

To appear in Qualitative Consciousness and Higher-order Thought, ed. Josh Weisberg, Cambridge University Press, August 2022

## **MENTAL APPEARANCE AND MENTAL REALITY**

David Rosenthal

Philosophy and Cognitive Science

Graduate Center, City University of New York

**Abstract:** Conscious states all have some subjective appearance. But do conscious states also have some mental reality apart from that subjective appearance? How one answers reflects a fundamental divide in thinking about consciousness, and affects many issues. I'll argue that there are strong reasons to think conscious states do have mental reality distinct from their subjective appearances, and no good reason to think otherwise.

### **I. Appearance and Reality**

Consciousness is the way our mental lives appear to us subjectively. The stream of consciousness is a stream of mental appearances of perceptions, feelings, desires, and thoughts that one seems subjectively to have. Consciousness is the subjective appearance of being in those states. Take away that subjective appearance and there is no consciousness.

Because consciousness is the subjective appearance of having various perceptions, - thoughts, and feelings, those perceptions, thoughts, and feelings are the mental reality that corresponds to the appearances of consciousness. They are the mental states that one subjectively appears to be in.

Despite these straightforward considerations, some deny that in the case of consciousness we can distinguish between appearance and reality. The subjective appearances, they insist, are all the mental reality there is for conscious states; there is no additional mental reality to distinguish from those subjective appearances. The mental reality of conscious states is exhausted by the way those states are for consciousness. Thus Thomas Nagel: "The idea of moving from appearance to reality seems to make no sense" in connection with conscious experience (1974, p. 444).

I'll argue in what follows that there are substantial explanatory and theoretical disadvantages to holding that the subjective appearance do exhaust mental reality in that way, and significant explanatory benefits to rejecting that claim. But even aside from such consequences, it's unclear why one would think that the subjective appearances do exhaust the mental reality of conscious states.

One might urge that consciousness is simply pure appearance, and that pure appearance stands alone, independent of any corresponding reality. But appearance is normally an appearance of something. So we would need some compelling reason to deny that consciousness also works that way, especially since it seems to do so when taken at face value. In our conscious mental lives we subjectively appear to be in various mental states. So it's natural to take the mental states we subjectively appear to be in to constitute the mental reality of the appearances of consciousness.

It's sometimes held that the conscious states one appears subjectively to be in cannot occur without being conscious, or that conscious states are intrinsically conscious. There is ample reason to doubt both claims; more in what follows. But since neither claim implies that the subjective appearances of conscious states exhausts their mental reality, neither can by itself support that view.

First-person access tells us what it's like for one to be in various conscious states. So if first-person access revealed the entire mental nature of conscious states, perhaps they would have no mental reality beyond their subjective appearances. But the only reason to think that first-person access does reveal the entire mental nature of conscious states would be that subjective appearance exhausts their mental reality. So the appeal to first-person access cannot provide independent support for that view.

Some find compelling a picture of first-person access as direct, unmediated acquaintance. And advocates of that picture might urge that unmediated acquaintance does deliver exhaustive information about the states we're acquainted with. But it's not obvious why that would be so. Acquaintance doesn't ordinarily reveal the entire nature of anything. It would do so for conscious states only if their subjective appearances exhausted their mental reality.

It might be urged that acquaintance could be unmediated only if conscious states are transparent in a way that does allow acquaintance with their complete natures. But it's unclear that our acquaintance with conscious states is actually unmediated. It appears subjectively to be unmediated, since consciousness reveals no mediating mechanism or process. But the failure of consciousness to reveal a mechanism or process wouldn't show there isn't any unless the mental reality of conscious states is exhausted by what subjectively appears. So neither first-person access on its own nor the unmediated-acquaintance picture provides independent support for the claim that subjective appearance does exhaust mental reality.

Whatever mental properties conscious states have in addition to their subjective appearances, conscious states also have various nonmental aspects. Such states are neurally implemented and have neural ties to other states of various sorts. But for the purposes of a distinction between appearance and reality, the reality that matters is distinctively mental, described in mental or psychological terms.

Distinguishing the appearance of conscious states from their mental reality raises a question about whether those mental appearances are always accurate about the

corresponding mental reality. In ordinary, nonmental cases, the appearances of things are not always accurate. If that's holds also for the subjective appearances of conscious states, consciousness might sometimes misrepresent the corresponding mental reality.

Some are convinced that such misrepresentation by consciousness cannot happen. And if there were no mental reality distinct from the subjective appearances, there would be no distinct mental reality for the appearances to misrepresent. So those who deny the possibility of such misrepresentation might seek preemptively to close off that question by rejecting from the outset any distinction between appearance and reality.

I'll return to misrepresentation by consciousness in §III. But for now it's worth noting that it's an overreaction to block such misrepresentation by denying any distinction between appearance and reality for conscious states. One can readily retain that distinction whatever one holds about misrepresentation. So closing off misrepresentation is not a sound reason to reject the distinction.

### **1a. Intuitions**

There are few if any actual arguments that we cannot distinguish the appearance of conscious states from their reality, or that the subjective appearances of conscious states exhaust their mental reality. Rather, most who hold those things don't appeal to argument at all, but simply to an intuition that subjective appearance is all there is to conscious experience. Thus François Kammerer: "[W]e have a strong intuition that there is no appearance/reality distinction in the case of phenomenal consciousness" (2018, p. 7).

Appeals to intuition are widespread in the contemporary literature, perhaps nowhere more than in connection with consciousness. And it's sometimes held that such intuitions are not merely reliable, but something like the last word about things. As Saul Kripke puts it: "I think [having intuitive content] is very heavy evidence in favor of anything, myself. I really don't know, in a way, what more conclusive evidence one can have about anything, ultimately speaking" (1980, p. 42).

But there are serious concerns about such appeals to intuition. For one thing, why would whatever intuitions we have be accurate? What could it be about our thought processes or any other aspect of our psychological functioning that could deliver such remarkable reliability? The Cartesian idea that our minds are so constituted as to be right about basic issues has not fared all that well. What else could there be?

And intuitions are rarely universally shared; intuitions seen as compelling by some are often rejected by others. It is even doubtful that such intuitions occur outside the relevant academic circles. One might contend that even when such intuitions don't occur spontaneously, they "can be teased out of ordinary subjects" (Chalmers 2018, p. 13). But which intuitions can be teased out will likely depend on who is doing the teasing, and how.

Intuitions must be represented as being wholly independent of theorizing. If an intuition relied on theorizing, we would have to evaluate that theorizing before we could credit the intuition. And methodological issues can in any case arise about how to assess competing views when one is supported by intuitions and another rests instead on straightforward theoretical considerations.

One might urge that intuitions have a pretheoretic status because they specify the nature of the phenomena to be discussed, and that must be done prior to theorizing. But theorizing often leads us to revise what we pretheoretically took to be the nature of a phenomenon. Since we cannot insulate what we think a phenomenon is from our theorizing about it, that account can't give intuitions a theory-independent status.

And pretheoretic claims that are intuitively inviting turn out often enough to conflict with scientific findings, and then science overrides the intuitions. Such conflicts arise even with pretheoretic beliefs about conscious experience. It's intuitively compelling, for example, that conscious parafoveal vision is rich in detail and color, but it isn't (Knotts et al 2019; Cohen et al, 2020).

If intuitions were truly independent of theorizing, there would be no way to adjudicate conflicting intuitions. And without a way to settle such conflicts, intuitions can come to be treated not as claims, but as data to which any further discussion must conform. It will then seem out of place to contest an intuition, so that appeals to intuition cut off debate. These are unfortunate consequences of seeing intuitions as independent of theorizing.

That attitude about intuitions may seem inevitable as long as we lack any account of their source. But there is a compelling hypothesis about their origin, which fits well with the foregoing observations. Though intuitions are represented as independent of theory, there is invariably a striking correlation between the intuitions somebody holds and that person's theoretical approach to the relevant phenomena. One can reliably predict somebody's theoretical position, at least in general terms, from the intuitions they hold, and conversely.

That close fit between intuitions and theoretical views points to an explanation of what intuitions are, how they arise, and why some find them decisive. Intuitions are appealing encapsulations of theoretical positions, packaged in ways that seem to skirt the need for independent support or assessment.

This explains why people have conflicting intuitions, since each person's intuitions reflects the theoretical position that the person finds inviting. And it explains why intuitions rarely occur spontaneously outside of contexts of theoretical debate. And because intuitions are disguised to look pretheoretic so as to sidestep any need for independent support, proponents tend to have an unwavering conviction in them.

But intuitions do channel theoretical views, and those views plainly require independent evaluation. Adapting Daniel Dennett's useful notion of an intuition pump (1991, pp. 282, 397), we can think of intuitions as theory pumps: one-liners designed to get others to adopt some theoretical position.

If intuitions were genuinely pretheoretic, we would have to demand some serious account of their origin before taking them seriously. Otherwise it would be just one person's undefended if strongly held hunch against another's. But if intuitions channel theoretical views, though disguised to seem independent of theorizing, we should see them as simply embodying those theoretical claims, which we can assess in standard ways.

## II. Appearance and Theory

The intuition we're considering is that there is no tenable distinction between appearance and reality in the case of conscious experience. If we see that intuition as genuinely pretheoretic, we can do nothing but simply endorse it or disregard it. But if we instead take the purported intuition to channel a theoretical claim, we can assess that claim in standard ways, invoking whatever considerations prove useful.

One consideration is that explanations of conscious phenomena often appeal to mental aspects of conscious experiences distinct from their subjective appearances. Consider again our subjective impression that conscious parafoveal vision is rich in color and detail. Explaining why we have that mistaken subjective impression will likely invoke mental aspects of those parafoveal states other than their subjective appearances (e.g., Knotts et al 2020).

Other psychological explanations will also appeal to some mental reality of conscious states that's distinct from their subjective appearances. And there seem to be no independent considerations that favor holding that subjective appearance exhausts mental reality. So we should see that claim, however seemingly inviting, as lacking serious substantiation.

If the subjective appearances did exhaust the mental reality of conscious states, those states would have no mental properties that aren't conscious. So one who held those claims might dismiss out of hand any account of what it is for mental states to be conscious that appeals to mental properties that aren't conscious. How, one might wonder, could unconscious mental properties help explain what it is for a state to be conscious?

But that would be a mistake. We can't explain what it is for mental states to be conscious by appeal to the subjective appearances themselves, since those conscious appearances are what needs explaining. Appealing only to conscious mental properties would at best enable us simply to describe the subjective appearances. It could not also help explain what it is for a state to have the property of being conscious.

So if one were convinced that the subjective appearances do exhaust the mental reality of conscious states, one might seek instead to explain consciousness in a way that does not appeal to mental factors at all, perhaps simply by identifying or correlating conscious

states with neural patterns. That would be informative about consciousness, just as identifying water with H<sub>2</sub>O is informative about the nature of water.

But though it would be informative, it could not help us understand the property of being conscious itself. And that is pivotal. Identifying water with H<sub>2</sub>O requires some prior grasp in commonsense terms of what water is; otherwise we wouldn't know what we're identifying. Similarly, we can't correlate consciousness with neural patterns unless we know when a state is conscious.

And knowing when a state is conscious requires some grasp, even if preliminary, of what it is for mental states to be conscious, that is, some grasp of the property of a state's being conscious. Neural correlations rely on that prior step, and so cannot help with it. And because being conscious is a mental property, we must understand what it is for a state to have that property in distinctively mental terms.

These considerations point to a potential impasse. Explaining what it is for a state to be conscious must appeal to mental factors. But the explanation cannot appeal to the subjective appearances themselves, since it's those appearances that need explaining. So the explanation must appeal to mental factors that are distinct from those subjective appearances. But if the subjective appearances exhaust the mental reality of conscious states, there are no such factors. The view that subjective appearance exhausts mental reality prevents any informative explanation of what it is for a mental state to be conscious.

Those who are convinced that the subjective appearances exhaust the mental reality of conscious states may be comfortable with that result. Thus Ned Block sees qualitative mental reality as constituted by what it's like for one (1995), that is, by the subjective appearances. And that may explain Block's dismissive appeal, in saying what conscious qualitative states are, to Louis Armstrong's famous remark about jazz: "If you gotta ask, you ain't never gonna get to know" (1978, p. 281).

And Block has recently been more explicit: "The best you can do is use words to point to a phenomenon that the reader has to experience from the first person point of view" (2015, p. 47). It's not clear how to understand the metaphor of pointing, or how words might point to an experience. But what matters is that if we confine ourselves to the subjective appearances, we can say nothing informative about what it is for a state to be conscious.

## **Ila. Explaining Consciousness**

But we need not acquiesce in that explanatory dead end. We can get leverage to explain what it is for a mental state to be conscious by accepting that conscious states do have some mental reality distinct from their subjective appearances.

Consider how conscious states differ from mental states that occur unconsciously. Perceiving occurs consciously, but also subliminally, without being conscious. This is evident from various experimental results. Subjects can be presented with stimuli they

sincerely take themselves not to perceive, even though priming effects and forced-choice guessing well above chance show that the stimuli did result in perceptual processing (Marcel 1983; Breitmeyer and Ögmen 2006; Dienes 2012). Such effects are also evident in blindsight (Weiskrantz 2009). The downstream perceptual effects show that perceiving occurs, and subjects' sincere denials show that it isn't conscious. (See Berger and Mylopoulos 2019 for more about unconscious perceiving. On eliminating confounds in assessing conscious versus unconscious cases, see Morales et al, 2022.)

Conscious and subliminal perceiving often differ in purely perceptual ways that aren't relevant to what it is to be conscious. Subliminal perceiving typically represents things more faintly and with less perceptible detail. But all that matters for being conscious is that in conscious perceiving there is some relevant subjective appearance, whereas there's none when a perception is subliminal.

The subjective appearances that occur in conscious perceiving pertain not only to what one perceives, but also to the perceiving itself. If one perceives something but one's perceiving does not appear to one subjectively, there is no associated subjective appearance. So that perceiving is not conscious. Perceiving is conscious only if it subjectively appears to one.

And appearing has a crucial connection with awareness. If one is wholly unaware of something, that thing doesn't appear to one in any way. So for something to appear to one, one must be aware of it in some way. And because a perception is conscious only if the perception subjectively appears to one, a necessary condition for a perception to be conscious is that one is in some way aware of it.

These considerations apply not only to perceptions, but to all mental states. Whenever a mental state of any sort is conscious, there is some way that state is subjectively for one, some way that state is for consciousness. The way the state is for consciousness is that state's subjective appearance. And since appearance implies awareness, it's a necessary condition for a state of any type to be conscious that one be in some way aware of that state.

This necessary condition is what I have elsewhere called the transitivity principle (e.g., Rosenthal, 2005). It's endorsed in one form or another by all higher-order (HO) theories of consciousness. And there is compelling reason to hold this principle even apart from the connection between awareness and appearance. If one is in some mental state but is in no way aware of being in it, we do not count that state as being conscious. And that is equivalent to the transitivity principle. The transitivity principle is simply an observation about which states we pretheoretically regard as conscious, and so underlies both commonsense judgments and experimental methodology. Thus even one prominent advocate of a global-workspace theory has in effect endorsed the transitivity principle (Naccache 2018).

Explaining consciousness requires saying how conscious and unconscious states differ. And because a state's being conscious is a mental property, we must specify that

difference in distinctively mental terms. The transitivity principle offers a satisfying proposal. It's not obvious what credible alternative there might be that's also cast in distinctively mental terms.

A state is conscious only if one has some higher-order awareness (HOA) of that state. These HOAs are themselves mental states that are distinct from the states they make one aware of. One might think there aren't two mental states but only one, since phenomenology typically reveals only one. But when phenomenology reveals only one, that's not because phenomenology sees both the HOA and its target and makes one aware of them as a single state. Rather, it's because HOAs are rarely conscious, and so phenomenology rarely sees HOAs at all.

Since HOAs are seldom conscious, phenomenology can't speak to whether a HOA and its target state are two states or one. But there's no independently well-founded way of individuating mental states that would count a HOA and its target as a single state (e.g., Phillips 2014). Phenomenology is the only reason one might see the two as a single mental state, and phenomenology would be reliable only if subjective appearance exhausted mental reality.

Being aware of a state in a suitable way constitutes the subjective appearance in virtue of which the state one is aware of is conscious. The conscious state itself is the state that one subjectively appears to be in, for example, a perception or a thought. And that state is a mental reality distinct from the subjective appearances. Distinguishing that mental reality from the subjective appearance enables an informative explanation of what it is for a state to be conscious, which would not otherwise be possible.

One might contend that a mental state that isn't conscious cannot make one aware of anything. If so, HOAs that aren't conscious couldn't make one aware of any first-order mental states, and so couldn't result in any states' being conscious. One could then save a HO theory only by stipulating, contrary to the first-person appearances, that the HOAs are always conscious.

But a mental state need not be conscious to make one aware of something. Perceptual objects have a significant effect on perceptual processing even when perceiving is unconscious. That could happen only if one were in some way aware of those objects. One is aware of these subliminally perceived objects, just not consciously aware of them. And that's also what happens when HOAs aren't conscious. Unconscious HOAs make one aware of their target first-order states, but not consciously aware of them.

Because HOAs are rarely conscious, being subjectively aware of them is not the reason to think they occur. Rather, we should hold they occur because of the theoretical considerations about subjective appearance and the transitivity principle. So it's no argument against such HOAs that they're rarely evident in everyday phenomenology. HOAs are justified in the way any good theoretical posits are.

Richard Brown (2015) adopts the HO theory that I have long advocated, but with a slight modification: that for the case of qualitative consciousness it's the HO state that exhibits qualitative consciousness. That modification confuses a state's being qualitatively conscious with a necessary condition for qualitative consciousness. There is no qualitative consciousness without a HOA. But one is rarely aware of any HO state. So there's rarely anything it's like to be in a HO state, and HO states are almost never conscious. Being aware of a first-order state explains why it subjectively appears that one is in that state, so that there's something qualitative it's like to be in it. So it's that first-order state that's qualitatively conscious. If the HO state were qualitatively conscious, the HO apparatus would do nothing to explain why qualitative consciousness occurs. On Brown's account that just remains unexplained.

Because first-person access rarely reveals any HOAs, one might conclude that we are never aware of our conscious states. Thus Gilbert Harman (1990) urges that the only properties one is ever aware of in conscious perceiving are properties of the things one perceives. But perceptions are not conscious unless one is in some way aware of the perceiving. So when they are conscious, we must be aware of them and their mental properties. We could ignore these theoretical considerations in favor of first-person access only if the subjective appearances exhausted the mental reality of conscious perceiving.

When mental states are conscious, their being conscious is typically inattentive and unfocused. But sometimes a state is conscious in the deliberately focused way we think of as introspective consciousness. In those cases the HOA is typically itself a conscious state. One is aware of the introspected state, and aware in addition of being aware of that state. So one is not only aware of the introspected state, but also consciously aware of it.

Still, 'introspective consciousness' is something of a term of art, and we can reasonably describe different things as falling under that heading (Giustina 2019). Thus one can attend to a conscious state even when one's HOA remains unconscious, and one might also count that as a type of introspective consciousness.

Being aware of a state is necessary for that state to be conscious, but it is not sufficient. Still, we can begin to close in on a sufficient condition by determining how one must be aware of a state for that state to be conscious. I've argued elsewhere (e.g., 2005, 2012b, 2018) that the required type of awareness consists in one's having a thought that one is in that state.

Such higher-order thoughts are states with intentional content and an assertoric mental attitude. They are rarely themselves conscious, they needn't satisfy requirements of rationality or other epistemic demands, and they will be relatively minimal conceptually. And because one is aware of one's conscious states in a way that subjectively seems unmediated, such higher-order thoughts must not seem subjectively to rely on any inference. Some unconscious inference may be operative, just none that's conscious.

Positing higher-order thoughts helps explain why we can typically report our conscious states in a way that seems subjectively independent of inference (Rosenthal 2005, chs. 2,

10-12; 2018). Reporting something always expresses a thought one has about that thing, so that the report and the thought it expresses have the same content. So if one is able to report being in some mental state, one must have a thought with that content, a thought one would express if one made the report. In addition, positing agentive thoughts that one is acting, on analogy with these higher-order thoughts, results in a compelling account of the phenomenology of agency (Mylopoulos 2017).

The appeal to subjectively noninferential reporting to argue for higher-order thoughts may recall Wilfrid Sellars' proposal about how conscious intentional states differ from those that aren't conscious (1956, §XV). But the present proposal, unlike that of Sellars, applies to qualitative as well as to intentional mental states. The appeal to higher-order thoughts also echoes W. V. Quine's idea that self-construal is in significant ways on a par with the construal of others (e.g., 1969, p. 46). And the appeal to higher-order thoughts reflects the use by both of theorizing in explaining commonsense phenomena. But in what follows I won't appeal specifically to higher-order thoughts, but will instead rely on generic HOAs.

For a state to be conscious, one must be aware of oneself as being in that state. And that awareness must be essentially indexical (Perry 1979). It must be awareness of oneself as such; it won't do for one to be aware of oneself under some irrelevant description. I've argued elsewhere that such essentially indexical self-reference consists in one's being disposed, should the question arise, to identify the individual one is aware of as identical with the individual that is thus aware (Rosenthal 2012a). (For more on HO theories and the self, see Weisberg 2019.)

Miguel Ángel Sebastián (2019) notes that one can't identify the individual a HOA is about unless one is aware of that HOA. And the relevant kind of third-order awareness of a second-order awareness would again make essentially indexical reference to oneself. So one could then identify the individual such a third-order awareness is about only if that third-order awareness were itself conscious.

Sebastián argues that this leads to regress. But my account of indexical self-awareness in Rosenthal (2012a) is dispositional. And the disposition for identifying the individual a HOA is about can obtain even if that HOA never becomes conscious and the question about which individual the HOA is about never arises. So a slight adjustment avoids Sebastian's regress: For any HOA, one is disposed, should the question arise for some HOA and that HOA does become conscious, to identify the individual the HOA is about with the individual that has that HOA. That disposition is itself an aspect of the mental reality that obtains independently of consciousness and subjective appearance.

## **IIb. Qualitative Consciousness**

Conscious perception differs from subliminal perception because conscious perceiving is accompanied by a HOA of the perception that's absent in the unconscious case. But perceptions, both conscious and unconscious, also differ among themselves in ways that are purely perceptual. Seeing a red square or a green triangle and hearing an oboe or a

trombone all differ perceptually, whether or not the perceiving is conscious. So there must be other properties in virtue of which perceptions differ in those ways.

It's sometimes assumed that the mental properties responsible for perceptual differences in conscious perceiving are intrinsically or essentially conscious. If that were so, those properties couldn't also explain unconscious perceptual differences. And then the properties responsible for unconscious perceptual differences might not be mental properties at all, but just subpersonal.

But it's unclear what reason there could be to think the properties in conscious perceiving are intrinsically or essentially conscious other than the assumption that subjective appearance exhaust mental reality, so that any mental properties in conscious perception would be an aspect of the subjective appearances. And that would prevent explaining not only what it is for a state to be conscious, but also the nature of the properties that explain conscious perceptual differences, since we would know nothing about those properties except how they are for consciousness.

It's also unclear what reason there could be to deny that the same properties explain perceptual differences in conscious and unconscious perceiving. Conscious and unconscious perceptions both differ among themselves in much the same ways, visual perceptions in respect of color and spatial aspects, auditory perceptions in respect of pitch, loudness, and timbre, and so forth. That suggests that the same properties explain perceptual differences in both. And if perceptual differences in unconscious perceiving were due just to nonmental, subpersonal properties, those nonmental properties could presumably also explain conscious perceptual differences, so that the conscious perceptual properties would be explanatorily idle.

We will do better not to see conscious perceiving as the primary form of perceiving, but rather to regard it as an important special case. Then we can focus on what conscious and unconscious perceiving have in common. And the most salient common factor is that both enable the discrimination of perceptible objects and properties. So perceptual states, whether conscious or unconscious, must have perceptual content properties that differ in ways that reflect those discriminable differences.

And the role in perceptual discrimination played by these perceptual content properties points to an informative account of their nature. We can taxonomize these properties by the type of stimulus each content property enables one to discriminate from its perceptible neighbors. This taxonomizing works in exactly the same way for conscious and unconscious discrimination.

And we can make this precise. When one can discriminate two stimulus properties but would be unable to do so were they any closer physically, those stimulus properties are just noticeable different (JND). We can test for JNDs in both conscious and unconscious perceiving, and use the results to construct a space of stimuli that an individual can just barely discriminate. Each discriminable stimulus will have a unique relative location in that

quality space. When physically distinct stimuli are indistinguishable for an individual, as often happens, they're assigned the same relative location.

Discriminating two stimuli requires an individual to be in states with distinct perceptual content properties, whereas when stimuli are indistinguishable an individual responds to both with perceptual content properties of the same type. So the space of JND stimuli also fixes an individual's perceptual content properties.

This quality-space theory provides an informative account of the perceptual content properties. Each content property is identified by its role in enabling an individual to discriminate some stimulus type from its perceptible neighbors. That specifies the nature of those content properties. And since we discriminate stimuli both consciously and unconsciously, the account is independent of first-person access (Rosenthal 1991; 2005, chs. 5-7; 2010; Clark 1993, forthcoming). The perceptual content properties constitute a mental reality distinct from the subjective appearances. As an extra, quality-space theory underwrites an informative way to individuate the various sensory modalities (Rosenthal 2015).

Because the same perceptual content properties occur in both conscious and unconscious perceiving, those properties aren't intrinsically conscious. And empirical findings support this conclusion. An elegant use of metacontrast masked priming for color constancy by Liam Norman and colleagues (2014) has shown that surface colors are perceptually registered unconsciously, independently of conscious color experience. These unconsciously registered properties likely figure in conscious vision as well.

The properties posited by quality-space theory enable discrimination in both conscious and unconscious perceiving. But when those properties occur consciously, they are the properties typically described as conscious mental qualities. And some insist that mental qualities, properly so called, cannot occur without being conscious.

But the same mental properties enable both conscious and unconscious discrimination. Since the properties we regard as mental qualities in conscious perceiving also occur unconsciously, we should count the unconscious properties as mental qualities as well. But whether one does so is entirely a matter of terminological preference; it makes no substantive difference. And in what follows I'll often refer to both as mental qualities.

The idea that mental qualities cannot occur without being conscious is built into Block's notion of phenomenal consciousness (1995), on which qualitative mental reality is constituted by what it's like for one. On that notion, subjective appearance is all there is to qualitative mental reality. So Block's notion of phenomenal consciousness is in that way theoretically tendentious. And it arguably faces significant problems as well (Weisberg 2011a, 2011b).

Block's notion of phenomenal consciousness is widely adopted, perhaps in part because there's no other accepted terminology. But Block's notion is wholly optional for thinking about and describing qualitative consciousness. Quality-space theory provides an

alternative way to do so on which the properties that explain conscious qualitative differences also explain unconscious perceptual differences.

Quality-space theory identifies each mental quality by the type of stimulus it enables one to discriminate from its perceptible neighbors. So it's natural to see each type of mental quality as representing that type of stimulus property. And one might take that to support the representationalist view that the perceptual content properties are wholly a matter of which stimulus properties various perceptions represent, as Jacob Berger (2018) has forcefully argued.

But on Berger's representationalist proposal, discriminative ability determines which stimulus property each perceptual content property represents. As with quality-space theory, discriminative ability is explanatorily basic. So we should see the perceptual content properties as primarily a matter of discriminative ability, not representation.

The stimulus properties that figure in quality-space theory are taxonomized by an individual's ability to discriminate them, not by their physical nature alone. That's what generates the nice representational match between mental qualities and stimulus properties so taxonomized (Berger 2021; Rosenthal 2010). So quality-space theory avoids problems that would arise if the stimulus properties were taxonomized solely by their physical nature, problems nicely highlighted by Kathleen Akins (1996).

A HO theory explains what it is for mental states to be conscious by appeal to HOAs. So such a theory requires explaining other mental properties independently of consciousness. Those properties will constitute a mental reality distinct from the subjective appearances. And because quality-space theory explains perceptual properties without appeal to the subjective appearances, that theory requires an independent explanation of why some mental states are conscious. Each of the two theories is independent of the other; any theory that provided the required complementary explanation would do.

Still, a HO theory and quality-space theory work very well together. Mental qualities are individuated by relative location in a space constructed from JND perceptible stimuli. And a state is conscious only if one is in some suitable way aware of that state. When a perception is conscious, one is aware of that perception in respect of perceptual properties fixed by their relative location in a quality space. So HOAs represent perceptual states in respect of mental properties individuated by such relative locations. That enables HOAs to capture and reflect the phenomenological richness and fineness of grain characteristic of conscious experience, addressing an important concern of Alex Byrne's (1997).

And the idea that HOAs represent perceptual states by relative location in a quality space fits strikingly well with the way we actually think in commonsense terms about conscious perceiving. We are subjectively aware of our perceptions in ways that reflect differences among the stimulus properties we can discriminate. So if one is asked to say what it's like to see some colored object, for example, one will typically compare one's experience with experiences of other shades that are available to those present or that are familiar from known types of object. We describe what it's like for us in terms of relative location in a

quality space. We are not confined to "us[ing] words to point to a phenomenon that the reader has to experience from the first person point of view" (Block, 2015, p. 47).

Third-person access to the mental states of others relies on causal connections the states have with behavior, perceptual inputs, and other mental states. Those causal connections reveal mental properties that determine the type of mental state one has third-person access to. And the mental properties that determine those types cannot be the subjective appearances of the states, since unconscious states of the same mental type would result in very similar causal connections. For that reason, third-person access typically doesn't reveal whether somebody else's state is conscious (Rosenthal 2018, §1).

So first-person access can pick out states of the same mental type as third-person access only if we have first-person access to the mental properties used by the third-person access that others have. Subjective appearance must connect with the mental properties that enable third-person access, which are distinct mental realities.

We can explain how that happens by appeal to the way HOAs represent their targets. HOAs constitute the subjective appearances. And HOAs represent qualitative states in respect of relative location in a quality space. So the subjective appearances make one aware of qualitative states in respect of the same discriminative properties that causal connections reveal to others. We have first-person access to subjective appearances that incorporate information about the same mental properties that make third-person access possible. Things work the same way for states that have other mental properties, such as intentional contents (Rosenthal, forthcoming), since the HOAs in those cases also represent the states in terms of the way those other mental properties operate psychologically.

Quality-space theory implies that undetectable inversion is not simply impossible, but also inconceivable. It isn't possible because uniquely fixing mental qualities by relative location in a quality space precludes that space from having any axis of symmetry, since qualities on opposite sides of such an axis would have indistinguishable relative locations. And because quality-space theory tells us how we conceive of mental qualities, undetectable inversion is not even conceivable. (See Clark 2021 for more on undetectable inversion.)

Andrew Lee (2021) has argued that mental qualities differ in respect of precision, and that a quality-space account must accommodate such differences. His main example of an imprecise mental quality involves seeing something red parafoveally or at a distance, so that the experience doesn't seem subjectively to determine a unique shade of red (cf. Block 2015).

But we must distinguish mental qualities from our subjective awareness of them. So it may be that such cases of degraded vision exhibit imprecision not because the mental qualities are themselves imprecise, but rather because one's subjective awareness of those mental qualities is imprecise. The mental quality may accurately reflect a precise shade, but do so too weakly to result in a precise subjective awareness. And an imprecise subjective awareness would not imply an imprecise mental quality unless subjective appearance exhausted mental reality.

To control for this confound involving subjective appearance we must taxonomize mental qualities in the first instance by discriminability of stimuli in optimal conditions. We can then seek to extrapolate to degraded perceptual cases by appeal to neural correlates or to measures that are independent of subjective appearance.

The mental properties that enable perceptual discrimination are theoretical posits of quality-space theory, just as HOAs are theoretical posits of a HO theory of consciousness. Together these complementary theories explain the subjective appearances of conscious experience. The two posits are established by that explanatory success, not by subjective awareness. Indeed, theorizing about mental phenomena cast in distinctively mental terms inevitably appeals to some mental reality that isn't accessed by first-person awareness.

### **III. Appearance and Misrepresentation**

If there is more to the mental reality of conscious states than their subjective appearance, the question arises about whether those appearances always accurately reflect that mental reality. Does one's being subjectively aware of seeing something red, for example, ensure that one is actually in a visual state of seeing something red?

Consciousness is the first-person appearance of being in some mental state. So if the subjective appearances can sometimes be inaccurate about mental reality, consciousness might misrepresent what mental state one is in. And many insist that such misrepresentation is impossible.

Misrepresentation couldn't happen if the subjective appearances of conscious experiences did exhaust mental reality. If conscious states had no mental reality apart from their subjective appearances, there would be nothing mental for consciousness to misrepresent, and the issue about misrepresentation would be idle. But we've seen no serious reason to think subjective appearance does exhaust mental reality. And it's unclear what else might preclude even the possibility of misrepresentation.

One might urge that consciousness cannot misrepresent because consciousness is simply what it's like for one. That misses the point. The question is whether one can subjectively appear to be in a mental state that one is not actually in. If that happened, what it would be like for one would be that one is in that state, even though one is not. And that is not automatically precluded if there is more to the mental reality of conscious states than their subjective appearances.

The issue about misrepresentation by consciousness has been discussed mainly in connection with HO theories, since those theories leave open that possibility. But HO theories also deny that the subjective appearances exhaust mental reality. And consciousness could misrepresent only if conscious states have some mental reality distinct from their subjective appearances, which those appearances could get wrong. So

some might see focusing on misrepresentation as a useful way to get leverage on the deeper issue about mental appearance and reality.

HO theories do allow that HOAs could sometimes misrepresent the states they make one aware of. But on HO theories there is reason to think that if misrepresentation does occur it would be rare. Because the first-order states that HOAs make one aware of will typically play a dominant role in causing HOAs, those HOAs will usually be accurate. Still, causal connections can go astray and other causal factors can intrude. So HO theories cannot guarantee accuracy.

It's often taken as a decisive objection to HO theories that they leave open the possibility of misrepresentation (Byrne 1997; Neander 1998; Levine 2001; Block 2011; for useful replies see Weisberg 2011b, Coleman 2018, §2, and Lau and Brown 2019). But this is odd. HO theories are neutral about misrepresentation; they don't predict that it ever occurs. So if one were convinced that consciousness never misrepresents or that it cannot, one could accept a HO theory and simply add a provision that the subjective appearances never misrepresent, or indeed that they cannot. Even if one would prefer one's theory to preclude misrepresentation without that added provision, not doing so doesn't by itself seem all that serious.

Why then is being neutral about misrepresentation seen as a telling objection? Objectors don't explicitly address this. So perhaps, as suggested above, they find a theory's being neutral as objectionable because it's a symptom of the deeper complaint that the theory accommodates a mental reality distinct from the subjective appearances. The concern about subjective appearance and mental reality may underlie the objection from misrepresentation.

And that diagnosis is supported by the way Karen Neander and Joseph Levine develop the objection. Neander argues that it would be incoherent to have a sensation of red but a HOA of having a sensation of green. What, she asks, would such a situation be like for one subjectively? If what it's like for one were red the HOA would play no role, but if it what it's like for one were green the first-order sensation would play no role (1998, p. 420).

Levine goes further, arguing that since either the sensation or the HOA would play no role, the two would collapse into a single state (2001, p. 108). The only role Neander and Levine see for any mental state in this situation is to determine what it's like for one. And given that, they argue that the mere possibility of misrepresentation implies an incoherent result.

But there are two distinct mental roles, one played by the HOA and the other by the first-order sensation. The HOA would determine what first-order state one is aware of being in, and so what it's like for one. The first-order sensation would affect what it's like only indirectly, by causally influencing what HOA one has. But the first-order sensation would have a rich mental role that's independent of what it's like for one. It would determine downstream psychological effects that depend on perceptual content, such as the causing of perceptual beliefs, priming effects, and the like.

So there are two mental roles, one determining what it's like for one, and the other determining perceptual processing. Why, then, do Neander and Levine see the entire mental role as solely a matter of what it's like for one? They must be assuming that subjective appearance exhausts mental reality, so that there's no room for any other role. The reason they see neutrality about misrepresentation as objectionable rests on denying any mental reality distinct from mental appearance. And lacking independent support for that denial, we should disregard the objection.

There are two ways we might describe a HOA as misrepresenting. We can say that the HOA ascribes to a first-order state mental properties that the state doesn't have. Or we can say instead that the HOA represents one as being in some first-order state that one is not actually in.

But the line between these two ways of describing things is arbitrary. Suppose one is in a visual state of seeing something red but one's HOA represents one as seeing some other color. Does the HOA misrepresent the actual visual state? Or misrepresent one as being in a visual state that doesn't actually occur? Which way one chooses to describe things will depend on what seems natural in the situation. There is no nonarbitrary cutoff. And by hypothesis the two situations would be subjectively indistinguishable.

Still, there might seem to be an advantage to saying that the HOA misrepresents an actual first-order state. On that description, the actual first-order state is a conscious state, though it's conscious as having mental properties it doesn't actually have. If instead one described the HOA as representing oneself as being in a first-order state that doesn't occur, there would be no relevant first-order state. What, then, would be the conscious state?

Consciousness is mental appearance. It is the way one's mental life subjectively appears to one. So the conscious state is always the state that one seems subjectively to be in. If one is aware of being in a state that one is not actually in, the state one is aware of being in is still the conscious state. It's just that the conscious state is then notional. That's how appearance operates; nonexistent things sometimes appear to one. Still, if one is uncomfortable about notional conscious states, one can describe the misrepresentation in terms of some actual first-order state. There will always be some suitable first-order state one can take to be misdescribed.

### **IIIa. More on Misrepresentation**

HO theories do not predict that consciousness ever misrepresents one's mental life. But there is compelling evidence, independent of such theories, that it sometimes does. Change blindness occurs when one fails to notice a reasonably salient, clearly visible change in one's visual field. And even when subjects don't consciously perceive a change, that change is sometimes visually represented, albeit unconsciously (e.g.,

Fernandez-Duque and Thornton 2000). Consciousness then misrepresents one's mental life by failing to reflect that visually represented change.

An especially dramatic example occurs with a type of change blindness pioneered by John Grimes (1996). Almost no retinal information reaches primary visual cortex during saccades. So Grimes used eye trackers to generate changes during saccades. Some changes were quite striking. In one a parrot changed color from green to red, a change that 18% of subjects failed to detect. Attention was not a factor, since the parrot was the central object in the presentation, occupying over 25% of the display.

A subject presented with a green parrot will be subjectively aware of seeing green. If the subject then doesn't notice the change of color, the subject's subjective awareness of seeing green will persist; if the subjective awareness changed to one of seeing red, the subject would plainly notice the change of color. But red wavelengths stimulate the retina after the change, and after the saccade that information gets through to visual cortex. So a subject who failed to notice the change would then be in a visual state of seeing a red parrot, though still subjectively aware of seeing green. (Grimes's work is currently being replicated, with various improvements; Fallon et al [2020-2022].)

There are less exotic cases in which subjective awareness is not fully accurate about what mental state one is in. When one sees a scene filled with many colored objects, one's subjective impression is typically of seeing many relatively generic colors. One rarely has a subjective awareness of seeing specific shades, and rarely could describe any exact shades. Still, one's visual states do likely register specific shades, some information about which might well be available from priming effects and forced-choice guessing. The visual states register specific shades, and the subjective appearances misrepresent them as exhibiting generic shades.

Block doubts whether one can "experience generic redness without experiencing any specific shade of red." How, he asks, can there "be an experience of red but not of any shade of red" (2011, p. 444-5)? There are two issues here. One is whether the mental qualities of visual sensations can exhibit generic shades. That's not relevant to misrepresentation.

What is relevant to misrepresentation is whether one can be subjectively aware of seeing a generic shade even if one's visual state is of some specific shade. It does sometimes happen that one cannot subjectively pin down any specific shade, and one reports being subjectively aware only of a generic shade. The only reason to think that couldn't happen even when the visual state itself is specific would be that one assumed that the subjective appearances exhaust the mental reality.

There are other cases in which subjective awareness clearly misrepresents mental reality. Keeping a painful stimulus constant, one's subjective impression of pain is more intense when one perceives the pain as due to somebody's intention to cause harm (Gray and Wegner 2008). Since identical stimuli typically result in similar pains, we should conclude that psychological factors distinct from the pain itself are distorting one's subjective

awareness of the intensity. One could reject these cases and others (Rosenthal 2012b, §4) only by denying any mental reality distinct from the subjective appearances.

There is an uninteresting way in which consciousness cannot misrepresent our mental lives. Consciousness cannot misrepresent the subjective appearances themselves. But that's because consciousness consists in those subjective appearances. It doesn't represent them at all and so cannot misrepresent them. And one would construe the issue about misrepresentation that way only if one held that subjective appearance exhausts mental reality, so that there could be no distinct mental reality to misrepresent.

#### **IV. Appearance and Explanation**

Conscious states have unconscious aspects that figure in psychological functioning. Their role in psychological functioning warrants regarding those unconscious aspects as an unconscious mental reality, distinct from the subjective appearances.

One could still seek to save the view that subjective appearance exhausts mental reality by construing any unconscious aspects of conscious states as simply not mental. Any candidates for an unconscious mental reality of a conscious state would on that construal be subpersonal, what Daniel Dennett calls "events of content fixation" (1991, pp. 365, 457-8).

But without any reason for that construal other than a desire to save the claim about subjective appearance and mental reality, such a construal is merely a terminological move. If conscious states have unconscious aspects that lend themselves to being described in distinctively mental or psychological terms, we must regard them as mental. However we describe them, they constitute an unconscious mental reality distinct from the subjective appearances.

Must all the mental properties of a conscious state occur consciously? This question hinges on the issue about subjective appearance and mental reality. If conscious states have no mental reality apart from their subjective appearances, every mental property would be an aspect of the subjective appearance, and thus conscious. And any mental reality distinct from the subjective appearances will involve unconscious mental properties.

And this affects how we explain various phenomena. George Sperling (1960) very briefly presented subjects with 3 rows of 4 letters; subjects report consciously seeing all the letters. But a moment after the display vanishes, subjects can accurately report specific identities for only about 4 letters randomly located throughout the array. Sperling's innovation was to present an auditory tone after the display vanishes, which cues subjects to one of the 3 rows. And then subjects accurately can report specific identities for most letters in the cued row.

Since subjects don't know which row will be cued, the ability to report most identities in whichever row is cued shows that specific identities have been registered and retained for

most of the 12 letters. One question is how that information is registered and retained. And since subjects do retain identities for most of the 12 letters, a second question is why they can't report most of those 12.

Discussing a closely related experimental paradigm, Block (2007) urges that the specific information persists consciously, but a bottleneck prevents subjects from accessing much of that conscious information. Consciousness, he argues, overflows cognitive access in that way.

But there is an alternative possibility, which is arguably more natural. Information about specific identities might instead be retained unconsciously, and subjects can retrieve some but not all of that unconscious information. On that hypothesis, the information that overflows cognitive access is largely unconscious. Unless that alternative is ruled out, Sperling-style experiments cannot support Block's version of an overflow hypothesis. (On overflow, see also Knotts et al 2019.)

So Block needs more to support his version of overflow. And he relies on subjects' own ideas about how they retain the information. On being informally asked, subjects say they retain it by way of conscious visual images. And Block takes those anecdotal reports "at face value" (2007, p. 488).

But we shouldn't. There's no reason to think subjects would be at all reliable about how they retain the information. Indeed, they might not be reliable even about whether they have a conscious image, as against some nonimagistic sense about things. And response bias would likely lead them to say it's a conscious image, since other possibilities wouldn't occur to them.

Still, there is a theoretical consideration that does favor Block's overflow hypothesis over the alternative. Subjects have a conscious experience of the whole array, and they register most specific identities in some way. So if conscious experiences must be conscious in respect of all their mental properties, subjects would have to register specific identities consciously. And then they might also retain them consciously. Holding that conscious states must be conscious in respect of all their mental properties supports Block's version of overflow.

But the only reason to think that conscious experiences must be conscious in respect of all their mental properties is the claim that subjective appearance exhausts mental reality. So without that claim, we also have no reason to think subjects do register or retain specific identities consciously. It's likely that much of what overflows cognitive access in Sperling-type experiments is unconscious.

There is experimental evidence that conscious experiences are not always conscious in respect of all their some mental properties. Stimuli that are consciously discriminable can be degraded so that they remain consciously detectable, but are no longer consciously discriminable. Subjects can consciously detect whether the stimuli are present, but can no longer consciously discriminate them.

But forced-choice guesses about discrimination are still accurate well above chance (Mealor and Dienes 2012). The best explanation is that subjects' conscious perceptions have unconscious mental aspects that enable that forced-choice discrimination. And since the unconsciously discriminated stimuli result from degrading consciously discriminable stimuli, we should regard the properties that enable that unconscious discrimination as mental. These unconscious mental properties constitute a mental reality of the conscious perceptions that's distinct from their subjective appearances.

#### **IVa. Utility**

Conscious states often have considerable utility for behavior and for subsequent psychological functioning. So it may be tempting to think that a mental state's being conscious by itself adds significant utility to the state over and above whatever utility it has independently of being conscious. And if consciousness did add utility, always or at least often, perhaps that would provide leverage for understanding the nature of consciousness.

But it's unlikely that a state's being conscious does add much utility if any to that which the state has independently of being conscious. The reason is straightforward. The utility a conscious state has is due mainly to its perceptual character and intentional content. Change those content properties and the utility changes accordingly. And perceptual character and intentional content are distinct properties from the property of being conscious. We can accurately assess what's responsible for the utility of conscious states only if we distinguish the property of being conscious from the other mental properties that conscious states have

Indeed, we should expect that the utility of conscious perceptions, thoughts, and desires is due mainly to their perceptual and intentional properties, and not to their being conscious. When mental states aren't conscious, they often have downstream psychological effects that confer significant utility, as with subliminal perceiving and with thinking or wanting something unconsciously. So it's unlikely that utility is due in any measure to a state's being conscious (Rosenthal 2008, 2012b, §5).

One might object that when conscious and unconscious states have the same perceptual and intentional properties, the unconscious states are typically less efficacious than conscious states, and so have diminished utility. But the enhanced efficacy of conscious states likely does not result from their being conscious. Unconscious states are typically weaker causally than corresponding conscious states. And being weaker causally has two results. It prevents them from making it to consciousness. But it also results in their being less efficacious, and so having less utility. The enhanced utility of mental states when conscious is likely due not to their being conscious, but to their being more robust causally, which results both in the greater utility and in their being conscious.

The reasons for thinking that consciousness adds little if any utility don't derive from any theory of consciousness, and in particular not from a HO theory. But the conclusion about

utility does give reason to reject a global-workspace theory. On such a theory, a state's being conscious consists in its being available for downstream psychological processing, which would often confer substantial utility.

That results in counterexamples to any global-workspace theory. Many peripheral visual states are conscious without being globally available, and those states have marginal if any utility. And unconscious thoughts and desires sometimes have significant downstream effects, which can have considerable utility.

The issue about utility has implications for what Dennett has called the hard question of consciousness: "[O]nce some item or content 'enters consciousness', what does this cause or enable or modify?" (2018, p. 1, emphasis Dennett's; cf. 1991, pp. 255, 266; forthcoming). I've argued that the likely answer is very little. Since a state's being conscious by itself has little if any utility, there won't be much that results solely from its entering consciousness. Any enhanced utility would likely result from the greater causal strength that independently leads to a state's entering consciousness, and not from entering consciousness.

If a state's being conscious did add significant utility, perhaps we could explain why some states are conscious by appeal to evolutionary reproductive advantage. But since consciousness adds little if any utility, some other explanation is needed.

For perceptions, perceptual error suggests a promising strategy. Noticing perceptual errors can help one avoid them and help compensate when error does occur. And noticing such errors does have utility. And since perceptual error consists in a disparity between a perceived object and the way one perceives it, noticing an error requires one to be aware of the way one perceives the object, and so to be aware of the perceptual state itself

If one became aware of one's perceptions reasonably often, becoming aware of them might come to be second nature. And then some perceptions would be conscious. Still, the utility in becoming aware of one's perceptual states is due to the role that awareness has in taking account of perceptual error, not because that awareness sometimes results in perceptions' becoming conscious. The self-organizing metarepresentational account (SOMA) of Axel Cleeremans and colleagues (2020) seeks to explain in a related way how mental states generally come to be conscious.

Dennett (forthcoming) raises a question about which creatures other than humans, if any, are likely to be in mental states that are conscious. The foregoing considerations about perceptual error offer some help. Noticing perceptual error is useful not only for humans, but also for creatures without language (e.g., Hirela et al 2020). So the perceptual states of some of those creatures may sometimes also be conscious.

But the appeal to perceptual error can't help with thoughts or other nonperceptual states. Some creatures without language do evidently ascribe intentional states to other creatures (Call and Tomasello 2008). And some of those creatures might even ascribe intentional states to themselves, though getting evidence of that would be difficult. But that self-

ascription might still not result in those creatures' being aware of their own intentional states in the subjectively noninferential way needed for states to be conscious.

So it may be that creatures will become aware of their own intentional states in that subjectively noninferential way only if they do have language and, in addition, a concept of a thought as expressible in language (Rosenthal 2005, ch. 10, §5; Sellars, 1956, §XV). It is in any case significantly more demanding for intentional states to become conscious than for perceptions to become conscious. And that fits with our pretheoretic hunches; it seems far more likely that nonlinguistic creatures have perceptions that are conscious than thoughts that are conscious.

Dennett (forthcoming) sensibly urges that a HO theory of consciousness should enable identifications of HOAs with neural processes. But he argues further that doing so will likely require appeal to utility that the HOAs by themselves confer. And he proposes that HOAs add utility by enabling the monitoring of one's own mental states. Dennett is clear that monitoring one's own mental states need not always be conscious. But he urges that such self-monitoring is in general possible only if some mental states are conscious.

I have doubts about both claims. We can identify HOAs with neural processes by appeal solely to straightforward correlations, independent of utility. And since self-monitoring does occur independently of consciousness (e.g., Reder et al 1996), it's unlikely that the ability to self-monitor in general requires some mental states to occur consciously.

Perceiving consciously typically generates confidence about what one perceives. And that might suggest that perceptual confidence is a reliable indicator of whether a perception is conscious. And since confidence has utility, if confidence were special to conscious perceiving, a perception's being conscious would add utility.

But unconscious perceiving also sometimes generates confidence, though confidence that typically itself remains unconscious. Forced-choice guessing in unconscious perceiving is not arbitrary. So it must rely on some measure of confidence, though guessing implies that the confidence remains unconscious. Unconscious perception can also guide action in blindsight (Danckert and Rossetti 2005), and the guiding of action requires some confidence, even if again wholly unconscious.

We tend in everyday contexts to think about confidence as being conscious, though a person's behavior can sometimes reflect confidence that they're unaware of. Still, there's no sound theoretical reason to deny unconscious confidence. So it's not confidence itself that has a special connection with consciousness, but only conscious confidence.

The conscious confidence in conscious perceiving is typically stronger than the unconscious confidence that accompanies unconscious perceiving, though experimentally assessing that would require factoring out conflicts that unconscious confidence has with conscious beliefs, which would be challenging. But the enhanced confidence in conscious perceiving need not be a result of consciousness. When a perception is causally strong enough to become conscious, it will also generate greater confidence. So it's likely that it's

that greater causal efficacy that results both in a perception's being conscious and in enhanced confidence, as with consciousness and utility generally.

Confidence, like utility generally, is driven not by consciousness, but by mental content whether conscious or not (Rosenthal 2019). The work is done largely by a mental reality that's distinct from subjective appearance. And though confidence can dissociate from perceptual discrimination (e.g., Maniscalco et al 2020), that may be because confidence and discrimination rely in part on different psychological processes, and not because of anything about consciousness.

Hakwan Lau has proposed a HO theory that appeals to the perceptual monitoring of reality. On that proposal, a perception is conscious "if there is a relevant higher-order representation with the content that, a particular first-order perceptual representation is a reliable reflection of the external world right now" (2019, p. 3; emphasis Lau's). For a state to be conscious as a perception, one must be aware of that state as being reliable about how things are, at least to some degree.

So it's important, as Lau argues, that there is machinery that assesses a state for such reliability. But that machinery is likely not part of any HO apparatus in virtue of which a perception is conscious. Unconscious perceiving often results in successful behavior, and to do so it must, like conscious perceiving, be assessed as reliable. So the machinery that assesses states for reliability is likely part of the perceptual process itself, and independent of consciousness.

Since perception, whether conscious or unconscious, functions psychologically as though it's reliable, being aware of a perception doesn't add reliability; it just adds awareness of the state as reliable. And a state can be conscious even if one is not aware of it as reliable, though it won't then be conscious as a perception. So Lau's provision about reliability affects only how a state is conscious, not whether it's conscious. And it's compatible with the standard transitivity principle, that a state's being conscious requires being aware of oneself just as being in that state.

#### **IVb. Problematic Properties**

Keith Frankish (2016) has argued that phenomenal properties, characterized as "ineffable, intrinsic, private, and immediately apprehended" (13), do not occur in our mental lives. He coins the term, 'illusionism', for his denial of conscious properties so characterized, which has a mildly odd ring, since illusion is itself typically thought to involve mental appearance.

But common sense does plainly countenance conscious qualitative properties (Kind 2008). So we should reject not those conscious qualitative properties, but only the theoretically tendentious characterization of them as "ineffable, intrinsic, private, and immediately apprehended." Frankish sees that characterization as built into the notion of a conscious mental quality (16), and believes there is no stable middle ground between accepting qualitative properties so characterized and denying them altogether.

But quality-space theory shows that we can construe qualitative mental properties without those problem characterizations. More generally, the problem characterizations will seem unavoidable only if one takes subjective appearance to exhaust mental reality. Conscious mental qualities need not be ineffable, intrinsic, private, or immediately apprehended if conscious qualitative states have some mental reality that's distinct from their subjective appearances.

The hard problem is "why and how do physical processes in the brain give rise to conscious experience" (Chalmers 2018, p. 6). The explanatory gap is the idea that we cannot trace an explanatory path between mental qualities as accessed by consciousness and any physical or functional-role properties associated with those mental qualities (Levine 2001, ch. 3). Both the hard problem and the explanatory gap concern the relation between a state picked out by appeal solely to its subjective appearances and an objective state of affairs associated with that state.

The issue about mental appearance and reality helps explain both worries. If subjective appearance were all the mental reality there is to a conscious state, it might indeed seem puzzling how a state so identified could connect with anything identified instead by its physical or functional properties. But if conscious experiences do have some mental reality distinct from their subjective appearances, that mental reality can bridge the gap and support an explanatory path between states identified by their subjective appearances and neural or functional properties.

So the hard problem and explanatory gap are of a piece with the difficulties noted earlier for saying anything informative about what it is for states to be conscious if conscious states have no mental reality apart from their subjective appearances. The hard problem and explanatory gap will seem compelling only if one holds that mental appearance does exhaust the mental reality of conscious experiences. That way of thinking about conscious mentality is shared by the hard problem, the explanatory gap, and illusionism.

The foregoing describes a number of ways in which holding that subjective appearance exhausts mental reality distorts how we think about the mind. We should firmly reject that view. Consciousness is but the tip of the mental iceberg, even for those mental states that are conscious. There is much mental reality to be found below, and much investigation and theorizing to be done in connection with it.

**REFERENCES**

- Akins, Kathleen (1996), "Of Sensory Systems and the 'Aboutness' of Mental States," The Journal of Philosophy, 93, 7 (July): 337-372.
- Baars, Bernard J. (1997), In the Theater of Consciousness: The Workspace of the Mind, Oxford: Oxford University Press.
- Berger, Jacob (2018), "A Defense of Holistic Representationalism," Mind & Language, 33, 2 (April): 161-176.
- Berger, Jacob (2021), "Quality-Space Functionalism about Color," The Journal of Philosophy, 118, 3 (March): 138-164.
- Berger, Jacob, and Myrto Mylopoulos (2019), "On Scepticism about Unconscious Perception," Journal of Consciousness Studies, 26, 11-12: 8-32.
- Block, Ned (1978), "Troubles with Functionalism," in Minnesota Studies in the Philosophy of Science, IX, ed. C. Wade Savage, Minneapolis: University of Minnesota Press, pp. 261-325.
- Block, Ned (1995), "On a Confusion about a Function of Consciousness," The Behavioral and Brain Sciences, 18, 2 (June): 227-247.
- Block, Ned (2007), "Consciousness, Accessibility, and the Mesh between Psychology and Neuroscience," Behavioral and Brain Sciences, 30, 5-6 (December 2007): 481-499.
- Block, Ned (2011), "Response to Rosenthal and Weisberg," Analysis 71, 3 (July): 443-448.
- Block, Ned (2015), "The Puzzle of Perceptual Precision," in Open MIND, ed. Thomas Metzinger and Jennifer M. Windt, Frankfurt am Main: MIND Group, doi: 10.15502/9783958570726.
- Breitmeyer, Bruno G., and Haluk Ögmen (2006), Visual Masking: Time Slices through Conscious and Unconscious Vision, 2nd edn., New York: Oxford University Press.
- Brown, Richard (2015), "The HOROR Theory of Phenomenal Consciousness," Philosophical Studies, 172, 7 (July): 1783-1794.
- Byrne, Alex (1997), "Some Like It Hot: Consciousness and Higher-Order Thoughts," Philosophical Studies, 86, 2 (May): 103-129.

Call, Josep, and Michael Tomasello (2008), "Does the Chimpanzee Have a Theory of Mind? 30 Years Later," Trends in Cognitive Sciences, 12, 5 (May): 187-192.

Chalmers, David J. (2018), "The Meta-Problem of Consciousness," Journal of Consciousness Studies, 25, 9-10: 6-61.

Clark, Austen (1993), Sensory Qualities, Oxford: Clarendon Press.

Clark, Austen (forthcoming), "Qualitative Character and Conceivable Inversions," this volume.

Cleeremans, Axel, Dalila Achoui, Arnaud Beauny, Lars Keuninckx, Jean-Remy Martin, Santiago Muñoz-Moldes, Laurène Vuillaume, and Adélaïde de Heering (2020), "Learning to Be Conscious," Trends in Cognitive Sciences, 24, 2 (February): 112-123.

Cohen, Michael A., Thomas L. Botch, and Caroline E. Robertson (2020), "The Limits of Color Awareness during Active, Real-World Vision," Proceedings of the National Academy of Sciences, 117, 24 (June 16): 13821-13827.

Coleman, Sam (2018), "The Merits of Higher-Order Thought Theories," Trans/Form/Ação, 41, edição especial: 31-48.

Danckert, James, and Yves Rossetti (2005), "Blindsight in Action: What Can the Different Sub-types of Blindsight Tell Us about the Control of Visually Guided Actions?," Neuroscience and Biobehavioral Reviews, 29, 7: 1035-1046.

Dennett, Daniel C. (1991), Consciousness Explained, Boston: Little, Brown & Co.

Dennett, Daniel C. (2018), "Facing up to the Hard Question of Consciousness," Philosophical Transactions of the Royal Society B: Biological Sciences, 373, 1755 (September 19): Article 20170342.

Dennett, Daniel C. (forthcoming), "A Challenge: Can Self-Monitoring Be Achieved without Consciousness?," this volume.

Dienes, Zoltán (2012), "Conscious versus Unconscious Learning of Structure," Statistical Learning and Language Acquisition: Studies in Second and Foreign Language Education, 1, ed. Patrick Rebuschat and John N. Williams, Boston: De Gruyter Mouton, pp. 337-364.

Fallon, Francis, Alan L. F. Lee, Brian Odegaard, Andrew Haun, and David Rosenthal (2020-2022), "Replication and Extension of Crucial John Grimes Experiment: Change Detection during Saccades," Templeton World Charity Foundation Project, TWCF Number 0455, <https://www.templetonworldcharity.org/projects-database/replication-and-extension-crucial-john-grimes-experiment-change-detection-during>

- Fernandez-Duque, Diego, and Ian M. Thornton (2000), "Change Detection without Awareness: Do Explicit Reports Underestimate the Representation of Change in the Visual System?", Visual Cognition, 7, 1-2-3 (January–March): 324–344.
- Frankish, Keith (2016), "Illusionism as a Theory of Consciousness," Journal of Consciousness Studies, 23, 11-12: 11-39.
- Gray, Kurt, and Daniel M. Wegner (2008), "The Sting of Intentional Pain," Psychological Science, 19, 12 (December): 1260-1262.
- Giustina, Anna (2021), "Introspection without Judgment," Erkenntnis 86, 2 (April): 407-427.
- Grimes, John (1996), "On the Failure to Detect Changes in Scenes across Saccades," Perception, ed. Kathleen Akins, New York: Oxford University Press, pp. 89-110.
- Harman, Gilbert (1990), "The Intrinsic Quality of Experience," Philosophical Perspectives 4: 31-52.
- Hirela, Marie, Constance Thiriaua, Inès Rohoa, and H  l  ne Meuniera (2020), "Are Monkeys Able to Discriminate Appearance from Reality?", Cognition, 196 (March): Article Number 104123.
- Kammerer, Fran  ois (2018), "Can You Believe It? Illusionism and the Illusion Metaproblem," Philosophical Psychology, 31, 1 (January): 44–67.
- Kind, Amy (2008), "How to Believe in Qualia," in The Case for Qualia, ed. Edmond Wright, Cambridge, Massachusetts: The MIT Press, pp. 285-298.
- Knotts, J. D., Brian Odegaard, Hakwan Lau, and David Rosenthal (2019), "Subjective Inflation: Phenomenology's Get-Rich-Quick Scheme," Current Opinion in Psychology, 29 (October): 49-55.
- Knotts, J. D., Matthias Michel, and Brian Odegaard (2020), "Defending Subjective Inflation: An Inference to the Best Explanation," Neuroscience of Consciousness, 2020, 1 (niaa025): <https://doi.org/10.1093/nc/niaa025>.
- Kripke, Saul A. (1980), Naming and Necessity, Cambridge, Massachusetts: Harvard University Press.
- Lau, Hakwan (2019), "Consciousness, Metacognition, and Perceptual Reality Monitoring," PsyArxiv, June 10, 2019, doi: 10.31234/osf.io/ckbyf.
- Lau, Hakwan, and Richard Brown (2019), "The Emperor's New Phenomenology? The Empirical Case for Conscious Experiences without First-Order Representations," in Blockheads! Essays on Ned Block's Philosophy of Mind and Consciousness, ed.

- Adam Pautz and Daniel Stoljar, Cambridge, Massachusetts: The MIT Press, pp. 171-197.
- Lee, Andrew (2021), "Modeling Mental Qualities," The Philosophical Review, 130, 2 (April): 263-298.
- Levine, Joseph (2001), Purple Haze: The Puzzle of Consciousness, New York: Oxford University Press.
- Maniscalco, Brian, Olenka Graham Castaneda, Brian Odegaard, Jorge Morales, Sivananda Rajananda, Megan A. K. Peters (2020), "The Metaperceptual Function: Exploring Dissociations between Confidence and Task Performance with Type 2 Psychometric Curves," PsyArXiv, July 6, 2020, doi: 10.31234/osf.io/5qrjn.
- Marcel, Anthony J. (1983), "Conscious and Unconscious Perception: An Approach to the Relations between Phenomenal Experience and Perceptual Processes," Cognitive Psychology 15, 2 (April): 238-300.
- Mealor, Andy David, and Zoltán Dienes (2012), "Conscious and Unconscious Thought in Artificial Grammar Learning," Consciousness and Cognition, 21, 2 (June): 865-874.
- Morales, Jorge, Brian Odegaard, and Brian Maniscalco (2022), "The Neural Substrates of Conscious Perception without Performance Confounds," in Neuroscience and Philosophy, ed. Felipe De Brigard and Walter Sinnott-Armstrong, MIT Press, pp. 285-323.
- Mylopoulos, Myrto (2017), "A Cognitive Account of Agentive Awareness," Mind & Language, 32, 5 (November): 545-563.
- Naccache, Lionel (2018), "Why and How Access Consciousness Can Account for Phenomenal Consciousness," Philosophical Transactions of the Royal Society B: Biological Sciences, 373, 1755 (September 9): Article 20170357.
- Nagel, Thomas (1973), "What Is It Like to Be a Bat?," The Philosophical Review, 83, 4 (October): 435-50.
- Neander, Karen (1998), "The Division of Phenomenal Labor: A Problem for Representational Theories of Consciousness," Philosophical Perspectives 12, 411-434.
- Norman, Liam J., Kathleen Akins, Charles A. Heywood, and Robert W. Kentridge (2014), "Color Constancy for an Unseen Surface," Current Biology, 24, 23 (December 1): 2822-2826.
- Perry, John (1979), "The Problem of the Essential Indexical," Noûs 13, 1 (March): 3-21.

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

Quine, W. V. (1969), Ontological Relativity and Other Essays, New York: Columbia University Press.

Reder, Lynne M., and Christian Dieter Schunn (1996), "Metacognition Does Not Imply Awareness: Strategy Choice Is Governed by Implicit Learning and Memory," in Lynne M. Reder, ed., Implicit Memory and Metacognition, Mahwah, NJ: L. Erlbaum, pp. 45-77.

Rosenthal, David (1991), "The Independence of Consciousness and Sensory Quality," in Consciousness: Philosophical Issues, ed. Enrique Villanueva, Atascadero, California: Ridgeview Publishing Company, pp. 15–36.

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

Rosenthal, David (2008), "Consciousness and Its Function," Neuropsychologia, 46, 3: 829-840.

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

Rosenthal, David (2012a), "Awareness and Identification of Self," in Consciousness and the Self: New Essays, ed. JeeLoo Liu and John Perry, Cambridge: Cambridge University Press, pp. 22-50.

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

Rosenthal, David (2015), "Quality Spaces and Sensory Modalities," in Phenomenal Qualities: Sense, Perception, and Consciousness, ed. Paul Coates and Sam Coleman, Oxford: Oxford University Press, pp. 33-65.

Rosenthal, David (2018), "Seeming to Seem," in The Philosophy of Daniel Dennett, ed. Bryce Huebner, New York: Oxford University Press, pp. 133-164.

Rosenthal, David (2019), "Consciousness and Confidence," Neuropsychologia, 128 (May): 255-265.

Rosenthal, David (forthcoming), "Content and Inference."

Sebastián, Miguel Ángel (2019), "Drop It Like It's HOT: A Vicious Regress for Higher-Order Thought Theories," Philosophical Studies, 176, 6 (June): 1563–1572.

Sellars, Wilfrid (1956), "Empiricism and the Philosophy of Mind," Minnesota Studies in the Philosophy of Science, vol. I, ed. Herbert Feigl and Michael Scriven, Minneapolis: University of Minnesota Press, pp. 253-329.

Sperling, George (1960), "The Information Available in Brief Visual Presentations," Psychological Monographs: General and Applied, 74, 11 (Whole No. 498): 1-29.

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

Weisberg, Josh (2011b), "Misrepresenting Consciousness," Philosophical Studies, 154, 3 (July): 409-433.

Weisberg, Josh (2019), "Higher-Order Theories of Consciousness and the Heidelberg Problem," Protosociology, 36: 340-357.

Weiskrantz, Lawrence (2009), Blindsight: A Case Study Spanning 35 Years and New Developments, 2<sup>nd</sup> edition, Oxford: Oxford University Press.