Consciousness might matter very much

Adam Shriver; Colin Allen

To cite this Article Shriver, Adam and Allen, Colin(2005) 'Consciousness might matter very much', Philosophical Psychology, 18: 1, 103 — 111

To link to this Article DOI: 10.1080/09515080500085437
URL: http://dx.doi.org/10.1080/09515080500085437
Consciousness Might Matter Very Much

Adam Shriver & Colin Allen

Peter Carruthers argues that phenomenal consciousness might not matter very much either for the purpose of determining which nonhuman animals are appropriate objects of moral sympathy, or for the purpose of explaining for the similarities in behavior of humans and nonhumans. Carruthers bases these claims on his version of a dispositionalist higher-order thought (DHOT) theory of consciousness which allows that much of human behavior is the result of first-order beliefs that need not be conscious, and that prima facie judgments about the importance of consciousness are due to confabulation. We argue briefly against his claim that ‘the moral landscape can remain unchanged’ even if all or nearly all nonhuman animals are taken to be incapable of conscious experience. We then show how a first-order representational (FOR) theory of consciousness might be defended against Carruthers’ criticisms. Finally, we argue that Carruthers’ appeal to confabulation undercuts his own arguments for an evolutionary explanation for consciousness, posing a greater epiphenomenalist threat to his DHOT theory than he concedes.

1. Introduction

According to Peter Carruthers, phenomenal consciousness might not matter very much. He pushes this claim in two directions, one ethical and the other ‘ethological’. Ethically, he maintains that ‘the moral landscape can remain unchanged’ even should it turn out that the pains and frustrations of animals are all unconscious ones. Ethologically, or with respect to comparative psychology, Carruthers suggests that there is likely to be a high degree of overlap in the explanations of humans and animal behavior. Phenomenal consciousness, he says, ‘might be almost epiphenomenal in its functioning within human cognition’ and he goes on to add that
‘its absence in animals may signify only relatively trivial differences in cognitive architecture’ (2005). To be sure, there are lots of hedge words in these claims, but let us take the suggestions at face value and see whether or not Carruthers’ views are as non-revisionary as he indicates.

Although the focus of his paper in this journal is on the psychological branch of his thesis, we can’t resist making a remark or two about the moral branch (Carruthers, 2004). It seems curious to us to speak of ‘sympathy’ when we have such a difficult time, as Carruthers points out, ‘imagining pains and disappointments that aren’t phenomenally conscious ones.’ Carruthers writes that we need only recognize that we share certain first-order states that are not consciously experienced by animals, and that this is sufficient to make animals morally significant to us. But even if such a view is consistent, it is revisionary insofar as most people do take themselves to be attributing phenomenally conscious states when they express sympathy for the plight of suffering animals. It is far from clear that ordinary intuitions about what is morally permissible would remain unchanged should people become convinced that the mental states of animals lack phenomenal consciousness entirely. So, even if there is no logical contradiction in Carruthers’ suggestion that animals lacking phenomenal consciousness are appropriate objects of sympathy, his claim that the debates about phenomenal consciousness will not matter practically for our treatment of nonhuman animals does not follow.

Furthermore, adding a conscious experience of frustration to any unconsciously frustrated desire would seem to create a morally worse state of affairs. Carruthers (2004) claims that desire-frustration is ‘the most basic kind of subjective harm’. Even so, it is still true that conscious agents will have more harm done to them for any given transgression, since both their original desire and their desire not to experience consciously the feeling of desire-frustration are not satisfied, whereas nonconscious agents would only be capable of the former type of desire-frustration. Carruthers (2000) previously acknowledged this point, but he asserts that ‘it remains true that the most basic, most fundamental, way in which desire-frustration is bad for, or harmful to, the agent has nothing to do with phenomenology’ (Carruthers, 2000, p. 207, footnote 13). His defense of this claim rests on his intuitions about an imaginary case. Carruthers asserts that it would still be appropriate to feel sympathy for the pains of Phenumb, ‘a conscious, language-using, agent . . . who is unusual only in that satisfactions and frustrations of his conscious desires take place without the normal sorts of distinctive phenomenology.’ We do not share his intuitions, and think that without more discussion of how the two different types of desire-frustrations fit into our moral theories, it is hard to see how the presence or absence of conscious desire frustrations could be said to have no import in the question of how to treat animals.

What is more, Carruthers’ theory would require a drastic revision of our current legal practices in the area of animal welfare. Sympathy alone does not entail legal obligation. We may feel sympathetic towards someone stuck in a stalled car, but there are no laws requiring us to render aid. In contrast, there are currently laws designed to prohibit excessive infliction of pain and suffering on nonhuman animals, and these
laws were put in place by activists and a public who were operating under the assumption that the pains are conscious (see Rollin, 1998, for more details on the debate). To undermine the basis for those laws could require a wholesale revision of the legal code with regard to the protection of animals in science and agriculture. Perhaps Carruthers’ point is that such a revision might not amount to much of a difference in what would be considered permissible treatment of animals. But this is far from obvious, unless it is vacuously assumed that every instance of (unconscious) pain and suffering is matched by an unconscious desire not to be in that state.

2. Animal Consciousness

But enough of the ethical issues. Carruthers’ main point in this paper is that the question of animal consciousness matters little for psychology. In a sense, we couldn’t agree more: Allen and Bekoff (1997) argued that the topic of consciousness has been used much too frequently by critics of cognitive ethology to condemn a much larger body of work that can be seen as relatively neutral on questions about conscious states in animals. In another sense, however, we couldn’t agree less: if the question of animal consciousness is to be sidelined, as Carruthers suggests, then many of the approaches to finding neural correlates of consciousness in humans and other animals will turn out to be fundamentally flawed. For example, working memory in prefrontal neurons and increased neuronal activity in striate and extrastriate cortex during selective attention have both been suggested as playing central roles in the conscious experiences of humans (see Bickle, 2002, for more detail), but most research on the cellular mechanisms of these processes is performed on rhesus monkeys who have not, as of yet, been shown to possess a theory of mind or any other features that suggest they are capable of thinking about mental states as such. If Carruthers is correct in suggesting that availability to be targeted by a higher-order thought is a necessary condition for a mental state to be conscious, it would appear that the aforementioned research focusing on shared mechanisms of humans and nonhuman primates is barking up the wrong tree, at least in its suggested application to theories of consciousness. Perhaps Carruthers will say that this is exactly right—that his higher-order theory of consciousness does undermine the premise for a large amount of basic neuropsychological research. But if psychological investigations are to be integrated with neuroscientific investigations, then it can hardly be right to say that the question of animal consciousness does not matter very much for psychology.

If we are right that quite a bit hangs on whether or not phenomenal consciousness is a uniquely human trait, it is important to be sure about the arguments that purport to show that it is uniquely human (or nearly so). Carruthers’ general strategy is to argue that his higher-order-thought theory of consciousness can account for certain empirical data, but that competing theories, especially the first-order theories associated with philosophers such as Tye (1995) and Dretske (1995), either cannot do so at all, or cannot do so as well. We believe that Carruthers begs the question during his quick dismissal of first order accounts.
To buttress his own theory, Carruthers (2005) describes Milner and Goodale’s (1995) now well-supported hypothesis that two functionally and anatomically distinct visual pathways are at work in the human brain. After traveling through the optic nerve, the lateral geniculate nucleus (LGN), and the primary visual cortex (V1), optic information branches into two functionally distinguishable visual systems. Information in the dorsal system is used in the coordination and execution of movements after an agent decides on a course of action, whereas information in the ventral system produces beliefs about the environment and grounds desires for perceived items. Carruthers claims that the activities of the ventral system are unconscious while those of the dorsal system are not. After finishing his description, Carruthers (2005) states that ‘the dual visual systems hypothesis provides the grounds for one of the main arguments against first-order accounts of phenomenal consciousness’.

However, even using Carruthers’ own definitions, there is no reason to believe this last claim to be true. He writes that ‘first-order theories of the sort defended by Dretske (1995) and Tye (1995) claim that phenomenal consciousness consists in a certain sort of intentional content being available to the belief-forming and decision-making systems of the creature in question’. Using this definition, it is clear that both of these theories would predict that the contents of the ventral system would be conscious and those of the dorsal system would not, since the former but not the latter would provide content that is available to the belief-forming and decision-making systems of the person. Carruthers apparently acknowledges this when stating that ‘admittedly, only the outputs of the ventral system are available for belief formation and planning,’ but then brushes off the consequences by writing that ‘it remains unclear why this should make a difference’.

Perhaps it is true that there is something mysterious about how the availability of intentional content to a belief-forming or decision-making system could make the content conscious. In fact, if mysterious is taken to mean difficult to understand, most people would agree that under any of our current accounts, consciousness is mysterious. But the important point is that Carruthers is wrong to claim that the dual systems account of vision provides an argument against first-order accounts of consciousness, because both Dretske’s and Tye’s theories as defined by Carruthers would predict the results that were obtained from the cited research.

3. Phenomenal Consciousness

A more charitable interpretation of Carruthers’ argument is to assume that he is relying heavily on the more filled-out arguments of Phenomenal consciousness (Carruthers, 2000). The central attack in this book goes something like this: Since First-Order Representational (FOR) theories define consciousness as a certain kind of intentional content available to the belief-forming and decision-making processes of the brain, these theories have a problem if the two can be disassociated, showing that this kind of content can occur without producing consciousness.
Carruthers gives a number of examples (driving while preoccupied by conversation, selectively filtering perceptual information while sleeping, performing above chance on tasks that require discrimination in the blind areas of blindsight patients) that are supposed to show that nonconscious experiences happen quite frequently in humans. He then claims that the common sense notion of experience is equivalent to a state which is apt to give rise to a belief, and writes that the above facts show that it cannot be true that simply being in a state that gives rise to a belief can be sufficient for consciousness. This claim is his primary argument against first-order theories, and he states that it presents first-order theorists with a trilemma from which they cannot escape: they must either deny the data, accept that these examples of sensory experiences are not phenomenally conscious, or insist that sensory experiences are phenomenally conscious in a way that makes them inaccessible to their subjects (2000, p. 168).

We take the second option, while denying that the examples he uses are cases where sensorimotor experiences are appropriately available to belief-forming mechanisms. Carruthers’ argument hinges on the idea that the causes of behavior in the aforementioned examples are states that could be claimed to give rise to beliefs. He writes that because ‘both availability to thought and availability to movement-control form equally important parts of our folk-psychological idea of experiences,’ then ‘the best option is then to allow that either one is individually sufficient for a state to count as an experience’ (2000, p. 167). He later writes that if ‘both sets of experience are available to forms of belief-fixation and reasoning, their functional roles may well be isomorphic in all respects except that just one set is available to the beliefs of the highest-level executive system’ (2000, p. 170). Consequently, the argument goes, FOR theories are mysterious because they claim that one out of the isomorphic pair is conscious merely in virtue of being available to a particular system while the other is not.

But all of these claims are ignoring an important distinction that Carruthers himself brings out in his discussion of the dual visual systems; the dorsal system never is used in deciding a course of action for the organism but only helps to guide and coordinate actions already decided upon. (Actual blindsight cases provide no counterexample to this claim.) Thus it is open to the FOR theorist to claim that the putatively nonconscious experiences are nonconscious because they are never available to the belief-forming part of the brain. Furthermore, the claim that the two systems are isomorphic would be extremely implausible since we have no reason to suppose that the same kind of information that is used to make highly coordinated and precise movements would be made available to a system that could make no use of that information. In fact, there is no reason to believe the two visual pathways are isomorphic either at a functional or neurological level of description. Thus, Carruthers’ description of the dual visual systems actually produces evidence against his claims that the so-called nonconscious experiences create a problem for first-order theories.

A better tack for Carruthers to take might be to point out that Tye (1997) states that one of the consequences of his FOR theory is that even honey bees are likely
to be conscious. It is certainly not intuitively obvious that a more cognitively sophisticated explanation is required for the behavior of honey bees than for that of blindsight patients, and this may be good grounds for challenging Tye’s claim that he can explain why the experiences of blindsight patients are unconscious under his theory (thanks to Gary Varner for pointing this out). But we think it is likely that the honey bees’ behavior can be accounted for without attributing beliefs or desires, which in turn would mean that these states are not appropriately ‘poised’ to count as conscious on Tye’s view (contrary to what he actually claims about bees). The objection is therefore met by requiring a more sophisticated notion of belief than one that is satisfied by the behavior of bees (Allen, 1997).

4. Confabulation and Behavior

We now return to the central question raised by Carruthers’ paper: is his theory as nonrevisionary for psychology as he claims? According to Carruthers, dispositionalist higher-order thought (DHOT) theories would call for a drastic revision of comparative psychological research if the following three statements are true: (1) DHOT theories claim that most (if not all) nonhuman animals are not conscious, (2) comparative psychology attempts to find continuities in the explanations of behavior for humans and nonhumans, and (3) the explanation of most human behavior is dependent upon our conscious beliefs and desires. Carruthers claims that this reasoning doesn’t go through because (3) is false in most situations, and that the same nonconscious states that give rise to behavior in many nonhuman animals are also sufficient to produce similar behavior in people. Consequently, he maintains, no drastic revision is required.

To explain how one could reject an idea so common to our folk psychology as the claim that our conscious states usually determine our behavior, Carruthers (2005) describes what he calls the ‘in virtue of’ illusion. Consider the case of a person who is looking for a ripe tomato and believes that redness indicates ripeness. According to Carruthers, though it is true that the human who picks the red tomato off the vegetable rack will have a conscious experience of redness prior to her behavior, it is a mistake to assume that the action of reaching for the tomato was caused in virtue of her phenomenal experience. Rather, he claims, the behavior is better explained by the occurrence of a first-order experience of seeing red (combined with a desire for ripe tomatoes), which is a property that humans and many nonhumans have in common (2005). This explanation is preferable, Carruthers claims, because ‘it is reasonable to believe that if the redness of the tomato had not been perceived, then neither human nor monkey’ would reach for the red tomato, yet ‘even if the human lacked the higher order analog content seeming red, she would still have chosen and acted the way she did’ (2005).

Carruthers offers further arguments against the intuition that our conscious experiences play a large role in determining our behavior by referring to research on humans’ tendency to confabulate, or to produce false explanations of our own behavior. He refers to one of the most famous experiments on this subject by
Gazzaniga (1998) where a split brain patient is given an instruction ‘walk’ by flashing the instruction on a screen that can only be perceived by one hemisphere. When the patient is asked why he walked across the room, his reply (which is presumably produced by the opposite hemisphere) is something like ‘I wanted to get a drink’. These experiments suggest that our minds may be self-interpreting much of the time and, consequently, providing explanations of our behavior that do not necessarily map perfectly onto the actual explanations. Thus, if it is true that our conscious states are not a necessary component of much of our behavior, then it seems to follow that Carruthers’ theory does not have as dramatic implications for comparative psychology as the original argument suggested.

Although Carruthers alludes to the following point in his article, it bears mentioning again: the fact that confabulation occurs in pathological and other ‘unusual’ cases does not mean that it occurs often. After all, what purpose would be served by having a confabulator that consistently produced false explanations of our behavior when we are already equipped with a system that can more accurately predict our behavior, namely, the processes themselves? At the very least, much more empirical evidence is needed on the subject in support of this claim. It is also worth noting here that the patient described by Gazzaniga does not confabulate a phenomenally conscious experience to explain his behavior, so it is unclear whether his confabulation provides an appropriate model for the kind of confabulation that Carruthers accuses us all of doing. Nevertheless it is worthwhile (we hope) to examine further what the theoretical consequences of such widespread confabulation might be.

We believe that the issue of confabulation brings up a central dilemma for Carruthers’ paper. As was stated earlier, if confabulation does not explain our tendency to describe our behavior in terms of conscious beliefs and desires, then the most reasonable assumption is that our conscious experiences do play a central role in determining our behavior and, consequently, that DHOT theories would call for a massive revision of comparative psychology (and related neuroscience). On the other hand, if we grant that confabulation occurs often and that our conscious experience plays very little role in behaviors such as reaching for tomatoes, then much of the evolutionary evidence for Carruthers’ DHOT theory is damaged, as we shall establish below. In fact, the parts of Carruthers’ argument challenged by confabulation are precisely those that he uses to resist the claim that his version of DHOT is completely epiphenomenalist, so it seems as though both horns of the dilemma would have dramatic implications for psychology, provided that DHOT theories are true.

A crucial part of Carruthers’ theory is that our perception of red becomes conscious only when it is available to downstream consumer systems that can produce both red and seems red. Thus, he says, our conscious states have dual analog content. Carruthers claims that the dual-analog content system provides an evolutionary advantage because it can be used to distinguish between appearance and reality as well as evaluate other agents’ mental states. But there is almost certainly no direct empirical evidence that conscious processing is required for either of these
abilities, so our main reason for believing this to be the case is our introspective account of our behavior. But if we doubt that our conscious experience of the tomato is what causes us to reach for it, what entitles us to claim that our conscious experiences of the appearance/reality distinction or of mind-reading actually cause any relevant change in behavior in these cases?

To put a little more gray matter on this argument, consider some of Carruthers’ arguments in favor of DHOT theories in his *Phenomenal consciousness* (Carruthers, 2000) compared to his description of the meta-cognition studies in dolphins described by Smith et al. (2000). In *Phenomenal consciousness*, Carruthers states that one of the possible evolutionary functions of the ability to make appearance/reality distinctions that could help to explain the emergence of consciousness is ‘enabling people to learn about the reliability of their own experiences’ (2000, p. 226). But in commenting on the experiments described by Smith, Shields and Washburn (2003), where humans, dolphins and rhesus monkeys all exhibit the same response pattern by choosing to ‘opt out’ of reward/timeout trials when faced with ambiguous stimuli (arguably showing that these animals ‘know that they don’t know’), Carruthers (2005) says that no appeal to metacognition is needed. Rather, for all three animals, we can explain their behavior by saying that qualities such as beliefs and desires come in degrees, and when a certain low degree of belief is present, a heuristic is triggered that is something like ‘when in a state of that sort, opt for a less-favored third alternative if you can’ (2005). But notice that if we appeal to degrees of belief to explain this behavior, we no longer have any justification for claiming that the appearance/reality distinction given by consciousness necessarily has any role to play in behavioral explanations. It is open to the skeptic to claim that low-degree beliefs are responsible for our behavior in appearance/reality distinctions, and terms such as ‘illusion’ or ‘deception’ are only added later as part of our elaborate confabulation process. Thus, the claim that consciousness is required for appearance/reality distinctions is on the same footing as the claim that our conscious experience of a tomato is what causes us to reach for it.

Carruthers continues the above quote from *Phenomenal consciousness* by saying that higher-order thoughts may be necessary for producing deceit, which fits with his claim that higher-order thoughts are necessary for the possession of a theory of mind (TOM) which provides an enhanced ability to predict the behavior of other agents. But similar reasoning can be applied to this ‘function’ of higher-order thoughts. If we accept Carruthers’ account, we know that humans and other animals are capable of solving all sorts of complex problems (such as reaching for a tomato), which involve integrating information about one’s own body position, numerous environmental stimuli, and ‘knowledge’ about what happened in past situations, all of which can take place at the level of first-order thoughts. If this is true, what then is to stop us from claiming that it is merely a similar problem-solving mechanism and resultant first-order beliefs that are responsible for the prediction of other agents’ behavior? Our claim that ‘Sally reached for the ripe tomato because she thought it would taste good’ may be just as convenient a fiction as our belief that we reach for ripe tomatoes because of our experience of their redness.
The moral of the story is this: Carruthers’ use of confabulation runs the risk of becoming epidemic; once we allow it to cover our normal cases of experience in a theory, we need to be shown some principled way of preventing it from spreading to the rest of the theory. If confabulation is widespread, there is no reason to suppose that our accounts of appearance/reality distinctions or theories of minds are more immune to its reach than are our accounts of why we reach for red tomatoes. As such, evolutionary accounts of why consciousness exists would be extremely limited, and this would presumably be a bad thing for current trends in psychology.

So Carruthers is faced with a dilemma. If we accept that confabulation occurs often in human behavior, most, if not all, of the evolutionary arguments for Carruthers’ HOT account of consciousness are undercut and we are left with epiphenomenalism. However, if we deny that confabulation is a good explanation of non-pathological accounts of behavior, then his theory does have dramatic implications for behavioral psychology research (which was the antithesis of his paper). Coupled with the fact that his quick dismissal of FOR theories appears to ignore crucial details of the scientific account of differences between conscious and nonconscious ‘experiences’, the attraction of his DHOT theory may be more of an appearance than a reality.

Acknowledgments

We would like to thank Eric Saidel and Gary Varner for helpful comments.

References


