The Appearance of Unity: A Higher-Order Interpretation of the Unity of Consciousness

Josh Weisberg (jwsleep@aol.com)
CUNY Graduate Center
Department of Philosophy, 365 5th Avenue
New York, NY 10016 USA

Abstract

Recent developments in neuroscience and psychology have put pressure on the traditional philosophical idea of the unity of consciousness. Some have argued that split-brain cases and multiple personality disorder demand a rejection and elimination of the very notion of a unified consciousness. In this poster, I argue that David Rosenthal's higher-order-thought theory of consciousness allows for an explication of unity that provides for the subjective appearance of unity, but respects the actual and potential disunity of the brain processes that underwrite consciousness.

Introduction

Researchers often have several distinct phenomena in mind when they discuss the unity of consciousness. One issue, which occupies a large part of the neuroscientific focus on unity, is the "binding problem." How are features processed in disparate anatomical and functional locations in the brain brought together into one coherent experience? Closely related is the problem of accounting for the apparent "spatial" relations that hold between our sensations, relations that allow us to view one feature as next to another in the same experience. A third phenomenon associated with unity is the apparent clear and seamless nature of our conscious experience. Finally, there are issues of ownership and the self. Why is it that my conscious perceptions belong, as it were, to me? What is the nature of this "I" that the perceptions belong to?

In this poster, I propose to address the issue of unity by first introducing a theoretical model of the conscious mind, David Rosenthal's higher-order-thought hypothesis, and then seeing which aspects of unity can be explained in terms of nonconscious neuroscientific or psychological processes. Then I will attempt to show how the remaining elements of unity can be adequately dealt with by the theory. I will close by briefly considering some worries about eliminativism that often accompany discussions of unity and consciousness.[1]

The HOT Theory

To begin, I will outline Rosenthal's higher-order-thought hypothesis. The view distinguishes between a transitive and an intransitive use of the term "conscious." The transitive use occurs when we talk about being conscious of something. The intransitive use applies to mental states, which can be described as conscious or nonconscious. The theory holds that for a mental state to be intransitively conscious, we must be *transitively conscious of it* in a suitable manner. Arguably, this process is best explained by the presence of a "higher-order thought" (HOT) to the effect that we are in that mental state. HOTs are intentional states that represent the subject as being in mental states. Being the target of a HOT is what makes a state conscious.

Furthermore, the HOT must arise in the appropriate way. To account for the seeming immediacy of conscious experience, we must be unaware of any inference or observation that causes the tokening of the HOT. Mental states not targeted by HOTs are not conscious states. so the HOTs themselves are generally not conscious states, unless targeted by an additional HOT. To summarize, the theory holds that mental states are conscious when we represent ourselves as being in those states, and this representation is achieved by higher-order thought.[2] .

---

[1] It is important to note that my work on this topic is strongly influenced by David Rosenthal's writings and instruction. Rosenthal has recently completed several articles concerning Unity (Rosenthal, forthcoming a, forthcoming b), and my work herein is meant to compliment that treatment.

[2] Obviously, much more can be said concerning the theory. I hope that this brief outline will suffice for the purposes at hand, and I will fill in details as they become relevant. For a full description and defense of the HOT theory, see Rosenthal

Lower-level Unity

Neuroanatomy tells us that the brain processes that underwrite perception are widely distributed. This raises the question of how the disparate elements of perception are combined into a single coherent experience. How does the brain bring it all together? This is the famous binding problem. The layered explanatory model of the HOT theory allows for this issue to be dealt with largely at the level of nonconscious processes. Promising work has been done in laying out a *temporal* solution to the binding problem (see, e.g., Crick and Koch, 1990; Llinás and Ribary, 1993; Edelman and Tononi, 2000). Very roughly, it is hypothesized that when groups of neurons oscillate at the appropriate rate, they bind by firing in synchrony. Arguably, this solution does not involve consciousness mental states at all.

Unconscious perceptions, like those registered in priming or subliminal perception, influence behavior in virtue of their perceptible properties, and these properties need to be bound, just as in conscious perception. In order for a priming image to have an appropriate semantic effect (for example, in influencing the disambiguation of a sentence), we must perceive the image's features as unified, in order to make the proper identifications. Perceptions are bound whether they are conscious or not. The processes that bind the percepts occur independently from consciousness, so this aspect of unity is not one that a theory of consciousness has to explicate.

Another aspect of unity, our awareness of experienced objects as located in this or that portion of a sensory field, can also be dealt with at the level of nonconscious processing. In vision, for example, various functional locations in the visual cortex nonconsciously extract and organize information concerning spatial orientation. However, this feature of unity can be more fully explicated in terms of sensory qualities. In any event, the explanation can proceed independently of consciousness.[3]

Sensory qualities are properties of mental states grouped in families that exhibit a structural relationship with families of perceptual properties in the world. Our commonsense taxonomy of perceptible color properties, for example, includes similarity and difference relations. Red is more similar to orange then it is to green. The sensory states employed in color vision must maintain these relationships. They must be in families with similarities and differences homomorphic to the similarities and differences that hold among perceptible properties. The spatial properties of sensory states can be dealt with in the same way. Sensory states that underwrite visual spatial perception must possess similarities and differences homomorphic to those possessed by perceptible spatial properties.[4]

In this fashion, sensory qualities can be explained independently of consciousness. What it is for a sensation to be located next to, or above, or below, another sensation in conscious awareness can be explicated by referring to the sensory qualities already possessed by those states. States have the features that locate them in mental space independently of their being conscious. So this feature of unity can be explained without invoking a theory of consciousness.

Thus, these elements of unity, binding and location in a sensory field, can be dealt with without reference to conscious mental states. Our perceptions have these features independently of consciousness. When we do become conscious of perceptual states, our HOTs represent them in terms of these antecedently present features. The act of becoming conscious of these features does not bring them into being. Our mental states are bound and organized in sensory fields independently of consciousness. Our HOTs simply represent these states as they are: bound, unified percepts, occupying this or that portion of the relevant sensory field. In this way, we become conscious of a bound, unified experience.

Conscious Unity

However, even though nonconscious perceptual processes bind and organize our perceptual sensations, it is clear that in some respects our conscious awareness of the world outstrips what is delivered by our perceptual mechanisms. This is exemplified by Daniel Dennett's discussion of a conscious perception of wallpaper made up of many repeated images of Marilyn Monroe (Dennett, 1991, pg. 354-355). When standing in front of such an array, we seem instantly to see that the wallpaper is all Marilyns, and what's more, we see this with an apparent clarity and depth that seems to take in the whole scene with equal acuity. But appearances may be deceiving here. Outside of the small area of foveal vision, we do not actually process visual stimuli with the fine-grainedness that we are aware of in consciousness. But in conscious experience, we seem to be aware of a full, rich, clear, unified expanse. How are we to account for this?

Here, we can refer to the clarifying effect of HOT to explain the appearance of unity. According to the HOT theory, what our HOTs represent us as

---

1986, 1991, 1993, 1997, 1999b. For critical views, see Byrne, 1997; Dretske, 1995; Stubenberg, 1998.

[3] Rosenthal, 1991.

[4] See Rosenthal, 1991, 1999a, 1999b; Sellars, 1956; Shoemaker, 1975; Clark, 1993.

experiencing is what we consciously experience. Our nonconscious perceptual processes deliver a range of information about the visual scene, which grades off sharply as we move away from foveal vision. But in the absence of any alarming discontinuity detectable outside of the foveal area, our HOTs can simply represent the scene in a "more of the same" manner. HOT can represent to the effect that we are in a perceptual state with such-and-such sensory qualities, and those features repeat in a clear and unbroken pattern to the edge of the visual field.

It might be argued that if this is the case, then the homogenizing affect of the HOTs should be readily noticeable in consciousness. However, we will not ordinarily notice the smoothing over imposed by HOT because whenever we look more closely at a scene to see if it really is unbroken and repeating, we refocus our attention, and token a more detailed, though less broad, HOT. Then we become conscious of that particular detail, but are no longer conscious of the whole scene. So we won't be aware of the cleaning-up effect of the HOT when we try to attend to it. But given the *apparent* clarity and depth of the original perception of the Marilyns, and the absence of clear perception away from foveal vision, the smoothing-over effect of the HOT is the best explanation of our unified conscious experience.

Ownership and the Self

Finally, I will turn to the sorts of issues that most often arise in philosophical discussions of unity, the apparent fact that our perceptions are in a sense "owned" by us, and that this ownership points to the presence of a "self" to do the owning. Interesting evidence from neuroscience and psychology over the last several decades (split-brain cases, hemineglect, and multiple-personality disorder [MPD], for example) have put pressure on the idea that there really is a unified self in the brain which experiences all of our perceptions and thoughts. But it sure seems to us that we are unified individuals as we undergo conscious experience. How can we account for this, on the HOT model?[5]

First, we need a little more detail concerning the content of a HOT. A HOT represents to the effect that "I, myself, am in a state with such-and-such properties." By representing in this manner, employing the concept "I," the HOT ascribes the mental state to "I," the self. In doing so, the HOT serves to tag the conscious state to this self, and so the state is represented as being mine. In this way, I "own" the state.

[5] What follows is closely related to Rosenthal's recent treatment in (Rosenthal, forthcoming a, forthcoming b).

It is important to note the indexical features of the concept "I." "I" serves a function much like that of the term "here," which automatically refers to the location of the utterance, and gets its more specific content on the particular occasion its use from the location that it occurs in. "I" works in a similar way. When "I" occurs in a HOT, it refers back to the thinker that tokens the thought. It is the function of a HOT to pick out various states that the organism is in, and by employing the indexical "I" there is a thin sort of immunity to error present in the self-ascription by the HOT, reminiscent of the way that uses of "here" are immune to error by automatically referring to the location of utterance. Our conscious states are owned by us because in becoming conscious of them, we represent the states in conjunction with "I," which automatically refers back to the producer of the thought (cf. Shoemaker, 1968).

So, when we become conscious of a mental state, we token a HOT to the effect that we are in that state. This provides us with a sense of self. We are, so to speak, a creature that ascribes states as being present in that very creature, and ascribes them by employing the concept "I." In this way we become conscious of *our* states, belonging to us.

But more must be said about this "thinker" that "I" refers back to. The concept "I" is best seen as a theoretical term in the folk-psychological theory that enables us to ascribe mental states to ourselves and others. Folk psychology posits the self as the referent of "I," and it posits a variety of features of the self. The self takes in perceptions and initiates action. It has direct, infallible access to thoughts and sensations. It accounts for the continuity we have as individuals, and makes us who we are. But this notion of self has come under strong pressure from philosophy, psychology, and neuroscience. Is there anything left that can serve as the referent of the concept "I"?

I believe that several components of the folk-psychological notion of self can be preserved. First, there are the "biological" or "ecological" elements of the self. These are the features that allow us to navigate through the world, and to distinguish the boundaries of our bodies from the external environment. They are relatively low-level features of an organism's psychology. Primitive creatures like lobsters can respond differentially to their bodies and the environment, and so avoid eating their own legs. The processes that underwrite this ability can serve as the reference of "I." (See Dennett, 1996; Bermúdez, 1998.)

In more developed creatures like us, "I" also picks out those features that account for our psychological continuity. "I" refers to a collection of moods, memories, and abilities present in the individual. We possess a core group of states that define us as a subject. We have memories about who we are, various labels,

like a name, an address, a social security number, we have abilities, like the ability to play the guitar, or to drive, and we have deep seated moods and personality traits, like being lazy, mellow, or high strung.

This group of states shifts over time, and the boundaries between core states and more peripheral states is not a firm one. Furthermore, we may not be able to bring these states to consciousness with any clarity. We may in fact confabulate the content of the states as we become conscious of them, possibly altering their content as we do so. The self is like a novel that is constantly edited and revised as it is read. So the "I" present in HOT refers to those elements of an organism that allow it to negotiate through the environment and distinguish its body from the world, and it refers to the core psychological states that define us as an individual. (See Damasio, 1994, Chp. 10; Ramachandran and Blakeslee, 1998, Chp. 12.)

But what of the remaining elements of the folk-psychological notion of self? Why does it seem to us that we are free agents with direct, infallible access to the content of our own thought? Here, once again, the HOT that makes us aware of our states misleads. The concept of self that we employ in HOT is the unrevised folk-psychological one. It goes beyond what is really there in our minds. It seems to us that we are this sort of being because our folk-psychological concept posits such features, and our HOTs employ this concept in making us conscious of our thoughts and experiences. But on this score, we are in error.

So, some aspects of the folk-psychological conception of self which informs our intuitions about unity are mistaken. The self is not what it seemed to be. Does this entail that we don't really have selves, and that the unity of consciousness is a fiction, an illusion? I would argue that this worry rests on an overly-strong criteria of what marks a concept for elimination. In this case, we do maintain some aspects of our folk-psychological conception of self, namely its connection with autobiographical continuity and ability. We also can see how other aspects unity, like the presence of bound, spatially located perceptions "owned" by a subject, are maintained by our theory. I suggest that the self is still there, and that consciousness is indeed unified, but things are not exactly as they appeared prior to our investigations.

Acknowledgments

A version of this poster was presented as a paper at the ASSC 4 conference in Brussels, July, 2000. My thanks to the audience there for helpful comments. Thanks also to Roblin Meeks, Doug Meehan, and especially David Rosenthal.

References

Bermúdez, J. L. (1998). *The Paradox of Self-Consciousness*. Cambridge, MA: MIT Press.

Byrne, A. (1997). "Some Like it HOT: Consciousness and Higher-Order Thoughts." *Philosophical Studies* 24, 4 (Sept.) 1-27.

Clark, A. (1993). *Sensory Qualities*. Oxford: Clarendon Press.

Crick, F. H. C., and Koch, C. (1990). "Towards a Neurobiological Theory of Consciousness." From *Seminars in the Neurosciences*, 2, 263-275.

Damasio, A. (1994). *Descartes' Error*. New York: Avon Books.

Dennett, D. C. (1991). *Consciousness Explained*. Boston: Little Brown.

Dennett, D. C. (1996). *Kinds of Minds*. New York: Basic Books.

Dretske, F. (1995). *Naturalizing the Mind*. Cambridge: MIT Press.

Edelman, G. and Tononi, G. (2000). *A Universe of Consciousness*. New York: Basic Books.

Llinás, R. and Ribary, U. (1993). "Coherent 40-Hz Oscillation Characterizes Dream State in Humans." *Proceedings of the National Academy of Science*, 90, 2078-2081.

Ramachandran, V. S. and Blakeslee, S. (1998). *Phantoms in the Brain*. New York: Quill/William Morrow.

Rosenthal, D. M. (1986). "Two Concepts of Consciousness." *Philosophical Studies* 49, 329-359.

Rosenthal, D. M. (1991). "The Independence of Consciousness and Sensory Qualities." In Villanueva, ed. *Consciousness: Philosophical Issues, 1*. Atascadero, CA: Ridgeview Publishing Company.

Rosenthal, D. M. (1993). "Thinking that one Thinks." In M. Davies and G. W. Humphreys (Eds.) *Consciousness*. Oxford: Blackwell.

Rosenthal, D. M. (1997). "A Theory of Consciousness." in Block, N., Flanagan, O., and Guzeldere, G.,(eds.), *The Nature of Consciousness: Philosophical Debates*, Cambridge, MA: MIT Press, 729-753.

Rosenthal, D. M. (1999a). "The Colors and Shapes of Visual Experiences." In *Consciousness and Intentionality: Models and Modalities of Attribution*, D. Fisette (ed.), Dordrecht: Kluwer Academic Press, 95-118.

Rosenthal, David M. (1999b), "Sensory Quality and the Relocation Story," *Philosophical Topics*, 26, 1 and 2, (Fall and Spring), pp. 321-350.

Rosenthal, David M. (forthcoming, a), "Introspection and Self-Interpretation," *Philosophical Topics* (Winter 2001).

Rosenthal, David, M. (forthcoming, b), "Persons, Minds, and Consciousness," in *The Philosophy of Marjorie Grene*, in *The Library of Living Philsophers*, Hahn, L. E. (ed.), La Salle, Illinois: Open Court.

Sellars, W. (1956). "Empiricism and the Philosophy of Mind." *Minnesota Studies in the Philosophy of Science*, vol. 1, ed. Herbert Feigl and Michael Scriven, University of Minnesota Press, Minneapolis.

Shoemaker, S. (1968). "Self-Reference and Self-Awareness." *Journal of Philosophy*, 65/19, 555-567.

Shoemaker, S. (1975). "Functionalism and Qualia." *Philosophical Studies* 27: 291-315.

Shoemaker, S. (1991). "Qualia and Consciousness." *Mind* 100:4, 507-524.

Stubenberg, L. (1998). *Consciousness and Qualia*. John Benjamins Publishing Co., Amsterdam/Philadelphia.