

## Unconscious higher-order thoughts (HOTs) as pre-reflective self-awareness?

Commentary on [Rowlands](#) on *Animal Personhood*

**Rocco J. Gennaro**

University of Southern Indiana

**Abstract:** Rowlands argues that many nonhuman animals are “persons,” contrary to the prevailing orthodoxy which rests on a mistaken conception of the kind of self-awareness relevant to personhood. He argues that self-awareness bifurcates into two importantly different forms — reflective self-awareness and pre-reflective self-awareness — and that many animals can have the latter, which is sufficient for personhood. I agree that there is good reason to think that many animals can have pre-reflective self-awareness, but I think Rowlands is mistaken about its nature. His account runs the risk of leading to an infinite regress objection, and his notion of pre-reflective self-awareness actually sounds more like reflective self-awareness. If Rowlands treats pre-reflective self-awareness as itself conscious, then it is less likely that animals can have even this form of self-awareness.

[Rocco J. Gennaro](#), Professor of Philosophy, University of Southern Indiana, specializes in philosophy of mind/cognitive science and consciousness, with interests in metaphysics, ethics, moral psychology, and early modern history of philosophy. His most recent book is *Consciousness* (Routledge Press, 2017).  
<http://faculty.usi.edu/rjgennaro>



Rowlands (2016) argues that many nonhuman animals are “persons,” contrary to the prevailing orthodoxy which rests on a mistaken conception about the kind of self-awareness relevant to personhood. He argues for the following claims:

*“(1) Broadly speaking, self-awareness bifurcates into two importantly different forms. (2) Although there is little evidence for the claim that animals have the first of these forms of self-awareness [‘reflective self-awareness’], there are very good reasons for supposing that they have the second [‘pre-reflective self-awareness’]. (3) It is this second form of self-awareness – the form that many animals can have – that is crucial to the question of personhood....(4) at least some animals can be persons because (4a) they have a mental life and (4b) are aware of this mental life in the sort of way that can confer the requisite unity upon it” (p. 2).*

I mostly agree with (1) and (2), which is the focus of my commentary. I’ll touch only briefly on (3) and (4) later. I also agree with Rowlands that there is good reason to think that many animals have pre-reflective self-awareness such as by having episodic memory (Gennaro 2009, 2012, chapter eight), although I think he is mistaken about its nature. Like Rowlands, I do think it is important to understand that “self-awareness is not restricted to reflective self-awareness”

(p. 6). Reflective self-awareness involves a level of sophistication including introspection where one turns one's conscious attention toward a distinct mental state, whereas pre-reflective self-awareness is more intimately connected to conscious states in some sense. Rowlands cites Sartre and refers "to this second kind of awareness as *pre-reflective* self-awareness" (p. 7). But as a proponent of a version of the "higher-order thought" (HOT) theory of consciousness (Gennaro 2012), I think it is better to view pre-reflective self-awareness as *unconscious*, whereas Sartre and others seem to treat pre-reflective self-awareness as itself conscious in some way.<sup>1</sup> Let's back up and consider the HOT theory alternative.

One question that should be answered by any theory of consciousness is: What makes a mental state a *conscious* mental state? There is a long tradition that has attempted to understand consciousness in terms of higher-order (self-)awareness. This view has been defended by several contemporary philosophers over the past few decades (e.g., Rosenthal 1986, 2002, 2005; Lycan 1996; Gennaro 1996, 2012). The basic idea is that what makes a mental state M conscious is a higher-order representation (HOR) of M. A HOR is a "meta-psychological" or "meta-cognitive" state, that is, a mental state directed at another mental state. So, for example, my desire to write a good paper becomes conscious when I am (non-inferentially) "aware" of the desire. Intuitively, conscious states, as opposed to unconscious ones, are mental states that I am "aware of" being in. This overall idea is sometimes referred to as the Transitivity Principle (TP): "A conscious state is a state whose subject is, in some way, aware of being in it."

Conversely, the idea that I could be having a conscious state while totally *unaware* of being in that state seems like a contradiction. A mental state of which the subject is completely unaware is clearly an *unconscious* state, that is, there is no feel or sentience associated with it. For example, I would not be aware of having a subliminal perception and so it is unconscious.

There are various kinds of HOR theory, with the most common division between higher-order thought (HOT) theories and higher-order perception (HOP) theories. HOT theorists, such as Rosenthal (2005) and I (2012), think it is better to understand the HOR as a thought containing concepts. HOTs are treated as cognitive states involving some kind of conceptual component. HOP theorists urge that the HOR is a perceptual state which does not require the conceptual content invoked by HOT theorists (Lycan 2004). I will focus on HOT theory here.

According to HOT theory, when a conscious mental state is a first-order, world-directed conscious state, the higher-order thought (HOT) is *not* itself conscious. But when the HOT is itself conscious, there is a yet higher-order (or third-order) thought directed at the second-order state. In this case, we have *introspection* which involves a conscious HOT directed at an inner mental state. When one introspects, one's attention is directed back into one's mind. For example, what makes my desire to write a good paper a conscious first-order desire is that there is an *unconscious* HOT directed at the desire. In this case, my conscious focus is directed outward at the paper or computer screen, and so I am not consciously aware of having the HOT from the first-person point of view. When I *introspect* that desire, however, I then have a *conscious* HOT directed at the desire itself. It is thus crucial to distinguish first-order conscious states (with unconscious HOTs) from introspective states (with conscious HOTs). So I think of

---

<sup>1</sup> I have argued elsewhere that Sartre would also have been better off holding a view closer to HOT theory (Gennaro 2002, 2015).

“pre-reflective self-awareness” as unconscious HOTs and “reflective self-awareness” as conscious HOTs (= introspection).

This aspect of HOT theory also allows it to avoid charges of an infinite regress or circularity because a conscious state is not explained in terms of other conscious states. But Rowlands’s alternative, perhaps more like Sartre or Brentano, would seem to carry with it the increased risk of an infinite regress since the notion of an accompanying pre-reflective self-awareness is itself conscious. Sartre and Brentano are usually taken as *rejecting* the view that there are unconscious mental states and that any reductionist account of conscious states is viable. I won’t defend these two positions here but both are surely widely held today. One might of course *argue* that Sartre and Brentano are right, but to assume they are for the sake of this specific argument seems rather contentious. Now Rowlands may be alluding to something like HOT theory in the following passage:

*“Nevertheless, such is the pull of the idea of reflective awareness that an opponent might ... dig in her heels and claim that these thoughts must be there – merely in unconscious form ... I am quite happy to accept that these thoughts are there implicitly – and I shall shortly try to explain what this sort of implicit awareness amounts to. But the category of the implicit is not the same as that of the unconscious. There is no reason to suppose that whenever I see an object as an object, my seeing must be grounded in an indefinitely large array of unconscious thoughts and/or other intentional states” (p. 8).*

But of course, there are, at least according to many of us, many good reasons to suppose that whenever I consciously see an object, there is at least one unconscious HOT directed at that conscious perception. I mentioned TP above but obviously a full defense of HOT theory cannot be offered here (see Rosenthal 2005, Gennaro 2012).

What exactly is the relationship between a conscious mental state and pre-reflective self-awareness? Unlike cases of introspection or reflective self-awareness where the self-awareness is clearly distinct from the target mental state, there would seem to be a closer connection between a conscious state and pre-reflective self-awareness of that state.

Rowlands, citing Wittgenstein, explains that “knowing that I am in pain is part of what it is to be in pain rather than a separate act that has my pain as its intentional object (p. 6).” Again:

*“Unlike its reflective counterpart, pre-reflective awareness is not a separate act. Rather it is built into acts of conscious experience that are not, typically, about the self but about something else entirely: the book that I see as a book, the chair that I see as affording sitting, and so on....Pre-reflective self-awareness is, in this sense, not separate from these conscious acts of seeing but part of these acts” (p. 15).*

I agree with Rowlands that pre-reflective self-awareness is not separate from conscious states. Unlike Rosenthal who holds that the (unconscious) HOT is *entirely distinct* from its target, I have argued that first-order conscious states are better construed as complex states with two parts: a mental-state part directed at the world (M) and an unconscious “meta-psychological thought” (MET) part directed at M. This is what I have called the “wide intrinsicity view” (WIV) such that consciousness is an intrinsic aspect of conscious states and thus such states should be

individuated widely (Gennaro 1996, 2006, 2012, chapter four). Nonetheless, during “reflection” or “introspection,” there is a *conscious* HOT (or MET) directed at a mental state and, in this case, there is a greater epistemic “gap” between the HOT and its object such that the conscious HOT is entirely distinct from its target mental state.

I differ from Rowlands, however, in that I would treat the pre-reflective self-awareness in question as itself an unconscious (but embedded) part of a conscious state. It also seems to me that doing so makes it much more likely that animals are capable of having pre-reflective self-awareness as opposed to some kind of conscious self-awareness. This is because, as Rowlands acknowledges, there is far less evidence that animals are capable of having the more sophisticated reflective self-aware states. Further, one runs the risk of infinite regress worries if, as it seems on Rowlands’s view, pre-reflective self-awareness is itself conscious in some way. This is because we would then also need to explain what makes this self-awareness conscious, and so on. Thus, if Rowlands’s account of the distinction between pre-reflective and reflective self-awareness does not stand up to scrutiny, then relying on it for the remainder of his argument is problematic in the ways discussed here.

Is pre-reflective self-awareness enough for personhood? Although this is not the main focus of my commentary, I am not convinced that it is enough for personhood, partly because I think that personhood is tied more closely to moral (or, at least, legal) *rights* and the capacity to reason, which has much more to do with reflective self-awareness or introspection. Many other commentators have responded more directly along these lines.

## References

- Gennaro, R. (1996). *Consciousness and Self-Consciousness*. Amsterdam and Philadelphia: John Benjamins Publishers.
- Gennaro, R. (2002). Jean-Paul Sartre and the HOT theory of consciousness. *Canadian Journal of Philosophy*, 32, 293-330.
- Gennaro, R. (2006). Between pure self-referentialism and the (extrinsic) HOT theory of consciousness. In U. Kriegel and K. Williford (Eds.), *Self-Representational Approaches to Consciousness*. Cambridge, MA: MIT Press.
- Gennaro, R. (2009). Animals, consciousness, and I-thoughts. In R. Lurz (Ed.), *Philosophy of Animal Minds*. Cambridge, MA: Cambridge University Press.
- Gennaro, R. (2012). *The Consciousness Paradox: Consciousness, Concepts, and Higher-Order Thoughts*. Cambridge, MA: The MIT Press.
- Gennaro, R. (2015). The ‘of’ of intentionality and the ‘of’ of acquaintance. In S. Miguens, G. Preyer, and C. Morando (Eds.), *Pre-Reflective Consciousness: Sartre and Contemporary Philosophy of Mind*. New York: Routledge Publishers.
- Lycan, W. (1996). *Consciousness and Experience*. Cambridge, MA: MIT Press.
- Lycan, W. (2004). The superiority of HOP to HOT. In R. Gennaro (Ed.), *Higher-Order Theories of Consciousness: An Anthology*. Amsterdam and Philadelphia: John Benjamins.
- Rosenthal, D. M. (1986). Two concepts of consciousness. *Philosophical Studies*, 49, 329-359.
- Rosenthal, D. M. (2002). Explaining consciousness. In D. Chalmers (Ed.) *Philosophy of Mind: Classical and Contemporary Readings*. New York: Oxford University Press.
- Rosenthal, D. (2005). *Consciousness and Mind*. New York: Oxford University Press.
- Rowlands, M. (2016). [Are animals persons?](#) *Animal Sentience* 2016.101.