

Comments on David Rosenthal's “Consciousness, Content, and Metacognitive Judgments”

Kati Balog

Department of Philosophy, Yale University, New Haven, Connecticut 06520

E-mail: kbalog@rci.rutgers.edu

David Rosenthal's higher order thought theory of consciousness (HOT) is one of the best articulated philosophical accounts of consciousness available. The theory is, roughly, that a mental state is conscious in virtue of there being *another* mental state, namely, a thought to the effect that one is in the first state. The second mental state is “higher order” since it is about a mental state. Furthermore, the thinker does not think that the higher order judgment is arrived at by conscious inference.¹ For example, my belief (memory, thought, wish) that my child is sleeping in the other room is conscious in virtue of my having another, second order, thought to the effect that I have this belief (thought, memory, wish), where this thought is not the result of a conscious inference. Of course, there is a lot that turns on the exact details, which we will examine later.

Rosenthal's paper “Consciousness, Content, and Metacognitive Judgments” seems to be occasioned by an apparent problem for his theory. Metacognitive judgments are judgments about one's mental states, e.g., that one remembers something or does not know something and so on. He notices that metacognition, i.e., “the access we have to whether, or how likely it is that, we know something,” involves, similarly to consciousness, higher order access to mental states we are in. However, and here is the problem, we can make metacognitive judgments without consciously accessing the state's intentional content; i.e., we can judge that we know something without the content in question being conscious. I can, for example, judge that I know what the capital of Hungary is without being aware that it is Budapest. Rosenthal's example of this is the “tip-of-the-tongue” phenomenon. But this raises a question about HOT: after all, if one can have higher order access to mental states without those states being conscious, there might be something wrong with the theory of consciousness according to which the consciousness of a mental state consist in our having a higher order state about that mental state.

Commentary 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 Despite the appearances, there is no circularity problem here. Rosenthal does not wish to give a reductive *analysis* of the term consciousness. He takes it for granted that we have a workable, pretheoretical notion of consciousness; he merely proposes an empirical theory of it. This point is missed by some commentators; see e.g., Seager (1999, p. 74).

Rosenthal has an answer to this. He describes what is happening when some piece of information, say, a name, is “on the tip of one’s tongue” in the following way. When, e.g., the real name of the author of *Out of Africa* is on the tip of my tongue, I am aware of the particular state that has the intentional content that the author of *Out of Africa* is Karen Blixen, but I am not aware of it with respect to its intentional content, only *as* the state that has an intentional content that specifies who the author is. Rosenthal says that for a state to be conscious, we must be aware of it *with respect to* its intentional content, that is, we must have a higher order thought to the effect that this state, represented in this thought to exhibit its full content, occurs. By contrast, in metacognition the higher order judgment only needs to represent the content of the state as the state that has the appropriate intentional content—without representing what that content is.

Rosenthal claims that the existence of metacognitive judgments about unconscious states, instead of weakening HOT, actually supports it. He reasons as follows: he first claims that there are only two ways in which we can be conscious of things: we can sense them or we can have thoughts about them. For example, I may be conscious of the toast burning by sensing it (experiencing the smell) and conscious of Milan’s birthday by thinking that it is his birthday. On the so-called “inner sense theory” of consciousness we are conscious of our mental states by “sensing” them. He rejects this account since only states with representational content can make the requisite distinctions, i.e., represent mental states sometimes *with respect to their content*, and other times merely *as* states that have a certain intentional content. Any “inner sense” model of consciousness, according to which consciousness of a mental state consists in our “sensing” this state similarly to the way we sense the outside world and the state of our bodies, must lose out to HOT, as, according to Rosenthal, “sensing . . . involves no intentional content, and typically occurs independently of any.” So he concludes that we are conscious of our mental states by having higher order thoughts about them. To sum up, the argument goes: one of my mental states is conscious if I am conscious of it. But I can be conscious of one of my mental states only by having a higher order thought about it. So my mental states are conscious in virtue of my having higher order thoughts about them.

Whatever shortcomings the “inner sense” model of consciousness has, this criticism seems unjustified and somewhat puzzling. Metacognitive judgments of the sort Rosenthal is talking about *are* higher order thoughts on anybody’s account. The question is what conscious states are. It is true, HOT can easily handle the distinction between those higher order thoughts that represent mental states merely *as* states carrying certain information (metacognitive judgments) and those that represent them *in respect of* their content. However, this does not mean that an “inner sense” model could not handle the same distinction. Even if we accept the controversial claim that sensing involves no intentional content, there is no reason why the “inner sense” theory should deny that metacognitive judgments involve higher order *thought*, as opposed to sensation. In other words, the “inner sense” theorist might say that when we are conscious of a thought it is partly in virtue of sensing it (i.e., it is like having a perception), and that when we make metacognitive judgments about unconscious thoughts the relevant sensations that make a thought conscious are absent, though there is a higher order thought present. Then the only point of contention between

HOT and the “inner sense” model is their account of consciousness.² I am not recommending the “inner sense” theory as an account of how we know our thoughts. As Rosenthal notes, it is very implausible that we know our thoughts via accompanying sensations. However, I think that thoughts often are accompanied by sensations and feelings—events with qualitative content—and these figure in our consciousness, at least in one sense of the term. Comparing HOT with the “inner sense” account is instructive because it suggests that the sense of a mental state’s being conscious that Rosenthal has fixed on may not be the same that the inner sense theorist has in mind. And it is the latter sense that most philosophers are concerned with when they discuss consciousness.

Rosenthal gets off on the wrong foot when he sets up his theory. He seems to think that it is self-evident that a mental state’s consciousness consists in our awareness of that state, in the sense in which we can be aware of *external* objects (i.e., sensing them or having thoughts about them). But to equate awareness of mental states (in the above sense) with those states’ being conscious is not simply to unpack the meaning of what it is for a mental state to be conscious. It is useful at this point to get a little clearer about what exactly we are seeking an account of.

The problem of consciousness, as Ned Block (1994) has pointed out, is not just one problem. First, there is the problem of access consciousness: how is it that we can access certain of our mental states, that we are able to verbally report them, take them into account in the organization of our behavior, etc? Second, there is the problem of qualia: how is it that, in Thomas Nagel’s (1974) phrase, there is something it is like to hear a baby cry, or smell milk, etc? It is unclear which problem Rosenthal aspires to solve by proposing HOT.

According to Rosenthal,

All mental states, of whatever sort, exhibit properties of one [or both] of two types: intentional properties and phenomenal, or sensory, properties. (1986, p. 332)

Rosenthal believes that both properties can be had by unconscious (in the access sense of “unconscious”) mental states. This means that certain mental states can be phenomenally conscious—i.e., there can be something it is like to have them—but not access conscious. What makes mental states, i.e., both intentional states and phenomenal states access conscious is the simultaneous existence of a higher order thought to the effect that one is in the lower order state. This theory seems to have nothing to say about what phenomenal properties themselves are, as these properties can be had by states that are themselves unconscious. It only addresses itself to the question of what it is for mental states (including phenomenal states) to be access conscious.

² In a later part of his paper, Rosenthal even agrees with the “No-Magic” Hypothesis of metacognition suggested by Nelson and Narens, according to which, for example, feeling of knowing judgments (FOK), do “not reflect any monitoring of unconscious information”; instead, such judgments “utilize only suprathreshold information about remembered attributes of the item . . . , along with rules for how to utilize that information in the FOK judgments” (1990, p. 158). If a theory along these lines were true, then metacognition would not even involve any direct causal connection between the knowledge-states themselves and the judgments about those states, so sensing those knowledge-states as the basis of metacognitive judgments would not even come into consideration as a theory. And, to repeat, this is perfectly compatible with an “inner sense” account of *consciousness*.

So it may seem that Rosenthal's theory is a theory of access consciousness, and not of phenomenal consciousness.

However, something is amiss here. It is really phenomenal consciousness that poses, in Chalmers' words, the "hard problem." Theories of consciousness generally aspire to solve *it* or explain why it cannot be solved. Access consciousness, though it might bring up interesting empirical (and conceptual) issues, seems rather tractable, even if only on the contentious assumption that the problem of intentionality is solvable. And HOT does not seem to be a particularly promising account of access consciousness anyway. It certainly seems possible that there be inaccessible mental states in spite of the simultaneous occurrence of (similarly inaccessible) higher order states to the effect that one is in the first state. In any case, if HOT did not have anything to say about phenomenal consciousness, the "what it's like" feature of sensory states, it would not be a very exciting theory.

But in fact, Rosenthal seems to think that HOT does have something important to say about phenomenal consciousness. This is evident from the way he handles the question of whether there could be "confabulated" conscious states, by which he means a second-order judgment to the effect that one is in mental state *m*, while *m* in fact does not occur. Is this a case of consciousness involving being "aware" of states we are not in, or is it the case of no consciousness at all? He thinks the former is the case.

Confabulated conscious states are states we are conscious of ourselves as being in even though the states do not actually occur. We are, in this way, actually conscious of states we are not in, but subjectively seem to us to belong to our stream of consciousness. . . . Being conscious of a state does not imply that the state exists, nor if it does that its mental properties match the way we are conscious of it.

If he thinks, as is obvious from the examples he gives, that this can be the case even with phenomenal states, then it follows that higher order thoughts in themselves, even in the absence of the relevant sensations, can produce phenomenal feels. HOT then is relevant as (part of) a theory of phenomenal consciousness. But the claim that the phenomenal feel of, say, a toothache, can occur without the appropriate sensation occurring, is very puzzling at best. First of all, if the phenomenal feel belongs to the first-order sensory state, as he stated above, then pain could not occur without the sensory state occurring, unless the phenomenal property of pain can also be had by higher order thoughts as well as first order sensations, which is not very plausible. Second, this would allow for the possibility of a HOT-zombie, i.e., a creature whose inner life, from the point of view of her stream of consciousness, is identical to mine, except that she does not have any sensory states, she never has pains, and might not even ever had any.³ This, again, seems paradoxical.

As a matter of fact, there is no reason why HOT should be committed to the possibility of confabulated phenomenal states.⁴ Reflecting on the question of how

³ Whether this last is really compatible with HOT depends on what constraints HOT places on the sensory concepts utilized by the higher order thoughts. More on this problem shortly.

⁴ Even in the case of intentional states, it is controversial whether we can be conscious of having such states without the states themselves occurring. Rosenthal thinks that I can think that I am thinking that *p* (or believing that *p*) without thinking (believing) that *p*. Burge (1988) would partly deny this, as he seems to hold that to think a thought with the content that I am thinking that *p* I have to think a thought with the content that *p*.

sensory states are represented in the higher order thoughts that supposedly make them conscious, one can make a case that HOT even requires that there could not be confabulated sensory states. In the case of sensory states, the higher order thought presumably does not linguistically represent the state's content. That would make qualia too cheap: just by thinking that I am having the pleasant sensations of a back rub I could put myself into a state subjectively indistinguishable from experiencing such sensations. Alas, this is not the case. A more plausible view is that in some (but not all) higher order judgments that one is in a certain sensory or phenomenal state that very state forms part of the judgment and represents itself. In confabulated higher order states the phenomenal state is absent and so such states lack the qualitative feel of the phenomenal state. On this account one could not be conscious of the sensations of a back rub unless one were experiencing the sensations of a back rub (although of course one could falsely judge that one is having the sensations of a back rub).

The above emendation to the HOT avoids the problems of the duplication of phenomenal properties and the possibility of HOT-zombies. But it remains the case that HOT, as a theory, is irrelevant as far as the question of qualia is concerned. The problem is that HOT needs lower order *mental* states as subjects of the appropriate higher order thoughts. Mental states either have intentional properties, or they have phenomenal properties; sensations are mental in virtue of their phenomenal properties. So 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; unless one wants to hold a "double" theory of phenomenal properties. The only way to make higher order thought relevant to the problem of phenomenal consciousness would be to hold that sensory states have intentional properties, but, in themselves, no phenomenal properties. The theory then would be that what makes a sensory state have phenomenal properties as well, is for there to be an appropriate higher order thought to the effect that the sensory state occurred. In that case two intentional properties (the intentional properties of the sensory state, and the intentional properties of the higher order thought) would make for a phenomenal property. But this is not Rosenthal's view, who himself thinks that sensory states have no intentional content. This being the case, HOT, as far as qualia are concerned, is not so hot.

REFERENCES

- Block, N. (1994). On a confusion about a function of consciousness. *Behavioral and Brain Sciences*, **17**, 2.
- Burge, T. (1988). Individualism and self-knowledge. *Journal of Philosophy*, **85**, 649–663.
- Chalmers, D. (1996). *The conscious mind*. Oxford: Oxford University Press.
- Nagel, T. (1974). What is it like to be a bat? *Philosophical Review*, **4**, 435–450.
- Nelson, T. O., & Narins, 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.
- Rosenthal, D. (1986). Two concepts of consciousness. *Philosophical Studies*, **49**, 329–359.
- Seager, W. (1999). *Theories of consciousness*. London and New York: Routledge.