

# Consciousness is Not a Property of States: A Reply to Wilberg

Jacob Berger

City University of New York Graduate Center

Published in *Philosophical Psychology* 27(6): 829-842.

[Please quote the published version.]

## Abstract

According to David Rosenthal's higher-order thought ("HOT") theory of consciousness, one is in a conscious mental state if and only if one is aware of oneself as being in that state via a suitable HOT. Several critics have argued that the possibility of so-called targetless HOTs—that is, HOTs which represent one as being in a state that does not exist—undermines the theory. Recently, Jonah Wilberg (2010) has argued that HOT theory can offer a straightforward account of such cases: Since consciousness is a property of mental state tokens, and since there are no states to exhibit consciousness, one is not in conscious states in virtue of targetless HOTs. In this paper, I argue that Wilberg's account is problematic and that Rosenthal's version of HOT theory, according to which a suitable HOT is both necessary and sufficient for consciousness, is to be preferred to Wilberg's account. I then argue that Rosenthal's account can comfortably accommodate targetless HOTs because consciousness is best understood as a property of individuals, not a property of states.

## 1. Introduction

Most people would agree that if one is in a mental state, but in no way aware of being in that state, then that state is not conscious. A logically equivalent claim is that one is in a conscious mental state only if one is aware of being in that state. This fact is implicit in the folk conception of consciousness and central to most recent empirical work on it (for a review, see, e.g., Lau and Rosenthal, 2011). This feature of consciousness is also the underlying motivation for David Rosenthal's higher-order thought ("HOT") theory of consciousness, which holds that one is in a conscious mental state if and only if one is aware of oneself as being in that state via a suitable HOT (see, e.g., Rosenthal, 2005).

It is thus implicit in the folk conception of consciousness which HOT theory seeks to capture that one is in a conscious state only if it appears to one, in some suitable way, that one is in that state. Consciousness is, as Rosenthal often emphasizes (e.g., 2009, p. 166), a matter of how one's mental life appears to one. But appearance and reality can in general diverge. For

## A Reply to Wilberg

example, it can perceptually seem to me that there is a red apple even if the apple is green or if there is actually no apple. Likewise, thoughts often represent things that do not exist. For example, my belief that Santa Claus is jolly represents Santa, though Santa does not exist. HOT theory should therefore be able to accommodate the following kind of case: One has the suitable HOT that one sees a red apple and it thereby seems to one that one sees a red apple, even though one does not see anything at all. Many critics of HOT theory, however, have argued that the possibility of these so-called targetless HOTs—HOTs which represent oneself as being in states that one is not in—undermines the theory (e.g., Byrne, 1997; Neander, 1998; Levine, 2001; Kriegel, 2003; Mandik, 2009; Block, 2011a; Block 2011b). HOT theory holds that one is in a conscious state if one has a suitable HOT (whether or not that HOT is targetless), but the problem is that in the cases of targetless HOTs it would seem there are no states to be conscious. I will call this the ‘targetless-HOT objection’ to HOT theory.

Recently, Jonah Wilberg (2010) has offered a novel reply to this objection. At the heart of the issue, Wilberg claims, is the fact that consciousness is a property of mental state tokens that HOTs confer to their target states. A targetless HOT cannot thereby confer the property of consciousness to its target state because no such state exists. But Wilberg argues that HOT theory can nonetheless offer a straightforward account of such cases: One is simply not in a conscious state in virtue of a targetless HOT. Wilberg calls this the “No Consciousness Account” of targetless HOTs (2010, p. 626).

Wilberg’s account may seem sensible, but I’ll argue in this paper that it is not the best account of HOT theory. In section 2, I will provide an overview of HOT theory and a description of Wilberg’s account of targetless HOTs. I will then argue in section 3 how Wilberg’s view does unnecessary violence to what is implicit in our ordinary conception of consciousness—namely,

## A Reply to Wilberg

that consciousness is a matter of which mental states it seems to one that one is in, whether or not one is in those states.<sup>1</sup> If, as I'll argue, the subjective impression that one is in some state is all there is to consciousness, then a suitable HOT is not only necessary but also sufficient for consciousness. This is the view that Rosenthal himself has long defended (see, e.g., 2005, chapter 7, specifically pp. 209-213).

In his paper, Wilberg argues that Rosenthal's view cannot be sustained because consciousness is a property of states.

Although many theorists do assume that consciousness is a property of states, this assumption is questionable. Indeed, Rosenthal himself has been explicit that consciousness is not a property conveyed by HOTs to first-order states (e.g., Rosenthal, 2005, p. 185). It is thus not clear whether Wilberg intends his account to be a clarification of Rosenthal's view or a new way of understanding HOT theory designed to cope with the targetless-HOT objection. In any case, in section 4 I will defend Rosenthal's account of targetless HOTs by arguing that there are independent reasons to think that consciousness is best understood not as a property of mental state tokens, but as a property of individuals.<sup>2</sup> In disambiguating 'consciousness', Rosenthal does often distinguish what it is to be in a conscious state—instances of what he calls 'state consciousness'—from what he calls 'creature consciousness', which he describes as the property of being awake and responsive to stimuli (e.g., 2009, p. 157). This terminology may seem to imply that state consciousness is a property of states, but I'll argue that it attaches to individuals.<sup>3</sup> Once consciousness is thus properly understood, it will be clear how Rosenthal's account comfortably accommodates targetless HOTs and why it better captures our ordinary conception of consciousness than Wilberg's "No Consciousness Account."

## 2. HOT Theory and the “No Consciousness Account” of Targetless HOTs

The fundamental motivation for HOT theory is the claim that one is in a conscious state only if one is aware of oneself as being in that state. Rosenthal has called this fact about consciousness the Transitivity Principle (“TP”) because it explains what is for one to be in a conscious state in terms of one’s transitive awareness of being in that state (e.g., 2005, pp. 3-4). It is clear, however, that the TP offers a necessary, but not sufficient, condition for consciousness. Being informed that one is in a state or consciously inferring that one is in it, for example, may result in one’s being aware that one is in a state, but it need not result in one’s being in a conscious state. To be in a conscious state, it must subjectively appear to one that one is in that state—that is, it is not the case that one is aware of being in that state in a way that seems to be the result of inference or observation. According to HOT theory, suitable HOTs are the states in virtue of which there are these subjective impressions. A suitable HOT is an occurrent intentional state that asserts the content that I am in some state.<sup>4</sup> So, for example, one consciously perceives a red apple only if it subjectively seems to one that one see a red apple. In other words, one has the suitable HOT that one sees a red apple.

Though the TP can be put in terms of one’s awareness of oneself as being in a state as I have described it here, it is sometimes glossed as the claim that a conscious state is a state that one is aware of. As I’ll argue shortly, the latter way of characterizing the TP is fine as shorthand, but it can be misleading. This latter way of capturing the TP suggests that consciousness occurs when one is in some first-order state, which is also accompanied by a suitable awareness of that state. This appears to be the model of the TP that Wilberg adopts because it yields his reading of HOT theory, which holds that “[a] mental state is conscious if and only if it is accompanied by the [suitable] HOT that one is in that state” (2010, p. 618).

This is a very common way to understand HOT theory (see, e.g., Gennaro, 2006, p. 222; Carruthers, 2011, §4). On this way of understanding HOT theory, consciousness requires a creature to be in two actual and suitably related states—a suitable HOT and a target first-order state—which invites the interpretation that the HOT somehow confers the property of consciousness to the existing first-order target state. And Wilberg explicitly claims that the theory is committed to consciousness' being such a property. Wilberg writes,

[C]onsciousness as a property of token mental states is, I think, part of the implicit conceptual framework for much recent work on consciousness, and it certainly seems central to a clear articulation of the basic tenets of HOT theory (2010, p. 625).

If consciousness were a property of existing first-order states, then it would appear that HOT theory faces grave difficulties when explaining cases of targetless HOTs. As Wilberg observes, in the case of a targetless HOT, one's HOT represents there to be a first-order state, but that state cannot exhibit the property of consciousness because it does not exist. After all, nonexistent things cannot exhibit properties (cf. Mandik, 2009). Moreover, if a HOT confers the property of consciousness to a state only when one is aware of that state, then there appears to be no state that one can be aware of in cases of targetless HOTs. Awareness is arguably factive in terms of the existence of the thing that one is aware of. We ordinarily would not say that one is aware of a red apple if that apple does not exist. It would thus seem that a targetless HOT cannot make one aware of a first-order state because, again, no such state exists.<sup>5</sup>

Many have concluded that HOT theory is unworkable (e.g., Byrne, 1997; Neander, 1998; Levine, 2001; Kriegel, 2003; Block 2011a), but Wilberg argues that there is an underappreciated aspect of the account that reveals how it can accommodate targetless HOTs. Specifically, Wilberg maintains that HOT theory is committed to the claim that HOTs must represent their

## A Reply to Wilberg

target states accurately. Wilberg claims to find evidence for this feature of the theory in Rosenthal's own discussion of what is known as tip-of-the-tongue phenomena (e.g., Rosenthal, 2000a, p. 204; cited in Wilberg, 2010, p. 619). One is often aware, for instance, that one knows the name of an author of a book without being consciously aware of the author's name—the name is, so to speak, on the tip of one's tongue. Common sense holds that the state of knowing the author's name is not conscious. Wilberg claims,

Such states make manifest, according to Rosenthal, that for a mental state to be conscious in virtue of a HOT, it is not enough for that HOT merely to represent that state in just any way... it seems that the HOT must represent, at least roughly accurately, the individuating features of that state (2010, p. 619).

Wilberg sees this discussion of tip-of-the-tongue phenomena as evidence for the requirement that suitable HOTs must represent the individuating features of their target states in roughly accurate ways. Wilberg thus interprets HOT theory as holding that a conscious state is conscious only in respect of the features that one's HOT accurately represents. And this condition on suitable HOTs would seem to make sense. The requirement seems to rule out the possibility of HOTs accidentally conferring consciousness to the wrong states. It also guarantees that HOTs confer consciousness to existing first-order states only.

This condition also appears to reveal how HOT theory handles cases of targetless HOTs. Since a targetless HOT does not accurately represent any state, there is no state to exhibit consciousness. In such a case, one only seems to be in a conscious state. Hence Wilberg proposes his "No Consciousness Account" of targetless HOTs. On Wilberg's view, HOTs are necessary, but not sufficient, for consciousness.

## A Reply to Wilberg

Wilberg does not deny that consciousness is a matter of mental appearance, nor does he deny that how a HOT represents one's mental life to be to one can diverge from how one's mental life actually is. But because Wilberg maintains that a state's being conscious involves its acquiring the property of consciousness, he denies that the subjective appearance of being in a certain state is sufficient for consciousness. On Wilberg's view, one is in a conscious state if and only if it appears to one that one is in a state via a suitable HOT and, in addition, one is in that state. So while consciousness does concern mental appearance, Wilberg maintains that it is a "matter of (rough) correspondence between appearance and reality" (2010, p. 630).

As a result, Wilberg maintains that a uniform treatment does not apply to all so-called false HOTs. On one hand, if I have a HOT that I perceive a red apple, but I do not perceive anything, my HOT is targetless. But if I have a HOT that I perceive a red apple when in fact I perceive a green apple, my HOT is not targetless; my HOT only misrepresents my perception to some degree—that is, in respect of its color. In these cases, there is a state to which the HOT can confer consciousness. So Wilberg concludes that one's states are conscious in respect of the aspects accurately represented by one's HOTs. If one's HOT misrepresents one's perception of a red apple as a perception of a green apple, then one consciously perceives an apple, but the state is not conscious in respect of its greenness.

### **3. Why the "No Consciousness Account" is Problematic**

Wilberg's "No Consciousness Account" appears straightforward indeed. But whatever its initial attractiveness, it is problematic. Again, Wilberg grants that if one has a targetless HOT that one sees a red apple, it will subjectively seem to one that one sees a red apple. But Wilberg proposes that, in such a case, one does not consciously see a red apple. However, our ordinary

## A Reply to Wilberg

conception of consciousness arguably regards this kind of subjective impression as sufficient for consciousness.

Consider the fact that it is deeply ingrained in our ordinary practice of attributing consciousness, as well as central to most recent experimental work on it, that we can verbally report only our conscious states (see, e.g., Lau and Rosenthal, 2011).<sup>6</sup> If one claims that one sees a red apple, we take that report at face value—that is, as reliable evidence that the person consciously sees a red apple. As Rosenthal observes (e.g., 2011, p. 432), we typically do not question whether a person consciously sees what they report seeing pending further investigation into the mental reality underlying that mental appearance. If one reports seeing a red apple, such a report signals that it appears to one that one sees a red apple—and all there is to a conscious state, we think, is the subjective appearance that one is in some state. Since appearances in general need not correspond to reality, it is implicit in our folk conception of consciousness that consciousness is a matter of mental appearance, whether or not those appearances are accurate.

This is, of course, not to say that consciousness does not exist or that it is only a matter of appearances. Consciousness does also concern mental reality: HOTs are existing mental states and the subjective appearances reflected by them are genuine aspects of one's mental life. If one consciously sees a red apple, it really seems to one that one sees a red apple and that seeming is an actual aspect of one's psychology. Irrelevant to whether there is consciousness, however, is whether one actually perceives the red apple. To accurately capture our ordinary conception of consciousness, then, HOT theory should maintain that a suitable HOT is not only necessary but also sufficient for consciousness—and this is the version of HOT theory that Rosenthal defends.

On Wilberg's "No Consciousness" version of HOT theory, by contrast, the mere subjective appearance of a state is insufficient for consciousness. For Wilberg, if one has a

## A Reply to Wilberg

targetless HOT that one sees a red apple, seems to see a red apple, and on that basis reports that one sees a red apple, one's report does not indicate that one consciously sees the apple. In severing the connection between verbal reports and consciousness, Wilberg's view violates a feature central to both our commonsense and experimental approaches to consciousness.

Wilberg does, however, anticipate this kind of objection to his view. Wilberg urges that there is reason to think that our verbal reports concerning our mental lives need not always correspond to our conscious states. He writes that if "someone unwittingly makes up a story, or applies a theory, in trying to describe their conscious experience, it is plausible to think that what they say does not correspond to their conscious experience at all" (2010, p. 629).

But it is unclear what Wilberg's evidence is here. If by 'unwittingly', Wilberg means that one applies a theory to oneself in an offhand but nonetheless conscious way, then he is correct that one might believe on the basis of that theory that one is in some state and might thus be inclined to say that one is in that state. But it need not be the case that this theorized state is one's conscious state. It might also seem subjectively that one is in some other conscious state and so one might thereby also be inclined to report that one is in that other conscious state. Indeed, there would seem to be no grounds for applying a theory to oneself to determine one's mental state if it did not already seem to one subjectively that one was in some other conscious state.

This sort of scenario is commonplace: Everyone has doubtless had the experience of being asked what is wrong by a concerned friend just at a moment when one subjectively feels that nothing is wrong. The concerned friend's question might then prompt one to reply in puzzlement, "Why—nothing is wrong!" and then to consciously theorize whether or not one nevertheless harbors a sense that something is wrong after all. But such theorizing need not change the fact that, at the moment the question was posed, one did not consciously feel that

## A Reply to Wilberg

something was wrong. Though the conscious theorizing might then reveal a nonconscious state of distress, one was not aware of that state—subjectively or otherwise—until the friend posed the question. Thus it appears that reports about what states it subjectively seems to one that one is in do correspond to one's conscious states after all.

If instead Wilberg means by 'unwittingly' that one applies a theory about one's mental life in such a way that it does not seem subjectively that one has applied that theory, then there is no reason to think that one's conscious state would differ from the state that it appears to one that one is in. Of course, one's subjective appearances may be the product of the nonconscious applications of theory, but it does not follow that such theorized states cannot be conscious states. If Wilberg were to maintain that a theorized state differs from one's conscious state, then it would be simply unclear which of these states he considers to be the conscious state. So Wilberg offers no good reason to think that reports of our mental lives might diverge from our conscious states.

Wilberg might rejoin that the suitable appearance of a mental state is insufficient for consciousness because Rosenthal's discussion of tip-of-the-tongue phenomena establishes the requirement that HOTs must represent accurately the individuating features of their target states. But the evidence concerning tip-of-the-tongue phenomena does not support this condition. Wilberg is correct that, if I am aware that I have the name of a book's author on the tip of my tongue—and am thus not aware that I know that the author is Joyce—then my state of knowing the author's name is not conscious. But this only shows that I do not have a HOT about the state of knowing that the author is Joyce. Instead, I have a HOT that I know the name of the book's author. Such evidence only establishes that what conscious state one is in corresponds to the state

## A Reply to Wilberg

described by the suitable HOT, not that the state that the HOT describes must be accurately represented.<sup>7</sup>

All the same, one might think that the fact that the “No Consciousness Account” undermines these aspects of our commonsense and experimental approaches to consciousness is simply the price we must pay, insofar as Wilberg’s account is the only way for HOT theory to explain targetless HOTs. But Wilberg’s view is unmotivated and unnecessarily problematic. Apart from the assumption that consciousness is a property of states (which I’ll soon contest), Wilberg provides no compelling reason to think that the targets of HOTs must be accurately represented or must even exist.

### **4. How to Understand Consciousness**

As Wilberg notes, Rosenthal has long defended the view that a suitable HOT is both necessary and sufficient for consciousness. Speaking of cases of targetless HOTs, Rosenthal claims:

Higher-order theories face no difficulty about such cases. What it’s like for one on these theories is determined by the way the higher-order awareness represents the first-order state.... Consciousness is a matter of mental appearance... that mental appearance is due solely to the higher-order awareness (2009, p. 166; quoted in Wilberg, 2010, p. 624).

Wilberg argues, however, that Rosenthal’s view cannot accommodate targetless HOTs. The bottom line, Wilberg maintains, is that in cases of targetless HOTs there is no state that can exhibit the property of consciousness.

Wilberg does acknowledge that Rosenthal has provided elsewhere various defenses of his view from the targetless-HOT objection,<sup>8</sup> but Wilberg charges that these replies are inadequate.

## A Reply to Wilberg

For instance, Rosenthal has suggested that targetless HOTs confer consciousness to states that are merely notional—that is, states that do not exist (e.g., Rosenthal, 2000b, p. 232; cited in Wilberg, 2010, p. 625). It is true that we often describe nonexistent things as having properties. If I believe that Santa Claus is jolly, we might hold that my belief represents the notional Santa as exhibiting the property of being jolly. Thus Rosenthal proposes that if I have a targetless HOT that I see a red apple, it is the notional perception of a red apple that exhibits the property of consciousness. I will return to this proposal shortly. But Wilberg retorts that this account faces the same challenges discussed above. Wilberg argues that notional states cannot exhibit consciousness because nonexistent things cannot exhibit real properties. Likewise, since we arguably cannot be aware of nonexistent things, HOTs cannot make us aware of notional states and thus to confer consciousness to them. Wilberg thus concludes that Rosenthal’s view cannot be sustained. He writes,

Rosenthal’s response to the [targetless-HOT] objection can only succeed by rejecting the relevant notion of... consciousness [as a property of existing mental state tokens] here.

Yet such an approach would seem to require a fundamental revision of the basic conceptual framework of HOT theory (2010, p. 626).

Though many theorists do assume that consciousness is a property of mental states, this assumption is rarely defended and I’ll now argue that it is false. Denying that consciousness is a property of first-order states conferred to them by HOTs does not, however, constitute a revision of HOT theory, insofar as consciousness’ being such a property was never a plank of the theory from the outset.

## A Reply to Wilberg

First, it is worth noting that it is unclear what argument there could be for the claim that consciousness is a property of states. It is far from obvious, for example, that folk psychology takes any stand regarding what kind of thing consciousness is a property of.

Second, notice that once one has been persuaded that some higher-order account of consciousness is correct, conceiving of consciousness as a first-order property of states—properties that are somehow conferred to states by a higher-order state of awareness—is a strange way to envision consciousness.<sup>9</sup> According to the TP, consciousness arises when one is aware of being in a state. And according to HOT theory, such states of awareness are HOTs—ordinary thoughts that are about other mental states. But having an ordinary thought about something does not confer a property to the intentional object of that thought. If I am aware of an apple by having the thought that it is red, my thought does not confer any property to the apple. How could it? Wilberg's view of consciousness requires that a HOT is a unique kind of thought, capable of conferring a special property to its target state. But it is at best unclear how a HOT could do that.

There is thus independent reason to think that consciousness is not a property of first-order states conveyed to them by HOTs. Indeed, this is why Rosenthal himself has never claimed that consciousness is this kind of property.<sup>10</sup> Rather, Rosenthal has been explicit that HOTs “do not transfer the property of being conscious from themselves to their targets; indeed, they don't induce any changes whatever in those targets” (2005, p. 185). For obvious reasons, Rosenthal often stresses that consciousness is not an intrinsic or a nonrelational property of states. But while Rosenthal does sometimes write that consciousness is a relational property of states, he is also clear that this view of consciousness is not wholly accurate either (e.g., 2005, p. 179, fn. 8). In the cases wherein one's HOT accurately represents its target, there is a way in which the first-

order state can be understood to accrue a property—namely, the relational property of being represented by a HOT. But since suitable HOTs are alone sufficient for consciousness, such relations and relational properties are not necessary for consciousness. Thus Rosenthal writes,

[A] mental state's being conscious is not strictly speaking a relational property of that state. A state's being conscious consists in its being a state one is conscious of oneself as being in. Still, it's convenient to speak loosely of the property of a state's being conscious as relational so as to stress that it is in any case not an intrinsic property of mental states (2005, p. 211).

But what kind of property is consciousness? Rosenthal's remarks, I think, naturally suggest a view according to which consciousness is not a property of states at all—and I'll argue that an alternative way of understanding consciousness shows how Rosenthal's version of HOT theory can comfortably accommodate targetless HOTs. Again, all that matters for consciousness is a suitable impression that one is in a state. According to HOT theory, HOTs are the states in virtue of which one has the subjective impression that one's mental life is some way. And HOTs are, of course, states of individuals. It is thus compatible with the folk conception of consciousness that consciousness is not a property of states, but a property of individuals—namely, the property of being aware of oneself as being in a state.<sup>11</sup>

The idea that that consciousness is a property of states emerges, I believe, from a misunderstanding of the fundamental motivation for HOT theory. As noted before, Wilberg frames the TP in terms of one's awareness of states. Construing the TP in this way invites a misleading interpretation of HOT theory—an interpretation according to which HOTs somehow accompany the states that they represent and confer the property of consciousness to those existing states.

But this is not the best way to understand the TP. Simply being aware of a state, but not of its being a state of oneself, cannot result in one's being in a conscious state. My awareness of a state of someone else's mind cannot result in that state's being conscious. So though it is acceptable shorthand to say that one is in a conscious state only if one is aware of that state, it is more accurate to say that one is in a conscious state only if one is aware of oneself as being in that state. This way of casting the TP yields a version of HOT theory according to which one is in a conscious state if and only if one is aware of oneself as being in that state via a suitable HOT.<sup>12</sup> This explains why the content of a suitable HOT is that I am in a particular state, not merely that there is such a state.

How does this way of construing HOT theory help us to explain targetless HOTs? On Wilberg's view, HOTs make us aware of states. If that were so, as Wilberg argues, Rosenthal's account would be problematic because awareness is arguably factive and, in cases of targetless HOTs, there is no state to be aware of (let alone to exhibit the property of consciousness). On Rosenthal's view, by contrast, what a HOT makes one aware of is, strictly speaking, oneself. So if one has a suitable HOT, one's HOT always renders one aware of something that exists—namely, oneself.<sup>13</sup> If one has a HOT, then one exists; as Descartes said, a thought requires a thinker (see, e.g., Weisberg (2011a, pp. 418-419, fn. 20) for an argument to this effect).

Of course, HOTs do not make us aware of ourselves only; HOTs make us aware of ourselves as being a certain way. But we are often aware of things in some respect or other. For example, I can be aware of an apple as being red, but not as being delicious. So when one has a HOT, what one is aware of is oneself in respect of having a property—namely, the property of being in a state. And though it is arguably true that we cannot be aware of nonexistent things, it is clear that we are often aware of existing things, but that we take them to be ways that they are

## A Reply to Wilberg

not. In other words, awareness may be factive in respect of the existence of the thing that one is aware of, but awareness is not factive in respect of the properties of that existent thing. I may not be able to be aware of a red apple if the apple does not exist, but I can be aware of the apple as being green even if it is in fact red.

The same goes for HOTs (cf. Weisberg, 2011a, pp. 418-419). Sometimes one is aware of oneself as being in a state that exists (when one's HOT is accurate) and sometimes one is aware of oneself as being in a state that does not exist (when one's HOT is targetless). Thus the fact that one's HOT can make one aware of oneself in an inaccurate way poses no problem for HOT theory. Whether the target of a HOT exists and the degree to which it is accurately represented is irrelevant to whether or not there is consciousness.<sup>14</sup> When one has a suitable HOT, one is aware of oneself as being in a state. And the suitable impression that one is in a state is all there is to consciousness. In other words, a suitable HOT is both necessary and sufficient for one to be in a conscious state. And if this account of consciousness is correct, then it is the individual—not the state—that exhibits consciousness.

One might immediately object that the logical grammar of expressions such as 'is in a conscious state' implies that if one is in a conscious state one must be in a state exhibiting a property of consciousness. Wilberg asserts that "[t]o say that subject S is in a conscious mental state of type M at time t, can only mean that there exists a conscious mental state token [S,M,t]" (2010, p. 626). But, as Rosenthal remarks in a recent paper, "one's being in a conscious state does not imply being in the state one is aware of being in; the phrase 'is in a conscious state' is in that way nonextensional" (2011, p. 432). One way to unpack Rosenthal's claim would be to hold that when one has a suitable HOT, one exhibits the property of being-in-a-conscious-state,

wherein this is understood to be the property of an individual who is suitably aware of being in some state.<sup>15</sup>

Naturally, this is not to deny that we can attribute conscious states to ourselves and to others or to deny that the very same mental state can be conscious at one time and nonconscious at another. It is only to say that what one's being in a conscious state amounts to is one's being aware of oneself as being that state, not that a conscious state exhibits a property of consciousness that a nonconscious state lacks. What might seem to be a property of a state is actually a property of an individual's representing itself as being in a state.

Once consciousness is thus properly understood, we can say that, whenever there is a suitable HOT, there is a conscious state from the individual's point of view. Conscious states are whatever states one is subjectively aware of oneself as being in. So there is, after all, a way in which we can comfortably describe consciousness as attaching to states, even in cases of targetless HOTs. If we do so, as Rosenthal proposes, we apply the property of consciousness to notional states. But this too is accommodated by HOT theory. We often apply properties to notional objects. And it is unclear what evidence could be brought to bear upon the decision between the view that consciousness is a property of notional states and the view that it is a property of actual individuals representing themselves as being in those notional states.

## **5. Conclusions**

Wilberg's "No Consciousness Account" of targetless HOTs clashes with our ordinary conception of consciousness. Rosenthal's account better captures that ordinary conception, though one might have thought that the targetless-HOT objection would nonetheless propel us to accept Wilberg's view. But Wilberg's view is unnecessarily problematic. There are good reasons

to think that consciousness is a property of individuals and that Rosenthal's account thereby comfortably accommodates targetless HOTs.

What it is for one to be in a conscious state is for one to have the suitable appearance of a state, whether or not one is in that state. This is not to say that consciousness does not exist or that it is only a matter of appearances. HOTs and the appearances that they reflect are real, though it is the individuals, and not the states, that are conscious.<sup>16</sup>

## Notes

1 Many, of course, deny that mental appearance and mental reality can diverge in this way (e.g., Nagel, 1974; Kripke, 1980; Chalmers, 1996). For example, speaking of the difference between appearance and reality in general, Thomas Nagel writes, "Experience itself, however, does not seem to fit this pattern. The idea of moving from appearance to reality seems to make no sense here" (1974, p. 444). However, this claim—the claim that how one's mental life appears to one cannot diverge from how one's mental life is—is rarely defended and I'll argue that it is inconsistent with what is embedded in our folk conception of consciousness.

2 One might immediately object that since consciousness is a property of states and since states are properties of individuals, consciousness is, naturally, a property of individuals—that is, a second-order property of individuals. I'll argue, however, that consciousness is simply a property of individuals, and not in virtue of being a property of states.

3 Henceforth, I will use 'consciousness' to refer to state consciousness.

4 There are several conditions that HOTs must meet in order to result in consciousness (see, e.g., Rosenthal, 2005, chapter 7), but not all are relevant here. I will refer to the sorts of HOTs that result in one's subjective awareness of oneself as being in a state as suitable HOTs.

## A Reply to Wilberg

5 One might think that a targetless HOT itself is a state that exhibits the property of consciousness. But as Wilberg also observes, HOT theory rules out this possibility because the theory holds that for any state—including a HOT—to be conscious it too must be targeted by a yet higher-order HOT.

6 For discussion of how our verbal reports of our mental states coincide with those states' being conscious, see, e.g., Rosenthal, 2005, chapter 2, specifically pp. 55-63.

7 I thank Rosenthal (personal communication) for clarifying his views about the nature of tip-of-tongue phenomena. See also his (2012a).

8 Others have provided replies to the targetless-HOT objection as well. Some replies are more in line with the spirit of Rosenthal's approach (e.g., Weisberg, 2011a; Weisberg, 2011b; Brown, 2012; Brown and Lau, forthcoming) and others are less so (e.g., Matey, 2006; Gennaro, 2006; Gennaro, 2012). I will not explore these proposals here, though I'll recommend a solution to the targetless-HOT objection that I think is consistent with Rosenthal's replies to the objection.

9 The idea that consciousness is a property that mental state tokens somehow acquire or lose fits far more comfortably with a first-order view of consciousness—a view that rejects that consciousness consists in some kind of higher-order awareness (e.g., Block, 2007; Prinz, 2012). But Wilberg explicitly defends a higher-order view.

10 If one assumes that consciousness is a property of states that HOTs somehow confer to their targets, then one might object to HOT theory along these lines. For example, in a commentary on Rosenthal (2000a), Georges Rey (2000, p. 228) wonders how a HOT could render a state conscious if having an ordinary thought about something does not change what the thought is about. But, as Rosenthal (2000b, p. 241) replies, it is question begging to conclude that HOT

## A Reply to Wilberg

theory is therefore problematic—it is an open question whether or not consciousness is a property conveyed by a HOT to a target state.

11 This way of understanding consciousness was suggested to me in conversation by David Pereplyotchik.

12 As Wilberg observes (2010, p. 626), Rosenthal frequently puts HOT theory in these terms; this description is Rosenthal's preferred way of construing the theory.

13 One might object to this account along Humean lines, arguing that there is no such thing as a self to be aware of. But the notion of a self required for HOT theory is quite minimal. Suppose I have the HOT that I perceive a red apple. As Rosenthal has urged (e.g., 2005, chapter 13; 2012b), one way to understand how this HOT renders me aware of myself is to analyze it simply in terms of a disposition to identify the referent of "I" in the HOT with the thinker of that very HOT. But even if we can be or even often are aware of ourselves, one might nonetheless object that this account fails because there are experiences that do not involve any awareness of self. Some claim, for example, that mystics and others can have conscious experiences that do not involve awareness of anything, including oneself (e.g., Forman, 1998). The existence and nature of such conscious experiences is, however, a matter of controversy (for a review, see, e.g., Gennaro, 2008). Moreover, it is unclear how people could report on such experiences without being aware of them as their own experiences.

14 This is why, as Wilberg notes, Rosenthal holds that "the distinction between an absent target and a misrepresented target is in an important way arbitrary" (2004, p. 32; quoted in Wilberg, 2010, p. 620). As Rosenthal has argued (e.g., 2005, p. 354, fn. 34), any HOT that represents there to be a state that does not exist can be described as misrepresentational, because there is no reason to suppose there is not some state it could be said to be misrepresenting to some (perhaps

large) degree. Likewise, any misrepresentational HOT could be described as representing a state that does not exist.

15 One might also worry that if HOT theory holds that consciousness is a property of an individual's representing itself, then the theory is not really a higher-order theory of consciousness at all (I thank an anonymous journal reviewer for bringing this objection to my attention). On the view developed here, consciousness does not require a creature to be in two actual and suitably related states—it only requires an individual to be in one actual state, a suitable HOT. But even on this interpretation, the theory is decidedly higher-order. Again, HOTs render individuals aware of themselves as being in states. Since HOTs make (albeit oblique) reference to mental states distinct from themselves, they constitute higher-order forms of awareness. In that way, consciousness remains a higher-order phenomenon.

16 I thank Dean Chapman, Douglas Keaton, Myrto Mylopolous, Ben Phillips, Rohan Charles Prince, Jake Quilty-Dunn, Dan Shargel and three anonymous journal reviewers for their helpful comments on drafts of this paper. I also thank Richard Brown, Michal Klincewicz, Pete Mandik, David Pereplyotchik, Josh Weisberg, and audiences at the (E)SPP and SSPP for their discussion of this material. I especially thank David Rosenthal—many aspects of this paper were shaped and influenced by discussions with him.

## References

Block, N. (2007). *Consciousness, function, and representation: Collected papers, volume 1*.

Cambridge, MA: MIT Press.

Block, N. (2011a). The higher-order approach to consciousness is defunct. *Analysis*, 71 (3), 419-431.

## A Reply to Wilberg

- Block, N. (2011b). Response to Rosenthal and Weisberg. *Analysis*, 71 (3), 443-448.
- Brown, R. (2012). Review of Rocco J. Gennaro's *The consciousness paradox: Consciousness, concepts, and higher-order thoughts*. *Notre Dame Philosophical Reviews*, retrieved from <<http://ndpr.nd.edu/news/30848-the-consciousness-paradox-consciousnessconcepts-and-higher-order-thoughts/>>.
- Byrne, A. (1997). Some like it HOT: Consciousness and higher-order thoughts. *Philosophical Studies*, 86 (2), 103-129.
- Carruthers, P. (2011). Higher-order theories of consciousness. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy*, retrieved from <<http://plato.stanford.edu/entries/consciousness-higher/>>.
- Chalmers, D. J. (1996). *The conscious mind: In search of a fundamental theory*. Oxford: Oxford University Press.
- Forman, R. K.C. (1998). What does mysticism have to teach us about consciousness? *Journal of Consciousness Studies*, 5 (2), 185-201.
- Gennaro, R. (2006). Between pure self-referentialism and the (extrinsic) HOT theory of consciousness. In U. Kriegel & K. Williford (Eds.), *Self-Representational Approaches to Consciousness* (pp. 221–248). Cambridge, MA: MIT Press.
- Gennaro, R. (2008). Are there pure conscious events? In C. Chakrabarti & G. Haist (Eds.), *Revisiting Mysticism* (pp. 100-120). Newcastle, UK: Cambridge Scholars Press.
- Gennaro, R. (2012). *The consciousness paradox: Consciousness, concepts, and higher-order thoughts*. Cambridge, MA: MIT Press.
- Kriegel, U. (2003). Consciousness as intransitive self-consciousness: Two views and an argument. *Canadian Journal of Philosophy*, 33, 103-132.

## A Reply to Wilberg

- Kripke, S. (1980). *Naming and necessity*. Cambridge, MA: Harvard University Press.
- Lau, H. & R. Brown (forthcoming). The emperor's new phenomenology? The empirical case for conscious experience without first-order representations. In A. Pautz & D. Stoljar (Eds.), *Festschrift for Ned Block* (pp. xxx). Cambridge, MA: MIT Press.
- Lau, H. & D. M. Rosenthal (2011). Empirical support for higher-order theories of conscious awareness. *Trends in Cognitive Sciences*, 15 (8), 365-373.
- Levine, J. (2001). *Purple haze: The puzzle of consciousness*. Oxford: Oxford University Press.
- Mandik, P. (2009). Beware of the unicorn: Consciousness as being represented and other things that don't exist. *Journal of Consciousness Studies*, 16 (1), 5-36.
- Matey, J. (2006). Two HOTs to handle. *Philosophical Psychology*, 19 (2), 151-174.
- Nagel, T. (1974). What is it like to be a bat? *Philosophical Review*, 83 (4), 435-450.
- Neander, K. (1998). The division of phenomenal labor: A problem for representational theories of consciousness. *Philosophical Perspectives*, 12, 411-434.
- Prinz, J. J. (2012). *The conscious brain: How attention engenders experience*. Oxford: Oxford University Press.
- Rey, G. (2000). Role, not content: Comments on David Rosenthal's "Consciousness, content, and metacognitive judgments." *Consciousness and Cognition*, 9 (2), 224-230.
- Rosenthal, D. M. (2000a). Consciousness, content, and metacognitive judgments. *Consciousness and Cognition*, 9 (2), 203-214.
- Rosenthal, D. M. (2000b). Metacognition and higher-order thoughts. *Consciousness and Cognition*, 9 (2), 231-242.
- Rosenthal, D. M. (2004). Varieties of higher-order theory. In R. Gennaro (Ed.), *Higher-order theories of consciousness: An anthology* (pp. 17-44). Amsterdam: John Benjamins.

## A Reply to Wilberg

Rosenthal, D. M. (2005). *Consciousness and mind*, Oxford: Clarendon Press.

Rosenthal, D. M. (2009). Concepts and definitions of consciousness. In W. Banks (Ed.),

*Encyclopedia of consciousness* (pp. 157-169), Amsterdam: Elsevier.

Rosenthal, D. M. (2011). Exaggerated reports: Reply to Block. *Analysis*, 71 (3), 431-437.

Rosenthal, D. M. (2012a). Higher-order awareness, misrepresentation, and function.

*Philosophical Transactions of the Royal Society B: Biological Sciences*, 367 (1594),

1424-1438.

Rosenthal, D. M. (2012b). Awareness and identification of self. In J. Liu and J. Perry (Eds.),

*Consciousness and the Self: New Essays* (pp. 22-50), Cambridge,UK: Cambridge

University Press.

Weisberg, J. (2011a). Misrepresenting consciousness. *Philosophical Studies*, 154 (3), 409-433.

Weisberg, J. (2011b). Abusing the notion of what-it's-like-ness: A response to Block. *Analysis*,

71 (3), 443-448.

Wilberg, J. (2010). Consciousness and false HOTs. *Philosophical Psychology*, 23 (5), 617-638.