**“Misrepresentation, Empty HOTs, and Intrinsic HOTs: A Reply to Pereplyotchik”**

**Rocco J. Gennaro**

[Final Version is in *Philosophical Psychology* 2015, 28, 449-451.]

David Pereplyotchik (DP) addresses some aspects of *The Consciousness Paradox* (CP) focusing mainly on the issue of whether or not my “wide intrinsicality view” (WIV) version of HOT theory is superior to Rosenthal’s “extrinsic HOT” (EHOT) theory.[[1]](#endnote-1)

Of course, simply calling something a ‘fallacy’ doesn’t make it so and this goes for what DP calls my “intrinsic-ergo-accurate fallacy.” I never make an inferential claim from “intrinsic” HOTs to “necessarily accurate.” I do hold that, unlike EHOT theory, the HOT is better construed as intrinsic to an overall world-directed complex conscious state. Even in the early passage quoted by DP, I say that it “seems” (2012, p. 58) more difficult to make sense of misrepresentation on intrinsic theory, intentionally trying not to put too much weight on this line of argument. Perhaps I could have made this even clearer throughout the entire book, but I do argue that *no* HOT theorist should take Levine’s option 1 (“The conscious experience is of a *greenish* sort”) or option 2 (“The conscious experience is of a *reddish* sort”), and that option 3 (“no consciousness”) has significant advantages. In this situation, we must decide how to handle a case of mismatch between a lower-order (LO) state representing green and a HO state representing red. I also *argue for* the WIV *on many fronts*, including how it can better explain the way that HO concepts color the very phenomenology of first-order conscious states, how it better fits the neurophysiological facts, and how it handles the question of why some objects of thoughts (such as rocks) don’t become conscious whereas mental states do. When DP says that I hold that HOTs are “necessarily accurate” and that I “guarantee” a match between a HOT and its target, this is very misleading or at least oversimplified. So let’s look more closely:

First, if we think about the intrinsic/extrinsic issue from a third-person neurophysiological perspective, there is something odd about holding that what makes a mental state M conscious is *something else*, i.e. something entirely extrinsic to M. If and when the true neural correlates of consciousness are discovered, it seems far more likely that they will be treated as part of conscious brain states. There can still be a HOT-like structure to a conscious state, but both the mental state M and the metapsychological thought (‘MET,’ or HOT) about M will be parts of the state. I greatly elaborate on this idea, especially in chapters four and nine, with reference to various brain mechanisms and feedback loops.

Second, much of what I say about the misrepresentation problem could (and should) be adopted by standard EHOT theory. Option 1 is still much more problematic to me (as is option 2). In Levine’s green/red diskette case, somehow the HO representation *alone* matters in determining what a conscious state is like. But, as Levine says, doesn’t this defeat the purpose of HOT theory which is supposed to explain state consciousness in terms of the relation between *two* unconscious states? Moreover, as I point out in my book: “HOT theory is supposed to be a theory of (intransitive) state consciousness; that is, the *lower-order* state is supposed to be the conscious one” (CP, p. 60). Also, according to Rosenthal, HOTs themselves have no qualia, so how can a HOT alone have any “what it’s like” aspect to it?

Third, option 1 leads to even more puzzling consequences in “targetless” or “empty” HOT cases. How can a lone *unconscious* HOT be conscious? DP and Rosenthal often say that what is important is the way that our mental states “appear to us” which is meant as an inner analogue of outer-directed hallucinations. But I argue that this is a very poor analogy given the structure of conscious states according to HOT theory itself. I say: “I agree…that our *introspective* states are fallible and may misrepresent our …mental states. But this distinction applies at the introspective level, not *within* first-order world-directed conscious states. If there is an inner analogy to an…hallucinatory first-order conscious state…it must be a *conscious state* (= introspection) directed at a mental state. But then this is not a case of an appearance/reality difference between an *unconscious* HOT (or MET) and a mental state M.…A lone unconscious HOT without its target is *not* a case of fallible *introspection*” (CP, p. 97). So DP’s frequent appeal to “mental appearance” fails if it is meant to apply to *unconscious* HOTs.

Fourth, I specifically point out (on p. 64) how the following two claims are really two sides of the same coin:

(1) There is no resulting conscious state when a misrepresentation does occur, and

(2) Misrepresentations cannot occur.

As I say in the book, (2) should really be understood as:

(2′) Misrepresentations cannot occur between M and MET *and* still result in a conscious state.

Indeed, DP quotes (2’) from p. 158 in my book. Perhaps I should even clarify further here:

(2’’) Misrepresentations cannot occur between M and MET *and* result in a conscious experience reflecting incompatible concepts.

It is not that misrepresentations cannot occur *as such*, but rather that misrepresentations cannot occur between M and MET *and* still result in a conscious experience reflecting incompatible concepts. It is the *combination* of these claims that is most important. Notice that Levine’s case involves *incompatible* concepts applied to the diskette case, i.e. the case cannot be all red and all green.

Fifth, DP’s oversimplified characterization of my view ignores extensive discussion in chapter six. For example, I present my final nuanced view as follows:

“Whenever a subject S has a HOT directed at e, the content c of S’s HOT determines the way that S experiences e (provided that there is a *full* or *partial* conceptual match with the lower-order state, *or* when the HO state contains more specific or fine-grained concepts than the LO state has, *or* when the LO state contains more specific or fine-grained concepts than the HO state has, *or* when the HO concepts can combine to match the LO concept).” (CP, p. 180)

The reasons for the above qualifications are discussed at length. I cannot summarize here except to say that I aim to explain what happens in some abnormal cases (e.g. visual agnosia) and in some atypical contexts (e.g. perceiving ambiguous figures such as the vase-two faces) where different kinds of mismatches might occur between the HOT and LO state. The details matter and they also show that it is very misleading to say that, on my view, misrepresentations “cannot occur.” I also present a robust account of concept acquisition in chapter seven via the key notion of implicit learning.

Sixth, partly due to some of the confusion above, DP is naturally dissatisfied with my discussion of any alleged analogy between misrepresentations in pictures or sentences and in conscious states (CP, pp. 66-67). But I still fail to see the appropriateness of these analogies given that they are neither psychological nor neurophysiological analogues of conscious states. Finally, DP places significant weight on the way that we “normally” individuate mental states by mental attitude and the way that we express and report our mental states. I have little to say about this line of argument other than what is already in my book (in section 4.5.2).

I conclude that the WIV is superior to EHOT theory, or at least just as plausible. DP has not shown otherwise.

**References**:

Gennaro, R. (forthcoming) “Defending HOT Theory and The Wide Intrinsicality View: A Reply to Weisberg, Van Gulick, and Seager,” *Journal of Consciousness Studies*….

Gennaro, R. (2012). *The Consciousness Paradox: Consciousness, Concepts, and Higher-Order Thoughts*. Cambridge, MA: MIT Press.

Pereplyotchik, D. (2015). “Some HOT Family Disputes: A Critical Review of The Consciousness Paradox,” *Philosophical Psychology* 28: 434-448.

1. I lack the space to address all of DP’s remarks. If the reader would like a longer version of this reply, please contact me. The reader might also be interested in the *Journal of Consciousness Studies* symposium papers where many of these themes are addressed further (Gennaro, forthcoming). [↑](#endnote-ref-1)