Elizabeth Schier
Department of Philosophy, Macquarie University, Australia

Glenn Carruthers
Discipline of Philosophy, Adelaide University, Australia
Department of Philosophy, Macquarie University, Australia
glenn.carruthers@adelaide.edu.au
Ph: +61423974275
Fax: (61 2) 9850 8892

Our Intuitions about Consciousness are Inconsistent

Our Intuitions about Consciousness are Inconsistent

Abstract

Two compelling intuitions about consciousness are that i) there is no appearance/reality distinction for phenomenal character and ii) phenomenal character plays a causal role in cognition and behaviour. Process theories of consciousness, such as Dennett's multiple drafts theory, secure the first intuition but must give up the second. In contrast vehicle theories, such as that offered by O'Brien and Opie, secure the second intuition but must give up the first. These approaches exhaust the options available to the cognitive scientist for explaining consciousness. Hence on the cognitive science framework at least one of our intuitions about consciousness must be given up.

Key Words: consciousness, phenomenal character, intuitions, cognitive science

Introduction

In this paper we argue that our intuitions about consciousness are inconsistent. Our focus is on that particularly problematic kind of consciousness known variously as 'phenomenal consciousness', 'phenomenal character', 'qualia', 'phenomenality', 'phenomenal quality' et cetera. 'Phenomenal character'[1] suits our purposes best so it is the term we shall use. Two compelling intuitions about phenomenal character are i) that there is no appearance reality distinction for phenomenal character and ii) that phenomenal character plays a distinct causal role in cognition. We argue here that on the cognitive science framework for explaining consciousness it is possible that at most one of these intuitions about phenomenal character is true. On functionalist or process theories of phenomenal character it is true that there is no appearance reality distinction for phenomenal character; however these theories make phenomenal character epiphenomenal. Conversely on vehicle theories phenomenal character plays a causal role, however, on these theories there is a reality to phenomenal character beyond how it appears to the subject. The upshot of this is that intuition based

[1] Note that this discussion concerns phenomenal character per se. There are many emerging distinctions regarding our concept of consciousness which point to senses of 'consciousness' which are distinct from the phenomenal sense (Block 1995; Block 2001; Block 2005b; Kriegel 2005; Natsoulas 1991; Natsoulas 1992; Natsoulas 1994a; Natsoulas 1994b). There are also accounts which are out to explain these other phenomena (for example Kriegel 2005; Metzinger 2003 present explanations of the "mineness" of experience). Suffice it to say that these other phenomena are not the focus of the explanatory argument and so are not the focus of this discussion.

It is important to note, however, that our discussion of access below is not a discussion of access as a distinct type of consciousness. Rather it is a discussion of explanations of phenomenal character in terms of access (O'Brien and Opie 1997).

theorising about consciousness is no longer viable as it is clear that our intuitions about consciousness are inconsistent. We will begin with some more detail about these intuitions.

Two intuitions about Consciousness

One compelling intuition about phenomenal character is that it is constituted wholly by appearances to a subject. That is, that there is no appearance/reality distinction for phenomenal character. This intuition has been seen as problematic for materialism: Nagel (1974) and Searle (1992) argue that our scientific methodology is incapable of dealing with phenomenal character because it is incapable of dealing with appearances. Phenomenal character is how things appear to us. But, the argument goes; the first step in any scientific explanation is to leave behind how things appear to us in order to get at how they really are. In providing a scientific account of things we move towards greater objectivity by characterising the phenomenon 'not in terms of the impressions it makes on our senses, but in terms of its more general effects and in terms of properties detectable by means other than the human senses' (Nagel 1974 pg 444). To achieve a scientific explanation of a phenomenon it is claimed that we must be able to characterise it in a way that transcends its appearances to us.

Searle uses the example of heat to explicate this argument. He suggests that our pretheoretical notion of heat concerns perceived temperature: something is hot if it feels hot and is cold if it feels cold (Searle 1992 pg 119). With scientific progress we discovered that these experiences of heat are caused by molecular motions and so we redefined heat to identify it with the underlying cause of our experiences: '"real" heat is now defined in terms of the kinetic energy of the molecular movements, and the

subjective feel of heat that we get when we touch a hot object is now treated as just a subjective appearance caused by heat, as an effect of heat (Searle 1992 pg 119). We can see in this characterisation of our understanding of heat, a move from appearance to reality. From how things appear to us to how they "really" are, independent of their appearance to us.

Nagel and Searle's worry is that this method of leaving behind how things appear to get at how they really cannot be applied to the phenomenal character of experience. It seems that the phenomenal character of experience *is* how things appear, so in leaving behind appearances we have left behind the explanadum. Nagel asks: 'Does it make sense, in other words, to ask what my experiences are *really* like, as opposed to how they appear to me?' (1974, p.448). The intuitive answer to this question is no, it doesn't make sense to draw a reality-appearance distinction for experiences. As Searle puts it 'we can't make that sort of appearance-reality distinction for consciousness because consciousness consists in the appearances themselves. *Where appearance is concerned we cannot make the appearance-reality distinction because the appearance is the reality*' (1992 pg 122). But if it doesn't make sense to ask what our experiences are really like, then it seems that we cannot obtain a scientific explanation of our experiences because science aims to capture the reality behind the appearances. Scientific explanations works by leaving behind how things appear to us in order to capture how they are in themselves, but the phenomenal character of experience is how things appear, so in leaving behind appearances we leave behind the phenomenal character of experience.

Nagel and Searle have different opinions regarding the consequences of this argument. Nagel thinks that this argument means that we currently do not know how materialism can be true, and that ‘‘it seems unlikely that any physical theory of mind can be contemplated until more thought has been given to the general problem of subjective and objective.’ (Nagel 1974 pg 450). He asks for new concepts of consciousness in order to solve this problem. We will see below that there are two viable materialist theories of consciousness that provide the necessary conceptual step forward. However, taking this step requires giving up at least one of our pretheoretical intuitions about consciousness.

In contrast Searle thinks that this argument has no deep metaphysical consequences whatsoever and is simply a ‘trivial consequence of the pragmatics of our definitional practices ... Consciousness fails to be reducible, not because of some mysterious features, but simply because by definition it falls outside the pattern of reduction that we have chosen to use for pragmatic reasons.' (Searle 1992 pg 123). Searle thinks that phenomenal character cannot be captured by the standard reductive scientific methodology, but that this is no sign that phenomenal character is not physical because we have set the methodology up to leave phenomenal character behind. Thus, the inability of science to capture phenomenal character is a reflection of how we do science, not a reflection of the non-physical nature of phenomenal character.

We side with Nagel on this issue (see also Bringsjord 1994; Garrett 1995). Although it is notoriously difficult to provide a definition of the physical (Jackson 1982), it is widely accepted that being amenable to scientific explanation is important in appreciating that a phenomenon is physical (Churchland 1984; Cottrell 1995; Dennett

1991; Jackson 1982; Levine 1998; Nagel 1974; Stoljar 2001).[2] In particular, it seems that obtaining an explanatory connection or a 'perspicuous nexus' (Cottrell 1995) between the phenomenon and other clearly physical phenomenon plays a key role in us to knowing that a phenomenon is physical (Churchland 1984; Cottrell 1995). This is suggested by the history of the study of life. At one point in time, 'vitalism' was popular. Vitalism is the idea that the property of being alive is a non-physical property and that what turns inanimate matter into animate matter is the presence of a non-physical vital spirit (Robinson 2006). However, as biology progressed and we obtained a micro-explanation of life, vitalism fell out of favour. As we obtained a scientific explanation of life, and saw how to build living organisms out of physical, inanimate matter, we changed from thinking that the property of being alive was non-physical to accepting that the property of being alive was physical. In other words, obtaining a scientific explanation of life vanquished a dualist account of life. Being able to provide a scientific explanation of a phenomenon seems to be crucial to understanding that the phenomenon is physical. Conversely not being able, in principle, to provide a scientific explanation of a phenomenon seems to suggest it is non-physical. In particular, it seems that if we were unable in principle to achieve a scientific explanation of phenomenal character then we would have reason to worry about our ability to know that it is physical. In contrast if we can obtain an explanation of a phenomenon that gives us good reason to suppose it is physical.

If obtaining a scientific explanation does play this role in seeing that a phenomenon is physical, then this argument to the effect that we cannot obtain a scientific

[2] Although as Chalmers (1996) points out being amenable to scientific explanation may not be sufficient for us to accept something is physical.

explanation of phenomenal character is an argument that we cannot know that it is physical. The deeper worry about this position is that one simple reason as to why we cannot know that a phenomenon is physical is because it is in fact not physical. So although Searle may be correct to say that the irreducibility of phenomenal character stems from our current definitional practices, he seems to be mistaken to think that this is trivial.

Nevertheless, we suggest that Nagel is correct to hold hope that changing our conceptions of consciousness may enable us to develop a form of scientific explanation that can accommodate appearances.[3] Indeed, in the next section we will see how there can be a materialist theory of consciousness on which there is no appearance/reality distinction for phenomenal character.

Another powerful, but less tricky intuition about consciousness is the intuition that phenomenal character plays a distinct causal role. This intuition is just the belief that how things seem makes a difference to one's behaviour. This is the intuition that tells us that we don't put our hand in the fire a second time because it *hurt*, that I can recognise the person in the mirror because it *feels like* it's me (Carruthers 2007), that the Capgras patient believes that his wife has been replaced by a robot because it *seems like* it's not really her (Young and Leafhead 1996) or that I went to the fridge to get a drink because I *felt* thirsty. This intuition not only features in our everyday

[3] We also suspect that Churchland (1996) is right to say that this explanatory pattern is an artefact of what the theories are trying to explain – physicists are not out to explain how phenomena appear to us in consciousness. Nevertheless, as we will see below, we can discover some interesting tensions between our pre-theoretical intuitions regarding consciousness by examining Nagel's and Searle's worry in more detail.

folk psychological explanations of behaviour, but also in many scientific accounts such as those mentioned above.

This intuition is to be contrasted to epiphenomenalism. In this context epiphenomenalism about phenomenal character is the position that a mental state being phenomenal or not makes no difference to the causal role it plays. The natural explanations of behaviour offered above are not available to the epiphenomenalist. We take it that epiphenomenalism about phenomenal character is a strongly counter intuitive notion.

We can see here two obvious and compelling intuitions about consciousness. Pretheoretically it seems entirely reasonable to think there is no reality to phenomenal character beyond how it appears to the subject of experience and that phenomenal character plays a causal role in the cognitive system. However we will argue below that given the range of theories of consciousness open to the cognitive scientist it is not possible for both of these intuitions to be true. At least one of them must be abandoned.

Theories and Intuitions of Consciousness

Cognitive science provides the best available framework for an explanation of the mind. This is most evident in the fruit that the framework has borne throughout psychology and the neurosciences. A central tenant of this framework is the hypothesis that the mind is (some sort or other of) a computer. The most comprehensive of computers to date takes computation to be a causal process involving representing vehicles where the trajectory of that process depends on the

content of the vehicles (O'Brien and Opie 2006). On this hypothesis there are two ways to explain mental phenomena, either in terms of the representing vehicles of the mind or in terms of the processes that those vehicles enter into. Phenomenal character is clearly mental; indeed it is the paradigm example of something mental. Unless we are to reject the explanatory resources of what has proven to be a very powerful explanatory framework, phenomenal character should be explainable with reference to the representing vehicles of the mind or the processes that those vehicles enter into (Atkinson et al. 2000; O'Brien and Opie 1999b). Call these vehicle and process theories respectively (O'Brien and Opie 1999b). These two types of theories of consciousness can also be distinguished by the way in which they answer Nagel's question: does it make sense to ask what my experiences are really like, beyond how they appear to me (Nagel 1974 pg 448)? Process theories say that it does not make sense whereas vehicle theories say that it does.

With specific reference to Dennett's (1991) process theory of phenomenal character we will show how we can secure the intuition that there is no appearance/reality distinction for phenomenal consciousness. However, we argue that accepting a process account of phenomenal character requires that we give up the intuition that phenomenal character does causal work. This is because process theories are a form of Functionalism about phenomenal character. With specific reference to O'Brien and Opie's (1999b) vehicle theory, we argue that on vehicle theories phenomenal character does do causal work. However, vehicle theories force us to give up the intuition that there is no appearance/reality distinction for phenomenal character. We will see that there are materialist theories of consciousness on which we can and

cannot draw the appearance-reality distinction for consciousness, but that both of these types of account require us to give up some of our pre-theoretical intuitions.

Accepting the Intuition that there is no Appearance/Reality Distinction for Phenomenal Character

In this section we will consider what sort of theory is available to the cognitive scientist who accepts the intuition that there is no appearance/reality distinction for phenomenal character. We need to begin by understanding the intuition in more cognitive terms. When there is an appearance-reality distinction to be drawn, we talk about some object, say a ripe tomato, and how it appears to a subject, say its red appearance to me. In this instance the reality is the tomato, and the appearance is constituted wholly by the perceptions of the subject. A corollary of this analysis of appearances is that it makes no sense to talk about an appearance unless it is an appearance *to* a subject. How can there be a way things appear without there being a subject to whom things are appearing? If we took subjects out of the world we would necessarily take out appearances as well. On this analysis appearances necessarily involve the subject.

What could it mean to say that *all* there is to a phenomenon is appearances? Usually we talk about a subject to whom an object is appearing. But if all there is is the appearance, then it seems that all we can talk about is the subject, there is no object of the appearance, akin to the tomato, which is the reality behind the appearance. To say that there is no appearance/reality distinction to be drawn for phenomenal character seems to mean that there are only the appearances to the subject and that there is no reality behind the appearances. Thus, to say that there is no appearance/reality

distinction for phenomenal character is to explain that character in terms of the subject – to put it in Block's (1995) terms, to explain phenomenal character in terms of accessibility to the subject.

It is widely considered that in order to provide a complete account of consciousness two phenomena that need to be explained are the phenomenal character of a mental state, and the accessibility of a particular mental state to the rest of the system. Some theorists (Baars 1988; Baars 1996; Baars 2002; Baars and Laureys 2005; Dehaene and Naccache 2001; Dennett 1991; Dennett 1993; Dennett 1995; Jack and Shallice 2001) take this as a single project and attempt to *explain* phenomenal character in terms of accessibility. As we will see these "access theorists" are providing a process account of consciousness. Dennett provides a particularly clear example of this strategy.

Dennett is particularly concerned to get rid of the problematic "Cartesian Theatre" view of consciousness according to which consciousness is a process whereby the objects of consciousness are displayed to the subject of consciousness. On the theatre metaphor the phenomenal players presents a play to us, the subjects of experience. Dennett demolishes the Cartesian Theatre by "Quining qualia" (Dennett 1988); by removing the phenomenal players and keeping the subject of experience (Dennett 1993). For Dennett, all there is to consciousness is the "judgements" of the subject as to what has gone on, there is no definitive performance about which the subject reporting. Instead there are multiple drafts, some of which influence behaviour and make it to memory and some of which don't. It is those drafts which have wide effects on the cognitive system that the subject is able to report, but these drafts are no more or less definitive than those which do achieve such effects and therefore are not

reportable: ‘those contents are conscious that persevere, that monopolize resources long enough to achieve certain typical and symptomatic effects – on memory, on the control of behaviour, and so forth’ (Dennett 1994 pg 61). It is thus a process account of phenomenal character (O'Brien and Opie 1997). It explains phenomenal character as the outcome of specific process over representing vehicles; namely those that produce the characteristic effects. Phenomenal character, for Dennett is cerebral celebrity; those contents which the rest of the system “knows about” are the conscious contents. On this process account a state is conscious if it achieves certain characteristic effects on the organism as a system (Dennett 1991 chapter 5; Dennett 2001 pg 224). These effects are i) becoming reportable ii) being remembered and iii) being used in action control. For Dennett having these effects is *constitutive* of the phenomenal character of a mental state.

On access accounts more generally one is subject to mental states that have certain characteristic effects on the wider system (Baars 1997 pg 306). Specification of what access processes matter may vary, for example for Rosenthal (1993; 1997; 2002) it is access to a higher order thought that matters. As may the mechanisms involved, Baars (1988; 1996; 1997; 2002) for example gives much more detail as to the mechanisms of access then Dennett does. Yet these theories maintain a common feature. Namely that all there is to the phenomenal character of a mental state is that it has these effects. These effects are constitutive of phenomenal consciousness. Thus, like Dennett, access accounts more generally explain phenomenal consciousness in terms of the processes certain representational vehicles are subject to. Access accounts of consciousness are process accounts.

We can see that for Dennett (and others who explain phenomenal character in terms of accessibility) we cannot draw an appearance-reality distinction for consciousness. On access theories of phenomenal character, all there is to phenomenal character are the appearances to the subject in the form of the effects on memory and behaviour. Returning to the Cartesian theatre metaphor, Dennett demolishes the stage, so there are no players which are appearing to the subject. There is no definitive answer to what story the players were *really* telling because there is no stage for them to tell their story. All there is are multiple drafts, some of which hang around long enough to have the necessary effects on the system that is, to be accessed by the subject. The reality of consciousness is the appearances, the effects on the subject of some content becoming a cerebral celebrity.

Importantly, contrary to Nagel and Searle, Dennett's theory of consciousness, on which it is impossible to draw an appearance-reality distinction for consciousness, makes consciousness amenable to scientific explanation and is a materialist theory. Given that consciousness is cerebral celebrity, then to study consciousness we just need to study the "gossip columns". If we want to know how things appear to the subject, then we ask the subject; we take their phenomenological reports seriously. For Dennett there is no possibility of error in this methodology because there are no real phenomenal states, there are no phenomenal players, which are appearing to the subject about which the subject could make a mistake. All there is to consciousness is the "judgements" of the subject as to what has happened, based on which contents won the fame game. To explain these effects, which are constitutive of phenomenal character, we need an account of the processes of access, such as that offered by Baars (Baars 1988; Baars 1996; Baars 1997; Baars 2002). On the model provided by

Dennett there is the possibility of a materialist account of phenomenal character on which there is no appearance/reality distinction.

So we can see that if we explain phenomenal character in terms of accessibility then we can have a materialistic account of consciousness on which we cannot draw an appearance/reality distinction for phenomenal character. Thus it preserves one intuition about consciousness. However, this theory is counterintuitive in another sense. It makes phenomenal character epiphenomenal. We will argue that Dennett's theory makes phenomenal character epiphenomenal in the same way that logical behaviourism makes the mental epiphenomenal. This is not to say that the theory is committed to a form of behaviourism, just that it is subject to the same objection.

For the logical behaviourist mental states are just dispositions to act (Graham 2007; Ryle 1949). For them to say I am in mental state, such as feeling hot, is just to say that I am predisposed to act in certain ways. For example that I will say 'I'm hot', change my clothes, turn on the air conditioning etcetera (Clark 2001 pg 165; Fodor 1981 pg 120). The problem with this analysis of mental states is that it makes them epiphenomenal (Fodor 1981 pg 121). Consider a simple explanation of a behaviour. Suppose one of us was to change their clothes, have a shower, turn on the air conditioning and say 'I'm hot'. Why might they do these things? One obvious explanation is that they *felt* hot. However, this explanation is not available to the logical behaviourist. For the logical behaviourist all there is to feeling hot is the disposition to change clothes, have a shower, turn on the air conditioning, say 'I'm hot' etcetera. The feeling of being hot is constituted entirely by the disposition to act in these ways. Thus, the natural explanation that one turned on the air conditioning

because one felt hot becomes the vacuous one turned on the air conditioning because one was predisposed to turn on the air conditioning. The general problem here is that if something is constituted entirely by its effects then it cannot explain those effects. That would be circular. Call this the behaviourist circle. This is a major conceptual problem for logical behaviourism. Logical behaviourism rules out the use of the mental in the explanation of behaviour. This makes the mental epiphenomenal.

Historically Functionalism about the mental has been considered a viable alternative to logical behaviourism. However, it also faces problems with the behaviourist circle. Functionalism defines mental states by their causal relations to other mental states and behaviour (Clark 2001 pg 168-169; Fodor 1981 pg 123). For example a mental state is the belief that it is raining because it causes me to say 'it is raining', to take an umbrella when I go out, to be disappointed that I can't play cricket, to remember later that it was raining and the like. Such causal relations are multiply realisable, that is, they can be implemented in systems other than brains. All that is needed to have the belief that it is raining is some system that can enter into these same relations (Fodor 1981 pg 124). Hence when defining a mental state, for example loving Adelaide United Football Club, the functionalist cannot make any reference to what that mental state is made of. The reason for this is that *the very same* mental state could be implemented in a completely different system. This love could be implemented in patterns of activity across neural networks or in electric impulses on Von Neumann architecture. As long as the appropriate causal relations are in place it is the same love (Fodor 1981 pg 126). For the functionalist the causal relations are constitutive of the mental state. In the case of our love for Adelaide United FC some of the relations are causing joy at wins, going to games, yelling at goals, hating Melbourne Victory et

cetera. It is the inclusion of relations to other mental states that significantly differentiates this position from logical behaviourism.

We can now see why Functionalism is also subject to the behaviourist circle. Why do we take joy in Adelaide United wins and hate Melbourne Victory? One obvious answer is that we love Adelaide United. However, this answer is not open to the functionalist as the hate and joy are constitutive of the love of our team. The natural explanation that we took joy in the win and hate Melbourne Victory because we love Adelaide United becomes the vacuous explanation that we took joy in the win and hate Melbourne because we take joy in wins and hate Melbourne. An effect cannot cause itself, so Functionalism is subject to the behaviourist circle (see Cummins 1996 especially pg 38-39 for this argument applied to functionalism about content, i.e. conceptual role semantics). Functionalism about mental states thus precludes the use of mental states in the explanation of thought and behaviour for the same reason as logical behaviourism.

We saw above that for Dennett all there is to the phenomenal character of a mental state is its having certain characteristic effects on the subject.[4] These effects are constitutive of phenomenal consciousness. All there is to the phenomenal character of a mental state is these effects. As such Dennett's theory can be classified as a functionalist theory, as the author himself notes (Dennett 1991 pg 31). Given the above problem for Functionalism of mental states, we may worry that Dennett's Functionalist account of phenomenal character is subject to the same worry.

[4] Note that Dennett doesn't need to commit himself to general Functionalism of mental states to provide a Functionalist account of consciousness.

Far from being concerned about this Dennett takes it as central to his theory. In responding to commentaries on his theory he says:

> Shoemaker seems to go along with the natural but treacherous assumption that reactive dispositions must involve the person reacting to a quale, presented somehow to the reactor, and causing, by its presentation, the reaction. (The given is then taken.) For instance, here's how pain works: the pain-networks produce (somewhere central?) the awfulness quale, which is then the very property to which "one" reacts with abhorrence. My view is that this confuses cause and effect; it is the reactions that compose the "introspectable property" and it is through reacting that one "identifies" or "recognizes" the property (Dennett 1993 pg 927).

And in his initial presentation of the theory:

> There is no reality of consciousness independent of the effects of various vehicles of content on subsequent action (and hence, of course, on memory) (Dennett 1991 pg 132)

Here it is clear that Dennett intends to identify the phenomenal character of a mental state with certain characteristic effects on the system. There is nothing to the phenomenal character of this state beyond these effects. Thus, this model is a Functionalist model.

It is the fact that Dennett offers a Functionalist account of phenomenal character that makes it subject to the behaviourist circle.[5] Again why might one of us take a shower, turn on the air conditioning, say 'I'm hot' and recall at a later time that they were hot? The natural answer is that they *feel* hot. This feeling being a particular phenomenal character of a mental state that represents something about their body, namely that it is hot. This natural answer is not available on Dennett's theory as on this account the feeling is wholly constituted by these behaviours, reports and memories. So the natural explanation of their taking a shower, turning on the air conditioning, saying 'I'm hot' and remembering that they were hot in terms of their feeling hot becomes the vacuous explanation that they took a shower, turned on the air conditioning, said they were hot and remember they were hot because they took a shower, turned on the air conditioning, said they were hot and now remember being hot. Dennett's model of phenomenal character is thus subject to the behaviourist circle. Phenomenal character, such as the feeling of being hot, can do no work in the explanation of behaviour and verbal reports on the multiple drafts model. The reason for this is that on this model phenomenal character is constituted by the verbal reports, memories and behaviours. On Dennett's theory we must give up the intuition that phenomenal character does causal work. Dennett has secured the intuition that there is no appearance/reality distinction for phenomenal consciousness. However, this is at the expense of the intuition that phenomenal consciousness does causal work.

[5] O'Brien and Opie (1997) make a similar point. They argue that on process theories, such as Dennett's, phenomenal character is the effect of the process not it's cause. Seeing that an effect cannot cause itself they conclude that phenomenal character cannot be what causes these processes. When these processes are those that lead to reports and memories of feelings this seems to create a problem. These effects seem to be what is caused by the phenomenal character of a mental state, yet on process theories they cannot be what causes them.

It should be clear that any account that explains phenomenal consciousness entirely in terms of access will make phenomenal character epiphenomenal. This is simply because such access accounts explain phenomenal consciousness in terms of the accessibility of a mental state to the wider cognitive system. That is they explain phenomenal character in terms of the causal relations between representing vehicles and behaviour alone. Those vehicles which are available as use for premises in reasoning, reportable and involved in action control (Block 1995) are informationally promiscuous, that is, conscious- they have phenomenal character. It is having widespread *effects* on the on the wider cognitive system that is *constitutive* of phenomenal character of the vehicle. Like Dennett's model, access accounts more generally identify phenomenal character with information processing effects that intuitively phenomenal character seems to cause. Phenomenal character cannot be a cause of those effects. Thus such accounts are subject to the behaviourist circle and therefore make consciousness epiphenomenal.[6]

[6] It is not the case that all theories which invoke access in accounts of consciousness are subject to the behaviourist circle. Any theory that seeks to explain access as a type of consciousness, rather than explain phenomenal character in terms of access, will not make phenomenal character epiphenomenal. On such accounts it may be phenomenal character that is accessed. It is not easy to classify many existing theories because some important distinctions between the phenomena to be explained are only just emerging (Block 1995; Block 2001; Block 2002; Block 2005a; Block 2005b). For instance higher order thought theories may also be explaining phenomenal character in terms of access (Rosenthal 1993). However it is not clear that they are actually out to explain phenomenal character or whether they are actually intended to explain the mineness or what Kriegel (2005) terms 'subjective character' of an experience. That is the phenomenal character of an experience plus the sense of owning or being subject to that experience. It is only those accounts that seek to explain phenomenal character in terms of access that make that character epiphenomenal.

We have seen that by developing a theory of consciousness that is consistent with the intuition that all there is to phenomenal character is appearances we have to give up the intuition that phenomenal character plays a causal role. In the next section we will see that the converse holds for vehicle theories of consciousness.

Accepting the Intuition that Phenomenal Character plays a Causal Role

What would be the consequences of accepting that we *can* draw a reality/appearance distinction for consciousness? In the above terminology this would mean that access by or appearance to the subject of a mental state is distinct from the phenomenal character of that state. This is one idea behind vehicle theories of consciousness. It is compatible with such theories that there is some process by which experiences appear to the subject, they just deny that this is what makes a mental state phenomenal. In contrast, they identify phenomenal character with the vehicles of (explicit) representation, such as the patterns of activation in neural networks (O'Brien and Opie 1999b).[7] Such accounts opt for a different approach to the process accounts considered above. Process accounts explain the phenomenal character of a mental state in terms of processes performed over it. The processes that do the work are usually processes of access. Vehicle theories, in contrast, seek to explain the phenomenal character of a mental state in terms of what it is; such as an explicit vehicle of representation.

[7] See O'Brien and Opie (1999b) for an argument to the effect that a vehicle theory of consciousness is only plausible on a connectionist account of computation.

The key point for the current discussion is that on a vehicle theory of consciousness, representing vehicles are phenomenal independently of how they appear to the subject. These vehicles are phenomenal in and of themselves, and not because they are accessible to the subject. They are not phenomenal because they are famous; they are phenomenal because they are the things that they are. To return to Dennett's Cartesian theatre metaphor the vehicles of consciousness are the players on the stage which are intrinsically phenomenal regardless of whether or not there is an audience there to watch them. Importantly for a vehicle theory of consciousness, it is possible to draw a reality-appearance distinction for phenomenal states. The reality is the nature of the vehicle, the appearance is the "judgements" the subject makes about the vehicle by accessing it.

Vehicle theories thus give up the intuition that the phenomenal is *just* how things appear to the subject. There are states which are phenomenal in and of themselves, independently of their relations to the rest of the cognitive system. On such theories, consciousness is not how things appear, because things are phenomenal independently of their appearance to the subject. On these theories there is an appearance/reality distinction for phenomenal character.

To some, such as Dennett,[8] this is nonsense. For them the phenomenal simply is how things appear. To talk about phenomenal states which are not appearances to a subject seems to make no more sense than a talking about a bachelor that is married. But as

[8] In personal communication he suggested that it made no sense to talk about "real seemings" independently of their appearance to the subject because it makes no sense to talk about seemings unless they are "seeming to" someone.

we made clear above, the cost of accepting this intuition is that one must give up the equally powerful intuition that phenomenal character plays a causal role. Some find this cost too high (O'Brien and Opie 1999a) and thus it seems to them to be worth exploring this "nonsensical" option. The purported existence of a hard problem of consciousness suggests that such conceptual changes are required and that we need to be critical of any attempt to hold onto our pre-theoretical intuitions.

The intuitive upshot of the vehicle theory is that it is not threatened by epiphenomenalism. Representing vehicles are able to have a causal impact on information processing. Phenomenal experiences are representing vehicles, and the trajectory of information processing within the brain is affected by which representing vehicles are being processed, and so phenomenal experiences are causes of information processing effects. Contrary to the process theories, vehicle theories are consistent with the intuition that phenomenal character matters to the system. Explanations in terms of vehicles are not explanations in terms of effects of things, but rather in terms of things themselves. So we can explain how this phenomenal thing, say this explicit vehicle of representation is able to have these effects in the same general way as we can explain how this wet thing, say a glass of water, is able to have certain effects. However, vehicle theorists must give up the intuition that there is no appearance/reality distinction for consciousness.

Implications

It seems it is not possible to hold both the intuition that we cannot draw an appearance/reality distinction for consciousness and the intuition that consciousness is a cause of behaviour within the framework provided by cognitive science. On process

theories of phenomenal character there is no appearance/reality distinction for phenomenal character because phenomenal character is constituted wholly by the accessibility of a mental state to the wider cognitive system. However, process theories deny the intuition that phenomenal character plays a causal role. In contrast, on vehicle theories of phenomenal character that phenomenal character does play a causal role in the cognitive system. However, vehicle theories deny the intuition that there is no appearance reality distinction for phenomenal character. Our intuitions about consciousness are therefore inconsistent. One or other of these intuitions will turn out to be true, but it is not possible for both to be true[9]. The upshot of this is that arguing about which intuition of consciousness is true gets us exactly nowhere. As Fodor puts it arguing about intuitions is "vulgar" (Fodor 1987).

There is another more important implication of this inconsistency. Theorising based on intuitions alone can no longer be seen as a legitimate approach to consciousness. These intuitions are inconsistent and thus we cannot hope to build a theory that accommodates them all. Any attempt to do so will surely land us in familiar mysterian nightmares. Of course this is not to say that understanding our intuitions about consciousness has no value, for example recognising the counter intuitive

[9] Unless, of course, in turns out that consciousness isn't to be explained in the cognitive science framework. However, to think this would be to deny that consciousness is a cognitive phenomenon. In the context of wide spread success of cognitive accounts of mental phenomenon the onus is on anyone who wishes to show that clearly mental phenomena such as phenomenal character don't fall under its explanatory framework. If this could be shown it may be possible to provide an account of consciousness where both of these intuitions are true. Of course to do this proponent of this position has to defeat all the cognitive accounts of consciousness and more importantly provide an alternative explanation of the data used to generate such accounts.

consequences of Dennett's multiple drafts model lead to the development of vehicle theories (O'Brien and Opie 1997). What it does tell us is that theorising about consciousness should not be based on intuitions alone. This is a methodological consequence that has implications for how philosophers study consciousness.

There is also an argumentative consequence of this inconsistency. The above argument implies that every theory of phenomenal character open to cognitive science will have a counter intuitive consequence. Pointing out the counter-intuitive nature of these theories thus cannot, by itself, constitute an argument against the theory. Pointing out counter intuitive consequences of a theory of phenomenal character only constitutes an objection to that theory if there are *independent* grounds for thinking that intuition is true. For example, if it turns out that phenomenal character must play a distinct causal role this would constitute an objection to Dennett's model and process theories more generally. However, until it has been shown that this is the case this intuition cannot be used to decide between process theories and vehicle theories.

Conclusion

Nagel and Searle have argued that the powerful intuition that there is no appearance/reality distinction for consciousness is problematic for materialists. It seemed that if there was no such distinction for consciousness then we could never provide a materialist explanation of consciousness. The reason for this is that materialist explanations must move beyond appearances to the reality of the target phenomenon. We could not know that materialism is true. The response to this has been to try and develop theories of consciousness that provide a conceptual step

forward on consciousness thus providing the resources by which we could know that materialism is true and explain consciousness.

Two such theories of consciousness and their implications for the appearance/reality distinction have been discussed here. It has turned out to be possible to provide a materialist theory of consciousness on which there is no appearance/reality distinction for consciousness. Dennett's multiple drafts theory of consciousness is one example of the general process and Functionalist strategy needed. We have argued that if one were to accept this theory, one must give up the intuition that consciousness plays a causal role. The other theory of consciousness discussed here is consistent with the intuition that consciousness matters. This is the vehicle theory offered by O'Brien and Opie. On this theory there is an appearance/reality distinction for consciousness. Consciousness is not just appearances to a subject. However on this theory consciousness does play a causal role.

The upshot of this is that the intuition that there is no appearance/reality distinction for consciousness is inconsistent with the intuition that consciousness plays a causal role. Given the current theoretical state of play, it seems like we have to give up an intuition. But with the details of the theories before us, giving up an intuition does not seem so inconceivable. Thus, these theories may be thought of as bringing about the conceptual changes necessary to solve the problems discussed by Nagel. We don't want to adjudicate on which intuition should be rejected. Indeed the inconsistency of our intuitions suggests that any such attempt is misguided. Of course we are not saying that the way to decide between these theories is to decide between the

intuitions. In contrast the way to discover which of these intuitions is true is to decide between these theories.

References

Atkinson, Anthony P, Michael S.C Thomas and Axel Cleeremans (2000). "Consciousness: Mapping the Theoretical Landscape." Trends in Cognitive Science **10**(4).

Baars, B.J. (1988). A Cognitive Theory of Consciousness. Cambridge, Cambridge University Press.

Baars, B.J. (1996). "Understanding Subjectivity: Global Workspace Theory and the Resurrection of the Observing Self." Journal of Consciousness Studies **3**(3): 211-16.

Baars, B.J. and S. Laureys (2005). "One, Not Two, Neural Correlates of Consciousness." Trends in Cognitive Science **9**(6): 269.

Baars, Bernard J (1997). "In the Theatre of Consciousness: Global Workspace Theory, a Rigorous Scientific Theory of Consciousness." Journal of Consciousness Studies **4**(4): 292-309.

Baars, Bernard J (2002). "The Conscious Access Hypothesis: Origins and Recent Evidence." TRENDS in cognitive sciences **6**(1): 47-52.

Block, N. (2001). "Paradox and Cross Purposes in Recent Work on Consciousness." Cognition **79**: 197-219.

Block, N. (2002). "The Harder Problem of Consciousness." The Journal of Philosophy **99**(8): 391-423.

Block, N. (2005a). "The Merely Verbal Problem of Consciousness: Reply to Baars and Laureys." Trends in Cognitive Science **9**(6): 270.

Block, Ned (1995). "On a Confusion About a Function of Consciousness." Behavioural and Brain Sciences **18**(2): 227-247.

Block, Ned (2005b). "Two Neural Correlates of Consciousness." Trends in Cognitive Science **9**(2): 46-52.

Bringsjord, S. (1994). "Searle on the Brink." Psyche **1**(5).

Carruthers, Glenn (2007). "A Model of the Synchronic Self." Consciousness and Cognition **16**: 533-550.

Chalmers, D.J. (1996). The Conscious Mind: In Search of a Fundamental Theory. Oxford., Oxford University Press.

Churchland, P.M. (1984). Matter and Consciousness, MIT Press.

Churchland, P.M. (1996). "The Rediscovery of Light." The Journal of Philosophy **93**(5): 211-28.

Clark, Andy (2001). Mindware: An Introduction to the Philosophy of Cognitive Science, Oxford University Press.

Cottrell, A. (1995). "Tertium Datur? Reflections on Owen Flanagan's Consciousness Reconsidered." Philosophical Psychology **8**(1): 85-103.

Cummins, Robert (1996). Representations, Targets and Attitudes, MIT.

Dehaene, S and L. Naccache (2001). "Towards a Cognitive Neuroscience of Consciousness: Basic Evidence and a Workspace Framework." Cognition **79**: 1-37.

Dennett, D.C. (1988). Quining Qualia. Consciousness in Contemporary Science. A. B. Marcel, E. Oxford, Clarendon Press.

Dennett, D.C. (1991). Consciousness Explained. New York, Penguin Books.

Dennett, D.C. (1993). "The Message Is: There Is No Medium." Philosophy and Phenomenological Research **53**(4): 919-931.

Dennett, D.C. (1994). Real Consciousness. Consciousness in Philosophy and Cognitive Neuroscience. A. Revonsuo and M. Kamppinen. New Jersey, Lawrence Erlbaum.

Dennett, D.C. (1995). "Consciousness: More Like Fame Than Television."

Dennett, D.C. (2001). "Are We Explaining Consciousness Yet?" Cognition **79**: 221-37.

Fodor, Jerry (1981). "The Mind-Body Problem." Scientific American **244**(1): 114-123.

Fodor, Jerry (1987). Psychosemantics: The Problem of Meaning in the Philosophy of Mind, MIT Press.

Garrett, B.J. (1995). "Non-Reductionism and John Searle's *the Rediscovery of the Mind*." Philosophy and Phenomenological Research **55**(1): 209-15.

Graham, George (2007). Behaviourism. Stanford Encyclopaedia of Philosophy. E. N. Zalta.

Jack, A.I. and T. Shallice (2001). "Introspective Physicalism as an Approach to the Science of Consciousness." Cognition **79**: 161-96.

Jackson, F. (1982). "Epiphenomenal Qualia." Philosophical Quarterly **32I**: 127-136.

Kriegel, Uriah (2005). "Naturalizing Subjective Character." Philosophy and Phenomenological Research **71**(1): 23-57.

Levine, J. (1998). "Conceivability and the Metaphysics of Mind." Nous **32**(4): 449-80.

Metzinger, Thomas (2003). "Phenomenal Transparency and Cognitive Self-Reference." Phenomenology and the Cognitive Sciences **2**(4): 353-393.

Nagel, T. (1974). "What Is It Like to Be a Bat." Philosophical Review **83**: 435-450.

Natsoulas, T. (1991). "The Concept of Consciousness$_2$: The Personal Meaning." Journal for the Theory of Social Behavior **21**: 63-89.

Natsoulas, T. (1992). "The Concept of Consciousness$_3$: The Awareness Meaning." Journal for the Theory of Social Behavior **22**(199-225).

Natsoulas, T. (1994a). "The Concept of Consciousness$_4$: The Reflective Meaning." Journal for the Theory of Social Behavior **24**: 373-400.

Natsoulas, T. (1994b). "The Concept of Consciousness$_5$: The Unitive Meaning." Journal for the Theory of Social Behavior **24**: 401-424.

O'Brien, G. and J. Opie (1999a). "A Defence of Cartesian Materialism." Philosophy and Phenomenological Research **59**: 939-963.

O'Brien, Gerard and Jon Opie (1997). "Cognitive Science and Phenomenal Consciousness: A Dilemma and How to Avoid It." Philosophical Psychology **10**(3).

O'Brien, Gerard and Jon Opie (1999b). "A Connectionist Theory of Phenomenal Experience." Behavioural and Brain Sciences **22**: 127-196.

O'Brien, Gerard and Jon Opie (2006). "How Do Connectionist Networks Compute?" Cognitive Processing **7**.

Robinson, H. (2006). Dualism. The Stanford Encyclopedia of Philosophy. E. N. Zalta. URL = <http://plato.stanford.edu/archives/fall2006/entries/dualism/>.

Rosenthal, David (1993). "Multiple Drafts and Higher Order Thoughts." Philosophy and Phenomenological Research **53**(4): 911-918.

Rosenthal, David (1997). A Theory of Consciousness. The Nature of Consciousness Philosophical Debates. N. Block, O. Flanagan and G. Guzeldere, MIT Press.

Rosenthal, David (2002). Explaining Consciousness. Philosophy of Mind: Classical and Contemporary Readings. D. J. Chalmers, Oxford.

Ryle, Gilbert (1949). The Concept of Mind. London, Hutchinson.

Searle, J.R. (1992). The Rediscovery of the Mind. Cambridge Massachusetts., MIT Press.

Stoljar, D. (2001). "Two Conceptions of the Physical." Philosophy and Phenomenological Research **62**(2).

Young, Andrew and Kate Leafhead (1996). Betwixt Life and Death: Case Studies of the Cotard Delusion. Method in Madness: Case Studies in Cognitive Neuropsychiatry. P. Halligan and J. Marshall. London, Psychology Press.