

Higher-Order Thought Rendered Defenseless

Review of Consciousness and Self-Consciousness: A Defense of the Higher-Order Thought Theory of Consciousness by Rocco Gennaro

Drew McDermott

Department of Computer Science
Yale University
New Haven, CT 06437
USA

mcdermott@cs.yale.edu

Copyright (c) Drew McDermott 1997

PSYCHE, 4(1), January, 1998

<http://psyche.cs.monash.edu.au/v4/psyche-4-01-mcdermott.html>

KEYWORDS: Higher-Order Thought, introspection, mentalism, qualia.

REVIEW OF: Rocco Gennaro (1996) *Consciousness and Self-Consciousness: A Defense of the Higher-Order Thought Theory of Consciousness* (Advances in Consciousness Research, Vol 6) John Benjamins. 220pp. US\$34.95; pbk. ISBN: 1556191863

I found this book to be a big disappointment. It attempts to elaborate and defend David Rosenthal's (1986, 1990, 1993) "higher-order thought" theory of consciousness, the idea that to be in a conscious mental state is to be thinking (correctly) that one is in a mental state. To have a conscious sensation of red is to be perceiving red and simultaneously to be perceiving the perception *as* a perception, of a color classed as similar to other colors, such as the color of apples. The theory is attractive because it tidily reduces consciousness to something else, namely nonconscious mental events. It is less popular than it could be for two big reasons: 1) it seems to rule out the seemingly obvious idea that one could be consciously aware of something without thinking about one's awareness; 2) it depends on having a theory of nonconscious mental states, which we don't really have.

I think these obstacles can be overcome, so I was naturally eager to read Rocco Gennaro's book *Consciousness and Self-Consciousness: A Defense of the Higher-order Thought Theory of Consciousness*. Unfortunately, the book does not deliver on the promise suggested by its subject. Nor does it even deliver an example of passable English prose: it suffers from poor grammar and spelling that good editing eliminates.

My heart sank when the author announces that he is not going to take a reductivist view of thought:

I think it best to explain conscious states in mentalistic terms. Since we cannot explain what makes a mental state conscious in physicalistic or in (non-mentalistic) functional language, then the only other option is to explain it in mentalistic language (p. 14).

The problem with this approach in the context of the higher-order-thought theory should be obvious: the theory derives all its power from appeal to a primary, nonconscious level of thought. Conscious thought then appears because of trickery going on at the nonconscious level. But it is almost impossible to say anything enlightening about nonconscious thinking if you remain at a mentalistic level. Husserl or Heidegger might be able to bring off such a feat, but Gennaro is not up to it. Mentalist talk tends to be based on introspection, and introspection is mute, or at least highly mystifying, on the topic of nonconscious thought. Whole chapters of Gennaro's book are virtually unintelligible to me because of his insistence on speculating on his intuitions about the unconscious mental life of animals, both actual and merely possible.

Gennaro begins by describing his disagreements with Rosenthal's formulation, which are illustrated by some diagrams of thoughts surrounded by ovals with arrows pointing to other thoughts (p. 25). If an oval surrounds two objects, then they are both intrinsic parts of the same thought; if an arrow points from one entity to another, then the first is "about" the second. Supposedly, Rosenthal would draw the ovals differently from Gennaro, thus indicating that Gennaro thinks the relationship between two thoughts is intrinsic, whereas Rosenthal thinks it's extrinsic. But these are nonconscious entities, not accessible to introspection, and not (so far) explained in terms of neurological or computational structures in brains. So what evidence could conceivably bear on this "intrinsicity" issue one way or the other?

After criticizing Rosenthal's view, one begins to want Gennaro to spell out his argument in favor of the basic higher-order-thought theory. But he never does. Instead, he spends a chapter (the third) arguing that the thing that "renders a mental state conscious" must be a thought, and not something else, like a belief. The difference between the two is that a belief is a disposition to behave in certain ways, whereas a thought *happens*. That's clear enough, but of course according to the theory the majority of thoughts are not conscious, so exactly what happens when a thought happens is murky. If a thought is not a conscious thought, and not an information-processing event, then I have no idea what it is. The best I can do is picture something that's just like a conscious thought, but isn't conscious. This picture is too vague to carry the load that Gennaro wants it to carry.

Then we get to Chapter 4, which is titled "Objections and Replies." This is probably the best chapter of the book; the objections are stated reasonably clearly and the replies are often quite lucid. But, as far as I can see, no opponent of the theory is under any compulsion to find a way to object to it, not until Gennaro provides an argument in favor of it, which, as far as I can tell, he never does.

After Chapter 4, the book becomes incoherent. The remaining chapters seem to be concerned with various arguments for the conclusion, "If a system is conscious, then it is self-conscious." It is often hard to tell what position Gennaro takes on the various arguments; nor is it explained what these arguments have to do with his theory. The arguments attempt to show that consciousness requires self-consciousness without appeal to the higher-order-thought theory's claim that consciousness *is* self-consciousness. But if the theory's claim is true, then these arguments are unnecessary, and if the theory's claim is false, then the arguments are beside the point.

In the course of these chapters, one begins to wonder whether Gennaro really understands the theory he is defending. He believes lions can consciously desire to eat deers, and therefore must have thoughts about their desires. But he goes on to say:

However, one can ... be aware of a token mental state M without being aware of it *qua* type-M. An animal can be aware of its desire to eat without being aware of it *qua* desire, which is partly to say that a system can be aware of a mental state without having the concept of that *type* of state.... When the lion is chasing the deer it is aware of its desire to eat, but it does not have the concept 'desire.' It can be aware of its hunger without recognizing it, or thinking of it as an instance of hunger.... Some creatures will be able to conceptualize, and so be aware of, their mental states *qua* their differences from other mental states. One might just be aware of a token-M as *different* from M', M'', and so on (p. 82).

At this point the theory is in danger of becoming silly. The force of Rosenthal's idea is that we explain *what a hunger is* by the role it plays in an organism's theory of itself. Of course, hunger is not just a theoretical entity; but conscious hunger is conscious precisely because it is perceived and categorized in a certain way by the metapsychological apparatus. But Gennaro's version appears to be quite different. The hunger is a hunger because it has a certain qualitative essence, some kind of hungerness that is analogous to the redness of red. It is "rendered conscious" by some mysterious metacomponent that makes this essence visible. It's like holding invisible ink up to a candle.

I am not sure this is what Gennaro is saying, but then it is hard most of the time to tell just what he is getting at. Here is an example (I could have chosen many others):

Some systems are capable of a higher degree of swift and controlled manipulations of their internal states (e.g. mental representations or symbols) than others. A system, A, might be able to more quickly process its own mental representations than another system B. One reason could

be that B has limited higher-order understanding of its states.... One explanation for why A can more swiftly process its internal states would be that it has some higher-order thought-awareness of those states. A will be more likely to display clearer signs of intentional and intelligent activity than B. Human beings and many other organisms are examples of A-type systems whereas present-day robots are examples of a B-type system (pp. 151-152).

He then uses the terms "A-type" and "B-type" for several more pages. But he hasn't actually said what "types" he is referring to. An A-type system has less "limited higher-order understanding of its states." What does this mean? What are examples of the "many other organisms" that pass this test? What is "understanding" in this context and why would it speed things up? Go back and read the paragraph again, and note how in the first sentence we have the phrase "swift and controlled manipulations of ... internal states"; in the second sentence this is explained (or elaborated?) by saying that A can "more quickly process its own mental representations"; in the third the "reason" is proposed that perhaps B has "limited higher-order understanding." Syntactically, it looks like a train of thought, but it's rather hard to say where it came from and where it is going. Paragraph after paragraph reads this way.

I could say more, but there isn't much point. Suffice it to say that Rosenthal's theory still awaits an adequate defense.

References

Rosenthal, D. (1986) Two concepts of consciousness. *Philosophical Studies*, 49, 329-359.

Rosenthal, D. (1990) *A theory of consciousness*. University of Bielefeld Mind and Brain Technical Report 40. Perspectives in Theoretical Psychology and Philosophy of Mind (ZiF).

Rosenthal, D. (1993) Higher-order thoughts and the appendage theory of consciousness. *Philosophical Psychology*, 6, 155-166.