John Cornwell (ed) *Consciousness and Human Identity* (OUP, 1998) 212-224.

Binding the Mind

Peter Lipton

Several of the essays in this collection discuss the `binding problem', the problem of explaining in neurophysiological terms how it is that we see the various perceptual qualities of a physical object, such as its shape, colour, location and motion, as features of a single object. The perceived object seems to us a unitary thing, but its sensory properties are diverse and turn out to be processed in different areas of the brain. How then does the brain manage the integration? Readers of the essays in this collection may find themselves suffering from an analogous binding problem about the study of consciousness, though this problem is conceptual rather than perceptual, and here the difficulty is to achieve the integration rather than to understand how an effortless integration is achieved. Consciousness is the ideal topic for inter-disciplinary investigation. It is a central concern of such diverse disciplines as neurophysiology, evolutionary biology, psychology, cognitive science, philosophy and theology, among others, yet none of these disciplines has come close to providing full answers to the central questions that consciousness raises. Inter-disciplinary investigation seems an obvious way forward, but it generates the conceptual binding problem that this collection displays. The standard of the essays is very high, but it is extraordinarily difficult to integrate their content into anything like a single picture. We are all apparently talking about the same phenomenon, the conscious awareness of the world that each of us enjoys first-hand, but it is quite unclear how to see the very different things we say about this phenomenon as part of a single picture, or even as parts of different but compatible pictures.

Having raised the binding problem for the inter-disciplinary study of consciousness, I hasten to say that I will not attempt even a partial substantive solution here: that is left as an exercise for the readers of this book. What I would like to do, however, is to suggest some ways we may make progress on the problem by seeing the sort of structure its correct solution might have. Since my own specialty is the philosophy of science, this essay will focus on the relation between philosophical and scientific

approaches to consciousness. One moral will be that part of the binding problem may be solved through unbinding: there should be a division of labour between scientists and philosophers, where scientists focus on the causal explanations of consciousness, leaving the recalcitrant problem of what consciousness itself is to the philosophers, who have had more practice at bumping their heads against brick walls. Another moral will be that apparently incompatible explanations of consciousness offered by different disciplines are sometimes compatible after all. The illusion of incompatibility arises because what appear to be different and incompatible answers to the same question are really different but compatible answers to different questions. I will illustrate one way this situation arises by describing some philosophical work on the structure of scientific explanations, work that brings out the importance of the contrastive structure of many why-questions.

Explaining Correlations

We can begin with something that might be accepted by every contributor to this volume: there are correlations between types of brain activity and types or aspects of conscious states. Neurological research has uncovered some of these correlations, and there are presumably many more to come. A correlation, however, can be explained in many different ways. Here are four possibilities. First, it could be a mere coincidence, though the likelihood of this will fall as the known extent and frequency of correlation increases. I will assume that many mind-brain correlations are not mere coincidences. Second, the correlated states may be effects of a common cause. Thus the sound of thunder is correlated with the flash of light, and this is no coincidence. The sound does not, however, cause the light; rather they are both effects of the electrical discharge in the clouds. Some correlations between brain states and conscious states are of this sort, but again I assume that not all of them are. The third possibility is genuine causation, where one of the correlated states causes the other. This is illustrated by the correlation between the electrical discharge and the sound, and also by the correlation between the electrical discharge and the light. The final explanation of correlation I will consider is identity, where the states are correlated because they are in fact one in the same state. There is an excellent correlation between the presence of photons and the presence of light, between molecular motion and temperature, between

H2O and water, and so on. The explanations of these correlations is not that photons cause light, that molecular motion causes temperature, or that H2O causes water. They are rather that light just is a stream of photons, temperature just is molecular motion, and water just is H2O. My observations about the division of intellectual labour on the problems of consciousness will focus on the differences between explaining mind-brain correlations in terms of causation and explaining them in terms of identity.

If we aim to explain consciousness, which sort of explanation should we seek? Should we aim for a causal explanation or an identity explanation, an account of what brings about conscious states or an account of what those conscious states themselves are? The traditional mind-body problem focusses on the identity question: are conscious mental states themselves physical or not? The physicalist claims that conscious states just are physical states; the dualist claims that they are not, however tight their causal links to the physical may be. Whatever answer we give to the identity question, however, the causal question remains. Whether mental states are physical or not they have, I take it, physical causes, and the question is what those causes are and how they operate. So what we really want are both sorts of explanation of consciousness. This is, however, where I suggest a division of labour. Scientists should tackle the causal question, leaving the identity question to torment the philosophers.

My reason for warning scientists off attempting identity explanations of mind-brain correlations is not that scientists are in general ill-equipped to discern identities. On the contrary, scientists are the preeminent discoverers of deep and informative identities, as my examples above of light, temperature and water illustrate. Moreover, both scientists and laypeople are in general surprisingly good at determining whether a particular correlation is due to coincidence, common cause, causation, or identity. Philosophers and psychologists are rather less good at saying just how we manage this cognitive feat, but some of our techniques are relatively clear. For example, if we want to determine whether a correlation obtains because of causation or because there a common cause, we may manipulate the one type of correlated state to see whether there is a reaction in the other. Barometer readings may be well-correlated with subsequent weather, but unfortunately we find that artificially changing the readings does not change the weather, so barometers don't cause the weather. We also have some relatively straightforward techniques for

3

distinguishing causal from identity correlations. For example, if we find that correlated states have different locations, we know the correlation cannot be a matter of identity, since `two' states can only be identical if they share all their properties. Causally connected states may be separated in space and time; identical states cannot be. Nevertheless, conscious mental states have a number of peculiar features that make the empirical determination of identity claims especially difficult. Many of these features are discussed in other essays in the volume, but it is perhaps useful to bring some of them together here. My aim is not to show that the physicalist's identification of conscious and physical states is mistaken, but just that the issue is contentious and strangely resistant to the normal empirical techniques that scientists deploy.

Identity Problems

There are several ways of bringing out the difficulties of a physical identification. One is to point to features of consciousness that no purely physical state could apparently possess. It is tempting to say here that there is one core feature that meets this description, and it is simply the consciousness of consciousness, the fact that experiences are, well, experiences. How could this experience I am now enjoying literally be a neural state? The problem, for those who take it to be a problem, is not just that the identity claim seems false, but that we do not understand what it would be for it to be true, which is to say that we do not understand what the identity claim means. This situation is quite different from that of the standard scientific identities mentioned above. You may be amazed to be told that temperature is molecular motion, you may even deny it, but you understand what is being claimed. The mind-brain identity claim, by contrast, does not seem even to make sense.

Or so say some. Other thoughtful people find no difficulty here, and accuse those who claim that the very state of being conscious could not be a feature of any purely physical system of begging the question. Why should consciousness not just be, say, the synchronous firing of certain neurons at a certain rate? In my view the claimed difficulty in understanding the identity claim does not in fact beg the question, but it obviously will not move anyone who does not feel it. So those who are troubled by the identity claim have looked for other more specific and more clearly recalcitrant features of conscious states. Some have pointed to the fact that many conscious states have content: they are representational, they are about something. Perceptual states are a good example. When someone sees a chair, we can distinguish between the perceptual states itself and what that state is about, its representational content. Strangely, perhaps, this contrast can be made even if the perception is an hallucination. Even here, the perception points beyond itself to a chair which, in the case of hallucination, just happens not to exist. Physical states, it is claimed, cannot have this feature. A piece of chalk, for example, has many properties, but it is not about anything, it does not represent, something only mental states can do. To this the physicalists may respond by pointing to the marks on the board that the chalk can be used to produce. Those marks are purely physical, yet being sentences, serve to represent. Then the dualist may reply that the physical representation is real but inevitably parasitic on the originating representational powers of the minds that perceive the marks. And so the debate continues.

A second claimed peculiarity of conscious states is that they are subjective and perspectival. Unlike physical states, conscious states can only exist in virtue of being experienced, and their existence thus incorporates a certain point of view, the point of view of the creature having the experience. A third and related peculiarity is the so-called asymmetrical access that we have to conscious states, again apparently unlike any physical states. The physical world includes things like chairs that are publicly observable, and things like electrons that are unobservable, and everything physical is either one or the other. Conscious states seem to be neither. My conscious states seem to be immediately accessible to me, but to no one else. To other people, my conscious states seem rather like electrons, whose existence can only be inferred, whereas I seem to have a more direct and quite different sort of access to my own consciousness, as you do to yours.

All these features of consciousness -- its very existence, its representational content, its subjective and perspectival nature, and our asymmetrical access to it -- have been claimed to stand in the way of any identification of conscious and physical states. Another way of bringing out the difficulty of identification is to stress the apparent contingency of any correlation between mental and physical states. No matter how much physical information is provided, it never seems necessary that consciousness should come along with it. Similarly, whatever behavioural or evolutionary story one describes, it never seems necessary that it be accompanied by consciousness. It always seems possible that the same behaviour or the same evolutionary history might have existed without any accompanying or resulting experience. The problem is completely general: we are given a sophisticated biological story, and then it is claimed that this or that behaviour could only be performed if the creature were conscious. That is the part that never seems convincing: it always seems possible that the behaviour (or the neural activity, or the evolutionary history) could exist without the existence of the conscious states. This persistent contingency stands in the way of seeing how we can identify mental and physical states, and because the feeling of contingency remains not just for the physical facts we now know, but for any physical facts we can imagine, the resolution of the difficulty does not appear to be one that further scientific research could resolve.

Keeping to Causation

The moral I draw from the formidable difficulties in making out a mind-brain identity claim is not that physicalism is false: perhaps they can be overcome. In any event, the dualistic alternatives to physicalism, which hold that mental states have an irreducibly non-physical component, face many difficulties of their own. The moral is rather that scientists should leave the identity question to the philosophers. It is not the sort of question that empirical inquiry is suited to solve, and it is precisely the sort of question that is the philosophers' business. This is not of course to say that philosophers will ever provide a definitive answer, but philosophers, unlike scientists, make their business out of the insoluble.

Many scientists are aware of the difficulties I have flagged concerning the identity of consciousness and are only too happy to leave them to the philosophers, but some overreact, taking the view that consciousness is not a fit subject for science. This is a mistake. Scientists should study consciousness, but they should seek causal explanations, not identity explanations. For the difficulties disappear or are at least substantially diminished if we shift our attention from the question of what mental states are to questions of their etiology.

Why should the situation improve so dramatically if we ask causal questions? The identity difficulties arise, as we have seen, because

conscious states seem to have features that no purely physical system could possess, and because any connection between physical and conscious states appears to be contingent, unlike the necessity that the identity of a thing with itself requires. Neither of these points is a barrier to causal connection. There is no presumption that causes resemble their effects. Pushing buttons and flicking switches cause an enormous variety of effects, with properties not to be found in the switch or button itself. Moreover, the fact that a cause might not have produced its effect provides no reason to say that it was not really a cause. The perennial possibility of interference and breakdown make this clear, as switches and buttons again illustrate. Perhaps there is no behaviour we perform which we could not imagine being performed without consciousness, but we know that we do these things with conscious states, and it is up to the scientists to tell us what in our brains causes those states to occur. By sticking to the causal questions, scientists will also avoid making identity claims that appear to explain consciousness only by explaining it away, denying its essential experiential component, and avoid the sort of implausible `nothing but' reductionism that is criticised in a number of the essays in this book.

John Searle is another champion of the scientific study of the causes of consciousness, as his essay shows. There is, however, an important difference between us. Searle holds that causes and effects need not be distinct, and that scientists ought to look for causes of mental states that also tell us what those states are, causes that answer the identity question. Thus he must reject my proposed division of labour, where the scientists leave the identity question to the philosophers. According to Searle, scientists ought to be answering both questions. So it may be useful for me to say just where I disagree with him and why.

Searle makes a helpful distinction between two versions of the subjective/objective distinction -- an epistemic version and an ontological version -- and uses that distinction to reveal a fallacy in a common argument against the possibility of scientific study of consciousness. The fallacious argument is that science is objective, consciousness is subjective, so science can't study consciousness. The source of the fallacy is an equivocation between the two versions of the subjective/objective distinction. Science is objective in the epistemic sense: its investigations lead to intersubjective agreement about objective facts. Consciousness is subjective, in the ontological sense, which is to say that consciousness states must be experienced to exist. Ontological subjectivity does not however entail epistemic subjectivity, so the conclusion of the argument doesn't follow. It remains possible to have an epistemically objective, scientific investigation into the nature of ontologically subjective conscious states.

I agree with Searle that the argument against the possibility of the scientific study of consciousness is fallacious, and for the reason he gives. I also agree about the nature of the subjectivity of consciousness: it is ontological, not epistemic. But it is this ontological subjectivity and its consequences that are at the basis of the identity difficulties I have mentioned. Searle, by contrast, holds that these are only difficulties in the context of one model of identities: there is another model for which they do not arise. He illustrated the first model with the example of the identity of heat and molecular motion, the second with the identity of solidity and the vibratory movement of molecules in lattice structures. I disagree with Searle here, because I think that the difficulties arise for both models.

Searle claims, rightly in my view, that the identity of heat and molecular motion is a bad model for consciousness, because the truth of that identity depends on conceiving of heat in a way that disassociates it from the feeling of heat. (This point has also been emphasised by Saul Kripke.) What makes this identity unproblematic is that we sharply distinguish heat or temperature from the conscious sensation of heat. It is not the sensation that is claimed to be identical to molecular motion. (For all I know, the molecules in my brain move more rapidly when I touch something very cold than when I touch something whose temperature is close to that of my own body.) The identity quite correctly treats heat or temperature as something out there, stripped of the feeling, and this makes possible the strict identification with the physical state. Many other scientific identity claims work in the same way. This strategy of stripping off the feeling is however obviously hopeless if it is the feeling that we are trying to identify, which is exactly what we are trying to do if we attempt a physical identification of consciousness. This is then our dilemma: if we strip off the feeling, we will lose the very thing we are supposed to be identifying; if we leave the feeling in, we cannot see how the identity could hold.

Searle, however, rejects the second horn of this dilemma, suggesting that we can understand a proper mind-brain identity if we use a different

analogy, such as the identity between solidity and vibration within a lattice. This is where we part company. The point that Searle and others have made about heat and molecular motion seems to me to apply equally to cases such as solidity. It may be that solidity is in some important ways a different case from heat. Perhaps solidity is a more `holistic' property than heat: for example, it makes no sense to say that a single molecule is solid. On the other hand, it may not make sense to say that a single atom has a particular temperature either; perhaps only an ensemble of atoms can have a temperature. An interesting difference between solidity and heat may be that solidity depends on physical structure (the lattice) in a way that heat does not. In any case, the salient similarity remains. Just as in the case of heat and molecular motion, we only understand the identification of solidity with vibration within a lattice by shaving off the feeling, in this case the feeling of solidity or resistance to pressure. The physical identity works by shaving off the feeling, whether it is the sensation of heat, or the sensation of resistance that solid objects afford. So the case of solidity does not help us to see how a physical identification of conscious states is possible. Nor would it help to chose another analogy, such as the identity between light outside the visible range and photons with certain specified energies, where there is no corresponding sensation to strip. This to brings us no closer to seeing how the peculiarities of consciousness could be tamed with a physical identification. This is the reason I continue to see a deep difference between the causal and identity questions, and support the division of intellectual labour.

I have argued that the metaphysical difficulties of the philosophers' mind-body problem are no bar to the scientific explanation of consciousness, where the explanations are causal. Some scientists have however held that such explanations are not to be had, for epistemological reasons. These are in my view bad reasons, and I now want to suggest why. One such reason is based on the equivocation over the claim that consciousness is a subjective phenomenon that Searle effectively exposes. Another concerns a feigned modesty about the scope of scientific inquiry. To see what this comes to, we may usefully return to the topic of correlations. As I have already observed, scientists are in general at least as sensitive as the rest of us to the distinction between the different explanations for correlations, such as coincidence, causation, and identity. Nevertheless, in their philosophical moments, some scientists claim that all

science can ever deliver are the correlations themselves, a position that would leave genuine causal explanations of consciousness outside their purview.

Why do they say this? Part of the explanation may be the legacy of logical positivism, a philosophical movement from the early part of this century with clear antecedents in the earlier history of philosophy. Positivism has attracted many philosophically inclined scientists, partially because, unlike most philosophical positions, it appeared to offer some useful methodological advice about scientific research, and perhaps also partially because it was a philosophy that held up science as the acme of human endeavour. Paradoxically, however, positivism has also lead some scientists drastically to understate what science can achieve.

The positivists maintained that only claims that have empirical consequences are really claims at all; they are the only sentences that say anything. If a sentence has no empirical consequences, it is just noise: it can not even rise to the level of being false. This is already an implausible view, but the positivists went on to say something much more implausible. They said that the meaning of a statement that does have empirical consequences just is those empirical consequences. One of the slogans of the time was, `The meaning of a statement is its method of verification'. This is a serious mistake, a confusion of the claim one is making with the evidence one might have for that claim. This confusion arises in many different areas in the history of science in this century, and certainly in psychology. As Mary Midgely points out in her essay, behaviourism is a particularly pure example of the confusion. The positivist view that all that a claim can describe is the evidence, combined with the view that the only evidence for other people's mental states is their behaviour, yields behaviourism. This is the position that talk about mental states is really just talk about behaviour. This position, along with one its major weaknesses, is summed up in the familiar story about the two behaviourists who meet in the street. One says to the other, 'You're fine. How am I?'

Behaviourism is not in fashion in psychological circles these days, but it is not a dead horse. The positivistic impulse remains among many philosophers and scientists, though it tends to find less obvious expression than it did in the case of behaviourism. It is the impulse to confuse the evidence with what it is evidence for. Only what is in some sense observable can be scientific evidence, but scientists can have evidence for what is unobservable. Here for once physics provides a good model. What physics shows us is that we can have good evidence for claims that go way beyond not only that we in fact observe but what is even in principle observable. Conscious states may not be inter-subjectively observable, and perhaps there is even an interesting sense in which causal relations are in general unobservable, but these are not good reasons to deny that science can study the causes of conscious states, any more than the unobservability of atomic interactions provides a reason to deny that physicists may profitably study them.

Explanatory Pluralism

Having made my pitch for the causal explanation of consciousness, I want to conclude with a few observations about the nature and structure of causal explanations. The papers in this book make a number of important contributions to this subject. Many of them bring out what we might call the pluralism of causal explanation. Behind every event, and so behind every conscious state, there is an long and dense causal history. No explanation can capture all of it, and no explanation needs to. But not just any cause will serve to answer a particular question about a given event. The big bang is part of the causal history of every event, but does not explain all of them. So we need some sense of what conditions a cause has to satisfy to answer a particular question. This is the problem of causal selection. A sensitivity to the plurality of causal explanations also helps us to see that different explanations of consciousness which may seem competitors are sometimes really compatible: they are just answering different questions.

The contributors to this book develop the themes of causal selection and explanatory pluralism in a number of ways. Margaret Boden brings out the contrastive `rather than' structure of many causal explanations, arguing that adequate explanations of conscious states must specify causes that explain why the state takes one form rather than another. Michael Sofroniew emphasises the importance of providing explanatory mechanisms, not just isolated causes. He also makes an important point about the `why-regress', a feature of virtually all explanations. Most of us discovered this feature as young children, although some of us have since forgotten it. Whenever somebody answers your first why-question, you can just respond by asking why again. That is one of the ways we learned to torture our parents. Sofroniew's point, however, is that the first answer can be a good answer on its own, even if an answer to the second or later why's is not available. There is a regress in explanation, but it is benign. Causes can explain effects even if we are ignorant about the causes of the causes. One of the many aspects of explanation that Stephen Rose discusses in his essay concerns the way the plurality of explanation stratifies into different levels of explanation, and the importance of finding the level appropriate to conscious states and the questions we ask about them. The idea the multiplicity of levels of explanation is also developed by Mary Midgley and by Fraser Watts, both of whom stress the importance of offering explanations at different levels and the difficulty of seeing how these explanations could be integrated into a single picture of the etiology of consciousness.

All of these observations raise important issues for the explanation of consciousness, but here I will just say a bit more about one of them, the importance of the contrastive structure of causal explanations. As I have already mentioned, different explanations of consciousness sometimes give the illusion of incompatibility when they are really compatible but answering different questions. One of the reasons two such explanations may seem incompatible is that they are both answers to questions that could be phrased as 'Why X?' for the same X, and so seem to be answering the same question. Yet they are answering different questions. The contrastive structure of explanation shows how this is possible. In many cases, what we are really asking is not just, `Why X?', but rather, `Why X rather than Y?', and this may be a different question, requiring a different answer, than the question, 'Why X rather than Z?'. This is another one of those deep truths about explanation that we learned as children. `Why do birds fly south in the winter? Because it is too far to walk.' It spoils the joke to explain it, but what this and many other children's jokes depend upon is an unexpected contrast switch. The intended question was, Why do birds fly south in the winter rather than stay where they are?', whereas the reply answers the different question, 'Why do birds fly south in the winter rather than walk?' Both questions cite the same fact -- birds fly south in the winter -- but they use different foils, requiring different answers. The two answers are compatible, but they answer different questions. In this example the joke works because the difference in contrast is so obvious, but for many of the more intellectually demanding

why questions we ask this is not so. We focus only on the fact, and then reject someone else's answer because it doesn't answer our real question, without realising that it may nonetheless be a legitimate answer for someone asking a different contrastive question about the same fact. We can see these different answers as compatible, once we appreciate the aspect of the plurality of causal explanations that the relativity of explanation to contrast reveals.

In addition to showing how apparently incompatible explanations can sometimes be compatible, an appreciation of the way contrastive explanation works shows how even a very partial causal story can sometimes provide an adequate explanation. The case of syphilis and paresis illustrates the point. Paresis can only be caused by tertiary untreated syphilis, but the vast majority of people with tertiary untreated syphilis fortunately do not go on to contract paresis. This causal situation has provided the scene for a raging dispute in the philosophical literature on explanation. The question is whether the fact that a person has paresis can be explained by pointing out that he previously had tertiary untreated syphilis. Some philosophers say yes, on the grounds that the syphilis is a cause of the paresis and causes explain effects. Other philosophers say no, on the grounds that the paresis could not be deduced, or even shown to be probable, on the basis of the syphilis, since most people with syphilis do not go on to get paresis. The syphilis is a cause, but it is too small a part of the total cause to be explanatory. The dispute can be settled if we bring in contrastive questions. If the question is, 'Why did Smith rather than Jones get paresis?', and Jones does not have syphilis, then Smith's syphilis provides a good answer. If on the other hand the question is, Why did Smith rather than Doe get paresis?', where both of them had syphilis, Smith's syphilis is not explanatory. Both questions cite the same fact but a different foil, and the cause that answers one of the questions will not answer the other. This again illustrates the sensitivity of explanation to contrast, and shows how even a very partial cause can be explanatory. In many cases, all it takes to explain a contrasts is a cause that makes a difference between the fact and the foil.

The contrastive structure of explanations thus reveals that causal explanation is sometimes easier than it may at first appear. One benefit of seeing this in the context of the explanation of consciousness is that it helps further to alleviate the worry about the apparent contingency of the link between mental and physical states. I have already suggested that this worry is reduced if we switch from identity to causal explanations. It is further reduced by an awareness of the way explanations are sensitive to contrasts. Perhaps none of the scientific accounts now available really do explain why an organism has a particular conscious state rather than being an unconscious robot. Even if this is so, however, they may explain why the organism has that particular conscious state rather than another one. The second question may be easier to answer, because it presupposes what the first question asks about, namely that the organism is conscious at all. An awareness of the contrastive structure of explanation should also encourage us to look more carefully at our questions before we assess the answers, and in particular to take care not to ask one contrastive question and then proceed to answer another, lest we inadvertently offer the grow-up equivalent of the joke about the birds.

My discussion of contrastive explanation takes me back to the conceptual binding problem with which this essay began, the problem of integrating the diverse inter-disciplinary approaches to consciousness. The request for integration has two parts and I have responded to them differently. The first part is a request to be shown how the different answers to questions about consciousness are compatible. Of course some of them just are not compatible, but I have suggested that there is more widespread compatibility that may at first appear, and that we can see this by distinguishing different types of explanation. We need to distinguish causal explanations from identity explanations, and we need to use the contrastive structure of explanation distinguish the many different causal questions that may all be questions about the same facts. By making these distinctions, we will find compatibility in some places where there initially seemed to be competition. The second part of the request for integration asks for more: it asks that we show not just that different answers are compatible, but that we fit them all together into a single, unified account of the mind. This is the request that I have resisted. At our present stage of understanding, what is needed is division of labour, not superficial attempts at a global picture. The disciplines concerned with consciousness have plenty to learn from each other, as the essays in this book show, but in my judgement the way forward for the foreseeable future is through the simultaneous development of diverse approaches at different levels, rather than through the attempt at a single, big picture.