

In Defense of H.O.T. Theory: A *Second* Reply to Adams and Shreve

Rocco J. Gennaro

Abstract: In Gennaro (2016), I had originally replied to Fred Adams and Charlotte Shreve's (2016) paper entitled "What Can Synesthesia Teach Us About Higher Order Theories of Consciousness?," previously published in *Symposion*. I argued that H.O.T. theory does have the resources to account for synesthesia and the specific worries that they advance in their paper, such as the relationship between concepts and experience and the ability to handle instances of 'pop-out' experiences. They counter-reply in Adams and Shreve (2017) and also raise further objections to H.O.T. theory which go well beyond the scope of their 2016 paper. In this paper, I offer additional replies to the points they raise in Adams and Shreve (2017).

Keywords: autism, concepts, conceptualism, consciousness, higher-order thoughts, synesthesia.

In Gennaro (2016), I had originally replied to Fred Adams and Charlotte Shreve's (A&S hereafter, 2016) paper entitled "What Can Synesthesia Teach Us About Higher Order Theories of Consciousness?," previously published in *Symposion*. I argued that H.O.T. theory does have the resources to account for synesthesia and the specific worries that they advance in their paper, such as the relationship between concepts and experience and the ability to handle instances of 'pop-out' experiences. I will not repeat those responses here. But A&S have counter-replied in Adams and Shreve (2017) and also raise further objections which go well beyond the scope of their 2016 paper. In this paper, I offer additional replies to the points they raise in Adams and Shreve (2017).

1. Worries about Introspection

A&S first say that "Gennaro (2017) takes issue with our seat-pressure case. In that case we say that on the H.O.T. theory, the H.O.T. is supposed to make one's experience of the pressure being exerted conscious. He says this is an appeal to 'introspection' and, on his view, it is not introspection" (Adams and Shreve 2017, 129).

Reply: I was partly merely pointing out that their description of the seat-pressure case (e.g. 'one's experience of the pressure') was somewhat ambiguous between (1) consciously attending to the felt pressure, and (2) being unconsciously aware of the seat pressure. If they meant the former, e.g. when they said that "one's attention turns to that pressure" (2016, 251), then it seemed like an appeal to introspection and thus not an example of a *first-order*

conscious mental state according to H.O.T. theory (or any theory, for that matter). If they meant the latter, then the conscious feeling is presumably more like an outer-directed perception, albeit aimed at one's body to some extent.

Recognizing the potential for ambiguity to some extent, A&S (2017, 129) say that "it is a kind of 'extrospection' upon the pressure being exerted on the seat. Attention is directed at least partially outwardly." But this also illustrates the other related reason that I raised a concern, namely, there is the more complicated independent issue of just how to characterize 'bodily experience.' Perhaps feeling seat-pressure is more akin to outer perception but sometimes it seems to involve attention to a bodily *sensation* which might be better construed as introspecting the feeling itself. It is also not clear to me how one's attention can be partly inner- and partly outer- directed *at the same time* but I won't elaborate on that here. A&S are certainly correct in saying that "turning one's attention on *something* is not necessarily introspecting" (2017, 130, my emphasis) but my point was merely that turning one's attention *to a mental state* (e.g. a feeling of pressure) would be introspecting. Thus, the example they used was perhaps just more problematic than other clearer cases of outer-directed attention.

A&S then remark that "...by the way, Gennaro accepts the TP principle. TP says 'A conscious state is a state whose subject is, in some way, aware of being in it' (Rosenthal 2005). To us that sounds a lot like introspection." (2017, 130)

Reply: But A&S are here clearly conflating two readings of 'aware of' in TP. There is unconscious awareness of a mental state and there is conscious awareness of a mental state – it is only the latter which is introspection, according to H.O.T. theory. But when one has a conscious outer-perception, one is unconsciously aware of the perception. By the way, the 'aware of' in TP is also itself neutral between H.O.T. and H.O.P. (higher-order perception) but H.O.T. theorists give reasons as to why it is best to interpret 'aware of' as H.O.T. (Rosenthal 2004, 2005, Gennaro 2012, chapter 3).

A&S then state that "Gennaro also maintains that H.O.T.s are somehow 'presupposed' by any experience or conscious mental state. This is partly why he thinks that every experience is conscious and why he thinks the contrary view makes no sense. We have already explained above, why the contrary view makes sense to us, so we are still struggling with the idea that every experience or conscious state 'presupposes' H.O.T.s. We just don't understand the view." (2017, 132)

Reply: Once again, I think that H.O.T.s are presupposed by all conscious states because the concepts that are applied in conscious states are in the H.O.T.s and because the concept application is rarely itself conscious. But I don't say that contrary views 'make no sense' but rather that they are wrong if I am right. Another way to make the point is as follows: If H.O.T. theory is true, then there is a H.O.T. present for each first-order conscious state. But we are not consciously aware of such a H.O.T.; that is, it is an unconscious H.O.T. In addition, my own

view is that it is best to think of a H.O.T. as part of a complex state which also includes its target state. This is what I call the 'wide intrinsicity view' (WIV). So it is simply in this sense that I say that every conscious state *presupposes* a H.O.T.

2. H.O.T.s and Sub-Personal States

A&S again object to my claim that H.O.T.s can be sub-personal, unconscious, and higher-order. "...we don't see how this can be true. They might be able to be non-conscious. Most versions of H.O.T. theory allow for non-conscious H.O.T.s. This is one of our problems with them, viz. how can something non-conscious bestow something it doesn't have (consciousness) on something else (an experience or thought) that also lacks it?" (2017, 132)

First: All (not 'most') versions of H.O.T. theory allow for non-conscious H.O.T.s – that is central to the theory in so many ways and also why it is a (mentalistic) reductionist theory of consciousness. Thus, a H.O.T. can also be 'sub-personal' – all I mean here is that the subject is not aware of it since it is an unconscious mental state. Of course, some sub-personal states may not be mental states at all or purely 'informational' in some sense, but I am not referring to those when referring to H.O.T.s. In the end, I do think that the structure of H.O.T. theory is realized in the brain (Gennaro 2012, chapters 4 and 9).

Secondly, the final question in the above quote from A&S ("how can something non-conscious bestow something it doesn't have (consciousness) on something else (an experience or thought) that also lacks it?") seems more like a request to solve the "hard problem" (Chalmers 1995) or perhaps to bridge the explanatory gap (Levine 2001) between consciousness and something non-conscious. I do spend a significant portion of Gennaro (2012, chapters 4 and 9) on these very important challenges in an attempt to show that H.O.T. theory is immune to Chalmers' hard problem and that H.O.T. theory offers a plausible reductionist theory. I cannot adequately summarize those chapters here but, for example, I argue that

The solution is that H.O.T.s explain how consciousness arises because the *concepts* that figure into the H.O.T.s are presupposed in conscious experience. Let us stick to first-order perceptual states. Very much in a Kantian spirit, the idea is that we first passively receive information via our senses. This occurs in what Kant (1781/1965) calls our 'faculty of sensibility,' which we might think of as early perceptual processing. Some of this information will then rise to the level of unconscious mental states, which can also cause our behavior in various ways. But such mental states do not become conscious until the faculty of understanding operates on them via the application of concepts....Kant (1781/1965) urges that it takes the cooperation of both the sensibility and understanding to produce conscious experience. Regarding the sensibility and understanding: 'Objects are *given* to us by means of sensibility. . . . They are *thought* through the understanding, and from the understanding arise concepts' (A19/B33). (Gennaro 2012, 77-78)

Rocco J. Gennaro

Still, it would be good to know if A&S are saying that *no* reductionist account of consciousness is plausible. If so, then any potential solution will remain unsatisfying to them for one reason or another. If not, then what reductionist theory of consciousness do they prefer? How does it explain consciousness or 'bestow consciousness' on some mental states but not others? How would it fare against H.O.T. theory? I certainly don't expect to convince A&S that they should become H.O.T. theorists. Still, there are advantages and disadvantages (or objections) to all theories of consciousness. The difficult part is trying to make the case that the one theory is, overall, better off than others.

3. Conscious Experiences One is Not Conscious of Having?

A&S explain that

Gennaro asserts: For example, if I am having a conscious desire or pain, I am aware of having that desire or pain. Conversely, the idea that I could have a conscious state while totally unaware of being in that state seems very odd (if not an outright contradiction) (Gennaro 2016, 444)...[but] There are alternative views to H.O.T. theories on which this is not only not contradictory, but quite plausible. As Gennaro knows, Dretske's (1993) view of conscious experience makes it possible to be in a conscious state (state of seeing something) and not know one is in that state. Dretske gives several examples of...'change blindness,' where one is presented with a visual array and later presented with another different array...it could be a missing dot or even a missing engine on Boeing 747 (Adams and Shreve 2017, 130-131).

Reply: First, it is unlikely that one is consciously aware of *every aspect* of, say, one's visual perceptions. Indeed, this is part of the point of change and inattention blindness, especially in cases where I am experiencing dozens of objects or dots in an array or, say, all of the spots on a leopard. It is unclear that one is actually *experiencing* each of the dots and spots. It is not just *what* one is *in fact* experiencing but *the way* that one is experiencing it. Surely, none of us should hold that we *consciously notice* every single aspect or property in a single given visual scene. (I do discuss change and inattention blindness in Gennaro 2012, chapter 6.) Consider also Dennett's (1991) case of the Marilyn Monroe wallpaper, where you walk into a room with wallpaper containing hundreds of her portraits. Your initial sense might convince you that you are seeing hundreds of identical Marylins. But are you really? Dennett persuasively argues that the real detail is not in your head but in the world. We simply assume that all the pictures are of Marilyn Monroe; that is, our brains 'fill in' the rest of the scene. We thus mistakenly assume that all of the Marylins are represented in our experience (Dennett 1991, 354-355). This likely occurs often when we experience a number of similar-looking objects at the same time, unless one object is so different as to 'pop out' in the experience. You obviously do not focus in (or foveate) on each and every portrait. Indeed, it would seem that you are only peripherally aware of the vast majority of portraits at any given time. It is

unlikely that you would notice if, say, six or seven of the portraits were altered to contain portraits of another blonde female. Still, there is a sense in which we still might *say* that you are ‘experiencing dozens and dozens of Marilyn portraits.’ Indeed, all of the above phenomena are sometimes used to show that conscious experience is *not* as ‘rich’ as it might seem (Gennaro 2012, chapter 6).

But, second and more to the point at hand, A&S’s use of Dretske’s argument is easily refuted by a H.O.T. theorist. Actually, William Seager (1994), no friend of H.O.T. theory, pointed out long ago that there is a crucial and defeating ambiguity in several of Dretske’s arguments against H.O.T. theory. For example, the two following claims are not equivalent and neither one entails the other:

- (1) One can have a conscious state without being aware that one is in it.
- (2) One can have a conscious state without being able to tell the difference between it and another very similar conscious state.

A H.O.T. theorist would of course deny (1) but could easily accept (2) which is likely to occur often, as A&S also agree. To put it more positively, consider:

- (3) When one has a conscious state one is aware of being in that state.
- (4) When one has two conscious states at different times, one is aware of the *difference(s)* between those states.

Again, (3) would be endorsed by a H.O.T. theorist and is really just the TP, but (4) simply does not follow. Why should a H.O.T. theorist suppose that one would always *consciously notice the difference* between two conscious states? Even if we granted that one consciously experiences every single aspect of a complex visual scene, it wouldn’t follow that one would therefore consciously notice the difference between it and another very similar scene.

Indeed, A&S (2017, 131) themselves rightly point out that “you can be in a state which is the conscious experience of the difference in arrays, but not be conscious that it is *the difference* that you are experiencing.” If this simply means (2) above, then I agree. But then they say that “you are in a conscious state that you are not conscious that you are in” which is clearly a different claim and one which would be denied by any H.O.T. theorist given that it contradicts TP.

Seager explains the objection as follows:

Dretske is equivocating between what is, in essence, a *de re* and a *de dicto* characterization of consciousness. Would H.O.T. theories demand that S be conscious of the difference between any two distinct experiences *as a difference*? Clearly the answer is no, for S may simply have never consciously compared them. In such cases – quite common I should think – S need not be conscious of the difference at all. Well, should H.O.T. theories require that if any two of S’s conscious experiences are different and S is actually conscious of the difference (i.e. conscious of what is different between the two experiences) then S must be conscious of this difference *as a difference*? This also calls for a negative answer. To say that S is conscious of the difference in this sense is to

say that there is something different about the two experiences of which S is conscious; this puts no, or very few, restrictions on *how* that experience will be characterized in S's belief about it which, according to the H.O.T. theory, constitutes S's consciousness. That is, to say that S is conscious of the difference is, on the H.O.T. theory, to say that S believes that he is experiencing the difference. But in the case envisaged this is true only on a *de re* reading of this belief. A more precise specification of this belief that brings out its *de re* character is this: of the difference (between the two experiences) S believes of it that he is experiencing it. It does not follow that S is conscious of the difference *as* a difference.... In short, the H.O.T. theories of consciousness can admit the phenomena that Dretske points out without succumbing to the objections he believes they generate. (Seager 1994, 275-276)

There can be a difference between *what* one sees ('seeing that') and *how or the way* that one sees it ('seeing-as'). In any case, A&S haven't shown that "I could have a conscious state while totally unaware of being in that state." At best, they point out (and I agree) that I could have two somewhat similar conscious states at different times and yet be totally unaware *of a difference* between them.

4. H.O.T.s and Conceptualism

A&S then explain that

it seems to us that there are cases where one has a conscious experience of a kind of thing for which one lacks a concept. Indeed, Gennaro himself (2012, 157) gives us this kind of case. He admits that one can see a whistle without seeing it as a whistle. Thus, one can have a basic visual perceptual experience of a whistle without applying the concept of a whistle. One can know what it (the whistle) looks like, even though one would not describe it as having the look of a whistle, because one lacks the concept of whistle or the concept of look of a whistle.....same with an infant in crib seeing a mobile.....(2017, 131)

Much the same from the previous section applies here. As A&S know, I use the whistle example in discussing visual agnosia where a subject (abnormally) is unable to recognize an object that would otherwise be very familiar. But these are simply more extreme cases of seeing or experiencing something which the subject does not recognize. Like the infant seeing a mobile or someone seeing an armadillo for the first time (more on this below), one can experience something *that is in fact* an X while not experiencing X *as an X*. Like Dretske, A&S trade on the ambiguity in the expression 'a conscious experience of a thing.'

A&S further explain that:

when one looks briefly at the words on a page one may have a visual presentation of each of the words. One consciously experiences them all. But one does not apply the concept 'word' to each and every single word on the page. Nonetheless, one consciously experiences every word. No higher order thought is required to generate the conscious visual presentation of the words on the page. There are too many words and too little time for higher order thoughts to produce each conscious element (2017, 132).

First, it is again doubtful that one consciously experiences every single word on a page when ‘one looks briefly at the words on a page.’ So there is no need to apply the concept WORD to each word on the page. It is more like briefly looking at the Marilyn Monroe wallpaper. Second, even when attending to a particular word, we need not consciously apply the concept WORD to the words on a page since the H.O.T.s are unconscious in such cases of outer-directed perception.

5. Concept Acquisition and Autism

A&S end with what they call two problems for my view: concept acquisition and autism. First:

On Gennaro’s view, when one consciously experiences X, one must have an H.O.T. of the form ‘I’m experiencing X.’ But this raises the problem of how can one acquire new concepts? Dretske (1993) gives the example of the first time he saw an armadillo. He had a conscious visual presentation of the armadillo, but didn’t know what it was. He used the incoming information about armadillos to form the concept of an armadillo. Gennaro’s view will require that to have a conscious experience of the look of an armadillo, one knows already what an armadillo is. Otherwise, one will consciously experience only an animal with a certain shape and moving in a certain way. But nothing specific to what it is to be an armadillo will be consciously experienced – because one doesn’t yet have the concept of what an armadillo is. So one can’t have an H.O.T. that one is having a visual experience of an armadillo (only of a creature or an animal or some such). So how does one ever consciously learn what an armadillo is or looks like? It seems to us this makes concept learning impossible for new empirical concepts. On our view, one must be able to receive new information about Xs and consciously experience Xs and their looks (perceptible properties) in order to form the concept of an X. Gennaro’s view might rely on some innate concepts (2012, 191-197), but none of those is going to be specific to what makes something an armadillo (as opposed to something else). So none of those innate concepts will generate the new empirical concept – armadillo. (Adams and Shreve 2017, 133)

Reply: As was explained in the previous two sections, one can see an armadillo for the first time without experiencing it as an armadillo. Still, I agree that concept acquisition is, in a way, the real hard problem. That is why in Gennaro (2012, chapter 7) I offer what I call the “TILT” (the implicit learning theory) of concept acquisition. The short answer to A&S concern here is that learning an empirical concept can take place both consciously and unconsciously, and that when one learns a new concept it can alter the very phenomenology of one’s perceptions. We can of course also learn about so many things via explicit conscious instruction and reasoning. But we can also unconsciously ‘receive new information about X’ and thus ‘form the concept of X.’ In some cases, this can happen fairly quickly. One can have a ‘conscious visual presentation of the armadillo’ but not initially experience the animal as an armadillo. With the help of the developmental psychology literature, I attempt to explain (at least to some

extent) how this can occur starting from a small set of innate concepts through to the formation of familiar empirical concepts (Gennaro 2012, chapter 7). The concept 'armadillo' is certainly not innate. Surely, however, infants and young children aren't explicitly taught every single concept that they possess. Something has to explain how they have *acquired* the concept but acquiring an empirical concept is not necessarily always done consciously.

Regarding autism, A&S say that

subjects with severe forms of autism are susceptible to pop-out synesthesia of the kind that we described in our initial paper (Adams and Shreve 2016). Now a hallmark of severe autism is what Baron-Cohen (1997) called 'mind-blindness.' This is the inability to apply mental concepts to self or others. People with severe autism have no trouble understanding people as physical systems with physical properties that are explicable in terms of natural physical laws. But when it comes to beliefs, desires, intentions, hopes, fears, wishes and other mental causes, severely autistic individuals simply do not understand behavior originating from these causes. Such purposive behavior is a complete mystery to them. Thus, they do not engage in applying mental concepts to themselves or others. Consequently, when a person with severe autism consciously experiences the pop-out of synesthesia, it cannot be the result of applying an H.O.T. to their experience because they don't employ H.O.T.s about mental states (of self or others). (2017, 133)

Reply: I think that the 'mind-blind' characterization of autism, even the more extreme cases, is mistaken or at least highly exaggerated. It is not at all clear to me that autistics cannot have or apply mental concepts to themselves or others (see Gennaro 2012, sec. 8.4). Much like with animals and infants, I think that critics of H.O.T. theory tend to make H.O.T.s seem more sophisticated than they need to be and also sometimes conflate reflection or introspection (= conscious H.O.T.s) with something more like 'pre-reflective self-awareness' or unconscious H.O.T.s. Here is a short quote from Gennaro (2012, 259):

One initial problem with the literature is that some authors who argue for a deficiency in 'self-consciousness' among autistic individuals leave the term undefined....self-consciousness, self-concepts, I-thoughts, and concept possession can come in degrees. At the most sophisticated level, there is introspection or reflection. Even if there are deficiencies in introspection, it does not follow that there are no I-thoughts or metacognitive states at all....it is one thing to suppose that autistic humans have *abnormal* or *impaired* self-consciousness, but quite another to claim that there is *no* self-consciousness at all. Indeed, even Frith and Happé (1999, 11-14) quote numerous cases of first-person reports from autistics [despite their skepticism regarding autistic self-consciousness].

A&S do say that "Gennaro (2012) thinks autistic individuals can have self-consciousness and that reflective self-awareness is not required for H.O.T.s. But how can it be a self-awareness if it is not reflective and self-aware? They must have something to make a thought self-referential (Adams and Shreve 2017,

133-134, fn 4).” Yes, non- or pre-reflective self-awareness still includes a self-referential element, namely, an ‘I’ as in “I am in mental state M.”

Overall, then, I think that H.O.T. theory can effectively handle the objections raised in Adams and Shreve’s 2017 paper though the focus has largely shifted away from synesthesia. To be sure, however, some readers may wish to look more closely at Gennaro (2012) and other related work (e.g. Seager 1994, Rosenthal 2004, 2005). Of course, many readers and A&S might still not be satisfied with my replies for various reasons. Nonetheless, this results in an interesting exchange of ideas and a better understanding of any deeper differences (and sometimes agreements) on these matters. (Indeed, many of these themes and related objections are also addressed in a 2013 *Journal of Consciousness Studies* symposium on Gennaro 2012, including my replies to William Seager, Robert Van Gulick, and Josh Weisberg.)

References:

- Adams, Fred and Charlotte Shreve. 2016. “What Can Synesthesia Teach Us about Higher Order Theories of Consciousness?” *Symposion. Theoretical and Applied Inquiries in Philosophy and Social Sciences* 3 (3): 251-257.
- Adams, Fred and Charlotte Shreve. 2017. “Reply to Gennaro.” *Symposion. Theoretical and Applied Inquiries in Philosophy and Social Sciences* 4 (1): 129-134.
- Baron-Cohen, Simon. 1997. *Mindblindness*. Cambridge, MA: MIT Press.
- Chalmers, David. 1995. “Facing Up to the Problem of Consciousness.” *Journal of Consciousness Studies* 2 (3): 200-219.
- Dennett, Daniel. 1991. *Consciousness Explained*. Boston: Little, Brown.
- Dretske, Fred. 1993. “Conscious Experience.” *Mind* 102 (406): 263-283.
- Frith, Uta, and Francesca Happé. 1999. “Theory of Mind and Self-Consciousness: What is it Like to be Autistic?” *Mind and Language* 14 (1): 1-22.
- Gennaro, Rocco J. 2012. *The Consciousness Paradox: Consciousness, Concepts, and Higher-Order Thoughts*. Cambridge, MA: The MIT Press.
- Kant, Immanuel. 1781/1965. *Critique of Pure Reason*. Trans. N. Kemp Smith. New York: Macmillan.
- Levine, Joseph. 2001. *Purple Haze: The Puzzle of Conscious Experience*. Cambridge, MA: MIT Press.
- Rosenthal, David M. 2004. “Varieties of Higher-Order Theory.” In *Higher-Order Theories of Consciousness: An Anthology*, edited by Rocco Gennaro. Amsterdam: John Benjamins.
- Rosenthal, David M. 2005. *Consciousness and Mind*. New York: Oxford University Press.
- Seager, William. 1994. “Dretske on H.O.T. Theories of Consciousness.” *Analysis* 54 (1): 270-276.