

Conscious Awareness, Higher-Order Theories, and Overflow

David Rosenthal



CUNY Graduate Center
Philosophy and Cognitive
Science

<https://wfs.gc.cuny.edu/DRosenthal/www/>

**TSC 2012:
"HOT or NOT"**



I. Higher-Order Theories

- All explaining must be based on a reliable, accurate description of the phenomena we want to explain. Otherwise our explanation may well simply miss those target phenomena.
- Scientific theorizing does often lead us to revise our take on the phenomena—as in coming to recognize that weight is a *relative* property, *not intrinsic*.
- But we must still take care to ensure that our explanations address the phenomena under consideration.

➤ In the case of conscious states, that means explaining how mental states that are conscious differ from those that are not conscious.

Otherwise we won't have said why it is that the conscious ones are conscious.

➤ This is a constraint on any type of theory of what it is for a state to be conscious—but this constraint on theories motivates higher-order theories. Let's see how.

➤ It won't do simply to say that we know perfectly well what conscious states are—from the inside, so to speak.

That won't capture the contrast between being conscious and not. Experience is no substitute for articulate description.

➤ A neural marker that occurs only when states are conscious can be highly useful. But even such a marker cannot, by itself, explain why those states are conscious.

➤ We must also say why our neural marker results in states' being conscious—and why mental states aren't conscious without it.

Compare: We must say why each atomic structure has its macroscopic properties.

➤ Neural markers may point to mechanisms in virtue of which some mental states are conscious and others not.

But we must still explain why it's those mechanisms that result in a state's being conscious—how they make the difference between a state's being conscious and not.

- We often get a useful handle on what a phenomenon consists in by seeing what's missing when that phenomenon is absent.
- That works well with consciousness:
The salient mark of a mental state that fails to be conscious is that the individual who's in the state is altogether unaware of it.
If we have reason to think a subject is in a mental state but the subject sincerely denies being in it—sincerity reflecting lack of awareness—that state is not conscious.
- That implies that a state's being conscious requires that one be in some way aware of being in that state.
And that's the core of any higher-order (HO) theory of consciousness.

Conscious Awareness, Higher-Order Theories, and Overflow

TSC 2012: HOT or NOT

5

- HO theories hold that a state is conscious only if one is in some way aware of the state. They differ about what that way is.
- On my higher-order-thought (HOT) theory, one is aware of one's conscious state by having a thought that one is in that state. That HOT need not itself be conscious.
And to capture the subjective immediacy of the way we're aware of conscious states, the HOT must not seem subjectively to rely on any inference or observation.
- But despite the elegant title of our session, I'll discuss HO theories generally and the higher-order awareness (HOA) they posit, rather than focusing specifically on my own HOT theory.

Conscious Awareness, Higher-Order Theories, and Overflow

TSC 2012: HOT or NOT

6



- First-order (FO) theories of consciousness, like Ned's, see a state's being conscious as independent of any substantive HOA of the state.
- FO theories vary in the accounts they give. Fred Dretske, e.g., provides a psychological account; Ned prefers a neural mark.
- But whatever the account—whether neural or psychological—it must explain how mental states that are conscious differ from those that are not.
Mere co-occurrence of a preferred feature with the states we count as conscious is not enough—unless it also explains why it's that feature that makes the difference between a state's being conscious and not.

Conscious Awareness, Higher-Order Theories, and Overflow TSC 2012: HOT or NOT 7



- One might urge that the HOA that occurs with conscious states is not substantive: Maybe being conscious of a conscious state is like smiling a smile (Sosa 2002, Block 2008).
- But smiling a smile is simply smiling—so it can't provide a contrast like that between a state's being conscious or not. No such "deflationary" account of the HOA can explain how conscious states differ from mental states that aren't conscious. For that we need a substantive awareness of the state. (More on that in a moment.)
- HO theories do not conflict with finding neural markers of mental states' being conscious. Indeed, they may well help guide the search for an NCC.

Conscious Awareness, Higher-Order Theories, and Overflow TSC 2012: HOT or NOT 8



- That's because of the constraint that HO theories place on any theory—and the NCC. HO theories reflect the observation that no state is conscious unless one is in some way aware of it. So a satisfactory NCC should point to or provide a mechanism that subserves or gives rise to that HOA.
- Hakwan and I (2011) have argued that PFC is likely implicated in any such mechanism. But on this occasion I'll leave the neural specifics to him.
- Instead, I'll say a bit more about why it is that HO theories do constrain the NCC. Why can't the search for an NCC proceed independently of the issues that divide HO from FO theorists?

Conscious Awareness, Higher-Order Theories, and Overflow TSC 2012: HOT or NOT 9

- A state's being conscious is a psychological phenomenon. So a merely neural marker won't do unless it points to a psychological way in which conscious states differ from mental states that aren't conscious.
- An NCC that explains why we're aware of mental states when they're conscious but not otherwise satisfies that condition.
- But FO theories, which preclude appeal to being aware of conscious states, arguably lack resources to explain in psychological terms how conscious states differ from mental states that aren't conscious.
- Why is that? Why can't FO theories explain that difference in psychological terms?

Conscious Awareness, Higher-Order Theories, and Overflow TSC 2012: HOT or NOT 10



- Since FO theories deny substantive HOAs, the only remaining *psychological* properties are those such as representational content, mental qualities, and attentiveness.
- And none of those helps, since they all occur with nonconscious states as well as conscious states.
Nonconscious thoughts exhibit intentional content, nonconscious mental qualities occur in masked priming and *subliminal perception*, and *attention* also occurs with nonconscious states (e.g., Robert Kentridge, 2012).
- What about global-workspace theory and Ned's access consciousness, on which a state is conscious if its content is accessible for use by many psychological subsystems?

Conscious Awareness, Higher-Order Theories, and Overflow TSC 2012: HOT or NOT 11



- Those are FO, cast in psychological terms. But such accessibility double dissociates from a state's being conscious. Conscious peripheral perceptions can be inaccessible to other systems, and nonconscious beliefs and desires can have global effects.
- Jesse Prinz appeals to *accessibility*—though to working memory (WM). But that's just a *subpersonal* mechanism. Also, it wouldn't explain *why* those states are conscious—unless being accessible to WM is *itself due to* there being a HOA, which might well be.
- So there's no FO account of a state's being conscious cast in distinctively psychological terms. For that, HO theories are arguably the only game in town.

Conscious Awareness, Higher-Order Theories, and Overflow TSC 2012: HOT or NOT 12

II. Objections

- Many theorists see consciousness as an *inseparable aspect* of qualitative states such as perceptions.
- But it's arguable that qualitative states can and do occur without being conscious.
- Subliminal perceptions—e.g., in masked cases—are not conscious in any ordinary sense. Nonetheless, such states evidently exhibit mental qualities:
We distinguish subliminal states in respect of the very same qualitative features as with conscious qualitative states.
Subliminal states have qualitative character.

Conscious Awareness, Higher-Order Theories, and Overflow

TSC 2012: HOT or NOT

13

- Since subliminal states exhibit qualitative properties, is there *something it's like* to be in subliminal states? Do those states exhibit *phenomenal consciousness*?
- If so, those notions *extend well beyond* any ordinary notion of consciousness. People in subliminal states sincerely deny being in them. So it's at best quixotic to regard those states as conscious. Intuitively, *there is simply nothing it's like for one* to be in such states.
- Since we do characterize subliminal states qualitatively, FO qualitative states can occur without phenomenal consciousness and "what-it's-like-ness." So *those properties can't be an inseparable aspect of FO states.*

Conscious Awareness, Higher-Order Theories, and Overflow

TSC 2012: HOT or NOT

14

- Ned denies this—holding that phenomenal consciousness and “what-it’s-like-ness” are aspects of FO qualitative states. And subjectively it does seem that way.
- But our explanations must look past the subjective appearances. The HOA is rarely itself a conscious state: We are seldom aware of being aware of a conscious state—perhaps only in the very special case of introspecting.
- Since we are typically unaware of any HO state, it will seem subjectively that only the FO state occurs. So the property of that state’s being conscious will seem—again subjectively—to be an aspect of the FO state.

Conscious Awareness, Higher-Order Theories, and Overflow

TSC 2012: HOT or NOT

15

- But only the FO state has mental qualities; so why is the HO state needed for one to have a conscious qualitative experience?
- A FO qualitative state won’t be conscious unless one is aware of oneself as being in that state. The nonconscious qualitative state still influences one’s psychological processing without any relevant HOA—but there isn’t anything it’s like for one.
- There’s no change in subliminal qualitative states themselves when they come to be conscious. Things light up for us because we become aware of ourselves as being in those states. Since we’re typically unaware of any HO states, we mistakenly think of the FO states themselves as “lighting up.”

Conscious Awareness, Higher-Order Theories, and Overflow

TSC 2012: HOT or NOT

16

- A FO state is conscious if a HO state makes one aware of it. But HO states are rarely themselves conscious.
So where does the consciousness come from? How is it that two states that aren't conscious result in a state that is conscious?
- That's the wrong question. A FO state's being conscious consists in one's being suitably aware of it. So the right question is whether a nonconscious HO state can make one aware in that way of a FO state.
- And it can. Just sensing subliminally makes one aware of the thing sensed, though not consciously aware of it; so a nonconscious HO state makes one aware of the FO state, though not consciously aware of it.

Conscious Awareness, Higher-Order Theories, and Overflow

TSC 2012: HOT or NOT

17

- HO theories can explain our subjective sense that the property of being conscious is inseparable from FO qualitative states.
- And we also have reason not to trust that subjective sense: Our best psychological handle on how conscious states differ from nonconscious mental states is that a HOA accompanies the conscious cases.
- Any acceptable theory must of course do justice to the subjective appearances.
- But it need not do so by taking those appearances to be veridical. It can do so by explaining why it is that we have those subjective appearances. (Cf. weight.) We must explain that anyway—and accuracy of the appearances couldn't help us do so.

Conscious Awareness, Higher-Order Theories, and Overflow

TSC 2012: HOT or NOT

18

- This bears on an objection Ned and others have raised: that HO theories don't rule out a HOA's occurring without the FO state that the HOA makes one aware of being in.
- This can't really be much of a worry. HO theories don't require that this happens. So if it turns out that it can't, one can just add a stipulation to that effect. This wouldn't be ad hoc, just an adjustment to accommodate the way things are.
- But it's unclear why a HOA cannot occur without the FO state one is aware of oneself as being in. It seems subjectively that it can't. But that shows that it really can't occur only on the highly implausible view that the mind is transparent to itself.

Conscious Awareness, Higher-Order Theories, and Overflow

TSC 2012: HOT or NOT

19

- And it will appear subjectively that HOAs cannot occur without their targets—since each HOA actually constitutes a subjective appearance that one is in a FO state.
- If what-it's-like-ness were an aspect of the FO state, as Ned (2011a) holds, there would be nothing it's like without a FO state. But if one is altogether unaware of being in that state, there will be nothing it's like for one to be in it. So having a HOA is at least necessary for there being something it's like—and, barring transparency, sufficient.
- And as noted earlier, we can explain our subjective sense that what-it's-like-ness is an inseparable aspect of FO states simply by our being unaware of (most) HOAs.

Conscious Awareness, Higher-Order Theories, and Overflow

TSC 2012: HOT or NOT

20

- But if a HOA *does* occur without a relevant FO state, what is the conscious state? It's not the FO state, since that's missing. And the HOA is seldom itself conscious. Doesn't this show the FO state must occur?
- Consciousness is mental appearance: It's how our mental lives subjectively appear to us. So conscious states are simply the mental states we appear subjectively to be in—even if occasionally it turns out that we're not actually in those states.
- Ned has argued (2011b) that perceptual consciousness overflows cognitive access—appealing to George Sperling's and Victor's work to argue that more is phenomenally conscious than we actually access.



Conscious Awareness, Higher-Order Theories, and Overflow

TSC 2012: HOT or NOT

21

- Ned will discuss these results in detail. I'll quickly describe Sperling's (1960) work and make a few comments. Sperling presented subjects very briefly with a 3 x 4 matrix of letters. After the letters have disappeared, subjects can identify only 3-4 letters of all 12. But if a subsequent tone directs subjects to just one row, they get most in that row.

K	S	M	R	← High tone
X	D	Q	G	← Medium tone
B	Z	O	H	← Low tone

- Since the tone occurs only after the letters vanish, subjects must somehow retain the relevant information. The question is how.
- Ned argues that they retain it consciously, appealing to Victor's work. My comments on Sperling apply also to Victor's work.

Conscious Awareness, Higher-Order Theories, and Overflow

TSC 2012: HOT or NOT

22

- Sperling subjects report having conscious perceptions of all 12 letters. But some theorists have urged that these conscious perceptions are generic—some letter or other, though not a specific one—or are fragmentary. The perceptions would then not overflow our limited cognitive access.
- Ned doubts that conscious perceptions are ever generic or fragmentary. But HO theories suggest they may well be.
- Even if the FO perceptions couldn't be generic or fragmentary, the HOA might still make one aware of the FO perception in a fragmentary or generic way. The perceptions would be conscious as fragmentary or generic.

Conscious Awareness, Higher-Order Theories, and Overflow

TSC 2012: HOT or NOT

23

- Subjects also report that they consciously see “more than they can cognitively grasp” (Block 2011b)—and that “[t]hey saw more than they remembered” (Sperling 1983). Ned argues that this helps confirm that the perceptions that overflow cognitive access are conscious.
- But these reports may just reflect subjects' assumption that the display contained more than they could identify—and so more than they were consciously aware of. They may just be assuming that the display had specific letters they couldn't identify.
- Conscious awareness would then coincide with cognitive access—i.e., with what subjects can describe and identify.

Conscious Awareness, Higher-Order Theories, and Overflow

TSC 2012: HOT or NOT

24

- Those who deny overflow sometimes urge that conscious perception is *sparse*—less rich than it seems. Ned insists that it's rich.
- Perceptions could be rich or sparse in two distinct ways. Perhaps, as I've suggested, the FO perceptions are in themselves rich but we're aware of them in a sparse way: *Our HOAs represent them sparsely.* But it might be that the FO perceptions are also less rich than we think—*independent* of how we're subjectively aware of them.
- We likely do subjectively overestimate the detail we *objectively* see—i.e., *in a FO way* (e.g., Rahnev et al 2011). If so, we have even less reason to hold that *conscious* perceiving overflows cognitive access.



Conscious Awareness, Higher-Order Theories, and Overflow TSC 2012: HOT or NOT 25

**Thank you for your
(conscious) attention!**

Conscious Awareness, Higher-Order Theories, and Overflow TSC 2012: HOT or NOT 26

References

- Block, Ned (2011a), "The Higher-Order Approach to Consciousness is Defunct," Analysis, 71, 3 (July): 419-431.
- Block, Ned (2011b), "Perceptual Consciousness Overflows Cognitive Access," Trends in Cognitive Sciences, 15, 12 (December): 567-575.
- Block, Ned (2008), "Consciousness, Accessibility, and the Mesh between Psychology and Neuroscience," Behavioral and Brain Sciences, 30, 5-6 (August): 481-499.
- Kentridge, Robert (2012), "Attention without Awareness," presentation at the Conference, Toward a Science of Consciousness, Tucson, April 2012.
- Lau, Hakwan, & David Rosenthal (2011), "Empirical Support for Higher-Order Theories of Conscious Awareness," Trends in Cognitive Sciences, 15, 8 (August): 365-373.
- Nagel, Thomas (1974), "What Is It Like to Be a Bat?", Philosophical Review, 83, 4 (October): 435-550.
- Rahnev, Dobromir, Brian Maniscalco, Tashina Graves, Elliott Huang, Floris P. de Lange, & Hakwan Lau (2011), "Attention Induces Conservative Subjective Biases in Visual Perception," Nature Neuroscience, 14, 12 (December): 1513-1515.
- Sosa, Ernest (2002), "Privileged Access," in Consciousness: New Philosophical Perspectives, ed. Quentin Smith & Aleksandar Jokic, Oxford: Clarendon Press.
- Sperling, George (1960), "The Information Available in Brief Visual Presentations," Psychological Monographs, 74, 498: 1-29.
- Sperling, George (1983), "Why We Need Iconic Memory," Behavioral and Brain Sciences, 6, 1 (March): 37-39.