temperature such as heat transfer and changes between solid, liquid and gaseous states could then be explained in terms of molecules' kinetic activity.

To be as successful as the thermodynamic account of heat, a theory of consciousness would have to perform a similar simplification. But the contents of consciousness are so much more *complex* than descriptions of molecular movements and the like that this seems like an unattainable goal. One can learn a lesson from the attempts of the introspectionist movement's to atomise human phenomenology at the start of this century. According to Güzeldere (1997, 14), there was an

apparently irreconcilable conflict between results coming out of different laboratories... Titchener's laboratory reported that they discovered a total of "more than 44,435" discriminably different sensations, largely consisting of visual and auditory elements. In contrast, Külpe's published results pointed to a total of fewer than 12,000...

Apart from their inability to come even close to agreement, the breadth of the taxonomy that each laboratory arrived at highlights the inordinate complexity of phenomenology in comparison with ordinary physical descriptions of events. It is difficult to imagine how a simple scientific theory could find connections between physical events and the vast heterogeneity of conscious experiences while remaining true to *what they are really like*.

⁷ Excluding *our experience* of temperature, of course.

The Problem of Mental Anomaly

The further objection *in principle* is outlined by Davidson in his well-known paper, 'Mental Events' (1970). Much of the piece is concerned with a form of the identity theory, but we are interested here in the argument against the possibility of psychophysical laws. Davidson says (89, 97–98):

Laws are linguistic; and so events can instantiate laws, and hence be explained or predicted in the light of laws, only as those events are described in one or another way...

There are no strict psychophysical laws because of the disparate commitments of the mental and physical schemes. It is a feature of physical reality that physical change can be explained by laws that connect it with other changes and conditions physically described. It is a feature of the mental that the attribution of mental phenomena must be responsible to the background of reasons, beliefs and intentions of the individual. There cannot be tight connections between the realms of each is to retain allegiance to its proper source of evidence... We must conclude, I think, that nomological slack between the mental and the physical is essential as long as we conceive of man as a rational animal.

Davidson is saying that the existence of laws connecting physical and mental events would entail a redefinition of the mental in which we would no longer recognise ourselves, since our concept of rational agency requires a certain degree of scientific anomaly. Clearly, he has raised a number of complex issues, particularly regarding the nature of scientific laws and what it means to say that an agent is rational. But much of Davidson's argument also comes from the assumption that the mental must *supervene* on the physical (88), implying that a comprehensive set of psychophysical laws would have to *explain* the mental by the physical. If, alternatively, there were laws linking *independent* mental and physical properties *to some extent*, the existence of "tight connections" would not exclude the possibility of "nomological slack". Much more could be said on this issue – for now, let us take Davidson's article as further highlighting the difficulties any naturalist account of consciousness would have to contend with.

Therefore God?

If, for argument's sake, we were to accept that no scientific explanation for the correlations between brain events and phenomenological content could ever be found, what then? Would it be sensible to "seek a personal explanation of mind-body correlations" (172), in other words, look to God for an answer?

Unless Swinburne unashamedly wants to embrace a 'god of the gaps' natural theology, the impossibility of finding a scientific explanation for a particular phenomenon can surely not provide justification for invoking a personal one⁸. If that principle were to have been applied 200 years ago, almost all of what we now believe to happen as a result of scientific laws would have been attributed to the activity of God. Colin McGinn, for one, embraces the following mysterian but naturalist position (1989):

Let us then say that there exists some property *P*, instantiated by the brain, in virtue of which the brain is the basis of consciousness. There exists some theory *T*... which fully explains the dependence of conscious states on brain states. If we knew *T*, then we would have a constructive solution to the mind-body problem... It is surely *possible* that we could never arrive at a grasp of *P*; there is... no guarantee that our cognitive powers permit the solution of every problem we can recognise. We *could* be like five-year-old children trying to understand Relativity Theory.

McGinn goes on to provide positive arguments for why our minds are probably cognitively closed with regard to the phenomenon of consciousness. For the purposes of the current discussion, it suffices to establish that the inability of the human species to understand a particular scientific principle does not imply that the principle does not exist.

But Swinburne would reply that it is our lot as reasoning humans to infer as best as possible from the information and faculties made available to us. In the earlier part of the same chapter (154–60), he goes to considerable lengths citing reasons why *God would want to cre-*

⁸ Clearly, I am considering the argument from consciousness in isolation from the putative cumulative case Swinburne puts forward in *The Existence of God*.

ate conscious beings. So if we recognise that consciousness is outside of our scientific understanding but that positing the existence of God could explain the phenomenon, would it not at least be *rational* for us to suppose that God indeed exists? Swinburne's C-inductive argument is structured similarly to others in the book (174):

For the reasons which I have given it does not look at all plausible to suppose that there is a scientific explanation of these phenomena. Once again, for reasons of simplicity, the most probably personal explanation is one in terms of the agency of God...

However, against all this, Mackie argues (1982, 126):

The question is not whether the materialist can *formulate* a theory that would explain the mind-body interaction, but whether he can reasonably believe that *there are* laws that would explain this... Analogously, though no one doubts that there are simple physical laws which account for all meteorological phenomena, no one expects to be able to predict the exact course of the next Caribbean hurricane...

We can round out Mackie's point as follows: based on the prior successes of scientific explanation with phenomena which appeared, *prima facie*, to be utterly scientifically inexplicable, we have good reason to infer that scientific explanation can also be extended to the current-day phenomenon which appears, *prima facie*, to be utterly scientifically inexplicable, even if that explanation will always lie beyond our cognitive reach. So to counter Swinburne's induction, Mackie is able to produce one of his own.

But the theological argument is not just that consciousness is *currently* scientifically baffling but that, as a matter of principle, it will probably always be so. Adams says (1987, 251):

It is not just that science has not *yet* found an explanation for the correlation between qualia and physical states. Science is headed in the wrong direction for finding such an explanation, and it would be silly to expect science to turn in another direction.

I think Adams is wrong about science's direction – it clearly *has* turned around in that a significant amount of theoretical and experimental work on the problem of consciousness is now being carried out – perhaps if we are stuck at the same place in 500 years' time, his

point will have greater force. More challenging for the naturalist are the *constructive* arguments against scientific explanation, specifically the problem of phenomenological complexity and Davidson's argument against psychophysical laws. Both suggest that no matter how far science develops, there will be no way for it to posit nomological connections between the mental and the physical while doing justice to our experiences or conception of ourselves.

Whether the problem of phenomenological representation can ever be solved or not, the trouble with the argument for scientific inexplicability is that it suffers all of the criticisms of any other 'god of the gaps' natural theology. To explain a gap in our understanding by reference to an entity who has no scientific role *other* than the plugging of that gap is to replace one mystery by another – if God is able to explain the connection between mental and physical events, why cannot some natural but incomprehensible principle do the same? On its own, therefore, the argument from scientific inexplicability is too weak to gain any firm ground for theism. I will now turn to consider whether there is a more positive reason to accept that consciousness implies the existence of a personal God.

THE ARGUMENT FOR SUBSTANCE DUALISM

In *The Evolution of the Soul*, after a lengthy categorisation of different classes of mental events as sensations, thoughts, purposes, desires and beliefs, Swinburne puts forward an argument for substance dualism, defined as (1986, 145)

the view that those persons which are human beings (or men) living on Earth, have two parts linked together, body and soul. A man's body is that to which his physical properties belong...
A man's soul is that to which the (pure) mental properties of a man belong.

Clearly, substance dualism does not in itself entail theism. However, if substance dualism can be established, theism would be strongly suggested for several reasons:

- (a) If we have reason to believe in a mental/spiritual realm separate from the physical world, we have reason to believe in a realm where an incorporeal God could exist.
- (b) It is considerably more difficult for science to explain interactions between two realms of existence than to explain interactions within a single realm, therefore a personal explanation (i.e. in terms of a personal God) would become more attractive.
- (c) Whereas we have some ideas about what physical stuff is made out of, we lack the same information when it comes to mental/spiritual stuff. Positing God as an example of immaterial substance would go some way towards providing the answer.

Points (a) and (c) may be considered as corollaries of one another – they both suggest that substance dualism and theism are natural metaphysical partners.

Where do *I* go?

Swinburne produces two arguments for substance dualism in *The Evolution of the Soul*. The first is explicated with the following thought-experiment (148–9):

The brain, as is well known, has two very similar hemispheres—a left and a right hemisphere... It might be possible one day to remove a whole hemisphere, without killing the person, and to transplant it into the skull of a living body from which the brain has just been re-

moved, so that the transplant takes. There would then appear to be two separate living persons. Since both are controlled by hemispheres originating from the original person *p*,... we would expect each publicly to affirm such apparent memories and to behave as if he had *p*'s character... But they cannot both be *p*. For if they were, they would both be the same person as each other, and clearly they are not—they have now distinct mental lives. The operation would therefore create at least one new person—we may have our views about which (if either) resultant person *p* is, but we could be wrong... However much we knew in such a situation about what happens to the parts of a person's body, we would not know for certain what happens to the person.

This intriguing argument draws two potential replies. One may be brought out by an analysis of the results of commissurotomy operations carried out during the 1960s and 1970s as a last resort cure for chronic epilepsy (Sperry 1968). The operation involved severing the corpus callosum bundle of 200 million nerve fibres linking the brain's two hemispheres resulting in an almost⁹ complete lateral localisation of information processing.

Since language control ordinarily resides almost entirely in the left hemisphere, only the left side of the brain can be *directly* interrogated about its phenomenology, which it claims it continues to have. And there is a debate as to whether the right hemisphere should be considered a conscious subject too, which would imply that two streams of consciousness now occupy a single human skull. Eccles' view is as follows (1977, 326–8):

The rigorous testing of the subjects who have been subjected to section of the corpus callosum has revealed that conscious experiences of the subject arise only in relationship to neural activities in the dominant [left] hemisphere... We can regard the minor [right] hemisphere as having a status superior to that of the non-human primate brain... It has many skills,... but it gives no conscious experience to the subject... Moreover there is no evidence that this brain has some residual consciousness of its own...

One can tell from Eccles' wording ("it gives no conscious experience to the subject") that

⁹ It is still possible for information to pass between the hemispheres by longer routes, such as via the thalamus, hypothalamus or cerebellum.

he shares Swinburne's views on the indivisibility of the self. Nonetheless, his interpretation undermines Swinburne's argument – if there is no evidence to suggest that the right hemisphere is conscious then the “person” in Swinburne's thought experiment would simply follow wherever the left hemisphere went.

Sperry, however, sees the right hemisphere as (1974)

a conscious system in its own right, perceiving, thinking, remembering, reasoning, willing, and emoting, all at a characteristically human level... though predominantly mute and generally inferior in all performances involving language or linguistic or mathematical reasoning...

On this issue, epistemic difficulties seem insurmountable. If we cannot ask the right hemisphere whether it is conscious, how are we to know? Nevertheless, even if we grant Swinburne's claim that we create “two separate living persons”, why should we not conclude that *neither* of the two persons is the same as the one before? If each hemisphere of the brain is specialised for certain forms of processing, surely the phenomenology of each will lack part of what used to make up that of the original connected brain. The left side may no longer appreciate music, and the right side will be incapable of conversation. It seems doubtful that each would “behave as if he had *p*'s character” as Swinburne suggests.

Swinburne is ready to accept the *possibility* of this account but his argument takes an epistemological turn (150–1):

Even if this notion of partial survival does make sense, it will in no way remove the difficulty, which remains this. Although it *may* be the case that if my two brain hemispheres are transplanted into different bodies, I survive partly as the person whose body is controlled by one and partly as the person whose body is controlled by the other, it may not be like that at all... Knowledge of what has happened to a person's body and its parts will not necessarily give you knowledge of what has happened to the person... so... persons are not the same as their bodies... It suffices to make my point to point out that the mere logical possibility of a person surviving with only half his brain (the mere fact that this is not a self-contradictory supposition) is enough to show that talk about persons is not analysable as talk about bodies and their parts.

This is similar to Chalmers' argument supporting the claim that *talk* about persons is not analysable as *talk* about bodies. But it does not establish that a person's *actual* identity is not constituted by their body (or brain). To conflate epistemology and ontology is to have allowed an argument from vitalism made 100 years ago to establish conclusively that a plant consists in more than its physical parts. The trouble with this mode of reasoning will be discussed further in the next section.

The Substantiality of Persons

Swinburne's second argument is based on the assertion that "persons are substances" (153) and the following 'quasi-Aristotelian' assumption and argument (154):

A substance S_2 at t_2 is the same substance as an earlier substance S_1 at t_1 only if S_2 is made of some of the same stuff as S_1 (or stuff obtained therefrom by gradual replacement)...

Given, that for any present person who is currently conscious, there is no logical impossibility... that the person continue to exist without his body, it follows that that person must now actually have a part other than a bodily part which can continue, and which we may call his soul—and so that his possession of it is entailed by his being a conscious being. For there is not even a logical possibility that if I now consist of nothing but matter and the matter is destroyed, that I should nevertheless continue to exist.

This seems to be an error. If Swinburne's logic were correct, we could substitute 'communication medium' for 'person', 'carrying a message' for 'being conscious', and 'wiring' for 'body' to obtain the following argument:

Given, that for any present communication medium which is currently carrying a message, there is no logical impossibility... that the medium continue to exist without its wiring, it follows that that medium must now actually have a part other than a wired part which can continue, and which we may call its soul—and so that its possession of it is entailed by its carrying a message. For there is not even a logical possibility that if a particular communication medium now consists of nothing but matter and the matter is destroyed, that it should nevertheless continue to exist.

The similarity between Swinburne's argument and this one will help clarify why neither can carry through. For a being to qualify as a 'person' in Swinburne's terms is for that being to be conscious. Since it is logically *possible* for any person to become disembodied, yet continue to be conscious, Swinburne infers that a particular person, himself, *must* have a non-material part, since otherwise were he to become disembodied he would not remain the same person (under the quasi-Aristotelian assumption). But, in my obviously false parallel argument, for an entity to qualify as a 'communication medium' is for that entity to be carrying a message. Now, it is clearly logically *possible* for any communication medium to lose its wiring, yet continue to carry a message (one could imagine the information disappearing at one end of where the medium used to be and reappearing at the other). Can I infer that a particular communication medium (for example, a telephone cable) *must* have a non-material part? For otherwise, under the quasi-Aristotelian assumption, were it become disembodied it would not remain the same communication medium!

In technical terms, Swinburne has used a premise about the secondary intension¹⁰ of the term 'person' to justify a conclusion about its primary intension. In virtue of the meaning (secondary intension) of the term, persons may or may not be nothing but matter. Perhaps some persons are and some other persons are not. But that does not imply that a particular person or indeed any *actual* person (primary intension) need consist of something but matter. His argument can demonstrate no more than the truth of one of its premises, i.e. that it is logically possible for there to be a non-material person.

Much debate, particularly in the pages of the *Faith and Philosophy*, has centred around this argument in *The Evolution of the Soul*. For example, Alston and Smythe (1994) challenge the quasi-Aristotelian assumption on which it is based. Stump and Kretzmann (1996) object to the notion of hard fact he uses in the appendix by way of clarification. Hasker (1998) accuses it of being epistemically circular, relying on its conclusion as a premise in its formula-

¹⁰ See Kripke (1972).

tion. Each of these objections has met with responses by Swinburne (1996b, 1996c, 1998). But I fail to see how it even gets off the ground – one simply cannot make an inference from logical possibility allowed by a particular description to a natural possibility supposedly upheld by a particular entity fitting that description. Reames (1999) makes a similar point, claiming that Swinburne’s notion of logical possibility is better described as an instance of natural or metaphysical possibility and so is not amenable to *a priori* arguments from conceivability.

An even more fundamental way to counter the argument is to deny that persons are substances in the first place. Hume raises the point elegantly (1739, 1.4.6):

When I enter most intimately into what I call *myself*, I always stumble on some particular perception or other, of heat or cold, light or shade, love or hatred, pain or pleasure. I never catch *myself* at any time with a perception, and never can observe anything but the perception.

But Swinburne rejects this on the following grounds (157–8):

One wonders what he supposes that the common subject would look like, and what he considered would count as its discovery... The self which he ought to have found in all his mental events is supposed to be the subject, not the object of perception. And finding it consists in being aware of different mental events as had by the same subject. Further... that certain simultaneous mental events are states of a common subject... It is among the data of your experience that these are all *your* mental events.

However, Hume’s point is precisely that this is not the case. He is not aware of “different mental events as had by the same subject” – he is aware only of a succession of mental events. James puts the point as follows (1892, 82–3):

The consciousness of Self involves a stream of thought, each part of which as “I” can remember those which went before, know the things they knew, and care paramountly for certain ones among them... The *I* which knows them... need [not]... be an unchanging metaphysical entity like the Soul or a principle like the transcendental Ego... It is a *thought*, at each moment different from that of the last moment, but *appropriative* of the latter, together will all that the latter called its own... The thoughts themselves are the thinkers.

There is no doubt that losing the folk-psychological notion of monadic selfhood requires considerable reflection and a battle against intuition. Nonetheless, nothing in the resultant transformation is contradicted by the raw data of experience. The phenomenology of personal identity can be understood in terms of James' stream of consciousness, where each thought picks up from the fringes of the previous one (1892, ch. 11), without being sustained by some unchanging substance¹¹.

So how successful are the combined arguments of Swinburne *et al*? In terms of the argument for scientific inexplicability, the early tasks of establishing correlations and producing a causal account do not seem to be beyond the wider vista of modern neurological science. However, the difficulties in positing and elaborating fundamental laws must be recognised by the naturalist as providing significant reason to doubt that science will ever be able to fully explain the phenomenon of consciousness. Nonetheless, we have noted that the science of the mind is in its very early infancy and that an argument from a gap in science (even if it be one *in principle*) to a divine explanation leaves much to be desired.

Swinburne's argument for substance dualism, while raising many interesting issues, fares worse. It relies both on a modal fallacy and the folk-psychological assumption that our thoughts pass through some unified indivisible substance separate from our bodies. By calling into question both the indivisibility and the substantiality of the self, many modern views on the mind/body problem challenge the basis on which the argument is constructed.

Despite the failings of these theological arguments, the naturalist corner would be significantly strengthened if it were able to put forward a workable account of consciousness that took both its existence and contents seriously. Just as God's existence cannot be proven by a gap in science, the possibility of scientific explanation cannot be assumed based on past success. To this end, I now turn to examine David Chalmers' theory, as expounded in *The Conscious Mind*.

¹¹ See Griffin (1998) for a modern process theory approach to the problem of consciousness.

CHALMERS' THEORY OF CONSCIOUSNESS

Underlying assumptions

It is beyond the scope of this paper to describe in detail the positive arguments that Chalmers provides for his theory of consciousness. In any event, in private conversation and at a recent talk, he has indicated a move away from the theory in *The Conscious Mind* towards considering some sort of quantum-physical account. Nevertheless, for current purposes, his book will serve as a useful candidate response to the natural theological argument. Before outlining the theory itself, I will briefly explain each of the five assumptions from which it emerges: naturalism, simplicity, causal closure, structural coherence and organisational invariance.

Naturalism

In the introduction, Chalmers states (xiii):

I take consciousness to be a natural phenomenon, falling under the sway of natural laws. If so, then there should be *some* correct scientific theory of consciousness, whether or not we can arrive at such a theory. That consciousness is a natural phenomenon seems hard to dispute: it is an extraordinarily salient part of nature, arising throughout the human species and very likely in many others. And we have every reason to believe that natural phenomena are subject to fundamental natural laws; it would be very strange if consciousness were not.

As part of this constraint, he also claims that “everything I say here is compatible with the results of contemporary science” (xiv). The justification for this assumption is identical to the point that Mackie makes above – that one can have reason to believe that there exists a scientific explanation for a phenomenon even if currently one has no idea how it may be reached.

Simplicity

In line with Swinburne's point in *The Existence of God*, Chalmers argues that (127):

Any lawful relationship must be supported by fundamental laws. The case of physics tells us that fundamental laws are typically simple and elegant; we should expect the same of the fundamental laws in a theory of consciousness.

Many hypotheses of consciousness suffer under this condition. Crick and Koch's well-known thesis that consciousness is subserved by 40Hz oscillations in the brain (1990) is one such view. To find it scientifically plausible, we need a reason why an oscillation at *that* frequency can cause consciousness to emerge. Similarly, a complex theory of consciousness that had hundreds of rules would remain scientifically unsatisfying unless those rules could be shown to follow from a few basic ones.

Causal Closure

Chalmers discusses the possibility of embracing some form of interactionism¹² in a section discussing epiphenomenalism. To do this, he claims (156)

requires a hefty bet on the future of physics, one that does not currently seem at all promising; physical events seem inexorably to be explained in terms of other physical events.

Problems with epiphenomenalism will be discussed later and Chalmers spends some time considering the issue. However, he maintains the view that physics is unlikely to move in the direction required to allow consciousness to make any difference to the development of physical events.

Structural Coherence

In arguing against materialism, Chalmers concluded that subjective experience ('con-

¹² Interactionism need not be Cartesian dualism – mental *properties* of physical substance could be granted a causal role.

consciousness') is not *logically* supervenient on functional cognition ('awareness'). Nevertheless, he does think that awareness is *naturally* supervenient on consciousness (220):

Where there is consciousness, there is awareness. My visual experience of a red book upon my table is accompanied by a functional *perception* of the book. Optical stimulation is processed and transformed, and my perceptual systems register that there is an object of such-and-such shape and colour on the table, with this information available in the control of behaviour.

The principle of structural coherence takes this one step further, suggesting not only that the *existence* of consciousness depends on that of awareness, but that the *structure* of consciousness mirrors that of awareness (224–5):

Similarities and differences between experiences correspond to similarities and differences represented in awareness; the geometry of experience corresponds to the geometry of awareness; and so on...

Thus, the contents of our subjective consciousness will very closely mirror those of our functional awareness. This seems to be borne out by empirical observation and is essential to any sense of rationality – if our subjective experiences diverged from our thoughts' functions, we would quickly lose all feeling of successfully interacting with the world.

Organisational Invariance

Chalmers produces a *reductio* argument against the claim that the existence and contents of conscious experience depend in any way on the *material* constituting our brains. He carries out several thought experiments to demonstrate that a silicon *functional* isomorph of a human brain would also produce exactly the same phenomenology. The *reductio* is as follows: if silicon brains lacked phenomenology, a complete replacement of neurons in one person's brain by silicon would result in 'absent qualia' (251–3), gradual replacement would result in 'fading qualia' (253–63) and the installation of a system to switch instantly between neural and silicon processing would allow for 'dancing qualia' (266–74).

Each of these possibilities is found problematic and implausible, thus Chalmers is led to

the principle of organisational invariance, stating that (247)

consciousness arises in virtue of the *functional organisation* of the brain. On this view, the chemical and indeed the quantum substrate of the brain is irrelevant to the production of consciousness. What counts is the brain's abstract causal organisation, an organisation that might be realised in many different physical substrates.

Functional organisation is best understood as the *abstract pattern of causal interaction* between various parts of a system, and perhaps between these parts and external inputs and outputs.

Chalmers calls himself a 'nonreductive functionalist', believing that "conscious experience is determined by functional organisation, but it need not be reducible to functional organisation" (275).

The Resultant Theory

As a result of the five principles outlined above, Chalmers has little room for manoeuvre when it comes to developing his theory. His demand for naturalism means that there must be psychophysical laws connecting physical brain events to experienced phenomenology. The principle of structural coherence means that these laws must match cognition closely to experience. The principle of organisational invariance entails that they can only depend on the functional form of the material subserving this cognition. His advocacy of physical causal closure means that causation runs in only one direction – consciousness has no explanatory role to play in the material world. Finally, the demand for simple, elegant laws means that he must locate some *fundamental* building block on which his theory may be built.

An obvious candidate for this primary element is *information*. Chalmers explains the notion of an *information space* at some length (277–280) – the core idea is that "information is a *difference that makes a difference*" (281) and that "information is as information does" (282). The skeleton theory he arrives at is as follows (286):

Information (in the actual world) has two aspects, a physical and a phenomenal aspect. Whenever there is a phenomenal state, it realises an information state, an information state that is also realised in the cognitive system of the brain. Conversely, for at least some physically realised information spaces, whenever an information state in that space is realised physically, it is also realised phenomenally.

This principle does not on its own come close to constituting a full psychophysical theory. Rather, it forms a sort of template for a psychophysical theory by providing a basic framework in which detailed laws can be cast. In fleshing out the principle into a theory, all sorts of questions need to be answered. For example, to just *which* physically realised information spaces does the basic principle apply?

Before considering some of the consequences of this approach, it would only be fair to quote his caveat (277):

I do not present a full-fledged theory with a comprehensive set of basic laws, but I put forward suggestions about the constructs involved in these laws, and about what the broad shape of the laws might be. This could be considered a *prototheory*: a skeleton around which a theory might be built.

The ideas in this chapter are much sketchier and more speculative than those elsewhere in the book, and they raise as many questions as they answer. They are also the most likely to be entirely wrong.

Nonetheless, for the purposes of this paper, we may note that a hard-headed naturalist has to accept much of what Chalmers has to say. Although his principle of organisational invariance is the most likely to be queried by others, denying the causal closure of the physical world would entail taking a big step away from the prevailing scientific world-view. Therefore, any criticisms that can be levelled at Chalmers' theory are also likely to detract from other naturalistic accounts. To the extent that there are successful, the natural theologians' position is strengthened.

Assessing the Theory

Firstly, we should note that Chalmers' theory, as far as it goes, seems to be coherent. However, that qualification is also earned by the view that there is only one mind, flipping from moment to moment between all the available streams of consciousness. Instead, we must ask whether the theory is *plausible* – to do so requires fleshing out some of its consequences, each of which are well discussed in the book.

Epiphenomenalism

Chalmers spends considerable time discussing the problem of epiphenomenalism, considering ways in which he can escape it by reconsidering the nature of causation (e.g. in Humean terms or by permitting overdetermination) (150–160). He concludes (160):

I do not describe my view as epiphenomenalism. The question of the causal relevance of experience remains open, and a more detailed theory of both causation and of experience will be required before the issue can be settled. But the view implies at least a weak form of epiphenomenalism, and it may end up leading to a stronger sort. Even if it does, however, I think the arguments for natural supervenience are sufficiently compelling that one should accept them.

Nonetheless, as I understand the term, he is clearly an epiphenomenalist. If the physical domain is causally closed, contra-causal free will is an impossibility. Events in the brain develop through either deterministic or random processes and there is no room for agency or personal explanation in the causal nexus. Although a compatibilist might call such a being 'free', she is clearly not free in the sense of being able to make *indifferent* choices regarding what her body does.

Are there any ways to conclusively refute epiphenomenalism, aside from a basic appeal to its unattractiveness or counter-intuitiveness? The main criticism of it seems to be that we make judgements and talk about our conscious experiences. Is it not absurd to claim that my being conscious is explanatory irrelevant to my saying "I do not see how it is possible that my physical brain causes me to be conscious"? As Searle says (1997, 48):

We can say that Chalmers wrote a book defending the irreducibility of his conscious states, but that, on his view, his conscious states and their irreducibility could have no explanatory relevance at all to his writing the book.

The point is by no means lost on Chalmers and he devotes an entire chapter entitled “The Paradox of Phenomenal Judgement” to it (172–209). When he comes to discuss his prototheory, he says (288–9):

A completed theory of mind must provide both a (nonreductive) account of consciousness and a (reductive) account of why we judge that we are conscious, and it is reasonable to expect that these two accounts will *cohere* with each other. In particular, we might expect that those features of processing that are centrally responsible for bringing about phenomenal judgements will also be those that are centrally responsible for consciousness itself...

If a theory shows how the explanation of phenomenal judgements centrally involves the explanatory basis of consciousness, then we will have woven the two together into a more unified picture of the mind, and some of the feeling of outrageous coincidence will be removed.

He goes on to expound how such a theory might develop. The crucial bridge is provided by the aforementioned notion of an information (292):

A conscious experience is a realisation of an information state; a phenomenal judgement is explained by another realisation of the same information state. And in a sense, postulating a phenomenal aspect of information is all we need to do to make sure those judgements are truly correct; there really *is* a qualitative aspect to this information, showing up directly in phenomenology and not just in a system of judgements.

Chalmers is saying that, instead of explaining judgements about consciousness by the consciousness itself, we explain *both* judgements *and* consciousness in terms of some lower-level entity – information. The suggestion is intriguing and is certainly worthy of consideration. Nonetheless, to embrace epiphenomenalism is to pay a high aesthetic price. In a sense, it is self-undermining, since my experienced belief in the plausibility of Chalmers’ suggestion is explanatorily irrelevant to my typing of this sentence. But ultimately the thesis cannot be *refuted* by its epiphenomenalist consequences.

Panpsychism

If conscious experience arises wherever there is information, a “difference that makes a difference”, then conscious experience is in a lot of places where we would not intuitively expect it to be. Does this constitute a *reductio ad absurdum* to Chalmers’ view? He approaches the difficulty as follows (293):

There are two ways that a support of the information-based approach might react to this situation. The first and most obvious is to look for further constraints on the *kind* of information that is relevant to experience. Not just any physically realised information space is associated with experience, but only those with certain properties... The alternative... is to bite the bullet and accept that all information is associated with experience. If so, then it is not just information that is ubiquitous. Experience is ubiquitous too.

In an endearingly-titled section, “What is it like to be a thermostat?” he describes the potential consequences of the second alternative (293):

Certainly it will not be very interesting to be a thermostat. The information processing is so simple that we should expect the corresponding phenomenal states to be equally simple. There will be three primitively different phenomenal states, with no further structure.

Discussion continues (293–301) and Chalmers avoids adopting a fixed position on the matter. The panpsychist consequences of his information theory may make it highly implausible to many readers but, like many other positions on the mind–body problem, panpsychism is irrefutable. Our lack of epistemic access to ‘other minds’ means that we will never be able to conclusively tell whether it is like anything to be a thermostat, a rock, or a hurricane. An objection along the lines of “but we have no reason to *believe* that rocks are conscious” can be assuaged by replying that the only reason we have to believe that *people* are conscious is that each of us is one of them, so we infer from our own predicament of sentience to that of others. Once again, a high price is to be paid in aesthetic terms but the possibility remains coherent.

Phenomenological Representation

The attentive reader will take note that, as explicated so far, Chalmers is yet to deal with the central problem discussed earlier, of how to represent subjective feels in a manner allowing them to be matched with physical events. To extend Chalmers' terminology, even if we solve his *hard* problem of explaining how consciousness arises, we still have an "even harder problem" – how to match the *contents* of consciousness with the information spaces from which it supposedly arises.

At the very end of the section discussing his prototheory (303–310), Chalmers addresses the issue (303–4):

We have something over and above a *pure* information space. Phenomenal properties have an intrinsic nature, one that is not exhausted by their location in an information space, and it seems that a purely information view of the world leaves no room for these intrinsic qualities.

He then goes on to make some suggestions about how phenomenology might be tied to intrinsic properties of matter itself. Perhaps (305)

the information spaces required by physics are themselves grounded in phenomenal or proto-phenomenal properties. Each instantiation of such an information space is in fact a *phenomenal* (or proto-phenomenal) realisation. Every time a feature such as mass and charge is realised, there is an intrinsic property behind it... The ontology that this leads to might truly be called a double-aspect ontology. Physics requires information states but cares only about their relations... phenomenology requires information states, but cares only about the intrinsic nature... Experience is information from the inside; physics is information from the outside.

Aside from noting just how speculative the book is now becoming, we might be able to take the proposition seriously. However, this still does not allay the theologians' misgivings. Chalmers wants to suggest that mass and charge might be very closely tied to some sort of intrinsic phenomenal properties but Swinburne would retort: How on earth are you going to get from simple physical properties to the full range of human conscious experiences? How are you going to explicate an isomorphism between, say, an aggregate of electron spin and

melancholy resignation? Chalmers himself recognises and discusses the difficulties in getting from such a protophenomenal atomistic schema to what we experience in our own phenomenology (305–8) and he concludes the section by leaving the problem open.

Therefore, *The Conscious Mind* fails to put forward a detailed or convincing answer to this strongest challenge of the natural theologians. And it remains difficult to see how science will be able to provide the necessary matching for all the reasons already discussed. Therefore, while his theory goes quite far towards positing a putative explanation of consciousness and its contents, Chalmers fails to substantively tackle the most difficult issue of all – phenomenological representation.

CONCLUSION

Reconsidering the Alternatives

Let us return to the metaphysical options enumerated in the introduction in order to see which remain viable in the face of all that has been said. Firstly, materialism has been eliminated by Chalmers' supervenience argument – in other words, by the brute, surprising existence of consciously experienced phenomenology. While the zombie seems logically impossible to some, I can only conclude from my own judgement that those people suffer from either a lack of metaphysical imagination or a dogmatic attachment to materialism.

Although Swinburne's argument *for* substance dualism fails, this does not constitute an argument *against* the possibility. But substance dualism does run into trouble if we make the naturalist assumption, for it is difficult to see how scientific laws could bridge events in two completely separate metaphysical domains. Nonetheless, one could certainly conceive of such natural laws, however unusual they may seem in comparison to the rest of science.

Idealism, despite its counter-intuitiveness, remains a live option for the broad-minded naturalist. Nothing about the world as empirically observed rules out the possibility that 'reality' is nothing but the interaction and exchange of ideas, sensations and impressions. However, once one has incorporated the laws of physics into the idealist metaphysic, it is doubtful whether the resulting picture is distinguishable from neutral monism – by comparison, a pure idealism such as that of Berkeley denies the existence of *any* nomological causal relationships between events.

For the naturalist, therefore, the most plausible remaining options are property dualism and neutral monism – that there is a single substance that has either two separate, disjointed sets of properties or a single set of properties whose nature is yet to be understood. The two positions differ only over whether physical and phenomenal properties are ultimately reducible to a single unified psychophysical property or not. While neutral monism's optimism results in a simpler ontological schema, it is left with the problem of explaining how

such apparently different classes of property could arise out of a single, more basic one and what the latter's nature could be.

Naturalism and Non-materialism?

But does naturalism not entail materialism? What right do naturalists have to start invoking thinking, feeling entities while claiming they remain within the scientific paradigm? J. P. Moreland, in a general article on the argument from consciousness (1998, 80) quotes David Armstrong (a materialist) as saying (1978, 72):

I suppose that if the principles involved [in analysing the single all-embracing spatio-temporal system which is reality] were completely different from the current principles of physics, in particular if they involved appeal to mental entities, such as purposes, we might then count the analysis as a falsification of Naturalism.

Armstrong is welcome to his own definitions but it seems somewhat fruitless to debate which theories do or do not deserve the particular title of 'naturalism'. A better question to consider is as follows: How on earth *could* mental entities be incorporated into a world-view which is otherwise materialist and scientific?

As many have noted, the quantum-mechanical view of matter leaves much room for precisely such a combination. Although any significant discussion is beyond the scope of this piece, some preliminary observations will be made. Several experiments in quantum mechanics have forced us to the conclusion that the world consists not of particles but of complex probability¹³ waveforms which, under certain conditions, 'collapse' into particles of either definite position or momentum (but never both at the same time). And all we know for sure about the 'collapse' of these waveforms is that it has to take place prior to, in some

¹³ *Complex* probability is expressed in the form $x+yi$ where i is defined as the square root of -1. See Penrose (1989, 306–20) for a good introduction to how the state of a quantum system is represented as a Schrödinger equation, whose value at each point is a complex probability.

sense¹⁴, the act of human observation since humans always see particles and not the underlying waveforms. Furthermore, it is possible for there to be instantaneous quantum effects (of a restricted variety) between particles which are light years apart from one another – some have suggested a connection with the binding problem in neuroscience. Still further, the quantum picture of the world incorporates a source of genuine randomness in the development of physical events, which some view as providing a window for mental events to exhibit influence on physical processes.

The interpretation of all of these maddening possibilities is an enterprise which has been under way since quantum theory was first developed, yet those engaged in it are far from reaching a consensus. Although it has invited much new-age quackery, many serious books, such as Hodgson (1988), Lockwood (1989), Penrose (1994) and Stapp (1993), explore the possibilities of incorporating consciousness into the scientific world-view via quantum mechanics. That two of the biggest mysteries facing science may be related to one another is a fascinating possibility. And while there is no question that the picture of the world that would be obtained by a revolution linking them would be vastly different to that currently held in the hallways of establishment science, there seems no *a priori* reason why it should no longer qualify as naturalistic.

The Old Problem

While all of the preceding offers the naturalist some hope, the problem of representing phenomenology remains. How will we ever be able to describe conscious experiences in such a way as to allow simple, law-like correlations with physical phenomena? A significant amount of work is now being carried out on the problem, with attempts being made to develop quantitative or at least objective schema for phenomenological representation. For ex-

¹⁴ The directionality of temporal causation is also questioned by quantum mechanics.

ample, Varela and Shear (1999) and Stanley (1999) suggest ways in which the vast range of experiences can be systematised and broken down. I will briefly describe two such avenues.

First, as mentioned above with regard to heat and cold, we noted that science is often able to proceed by positing continuity between what seemed, on first glance, to be unrelated phenomena. It would have been inconceivable several hundreds years ago for a *single* naturalistic conceptual framework to explain how contagious disease, plant life, insects, birds, fishes and mammals arose – the range of this planet’s life seemed so diverse and complex that only a personal explanation (i.e. God) could ever suffice. Yet, neo-Darwinism is, at the very least, a promising potential replacement candidate.

Could a similar thing be done with phenomenology? One point counting against is the fact that, with the contents of consciousness, there is no distinction between appearance and reality. Whereas the natural world’s complexity could be (and apparently was) an illusion, how could a subject be wrong about what *she experiences*¹⁵? If phenomenology cannot be re-defined, is there any other way forward? One alternative is to simplify by interpolation – to most of us, whose phenomenology includes every shade of grey, darkness seems a linear, one-dimensional property. But, if there are subjects who only ever experience charcoal black or fluorescent white, the two shades might well be considered as unrelated, much like we think of red and blue. Just as our phenomenal *range* is wider, yet its *representation* is simpler, perhaps we may discover through brain experimentation that certain seemingly heterogeneous phenomenal feels actually lie along a continuous scale.

However, this is a far less likely prospect when we come to consider thoughts, concepts and feelings, including those that are combined with our raw sense data to produce a meaningful internal image of the world around us. Here, we seem to need as many modalities of phenomenology as there are ideas that have occurred, or indeed could occur, to mankind – it is hard to see how there could be continuity between most human thoughts. Is

¹⁵ Although see Benson and Greenberg (1969) for evidence which may challenge this assumption.

there any way in which conceptual complexity could be reduced?

Consider an intentional action, such as the writing of a cheque to pay for some purchased groceries. Consider now a second action, described as follows: moving an object in a manner that past experience has suggested will make marks on a piece of paper, where it has been understood that these marks will be interpreted by an organisation of individuals in order to transfer a certain quantity of money from my 'pile' to their 'pile', which they deserve considering that I have taken some food from their 'place' to my 'place'. Are these two action descriptions recognisably identical? If so, we have gone some way towards intentional atomisation – breaking down a complex thought into simpler elements. And if this could be done for *all* our intentions then we may have a way of reducing their tyrannical complexity to perhaps a few dozen basic principles, each of which is likely to be grounded in the fairly mundane needs or experiences of a very small child. Such a representation might afford easier matching with events as physically described, both in terms of complexity and linguistic description – thus weakening Davidson's point about the "disparate commitments of the mental and physical schemes".

But all this is highly speculative. Even if these processes of interpolation and atomisation became effective, we would still be left with a fairly complex and anomalous representation of phenomenology and it *still* is not clear how this could be matched with physical events in any manner worthy of the title 'science'. Perhaps this is one part of science which will turn out to be interminably complex – by objectifying all phenomena until now, science has put off all the truly difficult problems until it felt ready to tackle consciousness itself. And now it is beginning to rise to the challenge, it faces the resurgence of all those old explanatory loans from which it had borrowed so much time.

The Theological Alternative

Perhaps, then, it is best for man to admit his limits and accept that science will never explain the phenomenon of consciousness. The only alternative, if we seek explanation at all, is

to posit a personal one, in terms of an all-powerful deity able to sustain the complex correlations between our bodies and our minds. I don't see how one could refute such a move, except by saying that an argument for God based on the inadequacy of today's science seems a highly risky theological proposition.

Kitty Ferguson compares the problem of consciousness with the cosmological argument for first cause (1994, 183):

Saving belief in God by talking about what science hasn't been able to explain, or looks unlikely to explain, is skating on thin ice... On the other hand it is no more intellectually viable to save *unbelief* solely on the assumption and hope that science will inevitably be able eventually to explain everything. We've allowed ourselves a stand-off on the grounds of 'It remains a mystery' in the First Cause contest between God, mathematical and logical consistency, and the universe... At present we also have no other choice but to allow a stand-off regarding an explanation for the human mind.

And if science *were* ever to explain consciousness, an argument from consciousness to theism would be no stronger than one from physics, chemistry or biology. And the force of the latter lies in the cosmological argument and the argument from design, not in any specific natural theology of the mind.

REFERENCES

- Adams, R. M. (1987) 'Flavours, Colours and God', in *The Virtue of Faith and Other Essays in Philosophical Theology*, Oxford: Oxford University Press, pp. 243–62.
- Alston, W. and Smythe, T. (1994) 'Swinburne's Argument for Dualism', *Faith and Philosophy*, vol. 11, pp. 127–33.
- Armstrong, D. (1978) 'Naturalism: Materialism and First Philosophy', *Philosophia*, vol. 8, pp. 261–76.
- Austin, J. (1998) *Zen and the Brain*, Cambridge, MA: MIT Press.
- Benson, D.F. and Greenberg, J.P. (1969), 'Visual Form Agnosia', *Archives of Neurology*, vol.20, pp. 82–9.
- Block, N. (1995) 'On a Confusion about a Function of Consciousness', *Behavioural and Brain Sciences* vol. 18, pp. 227–87.
- Chalmers, D. (1995) 'Facing up to the Problem of Consciousness', *Journal of Consciousness Studies* vol. 2, pp. 200–19.
- (1996) *The Conscious Mind*, Oxford: Oxford University Press.
- Cottrell, A. (1999) 'Sniffing the Camembert: on the Conceivability of Zombies', *Journal of Consciousness Studies*, vol. 6, pp. 4–12.
- Crick, F. and Koch, C. (1990) 'Towards a Neurobiological Theory of Consciousness', *Seminars in the Neurosciences*, vol. 2, pp. 263–75.
- Davidson, D. (1970) 'Mental Events', in Foster and Swanson (eds.), *Experience and Theory*, London: Duckworth & Co., pp 79–101.
- Dennett, D. (1991) *Consciousness Explained*, London: Viking.
- (1995) 'The Unimagined Preposterousness of Zombies', Commentary on T. Moody, O. Flanagan and T. Polger, *Journal of Consciousness Studies*, vol. 2, pp. 322–6.
- Descartes, R. (1985) *The Philosophical Writings of Descartes*, vol. 1, trans. Cottingham, Stoothoff and Murdoch, Cambridge: Cambridge University Press.
- Eccles, J. and Popper, K. (1977) *The Self and its Brain*, London: Springer-Verlag.
- Ferguson, K. (1994) *The Fire in the Equations*, London: Bantam Press.
- Griffin, D. R. (1998) *Unsnarling the World Knot*, Berkeley, CA: UC Press.
- Güzeldere, G. (1997) 'The Many Faces of Consciousness', in Block, Flanagan and Güzeldere (eds.), *The Nature of Consciousness*. Boston, MA: MIT Press, pp. 1–67.
- Hasker, W. (1998) 'Swinburne's Modal Argument for Dualism: Epistemically Circular', *Faith and Philosophy*, vol. 15, pp. 366–70.
- Hodgson, D. (1988) *The Mind Matters*, Oxford: Oxford University Press.
- Hume, D. (1739) *A Treatise of Human Nature*.
- Jackson, F. (1982) 'Epiphenomenal Qualia', *American Philosophical Quarterly*, vol. 32, pp. 127–36.
- James, W. (1892) *Psychology: The Briefer Course*, Allport (ed.). New York: Harper and Row 1961.
- (1950) *The Principles of Psychology* (2 vols). New York: Dover Publications.
- Johnston, P. (1993) *Wittgenstein: Rethinking the Inner*, London: Routledge.
- Kripke, S. (1972) *Naming and Necessity*, Cambridge, MA: Harvard Press.
- Levine, J. (1983) 'Materialism and Qualia: The Explanatory Gap,' *Pacific Philosophical Quarterly*, vol. 64, pp. 354–61.

- (1998) 'The Conscious Mind: Book Review', *Mind*, vol. 107, pp. 877–81.
- Locke, J. (1959) *An Essay Concerning Human Understanding* (2 vols). New York: Dover Publications.
- Lockwood, M. (1989) *Mind, Brain and Quantum*, Oxford: Blackwell Publications.
- Mackie, J. (1982) *The Miracle of Theism*, Oxford: Oxford University Press.
- McGinn, C. (1989) 'Can We Solve the Mind-Body Problem?' *Mind*, vol. 98, pp. 349–66.
- Moreland, J. P. (1998) 'Searle's Biological Naturalism and the Argument from Consciousness', *Faith and Philosophy*, vol. 15, pp. 68–91.
- Nagel, T. (1974) 'What is it like to be a Bat?' *Philosophical Review*, vol. 4, pp. 435–50.
- (1979) 'Panpsychism', in *Mortal Questions*, ch. 4.
- Olding, A. (1981) *Modern Biology and Natural Theology*, London: Routledge.
- Penrose, R. (1994) *Shadows of the Mind*, Oxford: Oxford University Press.
- Reames, K. (1999) 'A Response to Swinburne's Latest Defence of the Argument for Dualism', *Faith and Philosophy*, vol. 16, pp. 90–7.
- Ryle, G. (1949) *The Concept of Mind*, London: Hutchinson.
- Searle, J. (1980) 'Minds, Brains and Programs', *Behavioural and Brain Sciences*, vol. 3, pp. 417–24.
- (1997) 'Consciousness and the Philosophers', *The New York Review*, March 6, 1997, pp. 43–50.
- Shoemaker, S. and Swinburne, R. (1984) *Personal Identity*, Oxford: Blackwell.
- Sperry, R. (1968) 'Mental Unity Following Surgical Disconnection of the Cerebral Hemispheres', in *The Harvey Lectures*, New York: Academic Press, pp. 293–323.
- (1974) 'Lateral Specialisation in the Surgically Separated Hemispheres', in Schmitt and Worden (eds.), *The Neurosciences Third Study programme*, Cambridge, MA: MIT Press, pp. 5–19.
- (1994) 'Consciousness, Personal Identity and the Divided Brain', *Neuropsychologia* vol. 22, pp. 661–73.
- Stanley, R. (1999) 'Qualia Space', *Journal of Consciousness Studies*, vol. 6, no. 1, pp. 49–60.
- Stapp, H. (1993) *Mind, Matter, and Quantum Mechanics*, Berlin: Springer-Verlag.
- Stump, E. and Kretzmann, N. (1996) 'An Objection to Swinburne's Argument for Dualism', *Faith and Philosophy*, vol. 13, pp. 405–12.
- Swinburne, R. (1979) *The Existence of God*, Oxford: Oxford University Press.
- (1986) *The Evolution of the Soul*, Oxford: Oxford University Press.
- (1996a) *Is There a God?* Oxford: Oxford University Press.
- (1996b) 'Dualism Intact', *Faith and Philosophy*, vol. 13, pp. 68–77.
- (1996c) 'Reply to Stump and Kretzmann', *Faith and Philosophy*, vol. 13, pp. 413–4.
- (1998) 'The Modal Argument is not Circular', *Faith and Philosophy*, vol. 15, pp. 371–2.
- Taliaferro, C. (1994) *Consciousness and the Mind of God*, Cambridge: Cambridge University Press.
- Varela, F., Thompson, E. and Rosch, E. (1991) *The Embodied Mind: Cognitive Science and Human Experience*, Boston, MA: MIT.
- Varela, F. and Shear, J. (eds.) (1999) *The View From Within*, Thoverton, UK: Imprint Academic.