

- Rosenthal, D. 2000. Metacognition and higher-order thoughts. *Consciousness and Cognition* 92: 231–42.
- Rosenthal, D. 2002. Explaining consciousness. In *Philosophy of Mind: Classical and Contemporary Readings*, ed. D. Chalmers, 406–21. Oxford: Oxford University Press.
- Rosenthal, D. 2004. Varieties of higher order theory. In *Higher-Order Theories of Consciousness*, ed. R. Gennaro, 17–44. Amsterdam: John Benjamins.
- Rosenthal, D. 2005a. *Consciousness and Mind*. Oxford: Oxford University Press.
- Rosenthal, D. 2005b. Sensory qualities, consciousness and perception. In *Consciousness and Mind*, 175–226. Oxford: Oxford University Press.
- Rosenthal, D. 2009a. Concepts and definitions of consciousness. In *Encyclopedia of Consciousness*, ed. W. Banks, 157–69. London: Academic Press.
- Rosenthal, D. 2009b. Higher order theories of consciousness. In *Oxford Handbook of the Philosophy of Mind*, eds B. McLaughlin and A. Beckermann, 239–52. Oxford: Clarendon Press.
- Ross, W.D. 1961. *Aristotle, De anima*. Oxford: Clarendon Press.
- Smith, D.W. 1986. The structure of self-consciousness. *Topoi* 5: 149–156.
- Van Gulick, R. 2000. Inward and upward: reflection, introspection, and self-awareness. *Philosophical Topics* 28: 275–305.
- Van Gulick, R. 2004. Higher-order global states HOGS.: an alternative higher-order model of consciousness. In *Higher-order Theories of Consciousness: An Anthology*, ed. R. Gennaro, 67–92. Amsterdam: John Benjamins.
- Weisberg, J. 2010. Misrepresenting consciousness. In *Philosophical Studies*. Advance Access published 12 May 2010, doi:10.1007/s11098-010-9567-3.

## *Exaggerated reports: reply to Block*

DAVID ROSENTHAL

The report of my death is an exaggeration.  
(Mark Twain, *New York Journal*, 2 June 1897)

A state's being conscious is a matter of mental appearance – of how one's mental life appears to one. If somebody is in a mental state but doesn't seem subjectively to be in that state, the state is not conscious. This straightforward test for the consciousness of mental states is deeply entrenched in common sense and experimental work in psychology. Higher-order theories exploit this, arguing that a state is conscious only if one is subjectively aware of oneself as being in that state.

The higher-order-thought (HOT) theory explains that subjective awareness as due to one's having a thought that one is in the state. And to explain why

our awareness of our conscious states seems subjectively unmediated, that HOT must seem subjectively to be independent of inference and observation.

Ned Block<sup>1</sup> sees the theory as harbouring an incoherence. The alleged incoherence builds on the much discussed<sup>2</sup> possibility of one's having a HOT without actually being in the state the HOT represents one as being in. Concern about this possibility may seem especially pressing when the state the HOT represents one as being in is qualitative. I've argued (e.g. 2005: 217–18) that in such a case one will still be subjectively aware of oneself as being in whatever state the HOT describes one as being in. Such cases are presumably rare; but we can test for them, since the causal role of a HOT will differ to some extent from that of the first-order state it's about.

Still, there's a temptation to think such cases simply cannot occur. One reason is that when one is aware of oneself as being in some mental state, there is a compelling subjective sense that the state does occur. In addition, describing a state as conscious seems to commit one to that state's actual occurrence, and not just to its seeming to occur. Subjective appearance and surface grammar both suggest that being in a conscious state requires the occurrence of the state itself.

But neither reason holds up. A mental state is conscious if one is subjectively aware of being in it. So we take somebody to be in a conscious state if that individual seems, in a first-person way, to be in that state. We do generally assume that such first-person appearances are accurate. But we sometimes have no independent support for that assumption, either for oneself or for others. We typically rely solely on subjective appearances in determining that somebody is in a conscious state; we don't in addition require independent evidence that the person actually is in the state. So it might happen that one is in a conscious state but not in the state one seems subjectively to be in.

What counts for somebody's being in a conscious state is just the occurrence in one's stream of consciousness of the relevant subjective appearance, the appearance of being in the state in question. A theory of consciousness explains conscious subjectivity, not the underlying mental reality that's responsible for these appearances. Being in a conscious state is not a mental state's having some special monadic property; rather, it's an aspect of how one's mental life appears to one. So one's being in a conscious state does not imply being in the state one is aware of being in; the phrase 'is in a conscious state' is in that way non-extensional. And subjective appearances may mislead about what mental states one is actually in.

It's sometimes said that distinguishing appearance from reality is untenable when it comes to consciousness, since subjectivity is the last word about the mental reality it seems to present. But mental states that occur consciously can also occur without being conscious; even qualitative states such as

1 Block 2011; page references are to this article when not otherwise indicated.

2 For example, Neander 1998: 420, Levine 2001: 108 and Wilberg 2010.

sensations sometimes fail to be conscious, as in subliminal perceiving.<sup>3</sup> Since subjectivity doesn't reveal mental states that aren't conscious, it cannot be the last word about mental reality, but at best only about how that reality sometimes appears to us. We must distinguish one's seeming to be in a mental state from one's actually being in that state.

And since mental reality is to some extent independent of mental appearance, we have no reason to insist that it cannot occasionally seem subjectively that one is in some state that one is not actually in. A mundane and harmless example may often occur in consciously seeing something red. Even when the sensation that seems to occur in one's stream of consciousness is a sensation simply of generic red, the sensation that actually occurs may often be of a highly specific shade. One is aware of a sensation of generic red, though there is no sensation that strictly answers to that awareness. There is nothing startling or incoherent about this.<sup>4</sup>

Block sees the HOT theory as holding that a HOT at a particular time 'is sufficient for the occurrence of a conscious episode at' that time (Block 2011: 424). That would be so only if we construed conscious episodes as no more than the subjective appearance of those episodes. The theory does not hold that a HOT is sufficient for the actual occurrence of the state one is aware of oneself as being in. The theory says that a HOT suffices only for a mental episode to seem subjectively to occur, not also for that episode actually to occur.

But Block insists that the theory must after all be committed to the occurrence of the first-order state. To show that, he appeals to a particular construal of the phrase, 'something it's like to be in a conscious state', which he sees as 'common terminology' (2011: 420) between him and the HOT theorist.

As many, myself included, use that phrase, there being something it's like for one to be in a state is simply its seeming subjectively that one is in that state. Indeed, Block (2011: 424) quotes me to that effect: 'What it is like for

3 See, e.g. experimental work in masked priming (e.g. Breitmeyer and Ögmen 2006, Marcel 1983) and blindsight (Weiskrantz 1997).

4 More dramatic evidence of this dissociation of sensations from one's awareness of them occurs in John Grimes's (1996) change-blindness experiments. Grimes used eye trackers to switch displays during subjects' saccades, when no retinal signal reaches visual cortex. He found that subjects were not consciously aware of changes in salient visual features, e.g. a dramatic change between red and green. Since retinal input to visual cortex resumes after saccades, we can assume that the sensations in visual cortex changed to reflect display colour, despite subjects' unchanged awareness of those sensations. Qualitative states in this case diverge from the way individuals are subjectively aware of them.

Block has doubts, since other change-blindness researchers don't see their results that way. But since the researchers Block has in mind don't use Grimes's eye-tracker method, their results are different. Nor is it a mystery, as Block suggests, that subjects 'settle on one interpretation rather than another to maintain throughout those colour changes' (note 5). Subjects become aware of the initial sensation, but then fail through a number of changes to update their awareness of the sensation, as shown by their failing for some time to report any awareness of change.

one to have a pain, in the relevant sense of that idiom, is simply what it is like for one to be conscious of having that pain' (Rosenthal 1997: 733). And on that construal of 'what it's like', the theory does hold that a HOT is sufficient for there to be something it's like for one to be in the state the HOT describes, even if that state doesn't occur.

Block, however, has a substantially different understanding of that phrase. He equates '*there [being] something it is like for the subject*' with there being '*an episode or occurrence of what-it-is-like-ness*' (2011: 425; Block's emphasis), which he sees as a monadic property that attaches to a qualitative state. So on Block's construal, there cannot be anything it's like for one if the state one is aware of oneself as being in does not occur. Since the HOT theory asserts, with Block, that a HOT is sufficient for there to be something it's like for one, he concludes that the theory is committed to a HOT's being sufficient for the state it describes to exist.

But that commitment does not follow from the theory's assertion that a HOT is sufficient for there to be something it's like for one, since that assertion rests on a construal of 'what it's like' solely in terms of subjective appearances. The phrase 'what it's like' is not reliable common currency.<sup>5</sup>

As Block notes, what-it-is-like-ness construed as a monadic property of mental states is simply his well-known notion of phenomenal consciousness; indeed, he often proposes as the test of phenomenal consciousness there being something it's like to be in the state (e.g. 2001: 202). But Block's phenomenal consciousness is itself arguably just the occurrence of mental qualities, independently of whether those mental qualities occur consciously.

This is evident when Block describes the neural correlate of phenomenal consciousness as the 'neural basis of the *content* of an experience, that which differs between the experience as of red and the experience as of green' (2005: 46; emphasis Block's). Since those experiences differ solely in respect of their mental qualities, that's all their phenomenal consciousness consists in. But states with mental qualities sometimes occur subliminally, that is, when one subjectively takes oneself not to be in any such state. And it's quixotic to regard as conscious a state that one subjectively takes oneself not to be in. Since phenomenal consciousness consists in qualitative character independently of there being any way the state seems to one subjectively, the same holds for Block's what-it-is-like-ness. Neither is a case of a state's being conscious in any folk, intuitive sense.<sup>6</sup>

5 Indeed, William Lycan, whom Block (2011: 422) approvingly quotes as saying that 'the philosophical use of "conscious" is by now well and truly up for grabs' (Lycan 1996: 26), castigates the phrase 'what it's like' far more strongly, as 'now worse than useless: it is positively pernicious and harmful' (1996: 77; cf. 176, n. 3).

6 Block has sought to accommodate the observation that a state is conscious only if one is subjectively aware of it by adopting what he calls a same-order view, on which every phenomenally conscious state represents itself (2007: 485). But he has not offered a positive account of such same-order self-representation on which it differs from a higher-order

Block's distinction between modest and ambitious higher-order theories hinges on whether a theory seeks to explain phenomenal consciousness. I resist that taxonomy of higher-order theories. Since states with mental quality occur both consciously and not, mental qualities can occur without appearing in one's stream of consciousness. So one's being in a state with qualitative character is independent of one's being in a conscious state, and we need different theories to explain the two. A theory of consciousness will explain one's mental life subjectively appearing a particular way; a distinct theory must address what mental qualities are, independently of whether they occur consciously.<sup>7</sup>

This is so even when we explain conscious phenomenology, since conscious phenomenology consists in mental qualities, which need not occur consciously, together with those qualities' seeming to occur in one's stream of consciousness. Block's demand that a single theory explain both at once stems from his seeing mental qualities as by themselves constituting a kind of consciousness. That view echoes the traditional picture on which mental qualities cannot occur without being conscious, but as already noted subliminal perceiving and blindsight undermine that picture. So to reject both that picture and Block's demand for a single theory is simply to repudiate an outmoded theory, not the phenomenon itself.

Block seeks to reinforce his contention that the HOT theory is committed to HOTs' being sufficient for the states they describe actually to occur by appeal to the way many mental states matter to us. To be 'worth anything', he writes, a theory of consciousness 'must be about consciousness in a sense that matters to us in the way that conscious agony or ecstasy matters'. And mattering, he maintains, is tied to what-it-is-like-ness, which on his construal attaches to an actual first-order state. 'If a state of being conscious of agony is supposed to matter equally whether [the agony] exists or not, the supposed theory of consciousness is worthless' (2011: 427).

Pains do sometimes occur without being conscious, that is, without one's being in any way aware of them.<sup>8</sup> Such pains, despite not being conscious, may still interfere with mental and bodily functioning, and so adversely affect one's well-being. That's one way pains may matter to us, independently of their being conscious. Awareness of pains is another. A theory of what pain is would address the first; a theory of what it is for pains to occur consciously

---

content that refers to the target state. Without such an account, it's unclear that Block's same-order view constitutes a genuine alternative to higher-order theories.

7 The quality-space theory of mental qualities I've developed (e.g. 2005: chs 5–7, and 2010), which is independent of HOT theory but fits comfortably with it, is intended to do that job. That quality-space theory involves no higher-order states, since it explains mental qualities independently of whether they're conscious.

8 People sometimes favour a leg while walking without being in any way aware of the associated mild pain until somebody asks what's wrong.

tells us about the second. To draw this distinction is not to ‘abus[e] the notion of what-it-is-like-ness’ (2011: 427).

A phenomenon known as dental fear dramatically illustrates how the two kinds of mattering can come apart. Dental drilling sometimes leads patients to report feeling pain even though no relevant nerve exists. The standard explanation is that when a sensation of vibration occurs along with fear, patients mistake it for pain. Once this is explained and drilling resumes, patients are aware only of fear and vibration, though strikingly they continue to recall the earlier experience as pain; memory retains the way patients were aware of the earlier experience. And the subjective awareness of pain matters to them despite the absence of actual pain.

Block describes me as having retreated from an ‘aboriginal’ theory, on which the targets of HOTs always exist, to a ‘new version’ on which they need not (2011: 427). This is not so; in my earliest publication about consciousness I noted the possibility of absent first-order states (reprinted as [2005: 29]). For ease of exposition, I often introduce the theory by saying that a state is conscious when it’s accompanied by a HOT, noting that this characterization is not strictly accurate. And there’s no harm in putting things in those relational terms when the existence of HOTs’ targets is not under consideration.

All that matters for a state’s being conscious is its seeming subjectively to one that one is in that state. On the HOT theory, that’s determined by a HOT’s intentional content; so as Block notes, the causal origin of a HOT is irrelevant. He concludes that it’s ad hoc for the theory to hold that HOTs result in subjectivity only if they’re independent of inference. But that’s not what the theory provides; it says only that HOTs must subjectively seem to be independent of inference, ‘a fine point’ that Block says he ‘will ignore’ (2011: 421). And as noted earlier, this subjective independence of inference is intended to capture the sense of immediacy characteristic of the way we’re aware of our conscious states. That’s not ad hoc, and there’s no provision about causal origin.

Block also sees as ad hoc a stipulation that the first-order state occurs in the same mind or at the same time as the HOT. It’s reasonable to assume that these conditions are typically satisfied, but they aren’t part of the theory. All that matters for subjective appearances are the HOT’s content. That content ordinarily reflects the nature of the first-order state, but subjective appearances can occur without any first-order state at all.<sup>9</sup>

9 Block echoes Kati Balog’s (2000) suggestion that on the HOT theory we could forgo pleasurable experiences in favour of just thinking that we have them. But it’s rare that we can simply decide what to think – i.e. decide what assertoric thoughts to have – as against merely deciding what to imagine or the like. And the provision that HOTs are assertoric is not introduced to avoid that or other objections, but for the independent reason that other mental attitudes, such as doubting and wondering, do not by themselves result in one’s being aware of things. So that provision is also not ad hoc.

The HOT theory explains conscious subjectivity, and does so without ad hoc stipulations. The theory is not defunct; it is alive and well.

*Graduate Center  
City University of New York  
New York, NY, USA  
davidrosenthal@nyu.edu*

### *References*

- Balog, K. 2000. Comments on David Rosenthal's 'Consciousness, content, and metacognitive judgments'. *Consciousness and Cognition* 9: 215–19.
- Block, N. 2001. Paradox and cross purposes in recent work on consciousness. *Cognition* 79: 197–219.
- Block, N. 2005. Two neural correlates of consciousness. *Trends in Cognitive Sciences* 9: 46–52.
- Block, N. 2007. Consciousness, accessibility, and the mesh between psychology and neuroscience. *Behavioral and Brain Sciences* 30: 481–99.
- Block, N. 2011. The higher order approach to consciousness is defunct. *Analysis* 71: 419–31.
- Breitmeyer, B.G. and H. Ögmen. 2006. *Visual Masking: Time Slices through Conscious and Unconscious Vision*, 2nd edn. New York: Oxford University Press.
- Grimes, J. 1996. On the failure to detect changes in scenes across saccades. In *Perception*, ed. K. Akins, 89–110. New York: Oxford University Press.
- Levine, J. 2001. *Purple Haze: The Puzzle of Consciousness*. New York: Oxford University Press.
- Lycan, W. 1996. *Consciousness and Experience*. Cambridge, Massachusetts: MIT Press/Bradford Books.
- Marcel, A.J. 1983. Conscious and unconscious perception: An approach to the relations between phenomenal experience and perceptual processes. *Cognitive Psychology* 15: 238–300.
- Neander, K. 1998. The division of phenomenal labor: A problem for representational theories of consciousness. *Philosophical Perspectives* 12: 411–34.
- Rosenthal, D.M. 1997. A theory of consciousness. In *The Nature of Consciousness: Philosophical Debates*, eds N. Block, O. Flanagan, and G. Güzeldere, 729–53. Cambridge, Massachusetts: MIT Press.
- Rosenthal, D.M. 2005. *Consciousness and Mind*. Oxford: Clarendon Press.
- Rosenthal, D.M. 2010. How to think about mental qualities. *Philosophical Issues* 20: 368–93.
- Weiskrantz, L. 1997. *Consciousness Lost and Found: A Neuropsychological Exploration*. Oxford: Oxford University Press.
- Wilberg, J. 2010. Consciousness and false HOTs. *Philosophical Psychology* 23: 617–38.