

# Consciousness and Metacognition

**David Rosenthal**  
CUNY Graduate Center  
Philosophy and Cognitive  
Science  
<http://tinyurl.com/dr-cuny/>

ASSC 20, Buenos Aires  
15 June 2016



## I. Consciousness and Misrepresentation

- Metacognition is the cognitive registration of one's cognitive processes and states. Classic examples include the sense we sometimes have that we know or that we will recall something, and the sense that one has a word on the tip of one's tongue—though other types of higher-order cognition are also called metacognition.
- A family of theories holds that a state's being conscious consists in one's being aware of that state. Should we also count such awareness as metacognition?

THE GRADUATE CENTER  
OF THE CITY UNIVERSITY OF NEW YORK

Consciousness and Metacognition ASSC 20, Buenos Aires, 15 June 2016 2

- I'll say a very few words in support of such higher-order (HO) theories, and then ask whether they're best seen as metacognitive.
- HO theorists rely on the commonsense observation that a state of which one is wholly unaware is not conscious—an idea that underlies much experimental work. So a state is conscious only if one is in some way aware of it—only if one has a HOA. But the negative version is more compelling.
- Some urge reliance instead on a *neural* mark of a state's being conscious. But we can establish a neural mark only by correlations with states we know on independent grounds to be conscious—hence, on *psychological* grounds.

THE GRADUATE CENTER  
OF THE CITY UNIVERSITY OF NEW YORK

Consciousness and Metacognition ASSC 20, Buenos Aires, 15 June 2016 3

- And a HOA is a natural psychological mark of a state's being conscious. Reliance on verbal or other reports supports this—since reporting a state *expresses* a HOA of it. A report is a reliable symptom of a HOA.
- First-order (FO) theorists deny that any HOA figures in consciousness. Some urge that we just know *from the inside* what states are conscious. Yes—by being aware of the states! Otherwise that's just uninformative.
- Block (2007) has likened our being conscious of our conscious states to smiling a smile. But this deflationary move can't work. There is no smile without a smiling; so the analogy disallows any contrast between a state's being conscious and its not being so.

THE GRADUATE CENTER  
OF THE CITY UNIVERSITY OF NEW YORK

Consciousness and Metacognition ASSC 20, Buenos Aires, 15 June 2016 4

- One type of FO theory appeals to a global workspace (GW): States are conscious if their content is available for processing by—or actually broadcast to—many cortical subsystems (Dehaene & Naccache 2001; Dehaene & Changeux 2011; Baars 1988; cf. Block 1995 on access consciousness and Dennett 1993 on cerebral celebrity).
- Global effects do often coincide with states' being conscious, and it may be inviting to expect that they always do. We're aware only of mental states that are conscious; so it's only the conscious states we're aware of as having any effects at all.
- But global broadcast does not invariably coincide with consciousness. So global broadcast cannot be the neural correlate of consciousness (NCC).

THE GRADUATE CENTER  
OF THE CITY UNIVERSITY OF NEW YORK

Consciousness and Metacognition ASSC 20, Buenos Aires, 15 June 2016 5

- Beliefs and desires often have global effects without being conscious. This may even be so with very salient subliminal perceptions. Conversely, peripheral conscious perceptions typically have no global effects at all.
- So consciousness and broadcast may well coincide so often only because they result from a *common cause*, most likely high signal strength. Their having a common cause would also explain why they sometimes dissociate.
- This raises a question about what utility a state's being conscious has—*in addition* to its occurring *without* being conscious. I've argued elsewhere that being conscious adds little if any utility (2008, 2012).

THE GRADUATE CENTER  
OF THE CITY UNIVERSITY OF NEW YORK

Consciousness and Metacognition ASSC 20, Buenos Aires, 15 June 2016 6

- In brief: All conscious states have being conscious in common. So their *differential* causal efficacy is due to their content or qualitative character, and differential causal efficacy determines utility. Signal strength does add utility, but likely causes utility and consciousness independently. A state's being conscious adds no utility.
- This is wholly independent of a HO theory—though it does fit well with such a theory. But I won't pursue this any further here.
- Some FO theorists object that a HOA could misrepresent the state it's about, but—so it's claimed—consciousness itself simply cannot misrepresent in that way (Neander 1998, Levine 2001, Block 2011).

THE GRADUATE CENTER  
OF THE CITY UNIVERSITY OF NEW YORK

Consciousness and Metacognition ASSC 20, Buenos Aires, 15 June 2016 7

- This is an odd objection. HO theories don't say consciousness *does* misrepresent; they simply allow that it *may*. So if these critics are right that it can't, one could just tack that onto the HO theory of one's choice.
- The complaint is really just that HO theories don't *on their own*—as (most) FO theories do—require ruling out misrepresentation. The objection simply channels a FO view.
- Still, misrepresentation is of independent interest because of the question whether consciousness is metacognitive.

Metacognition does sometimes misrepresent: One may wrongly judge that one will recall something or that one has information one can't now retrieve (e.g., Siedlecka et al, 2016).

THE GRADUATE CENTER  
OF THE CITY UNIVERSITY OF NEW YORK

Consciousness and Metacognition ASSC 20, Buenos Aires, 15 June 2016 8

- So if consciousness also can misrepresent, it is in that way like metacognition.
- I'll return in a moment to whether HO theories are best seen as metacognitive. But let's first ask whether consciousness can misrepresent one's mental states.
- Consciousness is the way our mental lives appear to us. But it's often claimed that for consciousness, appearance and reality coincide (e.g., Nagel 1974). The mind simply is as it appears to be to consciousness. If so, consciousness can't misrepresent the mental reality that's conscious for one. No reason is ever given for this claim—it's supposed just to be "intuitively" obvious. But there are compelling counterexamples.

THE GRADUATE CENTER  
OF THE CITY UNIVERSITY OF NEW YORK

Consciousness and Metacognition ASSC 20, Buenos Aires, 15 June 2016 9

- One way consciousness does misrepresent is by being incomplete. We're often aware of an object as being red—even though consciousness doesn't capture the exact shade. But we nonetheless often register the exact shade perceptually, and can become aware of it by attending.
- Such incomplete awareness is a bit like the metacognitive tip-of-the-tongue feeling, in which it seems one knows something one can't access. One is aware of knowledge just as the answer to a question—not in respect of the content of that knowledge.
- But does consciousness ever misrepresent altogether—  
as against by simply being incomplete?

THE GRADUATE CENTER  
OF THE CITY UNIVERSITY OF NEW YORK

Consciousness and Metacognition ASSC 20, Buenos Aires, 15 June 2016 10

- John Grimes (1996), in George McConkie's lab, used eye trackers to switch displays during saccades. 18% of subjects were unaware of a salient change from green to red:



- So when the stimulus became red, they were actually aware of a perception of red as a perception of green—a very dramatic misrepresentation by consciousness indeed!
- Less exotic: Consciousness misrepresents in Nisbett & Wilson's (1977) finding that subjects denied preference among identical fabrics due to location, and confabulated perceptual reasons for those preferences.

THE GRADUATE CENTER  
OF THE CITY UNIVERSITY OF NEW YORK

Consciousness and Metacognition ASSC 20, Buenos Aires, 15 June 2016 11

- Why then do many hold that consciousness cannot misrepresent? Perhaps because it seems we can't test whether subjective appearances do accurately reflect mental reality—as we can test other appearances.
- But in the foregoing examples we can. We don't have first-person evidence about whether consciousness is accurate—but we sometimes do have third-person evidence.
- The conviction that consciousness cannot misrepresent is likely due to the idea that a state's being conscious is inseparable from that state's other mental properties. Can we explain this conviction—and the resulting insistence that for consciousness appearance and reality must coincide?

THE GRADUATE CENTER  
OF THE CITY UNIVERSITY OF NEW YORK

Consciousness and Metacognition ASSC 20, Buenos Aires, 15 June 2016 12

➤ A HO theory helps. HOAs are seldom themselves conscious—we're rarely aware of them. So it seems subjectively that consciousness is of a piece with a state's other mental properties. And that leads to a sense that misrepresentation can't occur.

➤ One might contend that this explanation doesn't do justice to the subjective sense that misrepresentation can't occur. No explanation can be right, one might insist, unless it *denies* misrepresentation.

➤ But that's a mistake. We must save the subjective appearances of consciousness. But we need not do so by holding that the appearances are accurate. We can instead simply *explain why we have them*.

THE GRADUATE CENTER  
OF THE CITY UNIVERSITY OF NEW YORK

Consciousness and Metacognition ASSC 20, Buenos Aires, 15 June 2016 13

## II. Metacognition and Confidence

➤ Consciousness, like metacognition, can misrepresent. And HO theories not only accommodate that, but can also explain why many insist that consciousness can't. Consciousness and metacognition also resemble in another way: Neither seems subjectively to rely at all on observation, inference, or any other mediating factor.

➤ Still, there is reason not to assimilate a HO theory of consciousness to standard types of metacognition. I'll finish by noting a few differences—and some consequences.

THE GRADUATE CENTER  
OF THE CITY UNIVERSITY OF NEW YORK

Consciousness and Metacognition ASSC 20, Buenos Aires, 15 June 2016 14

- As I noted, HOAs are seldom conscious states. Arguably they are only when we introspect—which is extremely rare.
- By contrast, metacognitive judgments are often, even typically, conscious—though by no means always (Kentridge & Heywood 2000). When one has a metacognitive feeling of knowing something (FOK) or one makes a judgment of learning (JOL) (Nelson & Narens 1994), one's feeling or judgment is typically conscious.
- Failure to note this may explain why some (e.g., Seth 2008) simply dismiss HO theories as implausibly metacognitive. Such theories would indeed be implausible if they did hold that HOAs are conscious.

THE GRADUATE CENTER  
OF THE CITY UNIVERSITY OF NEW YORK

Consciousness and Metacognition ASSC 20, Buenos Aires, 15 June 2016 15

- Metacognitive judgments, moreover, don't result in the cognitive states they are about themselves being conscious. One can, e.g., judge that one has learned something or feel one knows it without any relevant FO states' being conscious.
- A HOA results in a state's being conscious only if the HOA represents one as *being in that state*—in respect of the state's mental properties. Metacognition needn't do that. It's often just about what one *can* or *will* know or recall, or about mental processing.
- But one type of metacognitive judgment has been associated with a state's being conscious: *confidence* about whether one's perception of something is accurate.

THE GRADUATE CENTER  
OF THE CITY UNIVERSITY OF NEW YORK

Consciousness and Metacognition ASSC 20, Buenos Aires, 15 June 2016 16

- Confidence is inviting as an indicator of a state's being conscious, since total lack of confidence suggests guessing, and hence that the perceiving isn't conscious (Peirce & Jastrow 1884; Cheeseman & Merikle 1984; Merikle 1992).
- Appeal to explicit confidence ratings has been refined by having subjects wager on a perceptual decision (e.g., Persaud *et al* 2007)—though this may be no more accurate than explicit confidence ratings (Dienes & Seth 2010).
- High confidence and a willingness to bet, though affected by many factors, typically coincide with subjective reports (Persaud *et al* 2007; Dienes & Seth 2010)—presumably because lack of conscious awareness tends to result in low confidence and a reluctance to wager.

THE GRADUATE CENTER  
OF THE CITY UNIVERSITY OF NEW YORK

Consciousness and Metacognition ASSC 20, Buenos Aires, 15 June 2016 17

- Here would be an interesting case: Subjects in Sperling's (1960) partial-report task are confident they see all the letters as letters (even when some are not letters at all [Kouider *et al* 2010]).
- So they might at first also have confidence that they could identify all the letters, and be surprised that they can't. But after some trials they would likely lose confidence about identifying most—though remain confident that they are all letters.
- That suggests that their visual experience is conscious in respect just of the items' being letters—and not specific identities. Subjects' identifying the few they can relies largely on retrieving unconscious vision.

K	S	M	R
X	D	Q	G
B	Z	O	H

THE GRADUATE CENTER  
OF THE CITY UNIVERSITY OF NEW YORK

Consciousness and Metacognition ASSC 20, Buenos Aires, 15 June 2016 18

- Still, confidence and willingness to wager do not always coincide with the relevant perceptual states' being conscious.
- The blindsight subject, GY, is confident in some guesses about stimuli presented in his blind field (Sahraie, Weiskrantz, & Barbur 1998). And TN, with total blindsight (de Gelder et al 2008), visually navigates around obstacles—evidently with confidence in that ability despite having no conscious visual states.
- Peters & Lau (2015) report lack of evidence for anything like blindsight in normals, and they controlled for “hunches” based on unconscious vision. But arguably they did not consider a crucial feature of the blindsight situation.

THE GRADUATE CENTER  
OF THE CITY UNIVERSITY OF NEW YORK

Consciousness and Metacognition ASSC 20, Buenos Aires, 15 June 2016 19

- Rating one's confidence is a conscious response to perceiving something. And normals typically don't respond consciously to perceptions they're wholly unaware of. If pressed, their response to unconscious perceiving is to deny that they perceive anything. That's why confidence in normals coincides well with consciousness.
- But GY and TN were at first normals. And developing confidence for some blindsight cases, in addition to conscious states, must have taken some experience and learning.
- Since normals can overcome that limitation when afflicted with blindsight, there's every reason to expect they could also do so just with suitable training.

THE GRADUATE CENTER  
OF THE CITY UNIVERSITY OF NEW YORK

Consciousness and Metacognition ASSC 20, Buenos Aires, 15 June 2016 20

- Such training may not have to be lengthy. A bit of feedback will lead some to realize that their forced-choice guesses are often accurate, and so generate confidence.
- Even without feedback, some normals may, just with suitable encouragement, come to be able to tap into unconscious perceptions and so come to have some confidence.
- Confidence is, in standard cases, a reliable indicator that a perception is conscious, but only because in standard cases subjects ignore unconscious perceptions. But with training or encouragement, normals may develop confidence for nonconscious cases. The gold standard for a state's being conscious must remain subjective report.

THE GRADUATE CENTER  
OF THE CITY UNIVERSITY OF NEW YORK

Consciousness and Metacognition ASSC 20, Buenos Aires, 15 June 2016 21

- HO theories readily explain all this. Rating confidence in standard cases is explicitly about whether perceptions are accurate. Subjects *conceive* of their confidence ratings *in terms of* their perceptual states.
- So in standard cases one would withhold rating confidence unless one were aware of oneself as having the perception—which is having a HOA, enabling a subjective report. That's why confidence and consciousness go well together in standard cases.
- But a blindsight subject asked to rate confidence on a forced-choice guess is unaware of having any relevant perception, and so won't conceive of such confidence as being about a perception.

THE GRADUATE CENTER  
OF THE CITY UNIVERSITY OF NEW YORK

Consciousness and Metacognition ASSC 20, Buenos Aires, 15 June 2016 22

- A blindsight subject will instead conceive of confidence as being simply about the occurrence or nature of a *stimulus*. In blindsight, confidence is detached from any HOA of the perception, and so it occurs without the perception's being conscious.
- Similarly with normals who, with training or just encouragement, have confidence in their guesses even when the relevant perceiving is unconscious. They conceive of confidence solely in terms of stimuli.
- Still, subjects in these cases may come to be aware that there is a relevant perception—even though the perception isn't conscious. Subjects will assume that accurate guesses must be due to unconscious perceiving.

THE GRADUATE CENTER  
OF COLUMBIA UNIVERSITY  
IN THE CITY OF NEW YORK

Consciousness and Metacognition ASSC 20, Buenos Aires, 15 June 2016 23

- And on a HO theory a state is conscious if one is aware of it. So won't a HO theory count unconscious states whose content one has confidence about as conscious?
- No. A state is conscious on a HO theory only if the HOA seems subjectively unmediated. Being conscious seems *subjectively* to be a wholly unmediated property of a state.
- And a blindsight or normal subject with confidence about unconscious perceptions will be aware of those perceptions in a way that subjectively seems mediated. Consciously knowing that one's guesses are largely right will, from one's own subjective point of view, mediate any awareness of the unconscious perceptions.

THE GRADUATE CENTER  
OF COLUMBIA UNIVERSITY  
IN THE CITY OF NEW YORK

Consciousness and Metacognition ASSC 20, Buenos Aires, 15 June 2016 24

- HO theories are independently defensible, and help explain a number of otherwise puzzling aspects of consciousness.
- They help explain the connection between confidence and consciousness— why confidence in standard cases conforms to subjective reportability, and why the two can also come apart.
- And it helps explain how consciousness can misrepresent our mental lives— i.e., which mental states we are actually in. And it even helps explain why many find misrepresentation by consciousness so odd that—with no independent reason—they simply build it into their thinking that consciousness cannot misrepresent.

THE GRADUATE CENTER  
OF THE CITY UNIVERSITY OF NEW YORK

Consciousness and Metacognition ASSC 20, Buenos Aires, 15 June 2016 25

**Thank you for  
your attention**

THE GRADUATE CENTER  
OF THE CITY UNIVERSITY OF NEW YORK

Consciousness and Metacognition ASSC 20, Buenos Aires, 15 June 2016 26