Peter Carruthers

In Defence of First-Order Representationalism

Abstract: Carruthers (2000; 2005) provides a general defence of reductive representationalism about phenomenal consciousness while critiquing first-order theories of the sort proposed by Baars (1988), Tye (1995), Dennett (2001), and others (thereby motivating a form of higher-order account). The present paper defends first-order theories against that attack.

1. Introduction

One main goal of Carruthers (2000; 2005) was to critique mysterian and property-dualist accounts of phenomenal consciousness (Chalmers, 1996), defending the view that consciousness can be reductively explained in terms of active non-conceptual representations of a certain sort that occupy a specific kind of functional role. An important part in the account involved appealing to a set of purely-recognitional phenomenal concepts, since these were deployed to disarm (and explain away the appeal of) the various ‘hard problem’ thought experiments (zombies, explanatory gaps, and the rest). A second main goal was to show that dual-content theory (according to which conscious experiences have higher-order non-conceptual contents that mirror their first-order counterparts) is preferable to other forms of higher-order theory (specifically, inner-sense theory, of the sort proposed by Lycan, 1996, and higher-order thought theory, of the kind defended by Rosenthal, 2005). I still regard both sets of goals as having been successfully achieved, on essentially the grounds

Correspondence:
Email: pcarruth@umd.edu

Journal of Consciousness Studies, 24, No. 5–6, 2017, pp. 74–87
originally given. In any case I don’t propose to revisit those issues here.

Carruthers (2000; 2005) had a third main goal, however. This was to demonstrate that first-order representationalist accounts of phenomenal consciousness (as proposed by Baars, 1988; Dretske, 1995; Tye, 1995; Dennett, 2001; and now Dehaene, 2014) are inadequate. Two main arguments were given. One was that first-order theories cannot account for our possession of purely-recognitional phenomenal concepts. The latter need to be grounded, rather, in higher-order non-conceptual content (thus requiring Carruthers’ dual-content account). The other main argument was that first-order theories can in a sense accommodate, but not really explain, the distinction between conscious and unconscious experience. Neither argument now seems to me to have been successful. I propose to critique them both.

The present discussion will abstract away from disputes about the nature of intentional content itself, however. Some have held that a reductive explanation of phenomenal consciousness requires that we should also provide a reductive theory of intentional content (Dretske, 1995; Tye, 1995). And one can then debate whether such an account should be externalist or internalist, whether it should be information-theoretic, teleosemantic, or cast in terms of inferential role (or some combination thereof). Or one might claim, in contrast, that intentional content is already a naturalistically acceptable property, vindicated by the central role that it plays (and is likely to continue to play) in cognitive science (Botterill and Carruthers, 1999; Burge, 2010). While my sympathies lie in the latter direction, for present purposes I propose to remain neutral.

2. Phenomenal Concepts

One main argument mounted against first-order representationalist theories in Carruthers (2000; 2005) was that they cannot account for our possession of purely-recognitional phenomenal concepts. A background assumption of the argument was that such concepts are essential to explain away the appeal of the hard problem thought experiments. It is because phenomenal concepts aren’t conceptually tied to functional-role ones that one can imagine a zombie — that is, someone who is physically and functionally exactly like oneself but who lacks this sort of experience. Likewise, what Mary couldn’t know in her black and white room is the truth of thoughts employing such concepts. And so on. Then provided that phenomenal concepts
themselves are naturalistically explicable (without one needing to appeal to intrinsic qualia or non-physical properties when giving an account of them), everything that needs to be explained about consciousness would have been explained.

One major point of debate since has concerned the question whether phenomenal concepts really are naturalistically explicable (Alter and Walter, 2007). I continue to believe that they are, and that the ‘phenomenal concept strategy’ is the key to de-mystifying consciousness (Carruthers and Veillet, 2007). But I shall not try to defend those assumptions here. Moreover, although Tye himself was once a major proponent of the phenomenal concept strategy (Tye, 1999), he and others (using very similar arguments) now deny that there are such things as phenomenal concepts in the sense intended here (Ball, 2009; Tye, 2009). But the arguments presented are weak, I believe (Veillet, 2012). Once again, however, I don’t propose to argue for this here. I shall take for granted that there are phenomenal concepts, and will assume that the phenomenal concept strategy is the correct one for reductive representationalists about phenomenal consciousness to pursue. My focus is on whether I was right to claim that first-order representationalism can’t give an adequate account of such concepts.

Let us use the notation this R for a phenomenal concept that applies to experiences as of red things. Now the challenge made to a first-order representationalist was this: if the concept this R is applied recognitionally in response to non-conceptual first-order content representing the presence of red, then what makes it the case that the concept refers to the experience of red, rather than to redness itself? What makes the concept a higher-order one (referring to the experience), and distinct from the first-order recognitional concept red? Both concepts would be grounded in (and tokened in response to) the very same first-order non-conceptual contents. So in virtue of what is one of them higher-order while the other is first-order? (Carruthers’ own dual-content answer to this question, in contrast, was that the phenomenal concept is grounded in higher-order non-conceptual content whereas the concept red is tokened in response to first-order non-conceptual content.)

The only recourse open to a first-order theorist, it seems, is to say that it is something about the inferential role of the concept this R that

---

1 We can conclude from Tye’s mistaken change of mind, then, that one’s second thoughts on a topic aren’t always better. I hope that this doesn’t apply to the present paper also!
makes it a higher-order concept referring to the experience of red rather than a first-order concept referring to red. That is, the concept this R must be tacitly expanded, somehow, to have the role of this experience of red, or this seeming of red, or something of the sort. Since first-order theorists only have first-order non-conceptual content to work with, then it must be something about the phenomenal concept itself that constitutes it as a higher-order one. And it is hard to see what that something could be, except a tacit deployment of the concept experience of red or something similar.

In light of this move, however, it now appears that first-order theorists can’t accommodate the kinds of thought experiments that generate the ‘hard problem’ of consciousness in the first place. For it seems one can think, (1) ‘This R might have been of the sort to be caused by green rather than red.’ And it seems one can even think, (2) ‘This R might not have been an experience at all, but might have had some quite different functional role — say the role of a decision.’ Yet given the tacit expansion of the concept this R suggested above, the contents of the two thought experiments would then really be these: (1*) ‘This experience of red might have been reliably caused by green things’, and (2*) ‘This experience of red might not have been an experience.’ Both seem to be flat-out contradictory. But the datum to be explained is that inversion-type thought experiments like these aren’t conceptually impossible (although they might be metaphysically so). On the contrary, the thoughts deployed in those thought experiments are believed to be conceptually possible. How, then, might a first-order theorist respond? Let me consider the two cases in turn.

What might explain how (1) ‘This R might have been reliably caused by green things’ can seem thinkable, consistent with first-order representationalism? Well we are, of course, quite familiar with cases where one judges that, although something seems red, it is really not. These are cases where red-type experiences are caused by something of another colour in unusual lighting conditions, for example. And it is quite natural to think that what happens sometimes can happen always. Hence, knowing the truth of ‘This R is sometimes caused by non-red things’, we move to ‘This R might always have been caused by non-red things.’ What seems needed to render this move illicit is at least a partly-causal theory of intentional content. We need to accept that an experience can only represent what reliably causes it. But of course there is no suggestion that ordinary folk need endorse such a theory; and even if they do, they may fail to apply it in the case in
hand. Nor would anyone claim that such a theory is conceptually true. In that case (1*) ‘This experience of red might have been reliably caused by green things’ does, indeed, express a thinkable thought.

Another explanation is also possible, however. This is that the phenomenal concept this R might not implicate the concept red even tacitly (thus rejecting the expansion of this R suggested above, as ‘This experience of red’). Even if experiences of red will normally have the concept red bound into them (as Carruthers, 2000; 2015, suggests), it may be that when we form a phenomenal concept we succeed in focusing attention on, and designating, just the non-conceptual content of the experience. In that case this R should just expand to ‘This experience’ rather than ‘This experience of red’. And then there will be no difficulty at all in thinking the thought (1) ‘This R might not have been about red.’ For non-conceptual experiences don’t entail concept-applications. Rather, they ground them or warrant them. So there is no contradiction in supposing that a non-conceptual experience of red (but considered apart from its red-representing nature) might never have been about red, any more than there is a contradiction in judging green in the presence of such an experience. The result is a falsehood, not a contradiction. And there is no problem, in general, in supposing that what is false might have been true.

How, then, might first-order representationalism handle the other problematic case: that of thinking (2) ‘This R might not have been an experience”? How can this seem thinkable when (2*) ‘This experience might not have been an experience’ doesn’t? Here it may matter that the first tokening of the concept experience in (2*) is implicit, perhaps deriving from the way that the phenomenal concept this R was initially acquired. For in that case the thought in question may only actually activate one token of the concept experience. And then any appearance of contradiction will disappear.

Compare the following sort of case. Suppose one acquires the concept spaniel by learning to distinguish spaniels from other dogs identified as such. So in the learning phase, one deploys the concept dog. But thereafter the connection to dog is left implicit. Mightn’t one then be able to think, ‘Spaniels might not have been dogs”? There doesn’t seem to be any conceptual incoherence here. Of course we normally think that natural kinds belong to their superordinate kinds essentially. But this is metaphysical, not conceptual. So a first-order theorist just has to postulate that the concept experience is only deployed when
phenomenal concepts are initially formed (thus explaining their higher-order status), not when those concepts are used thereafter.

Again a further reply may also be possible. For it may be that although the mind-reading (or ‘theory of mind”) faculty distinguishes different categories of mental state from one another via their functional roles, it doesn’t commit to those roles being essential to, or definitional of, the states in question. Philosophers who are functionalists will believe this, of course. They think that mental states are individuated by their functional roles. But perhaps the mind-reading faculty makes no such commitment. Perhaps it deploys a set of causal generalizations when attributing mental states to oneself or to others, but without regarding those roles as defining the states in question. In that case there will be no incoherence in thinking, of a type of state that is normally attributed via its causal role (such as an experience), that it might have had some other role (that of a decision, say).

I conclude, then, that the first line of argument deployed by Carruthers (2000; 2005) against first-order representationalist theories of consciousness fails. There are a variety of acceptable ways for such theories to explain how one can have purely-recognitional concepts of experience, consistent with explaining the appeal of the ‘hard problem’ thought experiments.

3. The Conscious/Unconscious Distinction

The other main argument against first-order representationalist theories of consciousness put forward by Carruthers (2000; 2005) was that such theories are incapable of explaining the distinction between conscious and unconscious experience. As is now familiar, subliminal (unconscious) perceptual experience can nevertheless prime people’s responses and give rise to affective reactions (Dehaene, 2014). Similarly, people with damage to primary visual cortical area V1 exhibit ‘blindsight’. That is, while having no conscious experience in the relevant portion of their visual field, they nevertheless respond above chance in simple discrimination tasks (Weiskrantz, 1986). And most dramatically, Milner and Goodale (1995) present evidence that there is an entire visual stream whose contents remain unconscious, which is used for swift online control of action. In all of these cases it seems intuitive that the experiences in question are not phenomenally conscious. Yet they presumably consist of first-order non-conceptual
representations of the sort appealed to by first-order theorists of consciousness. So why aren’t they phenomenally conscious?

Tye (1995) has a readily available initial response, of course. This is that subliminal, blindsight, and dorsal-stream perceptual states don’t have the right kind of functional role to qualify as conscious. Specifically, they aren’t poised to have an impact on central reasoning and decision making systems in the way required by the theory. Indeed, adopting the language of Baars (1988), Dehaene (2014), and other cognitive scientists, a first-order theorist can say that phenomenal consciousness coincides with the ‘global broadcast’ of non-conceptual content in the brain. Carruthers (2000; 2005) acknowledged this response, and allowed that a first-order theory can provide an account that coincides with the distinction between phenomenally conscious and phenomenally unconscious experience. For this distinction does, indeed, seem to line up with the contrast between non-conceptual representations that are poised to have an impact on a wide variety of central thought processes and those that aren’t.

Carruthers (2000; 2005) objected, however, that this account fails to explain why the one set of contents should be phenomenally conscious while the other is not. Since both conscious and unconscious experiences can comprise non-conceptual contents of the same general sort, why is it that these contents should ‘light up’ and become phenomenally conscious when (and only when) they are made available to central thought processes? For the latter presumably effect no change in those contents. Why should the one set of contents possess a distinctive subjective dimension whereas the other does not?

In presenting this challenge, the claim was not that first-order theories have failed to close the explanatory gap between intentional and functional facts, on the one hand, and phenomenal ones, on the other. For Carruthers’ own dual-content theory likewise failed to close that gap. (Rather, it explained away the appearance of a gap by appealing to properties of purely-recognitional phenomenal concepts.) Rather, the charge was that first-order theories lack the resources to explain why states that possess the same sorts of non-conceptual content should nevertheless differ in their subjective properties merely because one set is globally available whereas the other is not.

So the questions posed to first-order theories were these: why should non-conceptual contents have feel or be like something to undergo when available to central thought processes, while lacking such properties otherwise? How can these differences in functional role confer on one set of contents a distinctive subjective dimension
that the other set lacks? In short: how does role (specifically, global broadcasting) create phenomenal character?

The seeming unanswerability of these questions motivated Carruthers to propose his dual-content theory, according to which one effect of global broadcasting is to make first-order non-conceptual contents available to a higher-order mentalizing or ‘mind-reading’ faculty capable of entertaining higher-order thoughts about those experiences. This, when combined with the truth of some or other kind of consumer semantics, was said to add a dimension of higher-order non-conceptual content to the first-order experiences in question. Every globally broadcast experience is then both a non-conceptual representation of the world or body (red, say) and a non-conceptual representation of that experience of the world or body (seeming red, or experiencing red, say). Globally broadcast experiences are thus not just world-presenting but also self-presenting. They thereby acquire a subjective dimension and become like something to undergo. Moreover, only globally broadcast experiences have this sort of dual content, on the assumption that only such experiences are available to the mind-reading faculty. Hence the conscious/unconscious distinction can genuinely be explained, it was claimed.

I now think, however, that this objection to first-order theories may have tacitly assumed a conception of phenomenal consciousness that all representationalists (whether first-order or higher-order) are committed to rejecting. For the objection seems to take for granted that phenomenal consciousness is some sort of intrinsic property that gets attached to an experience in certain circumstances, thereby ‘lighting it up’. If so, this would then be a version of intrinsic-qualia view that both Tye and Carruthers have always been committed to rejecting. So some other way needs to be found of raising the objection (if an objection can indeed be raised). It would just beg the question in favour of dual-content theory, for example, to say that an additional layer of content needs to be acquired by globally broadcast non-conceptual representations in order for them to become phenomenally conscious. So what one needs is some way of characterizing what it is for a state to be phenomenally conscious that is neutral between the two views (while not committing us to qualia or to property dualism). And it would then need to be argued that phenomenal consciousness (thus characterized) is not explained by a first-order representationalist account.

The point remains, however, that a first-order theorist needs to say something about why non-conceptual contents should be
phenomenally conscious if globally broadcast, but not otherwise. It is obviously true (almost by definition) that global broadcasting renders non-conceptual content *access* conscious. But what is it about global broadcasting that renders non-conceptual content *phenomenally* conscious? Even if one seeks to deny that these concepts pick out distinct properties, something needs to be said to explain why *what-it-is-likeness* and other properties distinctive of phenomenal consciousness should only co-occur with global broadcasting.

The way forward for first-order theorists, I suggest, is to co-opt the operationalization of phenomenal consciousness first proposed by Carruthers and Veillet (2011) to throw light on a different debate. This was the question of *cognitive* phenomenology. There is considerable dispute over the question whether cognitive/conceptual content makes a constitutive contribution to phenomenal consciousness (Bayne and Montague, 2011). Some claim that it is introspectively obvious that there is something it is *like* to hear a sentence with understanding rather than without, for example (Strawson, 1994; Siewert, 1998; Pitt, 2004). Those on the other side of the debate allow that there are phenomenal differences between the two cases, but argue that these can be differences in non-conceptual content *caused by* the presence or absence of conceptual understanding. This leads to different patterns of perceptual processing in each case (Tye, 1995; Tye and Briggs, 2011). Plainly, introspection alone cannot resolve this issue.

In order to address this problem, what Carruthers and Veillet (2011) proposed is that phenomenal consciousness can be operationalized as *whatever gives rise to the ‘hard problems’ of consciousness* (Chalmers, 1996). That is, a given type of content can qualify as phenomenally conscious if and only if it seems ineffable, one can seemingly imagine zombie characters who lack it, one can imagine what-Mary-didn’t-know scenarios for it, and so on. For the very notion of phenomenal consciousness seems constitutively tied to these issues. If there is a kind of state or a kind of content for which none of these problems arise, then what would be the point of describing it as *phenomenally* conscious nonetheless? And conversely, if there is a novel type of content not previously considered in this context for which hard problem thought experiments can readily be generated,
then that would surely be sufficient to qualify it as phenomenally conscious.\(^2\)

Once phenomenal consciousness is operationalized as whatever gives rise to hard problem thought experiments, however, it should be obvious that the initial challenge to first-order representationalism collapses. The reason why non-conceptual contents made available to central thought processes are phenomenally conscious, whereas those that are not so available are not, is simply that without thought one cannot have a thought experiment. Only those non-conceptual contents available to central thought are ones that will seem to slip through one’s fingers when one attempts to describe them (that is, be ineffable), only they can give rise to inversion and zombie thought experiments, and so on. This is because those thought experiments depend on a distinctively first-personal way of thinking of the experiences in question. This is possible if the experiences thought about are themselves available to the systems that generate and entertain such thoughts, but not otherwise. Experiences that are used for online guidance of action, for example, cannot give rise to zombie thought experiments for the simple reason that they are not available for us to think about in a first-person way, as this experience or something of the sort. They can only be thought about third-personally, as the experience that guides my hand when I grasp the cup, or whatever.

There is simply no need, then, to propose that dual higher-order/first-order non-conceptual contents are necessary in order for globally broadcast experiences to acquire a subjective dimension and be like something to undergo. Once possession of such a dimension/possession of phenomenal consciousness is operationalized as whatever gives rise to hard problem thought experiments, then the mere fact of global broadcasting provides the required explanation. For it is non-conceptual content made available to central thought processes, and which is thus available to be thought about in a distinctively first-personal way, that grounds those thought experiments.

If we suppose that this explanation on behalf of the first-order theorist is correct, however, then what should be said about

\(^2\) Deploying this operationalization of phenomenal consciousness, Carruthers and Veillet (2011) go on to argue that concepts make only a causal rather than a constitutive contribution to the phenomenal properties of our mental lives. They thus resolve the question of cognitive phenomenology in the negative. This is because, they argue, conceptual content does not give rise to the sorts of thought experiment that characterize the hard problem of consciousness.
phenomenally conscious experience in non-human animals? Presumably no animals have the conceptual resources to engage in hard-problem-type thought experiments. (Indeed, the same may be true of many humans.) Does that mean that their experiences aren’t phenomenally conscious ones? Surely not. For giving rise to hard problem thought experiments is not supposed to be constitutive of phenomenal consciousness. Rather, it provides a theory-neutral way of delimiting the class of phenomenally conscious states in ourselves: roughly, phenomenally conscious states are the ones that are especially philosophically challenging or puzzling. Instead (according to first-order representationalism), what constitutes phenomenal consciousness is being a globally broadcast non-conceptual state. And there is plenty of reason to think that many species of animal (perhaps all vertebrates) have states of that general kind (Carruthers, 2015).

Seen from this perspective, indeed, there isn’t any deep issue about the phenomenally conscious status of animal experience. Once we have established that an animal has a similar cognitive architecture to ourselves, with globally broadcast non-conceptual states that are made available to a range of different belief forming, affect generating, and executive decision making systems, then there is simply no further question whether its experiences are really like something for the animal, or whether its experiences genuinely possess a subjective — felt — dimension. For there is no further property that needs to be added in order to render an experience phenomenally conscious. All that needs to be shown is that the animal possesses states of the same kind that we identify as phenomenally conscious (that is, which give rise to hard problem thought experiments) in ourselves.

Indeed, from this perspective it also emerges that there isn’t really a deep divide between creatures capable of phenomenal consciousness and ones that aren’t. For instance, we know that bees have structured belief-like states that guide them in the service of multiple goals, informed by perceptual input from a number of different sense-modalities (Gould and Gould, 1988; Menzel et al., 2005; Cheeseman et al., 2014). So they seem to possess simple minds (Carruthers, 2004). But suppose it turns out that bees nevertheless lack globally broadcast perceptual states. This might be because different types of perceptual content are made available only to specific decision making systems, for example. Perhaps no perceptual states are broadcast to most such systems simultaneously. In which case they lack phenomenal consciousness, according to an account that identifies the latter with globally broadcast non-conceptual content. But so what?
This doesn’t mean that bees are all ‘dark on the inside’ or anything of the sort. Nor does it mean that there is any point in phylogeny when some special type of experience (one that is intrinsically like something to undergo) appears on the scene. Indeed, the question of when, precisely, phenomenal consciousness emerged in phylogeny makes no sense, from this perspective.

All that can be said is that there are a variety of kinds of non-conceptual perceptual state across creatures, some of which are available to inform more systems and some of which are available to inform fewer. These states thus differ in their functional roles, and some of these roles are more similar than others to the states in ourselves that give rise to hard problem thought experiments, that is all. Nothing special, or magical, or especially significant happened in evolution when global-broadcasting architectures first emerged on the scene. It was just more of the same, but somewhat differently organized.

I conclude that first-order representationalism has an adequate response to the second of the two main criticisms levelled against it by Carruthers (2000; 2005). It can, indeed, explain why globally broadcast states are phenomenally conscious whereas access-unconscious states are not. This is because the former, and only the former, are tokened in such a way that they can be the targets for hard problem thought experiments. Since the whole point of the notion of phenomenal consciousness is to pick out the class of experiences that are supposed to be deeply philosophically puzzling, this means that globally broadcast states, and only such states, can qualify.

4. Conclusion

I conclude that Carruthers (2000; 2005) rejected first-order representational theories of consciousness on inadequate grounds. As a result, since there is extensive evidence that conscious experience co-occurs with the global broadcasting of first-order non-conceptual contents in the brain (Dehaene, 2014), and since this evidence is most easily accommodated by first-order representationalism, the latter is preferable to any form of higher-order account. This leaves untouched, however, the twin claims that dual-content theory is preferable to other forms of higher-order theory, and that reductive representationalism is preferable to mysterianism and/or property dualism, as Carruthers (2000; 2005) also argued. On reflection and with hindsight, perhaps two out of three isn’t too bad.
Acknowledgments

I am grateful to Bénédicte Veillet and an anonymous referee for this journal for their helpful comments on an earlier draft of this essay.

References


Paper received June 2016; revised September 2016.