Papineau on the Actualist HOT Theory of Consciousness

Rocco J. Gennaro
Indiana State University

[final version in Australasian Journal of Philosophy, 2003]

In his wonderful book Thinking About Consciousness, David Papineau [2002: chapter seven] presents a criticism of so-called “actualist HOT theories of consciousness.” The HOT theory, held most notably by David Rosenthal [1986, 1990, 1993, 2000] and also by myself [Gennaro 1993, 1996, 2004], claims that the best explanation for what makes a mental state conscious is that it is the object of an actual higher-order thought directed at the mental state. This view should be contrasted with Peter Carruthers’ [2000] dispositional HOT theory, which holds that the HOT need not be actual, but can instead be a dispositional state of some kind. I will show that Papineau’s criticism of actual HOT theory not only fails, but also that it mischaracterizes and underestimates the resources of actualist HOT theory. I should say at the outset, however, that I am otherwise largely in agreement with Papineau’s views throughout the remainder of his book; for example, his generally materialistic approach and his criticisms of so-called “zombie” and “knowledge” arguments designed to show the falsity of materialism [see e.g. Jackson 1982; Chalmers 1996]. Nonetheless, Papineau’s argument against the actualist HOT theory is seriously flawed.¹

Papineau contends that actualist HOT theory faces ‘an awkward problem in relation to Higher-Order memory judgements [2002: 205].’ He first asks us to consider the following case:

¹ I will follow Papineau’s use of the acronym ‘HOT’ to cover all higher-order theories. It should be noted, however, that so-called ‘inner-sense’ higher-order theorists argue that it is preferable to characterize the higher-order states as more perception-like than thought-like [see e.g. Lycan 1996]. But for a more general argument in favor of a higher-order representational theory, see Lycan 2001.
I see a red pillar-box, form no Higher-Order phenomenal judgement about this at the time, but then later imaginatively recall the experience of seeing something red. This certainly seems initially possible. Moreover, in such cases subjects will presumably later report...’Yes, I consciously saw something red earlier’. So the standard methodology will count the earlier experience as conscious: after all, the subject has issued a phenomenal report to this effect... [2002: 206]

The above case sounds perfectly plausible and there is, indeed, little reason to deny that the earlier experience was conscious. However, Papineau then urges us to agree that ‘it is not at all clear that Actualist HOT theorists will want to count this earlier experience as conscious [2002: 206].’ But why? Papineau contends that the actualist HOT theorist faces a problem that he calls the ‘awkward backwards causation of conscious states’ [2002: 206] and says the following:

If no introspective Higher-Order phenomenal judgement was made at the time of the experience, then on their view the status of that experience as conscious will presumably have to depend on the occurrence of the later Higher-Order memory judgement. But this seems silly. How can an earlier state be rendered conscious by some later act of memory? What if the act of memory hadn’t occurred? Then presumably the earlier state wouldn’t have counted as conscious. But surely the status of some state as conscious must be fixed by how things are when it occurs, not by whether or not something happens later. [2002: 206, emphasis added]

Now, any actualist HOT theorist should certainly should agree with the last sentence in the above quotation. Indeed, Rosenthal says, for example, that the most plausible account of conscious mental states will ‘identify a mental state’s being conscious with one’s having a roughly contemporaneous thought that one is in that mental state [1986: 335, emphasis added].
So it is clear that if the earlier state was conscious, then what made it conscious was something, i.e. an actual HOT, which was present at the time of the earlier experienced state.

But it is also absolutely crucial to note the following important aspect of the actualist HOT theory: It might seem that the HOT theory results in circularity, i.e. defining consciousness in terms of HOTs. It might also seem that an infinite regress results because a conscious mental state must be accompanied by a HOT, which, in turn, must be accompanied by another HOT ad infinitum. However, the standard reply from the HOT theorist is to remind the objector that when a conscious mental state is a first-order world-directed state the higher-order thought (HOT) is not itself conscious; otherwise, circularity and an infinite regress would follow. Moreover, when the HOT is itself conscious, there is a yet higher-order (or third-order) thought directed at the second-order state. In this case, we have introspection which involves a conscious HOT directed at an inner mental state. When one introspects, one's conscious focus is directed back into one's mind. So, for example, what makes my desire to write a good paper a conscious first-order desire is that there is a (nonconscious) HOT directed at the desire. In such a case, my conscious focus is directed at the paper. When I introspect that desire, however, I then have a conscious HOT (accompanied by a yet higher, third-order, HOT) directed at the desire itself [Rosenthal 1986: 337-8; cf. Gennaro 1996: 16-21]. Much the same applies to the kinds of conscious visual experiences mentioned by Papineau, such as seeing the red pillar-box or watching children at play. When my conscious focus is outer-directed the HOT is not itself conscious, but when my conscious focus is inner-directed at the visual experience itself the HOT is conscious. So instead of threatening the HOT theory, this issue actually brings out an important subtlety of any viable higher-order theory.

With this distinction in place let us now critically examine what Papineau considers to be the only two possible escape routes for the actualist HOT theorist, which he argues are both unconvincing:

_option one:_ ‘...deny that the earlier state does qualify as conscious, since no
Higher-Order judgement was present at the earlier time [2002: 206].’

Papineau rightly dismisses this option as ad hoc and contrary to any standard methodology of consciousness research. Although we should not hold that all such reports of past conscious experiences are infallible, I do not wish to dispute Papineau’s contention that we can recognize this ‘without dismissing all phenomenal memories of non-introspected experiences [2002: 207].’ Notice already, however, that Papineau seems to be identifying all such ‘higher-order judgements’ with cases of introspection. But, as we have already seen, there are both conscious and nonconscious higher-order thoughts (or judgements). It is only the former, i.e. conscious HOTs, which are properly categorized as introspective states according to actualist HOT theory. At the least, Papineau’s use of ‘higher-order judgement’ is ambiguous between these two very different kinds of states. His conflation of the two gets him into serious trouble here, and, of course, we should also not confuse any such higher-order thoughts with their corresponding verbal reports. The actualist HOT theorist can agree with Papineau’s rejection of option one if by ‘higher-order judgement’ Papineau means ‘conscious HOT or introspective state.’ But, as we have seen, the actualist HOT theory also holds that outer-directed conscious mental states must be accompanied by a nonconscious HOT. Papineau is right that outer-directed conscious experiences do not entail introspecting those experiences at the time they occur. However, the actualist HOT theory also has the resources necessary to make sense of that claim and, contra Papineau, it provides us with no reason to give up the theory. Indeed, distinguishing between nonconscious and conscious HOTs is a fundamental distinction within the theory.

Option two: ‘The other way for Actualist HOT theorists to respond to the problem of experiences which are not introspected but are later remembered would be to argue that there aren’t in fact any such cases, because we humans don’t in fact ever phenomenally remember anything we didn’t introspect at the
It is true that the actualist HOT theorist should not take option two either, but, again, we have already seen that the actualist HOT theorist does not have ‘the problem of experiences which are not introspected’ in the first place. Such experiences are the typical first-order outer-directed experiences discussed earlier. These conscious states are accompanied by nonconscious HOTs. Thus, it would of course be silly for any actualist HOT theorist to ‘argue that there aren’t in fact any such cases’ for the reason Papineau gives or for any other reason. We often do have memories of conscious experiences which are not introspected at the earlier time, but, again, this is not a problem at all for actualist HOT theory. I can, for example, later recall my fully outer-directed conscious experiences of the concert or at the play. This happens frequently and so Papineau is right that ‘it just isn’t plausible that earlier introspection is empirically necessary for later phenomenal memory.’ However, it should now also be clear that actualist HOT theory can easily accommodate this fact once one carefully distinguishes between first-order conscious states and introspective states. So the actualist HOT theorist is not committed to the conclusion that ‘it would require an absurd amount of introspective activity for us to be able to imaginatively remember all the things we can so remember.’

Papineau has thus clearly ignored a third reply to his red pillar-box case:

*Option three:* The earlier outer-directed experience was conscious but not introspected, i.e. was accompanied by a nonconscious HOT. The later memory recalled that first-order conscious experience. The HOT theorist can acknowledge that one can imaginatively recall the experience of seeing something red without there being introspective consciousness at the earlier time.

So Papineau has set up, in essence, a false dilemma mainly fueled by an oversimplification of the actualist HOT model and an ambiguity in his use of the terms
‘introspection’ and ‘judgement.’ To come full circle, then, we can now see what is wrong with Papineau’s remark (quoted earlier) that ‘if no introspective Higher-Order phenomenal judgement was made at the time of the experience, then on their view [i.e. actualist HOT theory] the status of that experience as conscious will presumably have to depend on the occurrence of the later Higher-Order memory judgement [2002: 206, emphasis added].’ It should be clear by now why this claim is false according to actualist HOT theory. The antecedent can be satisfied when one has any first-order (non-introspective) conscious experience. Yet the consequent is clearly false; that is, the actualist HOT theorist can reject the notion that the ‘experience as conscious’ depends on the later higher-order memory judgement. The actualist HOT theorist accepts the view that what makes the experience conscious occurred at the earlier, not later, time. Papineau is thus similarly mistaken when he ultimately concludes that ‘...Actualist HOT theories can’t avert the threat of the backwards causation of consciousness by denying that we ever phenomenally remember experiences we didn’t introspect earlier [2002: 210].’ Yes they can.

Incredibly, however, Papineau has a similar reply ready at hand, but he neither recognizes it as such nor sees it an essential feature of the actualist HOT theory. In the immediately following section of his book [7.12], Papineau discusses what he calls ‘attention’ and distinguishes it from introspection. He means by ‘attention’ to our mental states ‘something less than introspection’ [2002: 209] and uses terms such as the ‘highlighting’ of experiences and ‘perceiving as.’ So Papineau is, after all, recognizing that we can attend to a conscious experience in a way that is less sophisticated than introspection. The actualist HOT theorist would agree, as we have seen. Presumably, like the HOT theorist, such attention to our mental states means that we are ‘aware’ (in some sense) of our conscious states in a way that is lacking for nonconscious states.²

² Actually, I think that Papineau’s use of the phrase ‘attending to our conscious experience’ can be rather misleading here and cause some confusion with ‘introspection.’ ‘Attention to our conscious experience’ might suggest that one’s conscious focus is inner-directed. Although this
Furthermore, Papineau then says much that is in line with actualist HOT theory and will sound familiar from the above discussion. For example:

I am happy to take it, then, that it is impossible phenomenally to remember experiences which did not involve attention at the time. The point I want to stress, however, is that this does not imply that it is impossible phenomenally to remember experiences which you did not introspect at the time....So even if all remembered experiences were attended to at the time, this doesn’t mean that they were then introspected. [2002: 209]

As I have argued above, an actualist HOT theorist could not agree more and it is puzzling why Papineau continues to treat such remarks as opposed to the HOT theory. To translate the above: It is impossible to remember experiences which did not at least involve a first-order conscious state accompanied by a nonconscious HOT. But this does not imply that it is impossible to remember experiences which were not introspected at the time. Much in the same spirit we have Papineau saying ‘an experience can be highlighted by attention without your forming any introspective judgement about that experience [2002: 209].’ Again, the actualist HOT theorist would concur and translate: one can be aware of an experience (via a nonconscious HOT) without introspectively thinking about that experience. And, once more, Papineau explains that ‘prior phenomenal introspection is not necessary for later phenomenal memory [2002: 209].’ Again, the actualist HOT theorist agrees, but yet Papineau mistakenly believes that the HOT theorist is committed to the opposing untenable view that prior phenomenal introspection is necessary for later phenomenal memory. Indeed, this is what initially generates does not seem to be Papineau’s intention (hence the contrast with ‘introspection’), we should be careful not to be misled by his choice of terminology. For the purposes of this paper, however, I follow his usage.
his objection to the actualist HOT theory, but we have seen how his entire argument has rested on this mistaken view about HOT theory.

So we can see that the actualist HOT theory has the resources to handle Papineau’s red pillar-box case. There is really no problem regarding higher-order memory judgements for the actualist HOT theorist. Nor are we committed to any ‘backwards causation’ of consciousness in such cases. Indeed, using his own distinction between attention and introspection, Papineau has himself, somewhat ironically, stumbled upon the actualist HOT theorists’ primary response. Unfortunately, Papineau not only does not initially acknowledge this crucially important subtlety of the HOT theory, but he then goes on to dismiss the actualist HOT theory even after making his distinction between attention and introspection. Papineau’s primary mistake, then, is his failure to recognize just how and why the actualist HOT theory can hold that the earlier experience was conscious but not introspected. If he had done so, then he probably would have seen why his criticism is really not a serious objection to actualist HOT theory at all.

Finally, we can then also reject Papineau’s contention that dispositional HOT theory has an advantage over actualist HOT theory on this matter [see section 7.13]. The alleged problem above for the actualist HOT theory is, according to Papineau, ‘no longer a difficulty [for dispositional HOT theory], since any state which is so remembered phenomenally will be one which the subject could have introspected phenomenally at the time [2002: 210-1].’ I favor actualist HOT theory over dispositional HOT theory for a number of reasons which go well beyond the scope of this paper, but it should at least be noted here that the dispositional HOT theory gains no advantage whatsoever due to Papineau’s argument.
REFERENCES:


