

Forthcoming in a special issue of
Trans/Form/Ação on consciousness,
ed. Alfredo Pereira

MISREPRESENTATION AND MENTAL APPEARANCE

David Rosenthal

Philosophy and Cognitive Science
Graduate Center, City University of New York

Abstract: I begin by discussing an objection often lodged against higher-order theories of consciousness. The objection is that those theories do not preclude consciousness from misrepresenting the mental properties of our conscious states. I argue that there are several reasons why this objection cannot succeed. Sam Coleman (2019) agrees that the objection fails, but sees it as pointing to a related objection, which he argues poses difficulties the higher-order theorist cannot avoid. His solution is a variant theory of consciousness that invokes mental quotation in the content of a higher-order awareness. I'll argue that mental quotation cannot be understood in any way that helps with the objection. I'll argue also that the objection Coleman articulates is question begging, since it in effect simply channels a first-order approach to consciousness, and has no independent basis apart from that.

Keywords: Higher-order theories of consciousness, misrepresentation by consciousness, mental quotation, same-order theories, mental appearance

In a thoughtful and probing article in this issue, Sam Coleman (2019) develops a difficulty for my higher-order-thought (HOT) theory of what it is for a mental state to be a conscious state. After offering a careful exposition of the theory, Coleman discusses one common criticism of that theory. He argues that though the HOT theory can withstand that objection, there is a related difficulty that he urges the theory cannot meet. Coleman goes on to develop an alternative theory, which he maintains can avoid what he sees as the more telling objection he advances.¹

The wide-spread objection to which Coleman calls attention is that higher-order theories do not on their own preclude the possibility that our higher-order awareness of conscious states may sometimes misrepresent some of the mental properties of those states. This feature of higher-order theories is seen by many as a decisive objection to such theories (e.g., Neander 1998, Levine 2001, Block 2011). Unlike many critics, Coleman regards my previous replies to that objection (e.g., 2011, 2012) as satisfactory. In section I, below, I discuss the objection at more length than I have in the past, and explain why it cannot undermine higher-order theories.

Though Coleman agrees with the bottom line of my assessment of the objection from the possibility of misrepresentation, he sees a related objection that my previous responses about misrepresentation leave unaddressed. And he regards that related difficulty as revealing the deeper issue that the misrepresentation objection stems from. To deal with that deeper difficulty, Coleman proposes an alternative theory that involves a notion of mental quotation. On that theory, the content of the state in virtue of which one is aware of a conscious state contains a quotation of that state, thereby “forming a larger composite structure wherein the first-order state is displayed” (—). In section II, I consider the related objection Coleman formulates and argue that it is without foundation, and in section III, I raise questions about whether mental quotation of the sort Coleman appeals to can be understood in any satisfactory way.

I conclude in section IV by discussing why the related objection Coleman formulates seems compelling. I argue that the objection is simply an expression of a first-order approach to consciousness, and that such first-order approaches are unable to provide any informative explanation in

¹ When not otherwise indicated, references are to Coleman’s article.

psychological terms of what it is for a mental state to be conscious. First-order approaches at best offer a descriptive phenomenology, but cannot, on their own, give any theoretical account of consciousness.

I. Misrepresentation by Consciousness

On higher-order theories of consciousness, a mental state's being conscious consists in one's being aware of oneself as being in the state in question. The basic motivating idea is that if one is in some mental state but wholly unaware of being in it, that state is not a conscious state. By contraposition, if a state is conscious, one is aware of it some suitable way.

There are various constraints one might impose on what that higher-order awareness consists in. On my own HOT theory, that higher-order awareness is a distinct state from the state that one is aware of oneself as being in. In addition, I argue that the higher-order awareness is a thought with intentional content, so as to explain the ability of creatures with suitable linguistic ability to report their mental states when those states are conscious. Reporting that one is in some mental state is expressing a HOT that one is in that state. And though inference might figure in how such HOTs sometimes arise, one would never be aware of any such inference; otherwise the awareness would not have the subjective appearance of being direct.

Many higher-order theorists share many of these conditions on the higher-order awareness, but not all. On the so-called same-order version (e.g., Kriegel 2009), the higher-order content in virtue of which one is aware of being in the target state belong to that target state itself. This raises a question about how we should individuate the relevant states (e.g., Phillips 2014), but I will not address that issue in this discussion.

Neither the distinct-state nor the same-order version, by itself, provides any means to preclude what we can call higher-order misrepresentation, on which an individual's higher-order awareness describes the target conscious state inaccurately. It may seem more obvious that this is so for the distinct-state version of the theory, which I'll sometimes call the standard version. If the higher-order awareness and target state are distinct occurrences, there can be no guarantee that the higher-order awareness will describe the target accurately.

But the issue arises equally for a same-order version. Since the higher-order content in virtue of which one is on a same-order theory aware of being in the target state is still a distinct mental property from the mental properties one is aware of the target state as having, there is again no guarantee that the higher-order content will be accurate in the way it describes those other mental properties.

Suppose the first-order state is a sensation of red and that there's a distinct higher-order state that makes one aware of oneself as having not a sensation of red, but a sensation of green (Neander 1998; Levine 2001). The higher-order state misrepresents the first-order sensation. Now suppose instead, as on a same-order theory, that there is a single state that has two mental properties: the qualitative character of a sensation of red, and a higher-order content that one has a green sensation. Despite there being only a single state, the higher-order content misrepresents that state. A same-order theory can do nothing, on its own, to preclude such higher-order misrepresentation.

This is crucial. The higher-order content in virtue of which one is aware of oneself as being in a particular mental state is a mental property that's distinct from the mental properties it represents the target state as having, and distinct from any mental properties it does have other than that higher-order content. It doesn't matter whether the mental properties all belong to a single state or the higher-order content belongs to a distinct state.

The point about distinct properties is indeed also independent of whether one casts things in higher-order terms at all. Consider a conscious sensation of red, and for argument's sake consider it to be a single state. The mental quality of red is a distinct mental property from the state's being conscious. This is evident because conscious sensations of red, green, and blue, effect, have in common the property of being conscious, but differ in that mental qualities they exhibit (Moore 1992: 22, 25).

And since the property of a state's being conscious is distinct from the qualitative mental properties it exhibits, it is unclear there can be any reason to deny that the property of a state's being consciousness could misrepresent the state's other mental properties. There may of course be some reason to insist that consciousness can misrepresent a state's

other mental properties. But if so, that reason must be independent of describing the issue in terms of a higher-order theory.

One might hope to contest that a state's being conscious is, after all, a distinct mental property from the state other mental properties. And it might seem that there is reason specifically in the case of qualitative mental states. Thus, one might contend that the property of being conscious is simply a determinable whose determinates are the various specific mental qualities. If that were so, the property of a state's being conscious would simply be a generic version of a state's qualitative mental properties.

But it is unlikely that this line can be sustained. For one thing, there are many conscious mental states that arguably lack qualitative character altogether, such as thoughts, desires, and other purely intentional states. If the property of a state's being conscious were a determinable of specific mental qualities, such nonqualitative mental states could not be conscious. An advocate of the determinable-determinate account might urge that the properties of a state's being conscious is different for qualitative and nonqualitative mental states. But there seems no reason to hold that other than a desire to sustain the determinable-determinate account.

But that account is questionable on grounds wholly independent of the current issue. The property of being a chemical element is a determinable of the property of being gold or of being oxygen, and the property of being an animal is a determinable of the property of being a horse or a dog. And in each case the determinable is a property distinct from the properties of its determinates.

The appeal to consciousness as a determinable may seem tempting if one denies that qualitative states can occur without being conscious. And one might seek to explain why such states couldn't occur without being conscious by construing the property of being conscious as a determinable of specific mental qualities. There is ample reason to hold that this denial is unfounded (Rosenthal 2005, chs. 5-7; 2010; Young et al 2014). But we can avoid this controversial topic here. Even if qualitative states couldn't occur without being conscious, that would not help show that being conscious is a determinable of the various mental qualities. We would need some independent reason for that.

Since being conscious is a distinct property from a state's other mental properties, we would need some special consideration if we want to preclude misrepresentation by consciousness. And it's not obvious what consideration would preclude that without simply being stipulative. Still, though higher-order theories do not on their own rule out such misrepresentation, they also do not imply that it can, or ever does, occur.

So if one did have suitable independent reason to think that it cannot or never does occur, one could simply add that onto a higher-order theory without any problem. What is important here is that adding such a provision onto a higher-order theory would doubtless not satisfy those who use misrepresentation by consciousness as an objection to such theories. The question is why not. Adding on such a provision, especially if it were based on a suitable independent reason, delivers they result the objectors ostensibly want. What more can they reasonably demand?

The quick answer is that these objectors regard no theory as satisfactory unless it rules out misrepresentation on its own, without any added provision. But that just pushes the question back one step. If there is independent reason to rule it out, why must a theory do so on its own, without appeal to any added provision?

It may help to consider the standard form of the objection, and the reply higher-order theorists give. Joseph Levine (2001, §4.4) casts things in terms of a case in which one has a sensation of red and a higher-order awareness of oneself as having a sensation of green. What, he asks, is it like for one in such a case? If what it's like is red, then the higher-order state is doing no work; if green, then the first-order sensation is doing no work.

The standard higher-order theorist's answer is plain. On a higher-order theory, the sensation is not, on its own, conscious at all. A sensation's being conscious is a matter of one's being aware of oneself as having a sensation of the relevant sort. So the higher-order state determines what it's like for one; the only role a first-order state could play in determining what it's like for one would be causal, if the first-order state caused a suitable higher-order awareness.

Though first-order sensation plays no noncausal role in determining what it's like for one, the sensation is not on that account mentally inert; there is

plenty for it to do. It figures in priming, in which one's being presented with a stimulus has various psychological effects downstream even when the perceiving is wholly subliminal, and so unconscious. It doubtless figures in Stroop effects, in which words for a color are harder to process if written in a different color. The sensation, independent of being conscious, would make a difference to reaction times for subsequent stimuli, and in very many other testable results. On a higher-order theory, the first-order state figures in psychological functioning other than consciousness, and the higher-order state constitutes subjective awareness.

Those who object to higher-order theories because they don't, by themselves, preclude misrepresentation by consciousness, are channeling the idea that motivates first-order theories of consciousness: that for qualitative mentality, the mental properties responsible for psychological functioning are the same as those that determine subjective awareness.

But the only reason ever offered for that insistence is that it seems that way subjectively. And that is not an acceptable reason. Subjective awareness cannot by hypothesis access a state that is not conscious on its own. So subjective awareness cannot determine what role is played by a state that is not conscious on its own. A reliance on subjective awareness begs the question against the standard, distinct-state version of higher-order theories. It is simply an expression of a preference for a first-order theory.

Things are somewhat more delicate for the same-order version of a higher-order theory, but the upshot is the same. The higher-order content on that theory does access all the mental properties of the state. But it is only that higher-order content that is responsible for what it's like for one, not the state's other mental properties. And though that is clear from the theory, it is not evident simply from one's subjective awareness.

One who advances the possibility of misrepresentation by consciousness as an objection to higher-order theories has no reason to complain that a higher-order theorist could meet the objection simply by stipulating that such misrepresentation cannot occur. But there is actually reason to hold that such misrepresentation does actually occur. And this reason is wholly independent both of higher-order theories and of the consideration that the property of being conscious is distinct from a state's other mental properties.

Coleman (n. 12) cites my appeal to so-called dental fear, in which dental patients subjectively take themselves to feel pain on drilling even though anesthesia or absence of any relevant nerve makes that impossible. When the patient learns that the sensation is likely instead just vibration plus fear, they subjectively experience when drilling resumes, though strikingly they also subjectively recall the earlier sensation as pain. (See Koyama et al 2005 and Anderson et al 1980 for some independent empirical substantiation.)

Dental fear is useful because it is a highly dramatic case in which our subjective awareness misrepresents what is actually going on in our mental life. Still, the case is exotic, and it would not be good if it were the main example of misrepresentation by consciousness. But it is not; as I have argued elsewhere, consciousness misrepresents our mental lives in plenty of everyday contexts.

When one glances at a group of people, one may see some shirts colored red and some colored blue. If one has simply taken in a crowd at a glance, what it's like for one is typically just seeing a generic red and a generic blue. It would be rare if just a glance led to what it's like for one being the seeing of a specific shade of red or blue. Despite that, the visual states in question are undoubtedly of whatever specific shades of red and blue result from the relevant stimuli. This is a form of misrepresentation; what it's like for one does not accurately represent the visual color states.

Ned Block (2011) disputes this. An "obvious problem" with this claim, he maintains, "is: why suppose that there can be an experience of red but not of any shade of red?" (445) But Block's rhetorical question equivocates on the term, 'experience'. We can use the term, 'expr' to refer either to the visual state on its own, or to what it's like for one. Block is doubtless right to doubt that there can be a visual state that is generically red; every visual state of red likely is of a particular shade. But that doesn't show that what it's like for one is not often quite generic.

Block would doubtless deny the equivocation, insisting that the visual state and what it's like for one are not distinct. But that is simply what is at issue, and without any reason to deny that the mental quality is distinct from the property of being conscious, there is no reason to accept his insistence.

Block offers a second objection, that if “there is no such thing as a generic experience as of colour, i.e. an experience as of colour but not any specific colour,” there couldn’t be a HOT of a generic experience (445). It sounds better to deny that there could be an experience that is generic in respect of color, say, red vs. blue, than to deny that an experience could be generic in respect of specific shades of, say, red or of blue.

But that aside, Block is still running together the visual state that figures in seeing a colored object with what it’s like for one to see that object. Block’s objections to cases of misrepresentation that derive from generic subjective awareness assumes that subjective awareness must match the nature of the visual state. And that assumption simply amounts to supposing question-beggingly that misrepresentation by consciousness cannot occur.

The case that involves generic subjective awareness of color shades plainly generalizes. But there is an abundance of other types of case. When pain results from the behavior of another that one perceives to be the intentional inflicting of a pain, the subjective experience is of a more intense pain than when pain results from the same stimulus but one does not perceive the behavior as an intentional inflicting of pain (Gray & Wegner 2008). Since identical stimuli presumably result in identical pains, psychological factors other than the pain itself must be affecting subjective awareness, which accordingly misrepresents the intensity of pain in one case or the other.

Misrepresentation by consciousness evidently does occur. And even if it didn’t or couldn’t, higher-order theories could readily accommodate an added provision to that effect. Why, then, does the objection seem so compelling to some? Coleman’s discussion helpfully points to how we might explain this.

II. Intimacy and Higher-Order Theories

Those who raise the objection from misrepresentation resist simply adding a provision that the higher-order awareness cannot misrepresent because they see patching the theory in that way as missing the difficulty that misrepresentation reveals. As opponents of the theory see it, the property of a state’s being conscious is not independent of the state’s other mental

properties in a way that allows for consciousness to misrepresent those other properties.

The possibility of misrepresentation strikes many as a telling objection to higher-order theories because they either reject the idea that the property of a state's being conscious is not distinct from the state's other mental properties or because they see a strong tie between those properties. So a theory on which the property of being conscious is independent of a state's other mental properties in a way that allows misrepresentation cannot, according to these objectors do justice to the nature of conscious states.

And patching the theory by providing that misrepresentation cannot occur would do nothing to redress the way theory misdescribes things according to these objectors. The objection from misrepresentation merely reveals, on this view, a symptom of the underlying difficulty. And that indeed is how Coleman sees the situation.

Some objectors doubtless believe that the property of being conscious is not distinct from a state's other mental properties; there is reason to think that this is Levine's view (2001) and Thomas Nagel's (1974). But there are powerful reasons rehearsed above for holding that the property of a state's being conscious is distinct from the state's other mental properties. And short of identity of being conscious with other mental properties, misrepresentation would also be precluded if there were some special tie between those properties that made it impossible.

And according to Coleman, such a special tie does obtain. Coleman uses Uriah Kriegel's term (2009, p. 108), 'intimacy', to refer to that special tie between a state's being conscious and the state's other mental properties. Indeed, Kriegel himself describes the issue about misrepresentation in terms of such intimacy. The objection about misrepresentation, he urges, is one way to call attention to the lack of intimacy between consciousness and other mental properties that standard higher-order theories imply.

Standard higher-order theories explain a mental state's being conscious as one's being aware of oneself as being in that state. And they explain that awareness, in turn, as a matter of one's being in a distinct mental state that represents one as being in the target state. The relevant intimacy between consciousness and the other mental properties consists, according to Kriegel, in there being a tie between subjectivity and the target mental state

that is closer in some significant way than can be captured by appeal to the relation between a representation and what it represents (107).

Kriegel's same-order theory, on which the higher-order content is a property of the same state that the higher-order content makes one aware of, is presumably motivated in part by the thought being props of the same state can capture this intimacy. But it is unclear how positing the higher-order content as a property of the target state can, by itself, capture any such intimacy. Even if the higher-order content is a property of the target state, it represents that state as having other mental properties. So the tie between a state's being conscious and the state's other mental properties will be representational, as it is if the higher-order content is a property of a distinct state.

Indeed, Kriegel acknowledges that the higher-order content could misrepresent the other mental properties of the relevant state. But he rightly notes that the higher-order content cannot occur in the absence of the state it is about, since the higher-order content is a property of that state. On the standard higher-order theory, by contrast, the higher-order content could occur in the absence of the state it is about, since that content belongs to a distinct state. But the difference here is trivial. On the same-order version of a higher-order theory, the higher-order content could still represent the state it is about as having other mental properties even if lacks all those other properties.

Both the same-order theory and standard higher-order theories that appeal to distinct higher-order states rely on higher-order content that represents one as being in a state with certain mental properties. That representation relation explains how it is we are aware of those of our mental states that are conscious. So neither theory posits a relation between consciousness and other mental properties that is more intimate than the relation posited by the other.

Coleman finds this unsatisfactory. His goal is a relation between a state's being conscious and other mental properties that is more intimate than the representation relation. He casts this dissatisfaction in terms of a distinction between one's being aware of something and one's merely being aware as of something. A representation relation, he urges, cannot capture one's being aware of a target mental state, but only one's being aware as of a state.

I find Coleman's his distinction between awareness as of and awareness of highly puzzling. Coleman uses vision as a model to explain the distinction. On one understanding of seeing, one sees something only if that thing is there to be seen. Similarly, Coleman urges, being aware of something requires that the thing be present for one to be aware of. By contrast, one can be aware as of something even if that thing isn't there at all.

I think Coleman is trading on a feature of the way we do actually speak and think about awareness, but that the way he construes that feature does not fit with higher-order theories in the way he thinks. If I am aware of a table, the table has to be there. So being aware of something has a factive aspect of the sort Coleman notes. But that factive aspect does not carry through to the properties of whatever it is one is aware of. I can be aware of a table as being brown even if it is green, or aware of it as being rectangular or tall even if it is oval in shape and short. Being aware of something is factive in respect of the object's being present, but not in respect of what properties that object has.

This works well for the HOT theory. On that theory, a state is conscious only if one is aware of oneself as being in that state. The object that one is thereby aware of is oneself, and that is conveniently guaranteed to be present for one to be aware of. One is, in addition, aware of oneself as being in some mental state, and being in that mental state is a property of oneself, that is, a property of the object one is aware of. And one can be in error about the properties of the object one is aware of, and hence whether one is in that state.

That representation relation is, moreover, perfectly suited to help capture being aware of. Take again, following Coleman, a visual case of awareness. If I see a table, my visual state represents a table as being present, though my visual state might also represent the table as tall when it is short or brown when it is green. The object I see might even be a box, though my visual state represents it as being a table. I won't count as being aware of the table if the table isn't actually there; nor on many accounts would I count even as seeing it at all if it isn't there. But I will count as being aware of it and as seeing it even if I misperceive some or even many of its visible properties.

Representation is of course weaker than being aware of or than seeing. I can represent a tree as being present even if it isn't, but I count as being aware of a tree or seeing it, on the construal of seeing under discussion, only if the tree actually is present. Having a thought about something is on a par in these respects with representing it. I can have a thought about something I take to exist, and it can go out of existence while I'm having that thought, without affecting the nature of my thought and, in particular, whether my thought is about that thing.

HOTs operate in exactly that way. But since the content of a HOT is that one is in the mental state in question, there is no issue about whether the item the HOT is about exists; the HOT is about the thinker of the thought, and that ensures its existence. This feature of HOTs ensures that they are suitable for the higher-order awareness that figures in a mental state's being conscious. Having a thought about something constitutes being aware of that thing only if one thinks about the thing as being present; having a thought about Julius Caesar or Saturn does not make one aware of those things. But when one has a thought that one is in a particular mental state, one automatically thinks of oneself as being present.

Coleman's contrast between being aware of and merely being aware as of misconstrues being aware of as factive not just in respect of the presence of the relevant object, but in respect of accuracy about that object's properties. Once we take into account that distinction and what the content of a HOT is, the theory does capture the way one is aware of those of one's mental states that are conscious. One is aware of oneself as being in the target state. The intimacy revealed by the relation of being aware of is intimacy between the individual who is aware of something and the object of which that individual is aware.

III. Mental Quotation

But this will not satisfy Coleman, who is after a kind of intimacy between consciousness and a mental state's other mental properties that the representation relation cannot capture. And he seeks to establish the kind of intimacy he finds lacking in any representation relation by appeal to a device of mental quotation. Putting aside whether this quest for intimacy is well-founded, I have independent doubts about this strategy.

Coleman likens the mental quotation he invokes to the mental quotation that advocates of so-called phenomenal concepts some appeal to. Phenomenal concepts are cast as concepts that apply to conscious qualitative properties, and whose mode of presentation for each concept is the very conscious mental quality that concept applies to (Balog 2012; Papineau 2002, 2007). Coleman's proposal is that a qualitative state is conscious if the higher-order content in virtue of which one is aware of being in that state contains a quotation of the mental quality in virtue of which the state is conscious.

But it is at all obvious what such mental quotation could consist in. Quotation is a linguistic device. So we need a linguistic model of quotation by appeal to which to understand mental quotation. There are a number of different types of quotation and various accounts of those types. (For a survey of the issues, see, e.g., Lepore 1999; Cappelen and Lepore 1997 and 2012.) But it's not clear that any of those types offer a suitable model for mental quotation that would serve Colman's purposes.

For a mental quotation of a qualitative property to serve his purposes, the qualitative property must figure in the higher-order content in which the mental quotation occurs. So the type of quotation typically used to capture the distinction between the use and mention of linguistic expressions won't do, since on standard accounts the quotation itself is syntactically atomic. The quotation simply mentions the linguistic expression, and that expression itself play no semantic role in the quotation. By analogy, if a mental quotation of a qualitative property occurred in a higher-order content, the qualitative property on that model of quotation would play no role in that higher-order content.

Perhaps ordinary direct-discourse quotation provides a more promising model, since in direct discourse the quotation of a linguistic expression is taken not just to represent the exact words, but also to exhibit the words as they are ordinarily used. But it is problematic how direct-discourse quotation plays that dual role. Simply stipulating that it does won't help, since we need a linguistic model to help understand how the relevant mental quotation is supposed to work. It is no clearer how mental quotation would both refer to and exhibit a qualitative property. And without understanding how the mental quotation of the sort Coleman appeals to could work, the appeal simply stipulates that an intimacy relation obtains, rather than explaining what such intimacy consists in.

Coleman presumably means for mental quotation to do two things. The mental quotation that occurs in any particular higher-order content must refer to the type of mental quality an instance of which is being quoted. Otherwise, the higher-order content won't make one aware of oneself as being in a state of the relevant type.

But when a mental quotation of a qualitative property occurs in a higher-order content, that qualitative property must also figure in its own right. Otherwise there is no intimacy of being conscious with a state's other mental properties beyond that which occurs in connection with the representation relation. It is the goal of having mental quotation do both of these things that makes the direct-discourse model of linguistic quotation appealing. But without an explanation of how direct discourse works that can transfer to the case of mental quotation, we are left without any informative account of the intimacy Coleman is after.

Coleman likens the mental quotation he wants to invoke to the notion of a display sentence that Kriegel describes. Kriegel's example is of seeing the words 'under construction' painted on the roadway of a bridge, and thinking as a result that this bridge is under construction. "[T]he combination of the words and the bridge whereupon they are painted that vehicles the full-blown proposition" (2009, p. 163). The vehicle of the proposition, Kriegel urges, is the bridge itself together with the words painted on its roadway. If either the words or the bridge were missing, or they were not physically related in the right way, the relevant content would not be tokened.

In Kriegel's case, the bridge as it occurs in the composite vehicle of the overall content refers to the bridge itself. But it's not clear on that account that we should describe this as a case of quoting the bridge. The overall content refers to the bridge, and it is physical contiguity of the words with the bridge that enables them all to constitute a vehicle for that content.

So working with Kriegel's display-sentence model suggests a model on which the qualitative property is somehow spliced into the intentional content of the higher-order awareness, not a model on which a quotation of the qualitative property occurs in that higher-order content. The higher-order intentional content would have a referential space that can be filled in by the mental quality itself.

Again, the bridge in Kriegel's example does not refer to itself; what refers to the bridge is the overall content of which the bridge is merely part of the vehicle. Similarly, when the higher-order content is completed by splicing in a qualitative property, it is that completed intentional content that would, on this model, refer to qualitative properties of the type spliced in. It is the splicing in of a token qualitative property that is supposed to yield the intimacy Coleman is after between consciousness and a state's other mental properties.

But the overall higher-order intentional content would, on this model, represent qualitative properties of the relevant type. Otherwise, the higher-order content could not make one aware of being in that type of qualitative state. So it is not right to describe Coleman's proposal, as he does, as "abandoning representation as the means by which the higher-order thought targets the first-order state" (—; emphasis Coleman's). Whether what's operative is mental splicing or mental quotation, unless the overall higher-order content represents qualitative properties of the relevant sort, such content cannot make one aware of oneself as being in a state with those qualitative properties. Neither quotation nor splicing can obviate the need for a representation relation.

It is not obvious that there is an informative account of a kind of mental quotation that would do the job Coleman sets for it. But splicing is hardly any better on that score. There is no informative or credible account of intentional content whose referential spaces could be filled in by mental qualities. Nor is it obvious what the relation could be between the mental quality that would fill in that referential space in the higher-order intentional content the mental quality of the target state. If the quality that figures in the higher-order content is a distinct token from that of the target qualitative state, some mechanism must be stipulated to ensure that the two tokens are of the exact same type; otherwise intimacy is not achieved. And if the same token figures in both target state and higher-order content, we need an explanation of how a single token quality could figure in both.

Mental quotation might seem to get around this last difficulty; the mental quoting mechanism, one might imagine, would simply reach from higher-order content down to the target state and quote that state's mental quality. But this raises questions for either a quotational or a splicing model. Quotation operates on a token of what is quoted, and if the token is a

property of a distinct state, machinery would have to be posited of a wholly unclear type.

Nor can it help to have the higher-order content and the quoted mental quality be properties of a single state. Since they are distinct properties, it is unclear how quotation could reach from one to the other. For any of this to be more than a mere stipulation of the intimacy Coleman is after, we need some independent understanding of how the mental quotation or splicing works, independent, that is, of the stipulation that that's what happens in these cases.

The questions posed about mental quotation and about mental splicing may have perfectly good answers. But to make good on the intimacy Coleman wants, we need to have those answers. And the answers cannot rest simply on the ability to generate the desired intimacy; we must have reason independent of that to think that the answers fit with a theoretically sound picture of how intentional content and mental qualitative character work. That is the heavy theoretical burden that a mental splicing or mental quotation model must discharge.

A quotational or splicing model may seem appealing to one who is already committed to some type of intimacy between the property of a state's being conscious and the mental properties in respect of which one is aware of that state. But it is unclear what the reason for that commitment would be. Coleman's distinction between being aware of and aware as of does not underwrite a commitment to intimacy. And without reason for holding that such intimacy occurs, it is unclear that we should hold a theory of consciousness hostage to the insistence that intimacy occurs, nor to the significant theoretical burden of explaining in an informative way what it is.

In the final section, I'll ask what the apparent appeal is of holding that such intimacy does occur. I'll argue that it seems inviting only if one tacitly adopts a central aspect of first-order approaches to thinking about consciousness. If so, the quest for intimacy and the objection from misrepresentation both in effect beg the question against higher-order theories of consciousness by taking for granted an assumption that underlies first-order approaches.

IV. Intimacy and Mental Appearance

Higher-order theories rely on two things. One is the recognition that the property of a mental state's being conscious is distinct from those other mental properties in respect of which a mental state is conscious. When one consciously sees a red square, the property of that state's being conscious is distinct from the state's mental qualities of color and shape. It is plain that this is so if such qualitative states can occur without being conscious. But as argued in section I, there are reasons independent of that to hold that these properties are distinct.

The other thing higher-order theories rely on is that if we have independent reason to think somebody is in some mental state, a reliable indicator that the state is not conscious is that the individual is not in a way aware of being in that state. If one is not aware of oneself as being in a mental state, that state is not conscious. And as noted at the outset of section I, contraposition delivers the conclusion that if a state is conscious, one is in some way aware of being in it. The higher-order state is posited to explain that awareness.

Coleman sees HOT theory as failing to capture the desired intimacy because of a shortcoming in the way it describes that higher-order awareness; the appeal to mental quotation is intended to remedy that. It is unclear that that device or a related device of mental splicing can be understood in a way that would help. But our current concern is instead about what the intimacy Coleman and others are after consists in, and why we should want or expect a theory of consciousness to accommodate it.

Coleman points in several passages to his answer. "The fabric of a Rosenthalian HOT," he writes, "is simply too 'thick' for first-order targets ever to penetrate to the subject's stream of consciousness" (—). "[S]ince the relevant appearances supplied to the stream of consciousness remain intact with the first-order state removed, it follows that the latter was contributing no appearances to subjective mental life even while it was present" (—). What Coleman urges is, instead, an "an account on which first-order contents are not screened off from consciousness by the contents of higher-order representations" (—).

Higher-order theories proceed by divide and conquer (Weisberg 2011). Since the property of a state's being conscious is distinct from its other

mental properties, we can posit a higher-order state that captures what it's like for one, and hence the property of consciousness. The first-order state, then, exhibits the other relevant mental properties. In the case just mentioned, the first-order state exhibits the qualitative character of redness and squareness, but independently of those mental qualities' being conscious. The property of being conscious results from the higher-order awareness of one's being in a state with those properties.

Even if one assigns all those mental properties to a single state, as on Kriegel's same-order theory, the distinct types of property operate independently. As Colman notes, assigning all the properties to a single state does not by itself achieve the intimacy he is after. At most it simply generates an appearance of intimacy. And by creating an appearance without the reality, it arguably does a measure of theoretical damage for one concerned to capture some type of intimacy.

If there are two distinct mental states, one exhibiting the property of being conscious and the other exhibiting the other mental properties, the higher-order state will bear the burden of there being something it's like for one to be in the relevant type of state and the first-order state will be responsible for the other mental properties of the relevant state. The desire for intimacy, as Coleman sees it, is for the first-order mental properties to play some suitable role in the state's being conscious. We want, he urges, "an account on which first-order contents are not screened off from consciousness by the contents of higher-order representations."

Coleman's mental quotation, or a related device of mental splicing, would undermine the recognition that the property of a state's being conscious is distinct from its other mental properties. Those devices seek to import into the higher-order content a token of the mental quality that characterizes the first-order state. In that way they go a step further than Kriegel's same-order theory. That theory assigns both higher-order content and the other mental properties to a single state; Coleman actually imports first-order mental properties into the higher-order content.

Can we prevent the screening off Coleman objects to while still accompanying the distinctness of the property of being conscious from other mental properties? A natural way to do so is to assume that in standard cases the first-order properties have a causal impact on which consciousness property is instantiated. Cast in higher-order terms, the

properties of the first-order state, such as qualitative red, will affect the content of the higher-order awareness that one is in a state of a particular type. In standard cases such a first-order property will cause one to have a higher-order awareness of being in a state of seeing red.

And something along these lines must occur. It cannot be an accident that we typically are aware of ourselves as seeing something red and square when presented with a stimulus that we can expect gives rise to a first-order visual state of seeing a red square. A robust causal tie between first-order and higher-order states must generally hold.

But Colman evidently finds a robust causal tie wanting, and demands instead a constitutive tie. It is unclear why. What reason is there to think that the intimate connection between first-order mental properties and the property of a state's being conscious consists in something more than a strong causal tie? And the constitutive tie Coleman wants plainly flies in the face of the evident distinctness of the property of being conscious from a state's other mental properties. If one imports first-order properties into the higher-order content that constitutes consciousness, the property of being conscious is no longer distinct from the first-order properties.

Coleman's distinction between being aware of and being aware as of cannot help establish a need for the constitutive intimacy he proposes. For one thing, being aware of is not different from being as of, on Coleman's construal, except for the guarantee that the thing one is aware of is present, and as Coleman sees things the additional guarantee that it has the properties one is aware of it as having. Put aside whether the provision about properties corresponds to our ordinary understanding of what being aware of something consists in. Even granting Coleman both guarantees, there is no need to adjust the nature of the higher-order awareness. The guarantees can simply be added provisions about the relation between higher-order and first-order states, leaving their mental properties intact.

One way to see this is to stress one of the challenges raised in section III about mental quotation or mental splicing. Suppose we understand how those devices operate, and a token of a mental quality is quoted in or spliced into the content of a higher-order awareness. What guarantee do we have that the token quoted or spliced is of the exact same type as the token that characterizes the first-order state?

And as noted in section III, we cannot say without additional explanation that the higher-order content quotes or splices the first-order token itself. Quotation and splicing both operate on a distinct token, and the first-order token is a distinct mental property from the higher-order content. Importing the first-order token, whether by splicing or quotation, into the higher-order content leaves open the question of whether that importation results in two tokens of the exact same mental type, the intimacy Coleman is after will not result.

On a first-order approach to consciousness (Block 1995, 2011; Nagel 1974; Levine 2001; Kripke 1982, pp. 144-155), the property of being conscious is not, after all, distinct from the qualitative mental properties of a qualitative state. The property of being conscious is built into those qualitative properties; being conscious is an aspect of any state's qualitative properties. I've argued in section I that this assimilation of properties is mistaken. I'll return in a moment to the apparent reasons the first-order approach is committed to that assimilation.

Coleman's approach goes somewhat in the opposite direction, but to much the same effect. Rather than insist that the property of being conscious is built into qualitative mental properties, Coleman holds that mental properties other than that of being conscious must figure somehow in the higher-order awareness in virtue of which a state is conscious. And because the way those other mental properties figure in the higher-order awareness is not merely causally, but in some way constitutive, the result is much the same as the assimilation of those properties by a first-order approach.

So Coleman's requiring a noncausal, constitutive intimacy of this type rests on the tacit endorsement of a central tenet of first-order approaches to consciousness. On a first-order picture, a state's being conscious is of a piece with a state's other mental properties. The nature of those mental properties is, on that approach, revealed by consciousness, so that the mental properties cannot occur without being conscious. Most first-order approaches of this type apply only to states with mental qualitative character (Nagel 1974; Block 1995; Levine 2001), though John R. Searle (1990) argues for that approach in connection with intentionality.

Coleman's demand for an intimacy that rests on the occurrence of first-order mental properties in the higher-order content itself is accordingly of a

piece with the way a first-order approach sees the property of being conscious as built into first-order mental properties. So the demand for an intimacy of that type, as against a robust causal intimacy, in effect channels the picture that the first-order approach counterposes against any higher-order theory. And without independent substantiation that demand simply begs the question at issue between higher-order and first-order approaches.

That Coleman's insistence on noncausal intimacy rests on a first-order approach to consciousness is already evident in a passage quoted above: "[S]ince the relevant appearances supplied to the stream of consciousness remain intact with the first-order state removed, it follows that the latter was contributing no appearances to subjective mental life even while it was present" (—).

Coleman sees the first-order state as itself having mental appearances that contribute to one's subjective mental life. That would be so on a first-order picture. And then a state's other mental properties must in some way be inextricably and constitutively bound up with the state's being conscious. The point of a higher-order theory is to enable us to treat the property of a state's being conscious independently of those other mental properties.

Coleman doesn't see it that way. On a higher-order theory, he writes, "all that a first-order state contributes to the stream of consciousness is a proxy that represents it. ... It does not appear in consciousness." "Therefore, on HOT theory, we are simply not conscious of our first-order states" (—; emphasis Coleman's).

This is puzzling. When I am aware of a physical object, the object does not itself figure in any mental state of mine. All that figures is some mental proxy that represents that object. Being aware of something does not require that the object one is aware of actually occur in the state of awareness, as against being represented by that state. This applies to being aware of one's own mental states no less than of physical objects. The demand that the first-order state appear in one's awareness of it reflects the first-order elision of the property of being conscious with a state's other mental properties.

A state's being conscious is a matter of mental appearance, of how one's mental goings on mentally appear to one. It is natural to assume that how

one's mental goings on actually are will play an outsized causal role in determining how those goings on appear to one mentally. What is going on in one's mind independently of consciousness will play a very substantial causal role in determining one's stream of consciousness, in determining, that is, how one's mental life consciously appears to one.

One might even hypothesize that one's mental goings on independently of consciousness would completely determine how those mental goings on consciously appear to one. If so, consciousness would never misrepresent those mental goings on. But there are reasons, such as those surveyed toward the end of section I above, to think this is not so. These reasons may not be decisive, but they are at the very least strongly suggestive.

But the core issue is whether a theory of consciousness should distinguish the property of a state's being conscious from that state's other mental properties. Confining attention here to a state's qualitative properties, what reason could there be not to distinguish the two types of property? The core reason seems to be that we learn the nature of mental qualities by way of their being conscious; any third-person understanding of mental qualities is indirect.

That reason is not compelling. When you speak of my qualitative states, say, my being in pain or seeing red, you are not talking about something accessible only by way of my first-person access. I am talking about something whose nature we both have an intersubjective grasp of. Only if that is so could we both talk about my conscious qualitative states. An exclusively first-person account of the nature of mental qualities cannot do justice to that.

And there is an alternative to an exclusively first-person account, an account of mental qualities that relies on the role they play in perceiving (Rosenthal 2005, chs. 5-7; 2010; Young et al 2014). Taking perceptual role into consideration underwrites the intersubjective grasp of mental qualities that figure in our ordinary understanding of their nature.

Still, there seems to be a temptation, compelling to many, to deny that there is after all any distinction to be drawn between the way our mental goings on actually are and the way they appear. As Nagel puts it, when it comes to conscious experience, "[t]he idea of moving from appearance to reality seems to make no sense" (1974, p. 444). And if that distinction

makes no sense, there may be no reason to distinguish the property of a state's being conscious from its other mental properties, and perhaps indeed no coherent way to do so.

Denying a distinction between mental reality and how it appears allows us to rely exclusively on the subjective aspect of mentality. But to rely on subjectivity is simply to rely on mental appearance, and that prevents us from even addressing the question of whether there is a mental reality distinct from the way it appears. We need a grasp of mental reality that is independent of subjectivity even to raise that question. For mental qualities, an appeal to perceptual role provides that grasp.

Coleman denies that his approach rests on denying a distinction between mental reality and the way it subjectively appears. His "theory is not motivated by the claim that mental appearance and reality cannot diverge," but "by the desirability of an account on which first-order contents are not screened off from consciousness by the contents of higher-order representations" (—).

But Coleman builds too much into his notion of being screened off from. First-order mental properties are not screened off causally from higher-order content on a higher-order theory; it is simply that the theory countenances that first-order properties are distinct from the higher-order content that constitutes subjective awareness. Demanding that first-order properties actually occur in the higher-order content, as mental quotation or splicing is intended to provide, is following a first-order approach in eliding the distinction between first-order mental properties and subjectivity.

The phenomenological appearances that constitute subjectivity are of great importance for fixing and describing the mental phenomena we want to explain, but phenomenology does not by itself explain anything. It simply consists of mental appearances to focus on; its job is to fix things descriptively, not to explain.

So a first-order approach to consciousness, which elides the distinction between our mental goings on and the subjective appearances we have of these goings on, is not friendly to theorizing. It instead rests largely or entirely with the kind of mental description that phenomenology affords.

This emerges in Block's colorful appeal, in saying what qualitative mental states are, to Louis Armstrong's remark about jazz: "If you gotta ask, you ain't never gonna get to know" (Block 1978, 281). There is on that first-order picture nothing informative to say about the nature of mental qualities, and so nothing on which to base theorizing of any sort.

This also emerges in Nagel's insistence that "experience does not have, in addition to its subjective character, an objective nature that can be apprehended from many different points of view" (443-4). On a first-order view, we cannot theorize about conscious experience; we can know about it only by way of consciousness. And ironically, consciousness seems on this first-order approach not to have anything articulate to tell us about the nature of experience; it leaves experience ineffable, and so not susceptible to theoretical treatment at all.

Block has argued that his approach does, after all, accommodate a theoretical treatment of mental qualitative character, since we determine in a theoretical way what the neural concomitants are of qualitative mental properties (2007). But the contrast between theorizing and mere phenomenological description pertains not to neural concomitants, but to the nature of the qualitative mental states in their own right. And Block has offered no informative account of that since his Louis Armstrong appeal. It is the nature of qualitative mentality itself that first-order approaches offer no theoretical explanation of, but at most just phenomenological description.

So the objection from misrepresentation and the insistence that any suitable theory accommodates the constitutive intimacy of the sort Coleman describes, arguably both rest on an atheoretical, first-order approach to consciousness. Nagel is upfront about this; conscious "experience does not have, in addition to its subjective character, an objective nature that can be apprehended from many different points of view" (1974, 443-4). The dependence of subjectivity on points of view in effect drives out objectivity, and with it any possibility of a theoretical treatment.

The intuitive sense that consciousness cannot misrepresent and the demand that a state's being conscious interact constitutively with the state's other mental properties both reflect the way a first-order approach ties consciousness to those other mental properties. So the issue about

misrepresentation and the demand for the kind of intimacy Coleman is after both in some measure channel a kind of resistance to theorizing at all about the nature of consciousness.

REFERENCES

Anderson, D. Bryant, and James W. Pennebaker (1980), "Pain and Pleasure: Alternative Interpretations for Identical Stimulation," European Journal of Social Psychology, 10, 2 (April-June): 207-212.

Balog, Katalin (2012), "Acquaintance and the Mind-Body Problem," in New Perspectives on Type Identity: The Mental and the Physical, ed. Simone Gozzano and Christopher S. Hill, Cambridge: Cambridge University Press, pp. 16-42.

Block, Ned (2011), "Response to Rosenthal and Weisberg," Analysis 71, 3 (July): 443-448.

Block, Ned (2007), "Consciousness, Accessibility, and the Mesh between Psychology and Neuroscience," Behavioral and Brain Sciences, 30, 4 (August): 481-499.

Block, Ned (1995), "On a Confusion about a Function of Consciousness," Behavioral and Brain Sciences, 18, 2 (June): 227-287.

Block, Ned (1978), "Troubles with Functionalism," in Minnesota Studies in the Philosophy of Science, IX, ed. C. Wade Savage, Minneapolis: University of Minnesota Press, 1978, pp. 261-325.

Cappelen, Herman, and Ernest Lepore (2012), "Quotation," Stanford Encyclopedia of Philosophy, <https://plato.stanford.edu/entries/quotation/>

Cappelen, Herman, and Ernie Lepore (1997), "Varieties of Quotation," Mind 106, 423 (July): 429-450.

Coleman, Sam (2019), "The Merits of Higher-Order Thought Theories," Trans/Form/Ação, —.

Gray, Kurt, and Daniel M. Wegner (2008), "The Sting of Intentional Pain," Psychological Science 19, 12 (December): 1260-1262.

Koyama, Tetsuo, John G. McHaffie, Paul J. Laurienti, and Robert C. Coghill (2005), "The Subjective Experience of Pain: Where Expectations Become

Reality,” Proceedings of the National Academy of Sciences 102, 36 (September 6): 12950–12955.

Kripke, Saul (1982), Naming and Necessity, Cambridge, Massachusetts: Harvard University Press (very slightly revised from “Naming and Necessity,” in Semantics of Natural Language, ed. Gilbert Harman and Donald Davidson, Dordrecht and Boston: D. Reidel Publishing Company, 1972, pp. 253-355 and 763-769).

Kriegel, Uriah (2009), Subjective Consciousness, Oxford: Oxford University Press.

Lepore, Ernest (1999), “The Scope and Limits of Quotation,” in The Philosophy of Donald Davidson, ed. Lewis Edwin Hahn, Peru, Illinois: Open Court Publishing Company, pp. 691-714.

Levine, Joseph (2001), Purple Haze: The Puzzle of Consciousness, New York: Oxford University Press.

Moore, G. E. (1922). “The Refutation of Idealism,” in G. E. Moore, Philosophical Studies, London: Routledge & Kegan Paul Ltd, pp. 1-30.

Nagel, Thomas (1974), “What Is It Like to Be a Bat?,” The Philosophical Review 83, 4 (October): 435-450.

Neander, Karen (1998), “The Division of Phenomenal Labor: A Problem for Representational Theories of Consciousness,” Philosophical Perspectives 12, 411-434.

Papineau, David (2007), “Phenomenal and Perceptual Concepts,” in Phenomenal Concepts and Phenomenal Knowledge: New Essays on Consciousness and Physicalism, ed. Torin Alter and Sven Walter, Oxford: Oxford University Press, pp. 111-144.

Papineau, David (2002), Thinking about Consciousness, Oxford: Oxford University Press, §4.8.

Phillips, Ben (2014), "Indirect Representation and the Self-Representational Theory of Consciousness," Philosophical Studies, 167, 2 (January): 273-290.

Rosenthal, David (2012), "Higher-Order Awareness, Misrepresentation, and Function," Philosophical Transactions of the Royal Society B: Biological Sciences, 367, 1594 (May 19): 1424-1438.

Rosenthal, David (2011), "Exaggerated Reports: Reply to Block," Analysis 71, 3 (July): 431-437.

Rosenthal, David (2010), "How to Think about Mental Qualities," Philosophical Issues: Philosophy of Mind, 20 (October): 368-393.

Rosenthal, David (2005), Consciousness and Mind, Oxford: Clarendon Press.

Searle, John R. (1990), "Consciousness, Explanatory Inversion, and Cognitive Science," Behavioral and Brain Sciences, 13, 4 (December): 585-696,

Weisberg, Josh (2011), "Abusing the Notion of What-it's-like-ness: A Response to Block," Analysis 71, 3 (July): 438-443.

Young, Benjamin D., Andreas Keller, and David Rosenthal (2014), "Quality Space Theory in Olfaction," Frontiers in Psychology (January), <https://www.frontiersin.org/articles/10.3389/fpsyg.2014.00001/full>