

Metacognition and Higher-Order Thoughts

David M. Rosenthal

*City University of New York, Graduate School, Philosophy and Cognitive Science,
365 Fifth Avenue, New York, New York 10016-4309*

E-mail: dro@rucss.rutgers.edu

Let me begin by expressing my gratitude to Kati Balog, Thomas Nelson, and Georges Rey for their thoughtful comments. It has been a pleasure reading and responding to their careful and provocative challenges.

Because there is a fair amount of overlap in the points by Balog and Rey, I will organize this response topically, referring specifically to each commentator as relevant. And, because much of the discussion focuses on my higher-order-thought (HOT) hypothesis independent of questions about metacognition, I will begin by addressing a cluster of issues that have to do with the status, motivation, and exact formulation of that hypothesis.

1. THE HOT HYPOTHESIS

The leading idea behind the HOT hypothesis is that a mental state is conscious only if one is, in some suitable way, conscious *of* that state. This is evident because we do not regard as conscious any mental state of which one is in no way whatever conscious. It decisively defeats an ascription of a conscious state to somebody for that person to be in no way whatever conscious of being in that state.¹ As Balog in effect notes (n. 1), this is a piece of pretheoretic folk psychology, put forth to help isolate the particular phenomenon of consciousness that needs theoretical explanation.²

I argue, still in folk-psychological terms, that we know of only two ways of being conscious *of* something: roughly, thinking about the thing as being present and sensing that thing.³ One reason I favor the HOT model of how we are conscious of our conscious states over the ‘inner sense’ model is, as Balog notes, because the HOT

Reply to Commentaries on David M. Rosenthal. (2000). Consciousness, content, and metacognitive judgments. *Consciousness and Cognition*, 9(2), 203–214. This article is part of a special issue of this journal on Metacognition and Consciousness, with Thomas O. Nelson and Georges Rey as Guest Editors.

¹ So it does not go against this to note, with Rey, that psychological theory has not as yet determined whether ‘the thoughts and feelings that flow by apparently unnoticed are conscious or not’ (n. 3) is beside the point. Moreover, it is arguable that more is involved in noticing something than in being conscious of it; if so, a state’s being apparently unnoticed would not show that one isn’t conscious of it.

² Hence, as she notes elsewhere, the connection is not intended to capture ‘the meaning of what it is for a mental state to be conscious.’

³ We need this folk-psychological conception of being conscious of something to capture the way in which a state that somebody is in no way conscious of does not count as a conscious state. So Rey’s observation that scientific psychology might come up with additional ways of being conscious *of* things (n. 3), is at right angles to my point.

model has the resources needed to differentiate among intuitively distinct cases; more on that below. But that is not my main reason for rejecting inner sense. The inner-sense model will not work because sensing always involves sensory qualities, no *higher-order* sensory qualities figure in the way we are conscious of our conscious states. This undermines the ingenious argument Balog suggests on the inner-sense theorist's behalf.⁴

Rey distinguishes two, "not obviously compatible," versions of the HOT hypothesis, both of which he sees me as putting forth in different places. On (TAR-C), the state a HOT is about cannot fail to exist, whereas on (HOT-C) it can. My HOT hypothesis is the second, not the first. I construe a thought's being about something so as to allow thoughts about nonexistent things; so the conscious state a HOT is about may be a merely notional state and may not actually exist.

There is, of course, a superficial awkwardness to this view. How can a state be conscious if that state does not even exist? And how, then, can I still claim just to be unpacking our commonsense, folk-theoretic views? The pressures these questions generate may well make it seem as though I must—or should—be putting forth instead Rey's (TAR-C).

But (HOT-C) seems actually to do better justice to the folk-theoretic phenomena. A conscious state is a state one is conscious of *oneself as being in*. What matters to consciousness is how one's mental states *appear* to one, not how they actually are. Ascriptions of mental states independently of consciousness trade on causal connections those states have to stimuli, behavior, and other mental states. For these connections to hold, the states must exist. But when one ascribes mental states as conscious states or report conscious states as one's own, all that matters is how the subject's mental life appears to that subject. Insofar as we describe a state as being conscious, that state need not actually exist, but can be merely notional.

It might seem as though the folk notion of consciousness does not allow for one's being conscious of something that is not there. Whatever the case about that, it is plain that being conscious of something is compatible with one's representing that thing inaccurately. And there is no nonarbitrary way to distinguish being conscious of some particular thing as having properties it does not have from being conscious of something that does not actually exist. These considerations affect the issue of confabulatory states, to which I will turn below.

Rey distinguishes (HOT-C) from (TAR-C) in terms of two ways we speak of one thing's being *of* another. A photograph can arguably be of something only if there is a causal link between them—Rey's "aetiological" 'of',⁵ whereas a thought can be of something that does not even exist—Rey's "intentional" 'of.' If one agreed with Balog (n. 2) that sensing something involves a causal link, the aetiological "of" would be relevant to the inner-sense model. But, if the higher-order state is a thought, it is hard to see the temptation of the (TAR-C) construal, which trades on the aetiological 'of.'

⁴ Some theorists, Rey (e.g., 1993) among them, have argued that the mental states involved in sensing are all a special case of intentional states. But, again, my argument here appeals only to the folk-psychological distinction between sensing and thinking.

⁵ Though perhaps I can appropriately describe a photograph of Olivier playing Hamlet as a photograph of Hamlet, or the result of trick photography as a photograph of a centaur.

It is worth stressing that I described the relevant HOT not in terms of its being *of* a lower-order state but, rather, in terms of its complete content. As I wrote, HOTs are “thoughts to the effect *that* one is, oneself, in a state of a particular sort.” This underscores that it is the intentional “of” that matters.

2. CONTENT VERSUS MENTAL ATTITUDE

This characterization of HOTs also bears on several other challenges raised by Balog and Rey. To say that HOTs are thoughts to the effect that one is in some particular state is to characterize them not just in terms of their content, but also in terms of their mental attitude, what Rey calls their “attitude role.”⁶ So I would reject Rey’s characterization of the HOT hypothesis as “largely a matter of content, and not role.”

Balog and Rey both urge that the HOT hypothesis makes it too easy to be in conscious states, and that this is especially implausible for conscious qualitative states, such as pains. Rey puts the point in terms of having thoughts *about* mental states; consciousness will be “too easy,” “[g]iven how easy it is for anyone to think about most anything.” Thinking by itself should not be able to make it so in the case of conscious qualitative states.

But thinking something is a lot less easy when some specific content is called for. It is easy to have thoughts about Santa Claus, but not so easy to form an occurrent, assertoric thought that Santa Claus is here in this room—really to believe it, as we might say. It is not the content that makes it hard, but the assertoric mental attitude. I can readily form many types of intentional state with the content that Santa is here, but not readily hold an occurrent, assertoric mental attitude toward that content; forming occurrent, assertoric thoughts is not just up to us in that way. Having HOTs with the content of our choice is not easy.

Can it really be, as Balog asks rhetorically, that “just by thinking that I am having the pleasant sensation of a back rub I could put myself into a state subjectively indistinguishable from experiencing such sensations”? Arguably yes, if we are talking not about imagining having back-rub sensations, but actually having an occurrent, assertoric thought that I have them. Self-suggestion is powerful and presumably takes place by way of getting oneself actually to think certain things; would that getting oneself to have such thoughts were easier. I will say more below in discussing confabulated states.

Balog and Rey both object that, as Balog puts it, “[i]t certainly seems possible” that a mental state might remain inaccessible despite its being accompanied by a roughly simultaneous HOT that one is in that state. Perhaps so. But the HOT hypothesis is an empirical hypothesis about what it is for mental states to be conscious. So we cannot expect to figure out just by thinking about it whether the situation Balog envisages is, after all, a real possibility than we could have figured out just by thinking about it whether water can occur in liquid state on the moon. We need theory and empirical discovery to determine what is really possible here, as opposed to merely conceivable or imaginable.

⁶ The target article did not make anything special of this point, but I have done so in many other articles, e.g., in (1986, p. 347) and (1993, p. 742).

Rey's argument that we cannot preclude such a counterexample appeals to Freudian repression. This arguably strengthens the objection, since some mechanism of repression seems tolerably well-grounded in theoretically sound empirical discovery.

Still, there are problems with the objection, even put this way. Consider the example Rey gives of a person having HOTs about a state that remains unconscious by being repressed. In Rey's case, Ann thinks of herself as selfish or vicious, but that is Ann's having a thought about her character, not a HOT to the effect that she is in a particular mental state. Nor, I think, will it be easy to come up with another example that is intuitively and theoretically convincing.

On Freudian theory, a person might desire to do something and that desire might provoke feelings of guilt that one has that desire. But, despite that higher order guilt directed on the desire, the desire might remain unconscious. But though higher-order guilt might have the right kind of intentional content, its mental attitude is not assertoric. Given the qualification about mental attitude, it is far from obvious that Freudian theory warrants any higher order intentional states that constitute counterexamples to the HOT hypothesis.

There is another difficulty with the appeal to Freudian repression. Suppose one represses a forbidden desire. Much repression, perhaps all, takes place by radical redescription of the content and object of one's desire, or by creating so much mental noise whenever the desire occurs that one is distracted from it and pays it no attention. But states can be conscious without one's paying any attention to them.⁷ And if one redescribes to oneself a forbidden desire as a desire for something acceptable, the desire may well be conscious, even though not in respect of its actual content. As noted earlier, a state can be conscious in respect of mental properties it does not have, and a state's being repressed, like its being conscious, is relative to a description.⁸

3. WHAT IS CONSCIOUSNESS?

The term "consciousness" applies to a number of distinct phenomena, whose interrelations are often less than obvious and subject to controversy. What specific phenomenon is the HOT hypothesis meant to explain?

Nelson takes my target article "to be focussed on *self-consciousness*," presumably because HOTs are in part about oneself. But one can be conscious of one's own mental states by having thoughts about them without self-consciousness, as we ordinarily understand that notion. Self-consciousness involves being explicitly conscious of oneself as the subject in which all one's conscious thoughts and experiences come together mentally; one is conscious of oneself as the mental focal point that seems to unify all one's conscious states. Such rich consciousness of oneself is not necessary for one to be conscious of one's mental states by having HOTs about them. HOTs can, instead, represent the self in some minimal way, for example, in terms simply of a distinction between oneself and everything else. In addition, one will not be self-

⁷ Indeed, it is doubtless just lack of conscious attention that characterizes the long-distance driver in D. M. Armstrong's (1980, pp. 59–60) frequently cited case, whom Armstrong describes as having no conscious perceptual states of the road.

⁸ I argue this in Rosenthal (1986, pp. 347–348).

conscious if the thoughts one has about the self are not conscious thoughts, and HOTs need not themselves be conscious.

Nelson asks what the difference is between a mental state's being conscious and its being monitored. A necessary condition, I have argued, for a state to be conscious is that one is conscious of that state in a way that seems unmediated, automatic, and spontaneous, and having a thought about something as being present is a way of being conscious of that thing. Monitoring is neither necessary nor sufficient for that. A mental mechanism or a cluster of mental states monitors for the occurrence of some target state if that cluster or mechanism is responsive to whether the target occurs. But the mental states involved in such monitoring need not have the content or mental attitude required for one to be conscious of the monitored state. Moreover, one can be conscious of a target without monitoring if, as may happen, there is no causal tie between the target and the higher-order state in virtue of which one is conscious of the target. Being responsive to something is distinct from being conscious of it.⁹

Relying on Ned Block's now well-known distinction between access and phenomenal consciousness, Balog raises a different issue about the kind of consciousness the HOT hypothesis seeks to explain. A state is access conscious, according to Block, when it is "poised to be used as a premise in reasoning, . . . [and] for [the] *rational* control of action and . . . speech" (1995, p. 231; emphasis Block's). By contrast, a state is phenomenally conscious when there is something it's like for one to be in that state. So phenomenal consciousness applies only to qualitative states. Balog believes that the HOT hypothesis fails to address phenomenal consciousness.

There are difficulties, however, with both of Block's two notions.¹⁰ For one thing, a state can be access conscious without its being in any intuitive way a conscious state. Many states that are not, from a commonsense, folk-psychological point of view, conscious states, are nonetheless poised for use in reasoning and the rational control of action; not all inference and rational control is conscious. Like metacognitive monitoring, access consciousness can occur without the target state's being conscious, and without one's being in any way conscious of the state. So the HOT model is not, as Balog suggests, "a theory of access consciousness."

Balog also suggests that the HOT model cannot do justice to what Block calls phenomenal consciousness. I do not agree. If a subject is in no way whatever conscious of a qualitative state, there is "nothing it's like" to be in that state—that is, nothing it's like *for the subject*. That is part of what seems initially inviting about the inner-sense model; it seems as though higher order qualitative states must be responsible for there being "something it's like" for one to be in conscious qualitative states.

Qualitative states sometimes occur without one's being in any way conscious of

⁹ Nelson "would like to know what 'additional separation' [between meta- and object-level states] Rosenthal has in mind that he is claiming to disagree with." My concern was with his suggestion (1996, pp. 109–110, 1997, p. 106) that meta- and object-level states occur in distinct modules or brain areas, whereas I believe that the two types of state may well occur functionally and neurophysiologically interspersed.

¹⁰ See Rosenthal (1997, unpublished a).

them, and so without there being anything it's like to be in them; this happens, for example, in subliminal perception. So being a conscious state is not necessary for a state to be qualitative. As Balog notes, the target article does not address the question of what properties distinguish qualitative states from others, though I have done so elsewhere. A state is qualitative if it has properties that resemble and differ from other properties in ways that parallel the similarities and differences among a range of perceptible properties of physical objects or, for the case of bodily sensations, a range of bodily conditions that we can sense. This characterization of qualitative properties is independent of whether the states in question are conscious.¹¹

But if qualitative states can occur without intuitively being conscious states—without, that is, one's being in any way conscious of them, then a state's simply being qualitative is not by itself being in some special way a conscious state. Phenomenal consciousness is the property of there being something it's like for a subject to be in a state; so if a qualitative state is not conscious, it has no phenomenal consciousness.

Balog argues that “the higher order states whose subjects are sensory states cannot play a role in the constitution of phenomenal properties since their very subjects already have phenomenal properties.” Yes and no. The target states have qualitative properties, but there is nothing that it's like for one to be in those target states unless one is conscious of the states. Qualitative properties do occur without HOTs, but HOTs are needed for there to be something it's like for one to be in such states. So if phenomenal properties are simply these qualitative properties that can occur nonconsciously, the target states do, on their own, have phenomenal properties. But if a state's having phenomenal properties means that there is something it's like to be in that state, the targets do not, by themselves, have phenomenal properties. HOTs do not contribute the qualitative properties themselves, but one's consciousness of them.

It might seem as though HOTs cannot achieve this. How could HOTs, which themselves have no qualitative properties, make the difference between there being something *qualitative* that it's like for one to be in some state and there being nothing that it's like? Balog's doubts that HOTs can do this job fuel her preference for a theory, like that of Brian Loar (1990/97), on which a thought represents a qualitative state by that qualitative state's occurring as a part of the thought itself. Putting aside the difficulty of telling in a non-question-begging way whether one state is part of another, if the qualitative state that is supposed to be part of the HOT is not conscious on its own, why would its being part of the HOT help? If the answer is that one is then conscious of the qualitative state, why cannot a HOT without qualitative parts do that? Whether or not one state is part of another should not by itself, make a difference to whether that state is conscious.

Balog doubts whether HOTs are even relevant to phenomenal consciousness. But there is compelling reason to think that they are. Qualitative differences among our experiences often become conscious only when we learn and master concepts fine-grained enough to draw the qualitative distinctions in question. The auditory experiences of oboe and clarinet may come to seem different only when we have words that

¹¹ See Rosenthal (1991, 1999, 2000, unpublished b). Though this account is, strictly speaking, independent of the HOT model, the two fit well together and lend support to each other.

capture the different experiences; similarly with some visual, gustatory, and olfactory experiences. Coming to have concepts of the qualities of sensory experiences could matter to how we are conscious of those experiences only if our thoughts about the experiences made a difference to how we are conscious of them.

One might object that we have no idea how HOTs could result in these differences. If what this means is that it is not evident from introspection how HOTs could make the difference, the objection is misplaced; on the current hypothesis, HOTs need not be conscious to do their work, and accordingly they need not be accessible to introspection. But in any case, there is no reason to expect that introspection or reason ever determines, independent of empirical discovery, which mechanisms will have which results, in the mental realm or elsewhere. It appears otherwise only when a theory of the connection is so well-entrenched as to be taken for granted.

Balog suggests toward the end of her discussion that HOTs would be relevant to phenomenal consciousness only if sensory states have intentional properties but in themselves no phenomenal properties. I have discussed the issue of phenomenal properties; if a state's having a phenomenal property means that there is something it's like for one to be in that state, I agree with that condition. But I do not see why a target sensory state's having intentional properties would help. Sensory states must exhibit some type of mental property. But intentional properties are not the only alternative to phenomenal properties, so construed; they can, and do, exhibit nonconscious qualitative properties.

4. CONFABULATED CONSCIOUS STATES

These considerations having to do with qualitative states connect with worries both Balog and Rey express about my countenancing confabulated qualitative states—about whether we can be conscious of ourselves *as being* in qualitative states that we are not actually in.

Rey sees adopting the HOT model (as opposed to his [TAR-C]) as committing one to the possibility of confabulated qualitative states, whereas Balog believes that it does not. Strictly speaking, no such commitment obtains. The HOT model allows for a HOT to occur unaccompanied by its notional target, but that occurrence might result in nothing, so far as consciousness is concerned. It could be that there just is not anything it's like for one to have a HOT unaccompanied by its notional target. Still, because HOTs and their targets are distinct and independent occurrences, the model invites the speculation that an unaccompanied HOT might, for a subject, be introspectively indistinguishable from a HOT accompanied by a state that matches its notional target.

There are two sorts of reason that might lead one to deny this possibility: reasons of theory and reasons of introspection. It might seem introspectively obvious that, if it seems to one that one is in a mental state, then one is. But introspection cannot help here, since it tells us only how things appear, not how they actually are. It is sometimes held that introspection cannot be erroneous, but plainly one cannot appeal to introspection itself to support this questionable doctrine. If we deny that confabulated conscious states are possible, it must be for reasons of theory.

One theoretical reason already discussed is that if we could confabulate conscious

states, it would be easier than it is to generate the appearance of being in mental states of our choice. But, as noted earlier, that is not so. Simply imagining (Balog) or being in some intentional state about (Rey) a notional target is not enough; one must have an occurrent, assertoric thought that one is in that state. One can pretend to think just about anything, but actually thinking it is another story. It is important to note, however, that we cannot determine by intuition or introspection just how easy it is to generate the convincing appearance of being in some conscious state. Research continues to reveal the surprising efficacy of self-suggestion and the placebo effect, in which genuine assertoric HOTs presumably occur.¹²

Rey also considers the opposite difficulty, that the HOT model would make it easier than it is to prevent unwanted mental states from being conscious. But just as it is not that easy actually to think something, as opposed to imagining or pretending to think it, it is also not that easy to stop thinking something one already thinks. Computational models of thinking, which Rey favors, may make it seem easier than it is to alter one's thoughts, since altering thoughts is then a bit like flipping a switch. But it may not be all that easy actually to flip that switch, and in any case it is notoriously not that easy to change one's thoughts.¹³

One might reject the possibility of confabulated conscious states because one holds that, for consciousness, there is no difference between appearance and reality, or that consciousness is an intrinsic property of mental states and so cannot occur in the absence of the state that seems to be conscious. Some theorists, indeed, seem to take these doctrines as in effect defining the subject matter of consciousness. But without independent support, these controversial claims are plainly question begging. Appeal to raw intuition on this issue does little more than articulate what assumptions a theorist is relying on.

Balog raises the interesting possibility of a HOT zombie, whose inner life is subjectively indistinguishable from ours despite the lack of sensory states. This Balog finds paradoxical. But the intuitive paradox rests on an ambiguity in 'sensory state.' The sensory states the HOT zombie would lack are only nonconscious states. Since conscious states are states one is conscious *of oneself as being in*, notional states are all that matter for the purposes of consciousness.¹⁴

¹² In related research Peter S. Staats (1998) reports findings in which rehearsed positive and negative thoughts whose content is independent of pain modify the effect of painful stimuli, both as subjective reported and by standard physiological measures.

¹³ Eliminativists of the sort Rey considers in this context are typically clear about how hard it would be, if possible at all, to get people to drop their folk-psychological thinking.

In any case, even if eliminativist training of the type Paul Churchland (1981) envisages did lead one to stop having HOTs cast in folk-psychological terms, one would not thereby become unconscious. A person's not being conscious is different from that person's mental states' not being conscious, as anaesthesiologists are well aware.

Churchland, moreover, envisages our "mutual understanding and even our introspection" being "re-constituted" (Churchland 1981, p. 67) within the new conceptual framework. And if that happened, new states might well play the role of HOTs, with a result subjectively indistinguishable from the current folk-psychologically described situation.

¹⁴ Moreover, nonconscious qualitative states arguably have characteristic causal connections with other states and with sensory stimuli (see references in footnote 11). If so, nonconscious qualitative states could not be absent without significant behavioral consequences.

Balog notes that if Tyler Burge is right that one cannot think that one is thinking something without also thinking that thing, we cannot confabulate such intentional states. But Burge's claim concerns

5. THE TOT PHENOMENON

In the target article I argued that the TOT phenomenon occurs when we are conscious of being in an intentional state that carries the desired information but not conscious of that state in respect of the relevant informational content. I welcome this chance to put my view account in a slightly different way.

Suppose I know Orwell's real name; then I am in an intentional, possibly not conscious, to the effect that Orwell's real name is 'Eric Blair.' We could also describe that very state simply as a state whose content specifies Orwell's real name. Even so, it is possible to be conscious of being in an intentional state whose content specifies Orwell's real name, but nonetheless fail to be conscious of being in a state whose content is that his real name is 'Eric Blair.' In the first case I am conscious of the state in respect of less of its intentional content than in the second case, but still in respect of some of its intentional content.

What matters in the TOT phenomenon is what information I want; since I want conscious access to the name itself, it is tempting, as I suggested in the target article, to say that if I assume not conscious of the state as having the content that Orwell's real name is "Blair," I am not conscious of that state in respect of its intentional content. But that is, in effect, shorthand for saying that I am not conscious of the state in respect of the content that matters to me.

Compare one's recalling that Orwell's real name begins with a 'b,' but not that it is 'Blair.' Described as the thought that Orwell's real name is 'Blair,' the state is not conscious, though described simply as the thought that the name begins with a 'b' that very same state is conscious. Our interests sometimes determine whether or not we regard a particular state as conscious.¹⁵

6. MISCELLANEOUS OBJECTIONS

I close by addressing a few objections not already covered. Not every mental state we are conscious of is a conscious state; one must be conscious of the state in a way that appears to one to be direct and spontaneous. I argue that we can provide for this by stipulating that the HOT in virtue of which one is conscious of the state relies on no conscious inference—no inference, that is, *of which* one is conscious.

Nelson believes that this conflicts with the "No-Magic Hypothesis" about metacognition that he and Louis Narens have advanced, on which FOK judgments rely not on actual monitoring but on inference (Nelson & Narens, 1990, p. 158). But the inference the No-Magic Hypothesis posits is presumably not normally, if ever, a conscious inference. Moreover, metacognition and consciousness, though similar, are not the same. And, if even the inference that sustains an FOK judgment is sometimes conscious, that judgment will be a conscious state only if one is conscious of it independently of that, or any other, conscious inference.

Rey argues that precluding conscious inference is not enough. One might have a HOT

higher-order judgments with performance force, such as the judgment, "I hereby judge that water is a liquid" (Burge 1988, p. 656). And it is not obvious that the HOTs invoked by the present model must have such force.

¹⁵ I welcome Nelson's remarks on the distinction between TOT and FOK, but by the phrase "involves having partial access" I meant no more than "has partial access."

passively, based solely on authority, or in a highly automatic way, based on internalized theory, and the target of such a HOT might still not be conscious. But these cases are not counterexamples; passive acquiescence and automatic appeal to internalized theory are not incompatible with reliance on an inference of which one is aware.¹⁶

Rey urges that we need not invoke HOTs to explain what it is for mental states to be conscious; instead, we can hypothesize that a state is conscious if it occurs in a cognitive buffer accessible to introspection and verbal report.¹⁷ And this, he says, would explain why conscious states are noninferentially reportable, which I have argued (1993) is best explained by appeal to HOTs.

One problem with explanations cast in terms of computational architecture is that they can get cashed out in different ways. Perhaps, when spelled out, the buffer Rey posits would simply implement the folk-psychological tools of the HOT model. A state is in that buffer if it is accessible to introspection and verbal report, and being conscious of a state by way of an inferential HOT also makes a state thus accessible. Still, such accessibility, by itself, is not enough; a state can be accessible to introspection and verbal report without one's being in any way conscious of it. So positing actual HOTs has an advantage over Rey's buffer.

This point recalls the requirement I impose that HOTs be occurrent, rather than merely dispositional. Rey misconstrues the reason for that condition. It is not to preserve the intuitively occurrent character of consciousness.¹⁸ Rather, a state is conscious only if one is conscious *of* that state, and being disposed to be conscious of something does not make one conscious of it, just as a state's simply being accessible to introspection and verbal report does not make one conscious of it.

Rey objects that "Vienna, Mars, the number three, and all of one's own and others' *past* and *future* mental states are not much affected, much less rendered conscious," by one's having thoughts about them. "Why should one's own occurrent mental state be?" (Rey's emphasis.)

This objection trades on several very different considerations. Having a thought about Mars or Vienna or the number three does not make those things conscious in part because the property of being conscious that we ascribe to objects is different from the property of being conscious that we ascribe to states. An object, such as a person or other creature, is conscious if it is awake and mentally responsive to sensory stimuli.¹⁹ Vienna, Mars, and the number three never are awake or mentally responsive

¹⁶ Rey writes that Freud holds that a state's being conscious requires "the patient's access to be of the more usual 'direct' sort one has to genuine desires, unmediated, consciously or unconsciously." I am unaware of any passage in which Freud rules out unconscious inference. Moreover, it is arguable that Freud allowed psychoanalytic interpretation to mediate a patient's access to the patient's states, as long it does not consciously seem to. And Freud plainly recognized that what matters is whether one's access *appears to one* to be unmediated.

¹⁷ Rey speaks simply of conscious contents, but I take it that conscious states, complete with mental attitude, are in question.

¹⁸ Rey thinks that the occurrence in the buffer of the target "is occurrency enough." Perhaps so, if theoretical occurrency is under consideration. But it is unclear why occurrency in the buffer would help with *intuitive* "occurrency," which depends on how things appear to one, not how they are. Intuitive occurrency is a matter of how our consciousness of our mental states represents them.

¹⁹ Conscious states need not figure in something's being mentally responsive, since not all mental states are conscious.

to sensory stimuli. And, in any case, having a thought about an object has no bearing on the property of consciousness applicable in these cases.

By contrast, the property we attribute when we say that a mental state is conscious is the property of one's being conscious of that state in a way that seems spontaneous and unmediated. Why, then, does not having thoughts about our own past and future mental states or those of others make those states conscious? Different considerations apply. Our thoughts about the states of others rely on inferences we are aware of. As for our own past and future mental states, having a thought about something makes one conscious of it only when one thinks of the thing as present to us.²⁰ We do not regard ourselves as being conscious of Caesar in virtue of our having thoughts about him. So having a thought about a state as being past or future will not make one conscious of it.

Rey notes that having thoughts about objects and states alike seems to leave those things unaffected, even one's own current mental states. How, then, can having a thought about one's current states result in their being conscious? Being conscious of one's own current mental states does leave their intrinsic, nonrelational properties unaffected. But it is question begging to assume without argument that a mental state's being conscious is a nonrelational property of that state. Indeed, the tie between a state's being conscious and one's being conscious of it shows that a state's being conscious is, at least in part, relational.

It is worth considering a difficulty related to these concerns. A state is conscious if one is conscious of that state in a way that seems spontaneous and unmediated. But even if one could have noninferential thoughts that one's liver is in various particular states, those states of the liver would still not intuitively count as conscious states (Block, 1995, p. 280).

The quick answer is that we count states as being conscious only if they are mental to begin with. Nor is this simply an arbitrary stipulation. Creatures are conscious *of* things and conscious *that* something is the case. And it is in virtue of their being in mental states that they are conscious of things and conscious that something is the case; mental states are simply states of being conscious of things and conscious that something is so. And we describe as conscious only those things which can be conscious *of* something and conscious *that* something is the case and the states in virtue of which they are conscious *of* or *that* something.²¹ Only creatures and their mental states can have the property of being conscious. Though the properties we ascribe to creatures and their mental states are different, it is not arbitrary that we describe only that limited range of things as being conscious.

²⁰ As noted in the target article, at the beginning of Section II.

²¹ And states of the liver are not states of being conscious of anything or conscious that something is the case.

Note that if we were mentally responsive to states of our livers, that would be so in virtue of our having sensations of our livers. So, if we were also conscious in a seemingly spontaneous and unmediated way of our being thus responsive to states of our livers, we would doubtless think of ourselves as having conscious sensations of our livers. The difficulty in imagining this situation is not due to any problem about having conscious sensations of bodily states; consider our conscious sensations of throbbing veins and states of our viscera. Rather, it is because we have virtually no intuitive idea of what distinctive states the liver is in to begin with.

REFERENCES

- Armstrong, D. M. (1980). What is Consciousness? In D. M. Armstrong (Ed.), *The nature of mind*, pp. 55–67. St. Lucia, Queensland: University of Queensland Press.
- Block, N. (1995). On a confusion about a function of consciousness. *The Behavioral and Brain Sciences*, **18**(2), 227–247.
- Burge, T. (1988). Individualism and self-knowledge. *The Journal of Philosophy*, **LXXXV**(11), 649–663.
- Churchland, P. M. (1981). Eliminative materialism and the propositional attitudes. *The Journal of Philosophy*, **LXXVIII**(2), 67–90.
- Loar, B. (1990/97). Phenomenal states. *PP*, **4**, 81–108; reprinted with significant revisions in (1997). N. Block, O. Flanagan, and G. Güzeldere (Eds.), *The nature of consciousness: Philosophical debates*, pp. 597–616. Cambridge, MA: MIT Press.
- Nelson, T. O. (1996). Consciousness and metacognition. *American Psychologist*, **51**(2), 102–116.
- Nelson, T. O. (1997). The meta-level versus object-level distinction (and other issues) in formulations of metacognition. *American Psychologist*, **52**(2), 179.
- Nelson, T. O., & Narens, L. (1990). Metamemory: A theoretical framework and new findings. In G. H. Bower (Ed.), *The psychology of learning and motivation*, pp. 125–173. New York: Academic Press.
- Rey, G. Sensational sentences. In M. Davies & G. W. Humphreys (Eds.), *Consciousness: Psychological and philosophical essays*, pp. 240–257. Oxford: Blackwell.
- Rosenthal, D. M. (1986). Two concepts of consciousness. *Philosophical Studies*, **49**(3), 329–359.
- Rosenthal, D. M. (1991). The independence of consciousness and sensory quality. In E. Villanueva (Ed.), *Consciousness: Philosophical issues, 1, 1991*, pp. 15–36. Atascadero, CA: Ridgeview.
- Rosenthal, D. M. (1993). Thinking that one thinks. In M. Davies & G. W. Humphreys (Eds.), *Consciousness: Psychological and philosophical essays*, pp. 197–223. Oxford: Blackwell.
- Rosenthal, D. M. (1997). Phenomenal consciousness and what it's like. *The Behavioral and Brain Sciences*, **20**(1), 64–65.
- Rosenthal, D. M. (1999). The colors and shapes of visual experiences. In D. Fissette (Ed.), *Consciousness and intentionality: Models and modalities of attribution*, pp. 95–118. Dordrecht: Kluwer Academic.
- Rosenthal, D. M. (2000). Sensory quality and the relocation story. *Philosophical Topics*, a special issue, in honor of Sydney Shoemaker, ed. Richard Moran, Alan Sidelle, and Jennifer Whiting, forthcoming, pp. 321–350.
- Rosenthal, D. M. (unpublished a). *The kinds of consciousness*. Paper presented at the University of Oxford Autumn School in Cognitive Neuroscience, Oct. 1998.
- Rosenthal, D. M. (unpublished b). *Sensory qualities, consciousness, and perception*. Paper presented at a symposium on sensation, perception, and sensory quality, at the 21st Annual Meeting of the Cognitive Science Society, Vancouver, Aug. 20, 1999.
- Staats, P. S., Hekmat, H., and Staats, A. W. (1998). Suggestion/placebo effects on pain: Negative as well as positive. *Journal of Pain and Symptom Management*, **15**(4), 235–243.