

## HOW TO THINK ABOUT MENTAL QUALITIES

David Rosenthal  
City University of New York

### I. Intuitions about Mental Qualities

It's often held that undetectable inversion of mental qualities is, if not possible, at least conceivable. It's thought to be conceivable that the mental quality your visual states exhibit when you see something red in standard conditions is literally of the same type as the mental quality my visual states exhibit when I see something green in such circumstances. It's thought, moreover, to be conceivable that such inversion of mental qualities could be wholly undetectable by any third-person means. And since first-person access is limited to a single individual, and so could not reveal a disparity in mental quality between us, third-person undetectability means undetectability *tout court*.

There is an extensive literature about whether such inversion is conceivable, and whether it's even possible. But my concern here will not be with the details of such discussions, but with the status of intuitions about undetectable quality inversion and what if any implications it has for theorizing about mental qualities.

It's important that the relevant quality inversion is undetectable. Humans do sometimes differ in what mental qualities occur when perceiving the same stimuli in the same perceptual circumstances. Some people, for example, are red-green color blind. And there are slight variations among individuals as to what color stimuli they can distinguish, which probably reflect slight variation in mental color qualities. These differences are all detectable by appeal to the physical stimuli people can discriminate. By contrast, the inversion of mental qualities that many take to be conceivable is hypothesized to be undetectable by any such means. And it's such undetectable inversion that will likely affect the way we theorize about mental qualities.

It's widely recognized that the quality space that defines the human ability to discriminate among perceptible physical color properties is asymmetric. So many theorists have retreated from the claim that inversion of mental qualities is possible to an ostensibly weaker claim that it's merely conceivable.<sup>1</sup> And it's arguable that since we can coherently describe such undetectable inversion, it must at least be conceivable.

I'll argue in what follows that undetectable inversion is not even conceivable. But as already noted, my concern with undetectable inversion is mainly with the status of claims about it and their implications for theorizing about mental qualities.

Why would the mere conceivability of undetectable inversion matter? Indeed, why would it matter if such inversion were also possible? We can't theorize in a vacuum; we always need data to theorize from and to test our theories' predictions. And it's often assumed that pretheoretic folk intuitions serve as data for theorizing in philosophy. What else, one may wonder, could serve as data for philosophy if that field is distinct from the empirical sciences? Also, it's often assumed that many of these pretheoretic folk intuitions in theorizing in philosophy will concern what's possible and what's conceivable, rather than what's actually the case.

The conceivability of quality inversion is often taken to serve as just such a datum for a theory of mental qualities; any satisfactory theory would then have to accommodate such conceivability. And it turns out that this imposes a remarkably strong condition on any theorizing about mental qualities. Such conceivable inversion would be undetectable. But if we could access or find out about mental qualities in any way that's independent of first-person, subjective access, any inversion would be readily detectable. So the mere conceivability of undetectable quality inversion implies that it's inconceivable that we have any third-person, intersubjective access to mental qualities.

This striking consequence is seldom noted, perhaps because those who take undetectable quality inversion at all seriously to begin with also simply take for granted that mental qualities allow only for first-person access. And if one builds the restriction to first-person access into one's conception of mental qualities, the conceivability of inversion is a straightforward consequence.

It's not obvious that there is any independent reason to suppose that undetectable inversion actually is conceivable. One could urge that it's conceivable in that one can describe such inversion and simply stipulate that it's undetectable. It's also often said that such inversion is part of our folk theory about mental qualities; many people, especially children, will say when presented with the inversion scenario that undetectable inversion is conceivable or even possible. Intuitions, moreover, arguably encapsulate aspects of our folk theories about things. So locating the conceivability of undetectable inversion as a plank in our folk theory of the mental arguably does justice to its alleged status as a pretheoretic intuition.

But these cases for conceivability are very weak. Consider the claim that it's conceivable that water is a simple, basic physical substance. Anybody who advanced that claim would be operating with too impoverished a concept of water, for example, a concept of something that simply looks and behaves like water in a relatively narrow range of common conditions. Given what water actually is and given anything like a satisfactory concept of water, it is simply not conceivable that water is a simple, basic substance.

The mere appearance that undetectable quality inversion is conceivable may in just that way mislead about what is actually conceivable. I'll argue in what follows that a satisfactory conception of mental qualities precludes the conceivability of undetectable inversion. But the paucity of substantive reasons for thinking that undetectable inversion is actually conceivable encourages the hypothesis that it seems conceivable only on the assumption that we can know about mental qualities solely by way of first-person access. The view that we can conceive of undetectable inversion is in effect simply an expression of the theory that consciousness is both the first and last word about mental qualities.

Similar considerations apply to several other intuitively inviting claims about mental qualities. It's often said to be conceivable that a creature physically and behaviorally indistinguishable from an ordinary human might nonetheless be in no states with qualitative character. Again, there is little to be said in support of the conceivability of such a creature, often referred to, following David Chalmers,<sup>2</sup> as a zombie. As with undetectable inversion, the conceivability of zombies is often taken simply to be a pretheoretic intuition that any satisfactory theory must accommodate.

Suppose, however, that we can know about mental qualities independent of consciousness, independent, that is, of any first-person access to them. That would have to be due to some difference mental qualities make to the physical makeup or behavior of the relevant creature; there's no other way to get knowledge of their presence independent of consciousness. So if the same physical makeup and behavior can occur sometimes with mental qualities and sometimes without, we cannot have third-person access to mental qualities. The possibility or conceivability of zombies rests on the view that mental qualities are accessible only in a first-person way.

The same considerations apply also to Joseph Levine's explanatory gap and Chalmers's hard problem. Levine, Chalmers, and others hold that a special difficulty affects any effort to explain why any particular physical process should subserve a particular conscious qualitative state rather than some other. Why should a particular brain state result in an experience of red rather than an experience of green? Indeed why it would result in any experience at all? Levine accordingly urges that an explanatory gap separates conscious qualitative states from whatever physical states underlie them;<sup>3</sup> Chalmers calls this the hard problem of consciousness.<sup>4</sup>

Levine argues for such an explanatory gap by urging a difference in what's conceivable about the mental quality of pain and what's conceivable about water. He writes:

While it is conceivable that something other than H<sub>2</sub>O should manifest the superficial macro properties of water, . . . it is not conceivable . . . that H<sub>2</sub>O should fail to manifest those properties (assuming, of course, that we keep the rest of chemistry constant).

By contrast, he maintains, "it is . . . conceivable that there should exist a pain without the firing of C-fibers, and the firing of C-fibers without pain" (548).

Since chemistry is needed for it to be inconceivable "that H<sub>2</sub>O should fail to manifest" the relevant properties, it's unclear why a satisfactory neuropsychological theory would not yield parallel results for pain. But put that aside for now. The idea that an explanatory gap separates conscious mental qualities from their physical underpinnings is again not a pretheoretic intuition that any satisfactory theory must honor. Rather, whatever plausibility there may be to an explanatory gap rests again on the tacit claim that mental qualities are accessible only by way of consciousness. Third-person access to mental qualities would provide the resources needed to build an explanatory bridge from physical makeup to mental qualities. We would encounter difficulties in bridging that gap only if mental qualities are accessible solely by way of consciousness.<sup>5</sup>

Chalmers's hard problem and Levine's explanatory gap both echo Leibniz's well-known mill argument, according to which "perception, and anything that depends upon it, cannot be explained in terms of mechanistic causation, that is, in terms of shapes and motions." Leibniz invites us to imagine "a machine, which was constructed in such a way as to give rise to thinking, sensing, and having perceptions." If the machine were expanded so that one could go inside, "like going into a mill," one would not, he maintains, see "anything which would explain a perception."<sup>6</sup>

Leibniz uses this intuition pump to argue that perceiving occurs only in simple substances. But that Cartesian claim to one side, the thought experiment encourages thinking that we can never explain how anything about physical makeup could result in any qualitative state, let alone result in one mental quality rather than another. But Leibniz's claim that we would see nothing inside the machine that explains perception simply captures in an especially memorable way the view that the mental qualities that occur in perceiving are accessible only by way of consciousness.

The same holds for Frank Jackson's well-known thought experiment about Mary, the scientist who knows everything there is to know from a third-person point of view about visual mental qualities, but has never consciously experienced red.<sup>7</sup> According to Jackson, when Mary first consciously sees red, she learns something new, namely, what it's like for her to see red. If

so, that new knowledge is presumably something about the mental quality that figures in her consciously seeing red, something she can, by hypothesis, know about only by way of consciousness.<sup>8</sup>

The ostensibly pretheoretic intuitions surveyed above all rest on the view, seldom explicitly articulated, that consciousness is our only source of knowledge about mental qualities. But that's not a pretheoretic intuition, but a theoretical claim about mental qualities, a claim that one might well contest. So the intuitions about quality inversion, zombies, the explanatory gap, and the hard problem, though at first sight they may seem independent of theory, all rest on a theoretical claim about mental qualities. Unless there is no alternative to theorizing about mental qualities in that way, those alleged intuitions cannot serve as data that any satisfactory theory must accommodate.

Recent work in so-called experimental philosophy has raised questions about the reliability of many ostensibly pretheoretic intuitions used to guide theorizing in philosophy, some relevant to issues about qualitative consciousness.<sup>9</sup> Much of this work convincingly suggests that these alleged intuitions are not actually endorsed by people who lack the relevant training in philosophy. And that suggests in turn that the alleged intuitions are often, or even always, induced by the adoption of a theoretical position.

It's not surprising that those of the folk innocent of philosophical debate fail to share many of the intuitions most cherished by those in philosophy and often regarded by them as beyond serious dispute. After all, there is substantial disparity even among those in philosophy about which intuitions to credit. And though endorsement of alleged intuitions typically matches theoretical proclivities, it's unclear whether intuitions influence theory or conversely. So it's likely that intuitions often follow substantive theoretical commitments, and are not theoretically neutral.

Perhaps the most significant contribution of experimental philosophy, then, is to underscore the likelihood that various intuitions are not theory neutral, but products of such theoretical commitments. If intuitions aren't widely shared by the folk and fit conveniently with theories espoused by those who invoke the intuitions, it's natural to ask whether the intuitions encapsulate those theories, rather than providing independent support.

But simply showing that the folk don't widely endorse those intuitions cannot by itself isolate the theoretical presuppositions that are likely to figure in forming the alleged folk intuitions. Nor can it settle whether the intuitions result from theoretical commitments or conversely. For that, we must examine the relation between intuitions and theory. It is that kind of consideration that provides reason to conclude that the intuitions most frequently invoked in recent discussions of mental qualities and qualitative consciousness are not neutral data, but stem from positions taken in those debates. Those intuitions rest on an unsupported and optional theoretical position about how it's possible to know about mental qualities.

In what follows, I examine whether there is any alternative to the theoretical position that underlies those intuitions. I argue in §II that an alternative theoretical approach is available, and in §III I sketch a specific development of that approach. And in §IV I argue that the theory advanced in §III fits comfortably with an independently supported explanation of the way mental qualities occur in conscious qualitative states.

## II. Two Concepts of Mental Quality

If we know about mental qualities only, or in the last instance, by way of first-person access, such knowledge will be immune to challenge. And there can then be no reason to think there's anything about mental qualities that we cannot know in that way. Holding that consciousness is the last word about mental qualities leads to what C. D. Broad aptly called the "curious superstition," that introspection "must give exhaustive and infallible information."<sup>10</sup>

Another consequence of adopting this view is that we will be able to say relatively little that's informative about the nature of mental qualities. Thus Ned Block's 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."<sup>11</sup> This echoes Justice Potter Stewart's famously uninformative remark about defining hard-core pornography: "I know it when I see it."<sup>12</sup>

Such epigrams are appealing in this context, since they vividly capture the idea that there's nothing more to know than what consciousness reveals.<sup>13</sup> But it raises a concern about whether it's all that clear what is being talked about, and the question of whether there might after all be something more informative to say about mental qualities. The conception of mental qualities that results from holding that consciousness is the only way we know about them is simply too thin to operate with.

One might suppose that we have no choice but to tolerate these uncomfortable consequences; after all, mental qualities are indisputably properties accessible by way of consciousness. But being accessible by way of consciousness does not by itself imply being accessible in that way only. And there is an alternative way of conceiving of mental qualities, based on another kind of access we have to them. We have access to mental qualities not only by way of consciousness, but also by way of their intimate and arguably essential role in perception. This second way of thinking about mental qualities provides an alternative to conceiving of them as accessible solely by way of consciousness.

I'll refer to theories on which we know about mental qualities only by way of consciousness as *consciousness based*, in contrast to *perceptual-role* theories, which appeal to the connection mental qualities have with perceiving. Perceptual-role theories don't deny the access consciousness

gives us to mental qualities, but only that consciousness is the first and last word about their nature. The role mental qualities play in perceiving gives us, according to those theories, an alternative, independent route to learn about mental qualities.

On a perceptual-role theory, mental qualities are the properties in virtue of which we make perceptual discriminations. So independent of anything that consciousness tells us, if an individual's perceiving involves such discriminations, that individual is in mental states that have mental qualities that pertain to those discriminations. Perceiving provides a third-person way of determining what mental qualities occur, independent of the first-person access that consciousness delivers.

Might perceptual role and consciousness conflict about what mental qualities an individual's perceptual states exhibit? There is no guarantee that this would never happen. But that's the situation with all knowledge about things; we have alternative ways of knowing about things, which may sometimes conflict. The job of a good theory is in part to help adjudicate such conflicts when they arise.

I'll say more in §IV about possible conflicts between consciousness and perceptual role. For now it's worth noting a striking way in which the two types of theory differ. Perceiving is not always conscious. Perceiving can occur subliminally, as demonstrated in masked-priming experiments, in which an individual reports not seeing a stimulus that nonetheless affects that individual's subsequent mental processing. This effect on mental processing is taken to show that perceiving did occur, and the individual's denial of seeing the stimulus as evidence that the perceiving was not conscious.<sup>14</sup>

Perceiving that isn't conscious also occurs in blindsight, in which cortical damage to a region of primary visual cortex, area V1, results again in visual input that the affected individual denies any awareness of.<sup>15</sup> Visual acuity in blindsight is impaired, since it relies on neural pathways that bypass area V1. But there is strong reason to think that mental qualities occur in blindsight as well as in subliminal vision in intact individuals.<sup>16</sup> Some partial blindsight patients can spontaneously respond or attend to input in the affected field,<sup>17</sup> and an individual with complete destruction of area V1 can spontaneously navigate relying solely on visual input.<sup>18</sup>

Nonconscious mental qualities likely occur also in commonsense contexts. One may be awaiting somebody's arrival in a crowded room and, without having consciously noticed the person, suddenly cast one's gaze straight to where the person is standing. There must have been visual input that enables one to shift one's gaze straight to the person, and hence mental qualities relevant to perceiving that person. But since one didn't previously notice the person consciously, the mental qualities occurred without being conscious.

So if mental qualities are properties essential to perceptual discrimination, they sometimes occur without being conscious. Perceptual-role theories not only rely on access to mental qualities that's independent of

consciousness; they rely on access to mental qualities in the absence of any consciousness of those qualities at all. On a consciousness-based theory, mental qualities have an essential tie to consciousness but not to perception; the opposite holds for perceptual-role theories, on which mental qualities have an essential tie to perception, but not consciousness.

One could simply dig one's heels in at this point and insist that if a mental state is not conscious, none of its mental properties are qualitative. It's now widely acknowledged that intentional states, such as thoughts, desires, expectations, wishes, and the like can occur without being conscious. But none of these states exhibits mental qualities. So one might maintain that mental states with only intentional properties can occur without being conscious, but not states with qualitative properties. Even Freud, who soundly rejected the traditional "equation . . . of what is conscious with what is mental,"<sup>19</sup> nonetheless insisted that "emotions, feelings and affects," that is, states with qualitative character, cannot occur without being conscious.<sup>20</sup>

But it's unclear what non-question-begging reason there could be to insist that states with mental quality cannot occur without being conscious. We count states that aren't conscious as nonetheless being intentional if there's good reason to describe them, as we do conscious intentional states, as having intentional properties, that is, content and mental attitude.<sup>21</sup> And perceptual role provides reason to describe the states that occur in perceiving as having mental quality whether or not the perceiving is conscious.

The mental quality of red occurs when one consciously sees something red under suitably standard conditions; so we can assume that the very same mental quality occurs when an individual subliminally sees a stimulus as being red. Indeed, priming evidence in laboratory settings often shows that an individual's nonconscious psychological processing distinguishes colors and other qualitative features of stimuli just as in conscious perceiving. There is every reason to conclude that mental qualities occur even when perceiving fails to be conscious.

Some might express puzzlement about what mental qualities could possibly be if they were to occur without being conscious. But this is just an expression of the consciousness-based view that access to mental qualities is solely by way of consciousness. On that view, we could in principle have no access to nonconscious mental qualities, and so no conception of them. But we cannot adopt the consciousness-based approach to mental qualities in evaluating the competing perceptual-role approach. We must evaluate each approach independent of any assumptions that rely on the other.

It's sometimes held that when theorists appeal to nonconscious mental states, those states are not occurrent mental states at all, but only dispositions to be in such states. On that view, the nonconscious states that figure in masked priming and subliminal perception generally and in blindsight exhibit no mental qualities; rather, they are simply dispositions for states with mental qualities to occur. But it's unclear what that insistence might

rely on other than the consciousness-based view; we would need independent substantiation of that claim to use it in evaluating the respective merits of the perceptual-role and consciousness-based approaches.

Consciousness-based objections to the perceptual-role approach to mental qualities will have no force without independent support. And we can't derive support for such objections from intuitions, such as the conceivability of inversion or zombies or the hard problem or explanatory gap, since they rely themselves on the consciousness-based approach to mental qualities.

Nonetheless, there are reasonable concerns to raise about a perceptual-role theory. Is there any developed perceptual-role theory that taxonomizes mental qualities as well as a consciousness-based theory? And is such a theory independently defensible? In §III I'll sketch a perceptual-role theory that does taxonomize mental qualities in as accurate and fine-grained a way as any appeal to consciousness. And I'll argue that that theory fits comfortably with both empirical findings and our commonsense views about mental qualities and perceiving.

The distinction between consciousness-based and perceptual-role theories differs from Block's distinction between theories that hold that there is an aspect "of conscious experience that goes beyond the intentional, the cognitive and the functional" and theories that deny that.<sup>22</sup> Outstripping "the intentional, the cognitive and the functional" is one thing; being accessible only by way of consciousness another. But a consciousness-based approach doubtless underlies the insistence by Block and others that an aspect of "phenomenal character" does outrun the intentional, cognitive and functional, since only consciousness could reveal such an aspect of conscious experience.

A perceptual-role theory need not do justice to the intuitions discussed in §I, since they rely on a consciousness-based approach. Still, a perceptual-role theory must in some way account for the way mental qualities do occur in conscious perceiving.

This is crucial. On a perceptual-role theory, mental qualities occur in both conscious and nonconscious perceiving, and they cannot on such a theory be intrinsically conscious. So a perceptual-role theory is unlikely by itself to be able to explain how it is that mental qualities sometimes occur consciously—why there is sometimes something it's like for one to be in states that exhibit mental qualities. But we can, it turns out, explain that by supplementing a perceptual-role theory with an independent theory of consciousness; that is the task of §IV.

### III. Quality-Space Theory

In perceiving things, we discriminate among various perceptible properties. As already noted, such discriminations occur both consciously and

not. We use vision to discriminate among colors and the spatial properties of visible size, shape, and location, audition to discriminate among pitches, timbres, and origins of sounds, and other sense modalities for perceptible properties special to those senses.

To discriminate among perceptible properties, a creature must be able to be in states that differ at least as finely as the perceptible properties being discriminated. And the relative differences among those states must reflect those among the perceptible properties. We can determine for any two properties a creature can perceive whether the creature can discriminate between them. And this allows us to construct a quality space in which the distance between any two perceptible properties is a function of how many properties between the two the creature can discriminate.<sup>23</sup>

The differences among the states in virtue of which the creature can perform these discriminations must reflect the differences among perceptible properties that the creature can discern. So the quality space that reflects the perceptible properties a creature can discriminate by a particular sense modality will also determine the perceptual states that make such discriminations possible. And if mental qualities play a role in perceiving, it's natural to identify them with the properties in virtue of which perceptual states differ. The quality space that captures the similarities and differences among the perceptible properties a creature can discriminate will also describe the mental qualities that figure in such discrimination. And this gives us an account of mental qualities in terms of the quality space that describes a creature's ability to discriminate.

Discrimination among the properties perceptible by a modality occurs even when the relevant perceptual states aren't conscious. So an account of mental qualities based on the quality spaces such discrimination determines does not require that the qualities occur consciously, and it does not therefore appeal in any way to first-person access to the qualities. Even when the mental qualities that figure in such discrimination are conscious, it is not their being conscious that enables such discrimination, but only the qualitative differences among them. And there is extensive evidence that individuals respond differentially to stimuli they're not consciously aware of.<sup>24</sup>

The use of such a quality space to describe the color properties perceptible by humans is standard, but the foregoing approach applies to every sensory modality. For any sense modality that discriminates among a range of perceptible properties, we can use matching or just noticeable differences to construct a quality space of all the discriminable properties in respect of how distant each is from every other. And that quality space of perceptible properties will also determine the mental qualities that enable such perceptual discriminations to be made.<sup>25</sup> This quality-space theory can account for the mental qualities special to any perceptual modality.

The quality spaces that describe the properties perceptible by each of the exteroceptive modalities, vision, audition, olfaction, gustation, and

tactition, determine the corresponding mental qualities. But quality-space theory also applies to the spatial properties of size, shape, and location within a perceptual field that are accessible by each modality.<sup>26</sup> The position of each perceptible spatial property in a quality space is determined by its perceptual discriminability from similar properties, just as with color and the other so-called proper sensibles.<sup>27</sup> And because quality-space theory describes these spatial perceptible properties, it can do so for proprioception and kinesthetic sensation as well.

The spatial properties perceptible by different sensory modalities are of course the same; the physical shapes, sizes, and locations we perceive by sight are the same as those we perceive by touch. But the corresponding mental qualities are not. Vision determines spatial perceptible properties as boundaries of color, whereas tactition determines them as boundaries pertaining to perceptible pressure and texture. The mental qualities that pertain to spatial properties are special to each of the sense modalities.<sup>28</sup> Cross-modal calibration of the spatial properties discerned by each modality must be learned.<sup>29</sup>

Quality-space theory also handles the mental qualities that figure in bodily sensations, such as pain. Pains are distinguishable in respect of intensity, bodily location,<sup>30</sup> and whether they are burning, throbbing, dull, sharp, and so forth. All these variations correspond to differences in stimuli that standardly cause pains. So we can locate the mental qualities in virtue of which we discriminate among pains in a quality space constructed from matching or just noticeable differences among the relevant stimuli.

It's important to forestall a particular misunderstanding about quality-space theory. The consciousness-based approach to mental qualities holds that our only route to mental qualities is the way they are conscious. If that were so, the only way to establish a correspondence between a quality space of perceptible properties and that of mental qualities would be to compare the two spaces and note that they correspond.

But quality-space theory adopts instead the perceptual-role approach to mental qualities, and rejects the claim that we have only first-person access to mental qualities. So quality-space theory does not establish correspondence between the two spaces by comparing them. Rather, it extrapolates from the quality space of perceptible properties to determine the space of the corresponding mental qualities. The two spaces match because mental qualities are the properties of perceptions in virtue of which an individual can discriminate among the relevant perceptible properties.

The quality spaces relevant here don't rely on the physical properties of the stimuli, for example, on the wave-mechanical properties of light or sound. Such a space for color would be one-dimensional along the single parameter of wave length or, equivalently, frequency. The quality spaces that figure here reflect instead only the relative discriminability of the stimuli for a particular individual, perhaps averaged over members of a species, as in a quality space

for humans generally. That requires for color a two-dimension space just to accommodate all perceptible chromaticities, that is, hues and saturations, and a three-dimensional space to accommodate brightness as well.<sup>31</sup>

Nor must the quality spaces for mental qualities reflect properties of the neural processes that subserve, or may even be identical with, those mental qualities. The relevant quality spaces reflect only the similarities and differences among mental qualities determined by a creature's ability to discriminate among various perceptible properties.

Traditional theories have sometimes held that mental qualities resemble the perceptible properties they pertain to. That's plainly not the case; mental color qualities, for example, aren't extended over surfaces of physical objects in the way perceptible colors are, and don't involve anything that literally resembles physical pigments. Gilbert Harman's phrase, 'mental paint', rightly lampoons the idea that experiences of color have properties that literally resemble the physical colors they're experiences of.<sup>32</sup>

It's sometimes held that the only alternative to seeing the properties of perceptual experiences as literally resembling the corresponding perceptible properties is to insist that only one type of property actually exists. Only literal resemblance, it's thought, could make sense of how the two ranges of property correspond to one another.<sup>33</sup> Perhaps then there are only physical colors and no mental qualities,<sup>34</sup> or perhaps only mental color qualities, and no physical colors, at least as we ordinarily conceive of them.<sup>35</sup> But without resemblance, it's held, we cannot accommodate both.

Quality-space theory goes between the Charybdis of resemblance and the Scylla of denying that there are two types of property. Mental qualities don't literally resemble perceptible properties; rather they occupy locations in quality spaces that correspond to locations in the matching quality spaces that define perceptible properties. There is no resemblance of individual properties, but there is a kind of resemblance at the level of entire families of properties.

Quality-space theory taxonomizes mental qualities by their role in such discriminations. So the theory taxonomizes mental qualities at least as finely as any consciousness-based approach can. Consider distinguishing two mental qualities by the way they appear to consciousness. Whatever difference consciousness reveals will correspond to a difference between stimuli that those mental qualities could result from. Consciousness reveals no difference among mental qualities that doesn't correspond to a difference among perceptible stimuli.<sup>36</sup>

Indeed, quality-space theory taxonomizes mental qualities more finely than consciousness often does. On quality-space theory, the mental qualities that occur in perception differ in respect of the most fine-grained discriminations an individual can make. Consciousness, by contrast, often fails to reveal such fine-grained differences; we seldom consciously discriminate between mental qualities as finely as we can.<sup>37</sup>

Quality-space theory taxonomizes mental qualities by relative location in a space constructed from an individual's ability to discriminate the corresponding perceptible properties. So if there were an axis with respect to which that quality space is symmetrical, it would be impossible to distinguish stimuli on one side of that axis from stimuli on the other. Quality-space theory would represent the properties on the two sides of the axis as indistinguishable. The qualities on one side of the axis would be redundant with respect to those on the other. But inversion around an axis would be undetectable only if that inversion preserved the relative locations of qualities on each side of the axis, and that could happen only if the space were symmetrical around the axis. So quality-space theory precludes undetectable inversion of mental qualities.

This is not surprising. The possibility of undetectable inversion rests on the consciousness-based approach to mental qualities, on which such qualities are accessible only by way of consciousness. A perceptual-role theory, by contrast, accommodates the occurrence of mental qualities in nonconscious perceiving, and thereby undermines any consciousness-based approach. So we can expect that no perceptual-role theory, such as quality-space theory, will accommodate undetectable inversion. And quality-space theory goes a step further in giving concrete reasons why such inversion is impossible.

But even if it's impossible, is it nonetheless conceivable? We can, it seems, conceive of many things that are not possible. But as noted in §I, the case for conceivability is weak. What's conceivable about some type of thing depends on what concept of that thing we're operating with. If our concept of mental qualities dictated accessibility only by way of consciousness, undetectable inversion would be conceivable, but not if that concept involves an essential tie with perceptual role, and hence a quality-space way of taxonomizing.

We can conceive of water's being a basic physical substance only given a highly impoverished concept of water. Similarly, undetectable inversion is conceivable only given a concept of mental qualities on which they have an essential tie to consciousness, but not to perceiving. Since mental qualities plainly do play a pivotal role in perceiving, that is arguably an impoverished concept of mental qualities. Indeed, the impoverished character of that concept is evident in the uninformative characterization, evoked by the Louis Armstrong's epigram, which a consciousness-based approach must give of mental qualities. There is little on a approach to say about their nature. Undetectable inversion is conceivable only on such an impoverished concept.

Though undetectable inversion is precluded on quality-space theory, a related issue arises, which seems to present a problem. Quality-space theory describes mental qualities by location in a particular quality space. But suppose there are two types of perceptible property a creature can discriminate among, and the discriminations the creature can make among qualities of each type define indistinguishable quality spaces. Suppose, for

example, that the perceptible colors and sounds a creature can distinguish determined indistinguishable quality spaces. If there is nothing to mental qualities but location in a quality space, the mental qualities in the visual space would be identical with those in the auditory space.

One way out would be to stipulate that quality-space theory precludes a creature from having two indistinguishable quality spaces. That's not altogether unreasonable; it seems independently unlikely that the perceptible differences among distinct types of stimulus would be the same. And perhaps that's good enough.<sup>38</sup> But rather than let the theory be held hostage to empirical findings, we can just add a condition for taxonomizing mental qualities. We can supplement location in a quality space by the dependence of the mental qualities in question on the operation of a particular sensory modality. Since quality-space theory is a type of perceptual-role theory, appeal to role in a particular modality is wholly within the spirit of the theory.

The appeal to perceptual role not only undermines the intuition about undetectable quality inversion; it also undercuts the conceivability of zombies and claims of an explanatory gap or a hard problem. Quality-space theory helps take that a step further. If mental qualities are defined by a quality-space constructed from a creature's discriminative ability with a range of perceptible properties, zombies are neither possible nor conceivable. Any creature that can discriminate in the relevant ways will be in states that exhibit mental qualities.

The hard problem concerns the apparent difficulty in saying why a particular neural occurrence should subserve the mental qualities it does, or indeed any at all. The difficulty in explaining these things allegedly results in an explanatory gap. But on quality-space theory, there is no difficulty. Each neural occurrence subserves the mental quality it does because of its role in the discrimination of the relevant perceptible properties.

But these consequences of quality-space theory arguably miss the point. The mental qualities that quality-space theory determines occur without being conscious. And all the intuitions about inversion, zombies, the hard problem, and the explanatory gap pertain just to conscious mental qualities. So to see whether quality-space theory tells against those intuitions, we must consider mental qualities in their conscious form. And that requires addressing the more general concern about whether quality-space theory can, supplemented by another theory, account for there being something it's like for one to be in conscious qualitative states.

#### **IV. Higher-Order Awareness**

Arguments for theories of consciousness typically appeal to the phenomenological appearances. Any theory of consciousness must of course

save those appearances, but saving the appearances need not involve taking those appearances to exhaust the relevant mental reality.

The consciousness of mental states is mental appearance; it is the way one's mental life subjectively appears to one. But appearances can be inaccurate, no less with the mental than in any other realm. A consciousness-based theory suggests otherwise; if we know about mental qualities only by way of what consciousness reveals, perhaps there is after all nothing to mental reality beyond mental appearance.<sup>39</sup> But we need not acquiesce in the dictates of consciousness-based theories, since a tenable perceptual-role alternative is available.

There are, moreover, disadvantages to consciousness-based theories when it comes to explaining consciousness. Identifying mental reality with the mental appearance that consciousness delivers results in a stipulative exclusion of the nonconscious mental states that occur, for example, in subliminal perception and blindsight. And by ruling that mental states are all conscious, it likely precludes explaining what that consciousness consists in and why mental states are sometimes conscious, since an explanation that appeals only to conscious mentality would be uninformative.

It's sometimes thought that a theory of consciousness must use the phenomenological appearances in explaining consciousness itself. But that's a mistake. The phenomenological appearances are crucial in providing a catalog of data that a theory of consciousness must explain. But such a theory need not appeal solely to those appearances in explaining consciousness. The appearances are the *explananda* of consciousness, not the *explanantia*.

When a mental state is conscious, one is aware of being in that state. One can report that one is in the state, and say something about its mental character. By contrast, when we have evidence that somebody is in a mental state but that individual is unaware of being in it and denies that it occurs, we have evidence that the the mental state that individual is in is not conscious. No state of which an individual is wholly unaware, and so cannot report, is conscious.

These considerations provide a necessary condition for a state to be conscious: A mental state is conscious if one is aware of that state, more precisely, if one is aware of oneself as being in the state. The more precise formulation is important, since it's unclear what it would be simply to be aware of a token state independently being aware of it as a state of some individual. And the individual that figures in a mental state's being conscious is plainly oneself. The more precise formulation will, moreover, turn out to be useful in what follows.

I've argued elsewhere<sup>40</sup> that the way one is aware of a state when that state is conscious consists in having a thought to the effect that one is in that state. Our awareness of conscious states seems subjectively to be unmediated and direct, and we can explain that by stipulating that these higher-order thoughts do not themselves seem subjectively to rely on any conscious

inference or observation. And we can explain why we are rarely aware of these higher-order thoughts; that would require a third-order thought about a second-order thought, and we can assume that these seldom occur. Invoking higher-order thoughts explains, moreover, the connection between consciousness and reportability, since reporting a state means verbally expressing a thought that one is in that state. Still, the specific way we are aware of our conscious states will not matter in what follows; so I won't speak in what follows of higher-order thoughts, but in a more generic way of *higher-order awarenesses (HOAs)*.

If one is not in any way aware of a mental quality, there is nothing it's like for one to be in a state that exhibits that mental quality. There being something it's like for one to be in a qualitative state is due to the HOA in virtue of which one is aware of that state. Qualitative states, moreover, are conscious in respect of the mental qualities they exhibit. So the HOA in virtue of which one is aware of a qualitative state must represent that state as having the relevant mental qualities. And since mental qualities are taxonomized by location in a suitable quality space, HOAs must represent them in respect of such locations. When qualitative states are conscious, one is aware of their distinctive mental qualities in terms of the location those qualities have in a suitable quality space.

There is evidence, independent of quality-space theory and a higher-order theory of consciousness, that this is so. It's well-known that we can identify and distinguish mental qualities that occur together much more finely than we can when the mental qualities occur in temporal succession.<sup>41</sup> This finding is commonly presented in connection with color qualities; one can discriminate mental qualities of similar shades in far finer grain if they occur together than one after the other. But the effect occurs with other modalities as well.<sup>42</sup>

This effect is often held to result from a limitation of memory<sup>43</sup>; it's thought that memory doesn't retain mental qualities in as fine-grained a way as perception presents them, and so cannot enable distinguishing very similar qualities. Whatever the case about that, the result shows that we're aware of mental qualities in exactly the way that's implied by quality-space theory combined with a HOA theory of consciousness. Mental qualities are taxonomized in respect of their relative position in a quality space. So comparing two concurrent mental qualities makes it possible to discriminate them in a far more fine-grained way than if one has access to them one at a time.

The result about simultaneous and successive occurrence will seem odd if one assumes that mental qualities are intrinsic properties of qualitative states; comparing the qualities would then make no difference to how they seem to one. But if quality-space theory is right, and mental qualities are taxonomized in respect not of their intrinsic nature, but rather their relative location in a quality space, the result is to be expected. Fine-grained differences among

mental qualities will then be a relative matter, and our awareness of them will follow suit. The effect about simultaneous and successive occurrence follows from the comparative nature of mental qualities and our awareness of them.<sup>44</sup>

The effect also fits with the way we ordinarily discriminate slightly different shades of color and are aware of them as different. When one sees two such shades, one is aware of them and differentiates and describes them comparatively, for example, in respect of relative brightness or how close one is to another color. We taxonomize mental qualities comparatively, not independently of one another.

It's intuitively inviting to hold that we taxonomize each mental quality noncomparatively as a qualitative atom, independent of any others. Doesn't Mary see a red object as being red, despite not having our full color quality space within which to locate her new sensation? Doesn't consciousness reveal each mental quality on its own, independent of every other? The noncomparative picture is especially tempting when we're taxonomizing qualitative states in a coarse-grained way as simply pertaining to red or blue, as we imagine Mary to do, and not in respect of finely differentiated shades.

There are two questions here, which must be distinguished. One is whether we can individuate token mental qualities one by one, independent of other tokens; the other is whether we can taxonomize those tokens noncomparatively in respect of what type of mental quality it is. Plainly we can individuate each token on its own; consciousness-based and perceptual-role theories agree on that. But if we have access to mental qualities only by way of consciousness, we would have nothing to go on in taxonomizing them except the way consciousness represents each quality. And since consciousness represents them one by one, we would have to taxonomize them in that way as well; comparative taxonomizing would be ruled out.

But that would not do justice to the way we actually taxonomize qualities that differ very slightly, and there is no reason apart from the consciousness-based approach to hold that mental qualities are taxonomized noncomparatively. We are aware of quality tokens one by one, though typing each quality depends on comparisons with others. The subjective impression that the type each mental quality belong to is a noncomparative matter stems from conflating the individuation of qualities with their taxonomy.

Many theorists have held that consciousness is intrinsic to mental qualities, and that it can't be wrong. These claims are natural on a consciousness-based theory. The best explanation of our having access to mental qualities only by way of consciousness is that consciousness is intrinsic to those qualities, and nothing could then contest what consciousness tells us about them.

But we can readily explain the subjective appearance that consciousness is intrinsic to mental qualities and the last word about them on a HOA

theory of consciousness. The HOA is seldom itself a conscious state; to be so, there would have to be an third-order awareness of the HOA. So when a qualitative state is conscious, one is subjectively aware only of a single state, and one's awareness of it seems not to be a matter of a distinct state, but to be intrinsic to the state one is aware of. And even though there are third-person considerations that may force an adjustment of what consciousness tells one about one's mental qualities, there are no first-person considerations apart from consciousness itself. So it seems subjectively as though consciousness is the last word about mental qualities.

The HOA in virtue of which one is aware of a qualitative state when it's conscious is independent of that qualitative state. Otherwise it would be difficult to explain how qualitative states can occur consciously at one time and at another time not.<sup>45</sup> So the combined HOA and quality-space theory allows for one to be aware of a qualitative state inaccurately, as exhibiting a mental quality distinct from the quality it actually has. Indeed, the HOA might even make one aware of oneself as being in a qualitative state that one isn't in at all.

There are mundane types of inaccuracy, as when one is aware of a mental color quality simply as a generic red, for example, and not as the specific shade of red that it is. But more dramatic inaccuracy will presumably be rare, since inaccurate awareness of one's own mental states would occasion a measure of cognitive dissonance. If a HOA awareness of a green sensation occurs in connection with a sensation of red, the sensations and HOA will exert conflicting psychological pulls, and we can expect well-ordered psychological functioning to make such occurrences rare.

Subjectively, it seems as though an inaccurate HOA never occurs. But we have nothing to go on subjectively except the way we're aware of of mental qualities; so subjective phenomenological appearances can't help in determining whether those appearances themselves are occasionally inaccurate.

And there is evidence that inaccurate HOAs do sometimes occur. Change blindness is the somewhat surprising tendency of people not to consciously notice salient changes in their visual fields. John Grimes demonstrated change blindness by relying on changes that occur during the eyes' saccadic movements. No retinal signal reaches primary visual cortex during saccades, and using eye trackers Grimes had changes occur during saccades. One particular striking result involved a salient shift in the color of a parrot between red and green, which 18% of participants failed to detect. Attention doesn't figure, since the parrot is the central object in the presentation, occupying over 25% of the screen. Similarly for many other switches of visible properties.<sup>46</sup>

What occurs subjectively for participants that fail to detect the change? The parrot starts being red, and changes to green. Since the stimulus is green, green is projecting from the retina to primary visual cortex. So we can assume

that a mental quality of green occurs in the participant. But when that change isn't detected, the subjective experience of such a participant continues to be that of seeing red. So the HOA is unchanged from the initial HOA, when the awareness was of a mental quality of red; otherwise the participant would consciously detect the change in color. So when a participant fails consciously to detect that the parrot has changed to green, there is a mental quality of green but a HOA of the relevant mental color quality as being of red. Subjective consciousness cannot reveal such a disparity between mental quality and HOA, but experimental tests can.

How about the possibility of a HOA without any relevant mental quality at all?<sup>47</sup> It's sometimes urged that such a case would pose a problem about what the conscious state is. Suppose I have a HOA of a red sensation and there's no relevant sensation. On the HOA theory, the conscious state is the state I'm conscious of; but that state doesn't exist. And the HOA can't, in most cases at least, be the conscious state, since we're rarely aware of our HOAs. There seems at first sight to be nothing reasonable to say about what the conscious state would be in such a case.

But as noted earlier, the HOA is not simply an awareness of a state; it's an awareness of oneself as being in a state. When one isn't actually in the relevant state, that's still the way it seems to one subjectively; one's phenomenological appearance is that one is in that state. The conscious state is the state one is subjectively aware of oneself as being in, whether or not that state actually occurs. Just as something need not exist for it to be an object of one's thoughts, so a state need not occur in order to be a conscious state, that is, a state one is conscious of oneself as being in.

Consciousness is mental appearance; it is the way one's mental life appears to one. So it is natural to regard as conscious a state one appears phenomenologically to be in, whether or not one is actually in that state. There is no special problem about a HOA that occurs without any relevant mental target.

As noted at the end of §III, quality inversion, zombies, the hard problem, and the explanatory gap all concern mental qualities when they occur consciously, not when they occur without being conscious. How do those things fare given quality-space theory supplemented by a HOA theory of consciousness?

Quality spaces cannot be symmetrical around any axis without making it impossible to distinguish qualities on one side of the axis from qualities on the other. And HOAs make one aware of mental qualities in respect of their relative position in the relevant quality space. So that lack of symmetry carries over to the way the HOAs make us aware of mental qualities. Undetectable quality inversion is accordingly no more possible for conscious than for non-conscious mental qualities. Since a consciousness-based concept of mental qualities cannot accommodate subliminal perceiving, quality-space theory and a HOA theory of consciousness likely reflect our commonsense, folk

concept of conscious qualities. So on that concept, undetectable inversion of conscious mental qualities is also not conceivable.

Mental qualities are detectable by way of their role in perceiving, though it's not obvious that there need be any perceptual impairment due simply to mental qualities' not being conscious.<sup>48</sup> But if an individual reports mental qualities in a way that does not so far as we can determine rely on conscious inference or self-observation, that indicates that the individual has HOAs in virtue of which those mental qualities are conscious. So the undetectable absence of conscious mental qualities characteristic of zombies is not possible. And if quality-space theory and a HOA theory of consciousness does capture our commonsense, folk concept of conscious mental qualities, undetectable absence of conscious qualities will not even be conceivable.

Can we explain why a particular neural state should subserve one conscious quality rather than another, or any at all? The roles the neural states play in perceiving enable us to explain why particular neural states subserve the mental qualities they do independent of consciousness. Similarly, the neural states that subserve HOAs result there being something it's like for one to be in particular types of qualitative state because those neural states subserve awareness of being in those states. Contrary to claims of a hard problem and an explanatory gap, there is no difficulty explaining why particular neural states subserve the conscious mental qualities they do.

Rejecting the unargued and unwarranted assumption that we have access to mental qualities only by way of consciousness makes room for an informative quality-space theory of mental qualities. Combining that theory with a HOA theory of consciousness explains how some mental qualities come to be conscious, and does justice to both our folk conceptions of conscious mental qualities and relevant experimental findings.

## Notes

1. E.g., Sydney Shoemaker: "The Inverted Spectrum," in Shoemaker, *Identity, Cause, and Mind: Philosophical Essays*, 2nd edn., Oxford: Clarendon Press, 2003, 327–357, p. 336, and "Intrasubjective/Intersubjective," in Shoemaker, *The First-Person Perspective and Other Essays*, Cambridge and New York: Cambridge University Press, 1996, 141–154, p. 150.

There is debate about whether conceivability is weaker than possibility, or perhaps coincides with it. See, e.g., David J. Chalmers, "Does Conceivability Entail Possibility?", in *Conceivability and Possibility*, ed. Tamar Szabó Gendler and John Hawthorne, Oxford: Clarendon Press, 2002, 145–200. I'll assume for present purposes that conceivability is indeed weaker.

2. *The Conscious Mind: In Search of a Fundamental Theory*, New York: Oxford University Press, 1996, 94.

3. "On Leaving Our What It's Like," in *Consciousness: Psychological and Philosophical Essays*, ed. Martin Davies and Glyn W. Humphreys (Oxford: Basil Blackwell, 1993), pp. 121–136; reprinted in Ned Block, Owen Flanagan, and Güven Güzeldere, eds., *The Nature of Consciousness: Philosophical Debates*, Cambridge, Massachusetts: MIT Press/Bradford Books, 1997, pp. 543–555, 548; page references below are to Block *et al.*
4. "Facing Up to the Problem of Consciousness," *Journal of Consciousness Studies*, 2, 3 (1995): 200–19; "The Puzzle of Conscious Experience," *Scientific American*, 237, 6 (December 1995): 62–68; *The Conscious Mind*, xii.
5. Indeed, Levine himself notes that the explanatory gap is of a piece with the conceivability of zombies, something he also endorses. *Purple Haze: The Puzzle of Consciousness*, New York: Oxford University Press, 2001, 79.
6. Gottfried Wilhelm Leibniz, *Monadology*, tr. George MacDonald Ross, available at <http://www.philosophy.leeds.ac.uk/GMR/moneth/monadology.html>, §17.
7. "What Mary Didn't Know," *The Journal of Philosophy* LXXXIII, 5 (May 1986): 291–295.
8. It's arguable that whenever there is knowledge that we describe using a 'wh' nominal, such as 'what it's like for one', that knowledge must be sustained by specific knowledge that something is the case. We can't know what or where something is, e.g., unless we know that is is a particular thing or at a particular place. Even learning who somebody is by acquaintance yields some new descriptive information. But it's unclear what new knowledge Mary could gain that would be describable with a 'that' clause. And if there is none, her new knowledge might be nothing more than the circumstance of Mary's first consciously experiencing something that's red.
9. Among the work that is relevant to mental qualities, see Justin Sytsma and Edouard Machery, "How to Study Folk Intuitions about Phenomenal Consciousness," *Philosophical Psychology*, 22, 1 (2009): 21–35; Bryce Huebner, Michael Bruno, and Hagop Sarkissian, "What Does the Nation of China Think about Phenomenal States?," *Review of Philosophy and Psychology*, forthcoming; Justin Sytsma and Edouard Machery, "Two Conceptions of Subjective Experience," *Philosophical Studies*, forthcoming; and Bryce Huebner, "Commonsense Concepts of Phenomenal Consciousness: Does Anyone Care about Functional Zombies?," *Phenomenology and the Cognitive Sciences*, forthcoming.
10. *The Mind and Its Place in Nature*, London: Routledge & Kegan Paul, 1925, p. 284.
11. Block writes: "You ask: What is it that philosophers have called qualitative states? I answer, only half in jest: As Louis Armstrong said when asked what jazz is, 'If you got to ask, you ain't never gonna get to know.'" "Troubles with functionalism," in *Minnesota Studies in the Philosophy of Science*, IX, ed. C. Wade Savage, Minneapolis: University of Minnesota Press, 1978, widely reprinted, §1.3.  
Block's move is endorsed, without the quasi-qualification, by Galen Strawson, "Realistic Monism: Why Physicalism Entails Panpsychism," *Journal of Consciousness Studies*, 13, 10–11 (October–November 2006): 117–128, n. 6; reprinted in *Consciousness and its Place in Nature: Why Physicalism Entails Panpsychism*, Galen Strawson et alii, ed. Anthony Freeman, Thorverton, UK: Imprint Academic, 2006, pp. 117–128.

12. *Jacobellis v. Ohio*, 378 U.S. 184 (1964) (Stewart, J., concurring) (discussing possible obscenity in “The Lovers”).
13. Recall the difficulty (n. 8) in saying just what it is that Jackson’s Mary might learn on first consciously seeing red. Presumably she learns what it’s like for her consciously to see red. But learning ‘wh’ implies learning that something is the case, and it’s unclear how Mary could indicate descriptively what new thing she has learned.
14. E.g., Anthony J. Marcel, “Conscious and Unconscious Perception: Experiments on Visual Masking and Word Recognition,” *Cognitive Psychology* 15, 2 (April 1983): 197–237, and “Conscious and Unconscious Perception: An Approach to the Relations between Phenomenal Experience and Perceptual Processes,” *Cognitive Psychology* 15, 2 (April 1983): 238–300; Bruno G. Breitmeyer, and Haluk Ögmen, *Visual Masking: Time Slices through Conscious and Unconscious Vision*, 2nd edn., New York: Oxford University Press, 2006; and Haluk Ögmen and Bruno G. Breitmeyer, eds., *The First Half Second: The Microgenesis and Temporal Dynamics of Unconscious and Conscious Visual Processes*, Cambridge, Massachusetts: MIT Press, 2006.
15. E.g., Lawrence Weiskrantz, *Blindsight: A Case Study and Implications*, new edition, Oxford: Clarendon Press, 1998; *Consciousness Lost and Found: A Neuropsychological Exploration*, Oxford: Oxford University Press, 1997; and “Unconscious Vision: The Strange Phenomenon of Blindsight,” *The Sciences* 32, 5 (September/October 1992): 22–28.
16. See, e.g., Weiskrantz, “Pupillary Responses With and Without Awareness in Blindsight,” *Consciousness and Cognition* 7, 3 (September 1998): 324–326.
17. James Danckert and Yves Rossetti, “Blindsight in Action: What Does Blindsight Tell Us about the Control of Visually Guided Actions?,” *Neuroscience and Biobehavioural Reviews*, 29, 7 (2005): 1035–1046.
18. Beatrice de Gelder, Marco Tamietto, Geert van Boxtel, Rainer Goebel, Arash Sahraie, Jan van den Stock, Bernard M. C. Stienen, Lawrence Weiskrantz, and Alan Pegna, “Intact Navigation Skills after Bilateral Loss of Striate Cortex,” *Current Biology*, 18, 24 (December 23, 2008): R1128–R1129. It’s likely that total blindsight results in pressure to make us aware of blindsight input, as in the prodigious vision-based behavior of an early blindsight monkey (Nicholas Humphrey, *Consciousness Regained*, Oxford: Oxford University Press, 1983, 38, 19–20).  
 Also of considerable interest in this connection is Petra Stoerig, “Cueless Blindsight,” *Frontiers in Human Neuroscience*, 3, 74 (January 2010): 1–8.
19. “The Unconscious” in *The Complete Psychological Works of Sigmund Freud* (henceforth *S.E.*), tr. and ed. James Strachey, London: The Hogarth Press, 1966–74, XIV: 166–215, p. 167.
20. “The Unconscious,” 177–8; also, e.g., *The Ego and the Id*, *S.E.*, XIX: 3–68, pp. 22–3, and *An Outline of Psychoanalysis*, *S.E.*, XXIII, p. 197. To describe feelings as unconscious, he maintained, is to “speak in a condensed and not entirely correct manner” (*The Ego and the Id*, 22) about cases in which the representational character of the feelings is repressed or misrepresented.
21. Again Freud: “[A]ll the categories which we employ to describe conscious mental acts [i.e., intentional states] . . . can be applied” equally well to various states that aren’t conscious (“The Unconscious,” 168).

22. "Mental Paint," in *Reflections and Replies: Essays on the Philosophy of Tyler Burge*, ed. Martin Hahn and Bjørn Ramberg, MIT Press, 2003, 165–200, p. 165.
23. There are two techniques for determining discriminability. One can test whether two stimuli are just noticeably different for a creature or whether they are indistinguishable for the creature. For discussion of the respective merits of these methods, see, e.g., Nelson Goodman, *The Structure of Appearance*, Cambridge, Massachusetts: Harvard University Press, 1951, 3rd edn., Dordrecht-Holland: D. Reidel Publishing Co., 1977, pp. 197–200, 226–7. But the differences between the methods won't matter for present purposes.
24. Even apart from masked priming, nonconscious discrimination occurs in so-called statistical learning (e.g., Nicholas B. Turk-Browne, Justin Jungé, and Brian J. Scholl, "The Automaticity of Visual Statistical Learning," *Journal of Experimental Psychology: General*, 134, 4 (November 2005): 552–564, and Gustav Kuhn and Zoltán Dienes, "Learning Non-local Dependencies," *The Quarterly Journal of Experimental Psychology*, 61, 4 (April 2008): 601–624), as well other paradigms (e.g., Bruno Berberian, Stephanie Chambaron-Ginhaca, and Axel Cleeremans, "Action Blindness in Response to Gradual Changes," *Consciousness and Cognition*, 19, 1 (March 2010): 152–171).
25. For an example of empirical work that sustains such a quality-space approach for olfaction, see James D. Howard, Jane Plailly, Marcus Grueschow, John-Dylan Haynes, and Jay A. Gottfried, "Odor Quality Coding and Categorization in Human Posterior Piriform Cortex," *Nature Neuroscience*, 12, 7 (July 2009): 932–938.
26. See, e.g., my *Consciousness and Mind*, Oxford: Clarendon Press, 2005, pp. 198–201, 206, and 222, and "Color, Mental Location, and the Visual Field," *Consciousness and Cognition*, 10, 1 (March 2001): 85–93; and Douglas B. Meehan, "Spatial Experience, Sensory Qualities, and the Visual Field," in *Proceedings of the Twenty-Third Annual Conference of the Cognitive Science Society*, ed. Johanna D. Moore and Keith Stenning, Mahwah, N.J.: Erlbaum, 2001, 623–627, and "Qualitative Character and Sensory Representation," *Consciousness and Cognition*, 11, 4 (December 2002): 630–641.
27. I.e., perceptible properties accessible only by one modality; Aristotle, e.g., *de Anima*, II.6, 418a11.
28. Aristotle held that they are the same. But he inferred that because he held that perceiving consists of the organism's literally taking on the perceptible properties of the object (*de Anima* II.5 418a4, II.11 423b31, II.12 424a18, III.2 425b23). On that view, mental qualities are literally identical with the corresponding perceptible properties; so since perceived spatial properties are the same across modalities, the corresponding mental qualities would be as well.
29. Most experimental work is done with older infants, in which some learning may already have occurred; for work on very young infants, see Kathleen W. Brown and Allen W. Gottfried, "Cross-modal Transfer of Shape in Early Infancy: Is There Reliable Evidence?," *Advances in Infancy Research*, 4, 1986: 163–170, 240–245. But see also Susan A. Rose and Esther K. Orlian, "Asymmetries in Infant Cross-Modal Transfer," *Child Development*, 62, 4 (August 1991): 706–718.

Andrew N. Meltzoff tested one-month-old infants on two types of pacifiers, both spherical in shape but one with eight or so small protuberances, and

concluded that they already associate tactile with visible shapes without prior learning. “Molyneux’s Babies: Cross-Modal Perception, Imitation and the Mind of the Preverbal Infant,” in *Spatial Representation: Problems in Philosophy and Psychology*, ed. Naomi Eilan, Rosaleen McCarthy, and Bill Brewer, Oxford: Blackwell Publishers, 1993, pp. 219–235. But Meltzoff’s dismissal of prior learning at one month is unfounded, since the small protuberances have roughly the shape of tiny nipples, with which one-month-olds will have had extensive, salient experience, both visual and tactile.

30. See my “Consciousness, the Self, and Bodily Location”, *Analysis*, 70, 1 (April 2010).
31. The 1931 International Commission on Illumination (CIE) chromaticity diagram, widely available online, is still relatively standard for the former; there are a variety of ways to incorporate brightness into a three-dimensional space.
32. “The Intrinsic Quality of Experience,” *Philosophical Perspectives* IV (1990): 31–52, p. 39. Harman also argues there that such experiences have no qualitative properties that we’re aware of, which I’ll contest in §IV.  
Block uses the idea of mental paint to argue for his view that there is an aspect of phenomenal character that outstrips representational content (“Mental Paint”); see also Block, “Attention and Mental Paint,” *Philosophical Issues*, this issue).
33. Frank Jackson, *Perception*, Cambridge: Cambridge University Press, 1977.
34. E.g., C. L. Hardin, *Color for Philosophers: Unweaving the Rainbow*, Indianapolis: Hackett Publishing Co., 1988; expanded edn., 1993.
35. E.g., Paul A. Boghossian and J. David Velleman, “Color as a Secondary Quality,” *Mind* XCVIII, 389 (January 1989): 81–103.
36. Indeed, it’s likely that quality-space theory will individuate mental qualities even more finely than a consciousness-based theory could, since it’s likely that subliminal perception sometimes draws qualitative distinctions unavailable to conscious perception.
37. It’s even possible that subliminal perception sometimes draws qualitative distinctions that outstrip an individual’s ability to discriminate consciously among mental qualities. But it’s difficult to test the fineness of subliminal discrimination, since subliminal perceiving typically results from somewhat weaker neural signal strength. Still, there is some evidence of exceptionally fine-grained nonconscious vision; see Bruno G. Breitmeyer, Tony Ro, Neel S. Singhal, “Unconscious Color Priming Occurs at Stimulus—Not Percept-Dependent Levels of Processing,” *Psychological Science*, 15, 3 (May 2004): 198–202; Bruno G. Breitmeyer, Tony Ro, Haluk Ögmen, Steven Todd, “Unconscious, Stimulus-Dependent Priming and Conscious, Percept-Dependent Priming with Chromatic Stimuli,” *Perception and Psychophysics*, 69, 4 (May 2007): 550–557; Tony Ro, Neel S. Singhal, Bruno G. Breitmeyer, Javier O. Garcia, “Unconscious Processing of Color and Form in Metacontrast Masking,” *Attention, Perception, and Psychophysics*, 71, 1 (January 2009): 95–103; and Bruno Berberian, Stephanie Chamberon-Ginhaca, and Axel Cleeremans, “Action Blindness in Response to Gradual Changes,” *Consciousness and Cognition*, 19, 1 (March 2010): 152–171.

Still, see Tony Ro, Neel S. Singhal, Bruno G. Breitmeyer, and Javier O. Garcia, “Unconscious Processing of Color and Form in Metacontrast Masking,”

*Attention, Perception, and Psychophysics*, 71, 1 (January 2009): 95–103, and Bruno Berberian, Stephanie Chambaron-Ginhaca, and Axel Cleeremans, “Action Blindness in Response to Gradual Changes,” *Consciousness and Cognition*, 19, 1 (March 2010): 152–171, for suggestive findings.

38. The mental qualities that pertain to spatial perceptible properties would not tell against that move. Even though the quality spaces for spatial mental qualities might be the same for distinct modalities, spatial mental qualities are tied to nonspatial mental qualities, such as those that pertain to color and pressure. And the connection with quality spaces that define those nonspatial mental qualities would differentiate the spatial qualities.
39. As Thomas Nagel, e.g., maintains; “What Is It Like to Be a Bat?” *The Philosophical Review* LXXXIII, 4 (October 1974): 435–450, p. 444.
40. E.g., *Consciousness and Mind*.
41. E.g., Rita M. Halsey and Alphonse Chapanis, “On the Number of Absolutely Identifiable Spectral Hues,” *Journal of the Optical Society of America*, 41 (1951): 1057–1058, and Joaquín Pérez-Carpinell, Rosa Baldoví, M. Dolores de Fez, José Castro, “Color Memory Matching: Time Effect and Other Factors,” *Color Research and Application*, 23, 4 (August 1998): 234–247. For additional references and discussion, see Diana Raffman, “On the Persistence of Phenomenology,” in *Conscious Experience*, ed. Thomas Metzinger, Exeter, UK: Imprint Academic, 1995, pp. 293–308.
42. Edward M. Burns and W. Dixon Ward, “Intervals, Scales, and Tuning,” in *The Psychology of Music*, ed. Diana Deutsch, New York: Academic Press, 1982. See Raffman for other results.
43. Raffman calls it the memory constraint (294).
44. Thomas Metzinger has argued that the memory constraint undermines any version of a HOA theory on which the HOA is conceptual, as on my higher-order-thought theory. He infers that if memory can’t capture the fine-grained differences among mental qualities, those differences can’t be captured conceptually. *Being No One: The Self-Model Theory of Subjectivity*, Cambridge, Massachusetts: The MIT Press/A Bradford Book, 2003, 70.

It’s not obvious that limitations of memory reflect conceptual limitations. But that aside, there need not be concepts for every distinct mental quality, since individuating them often depends on comparative concepts.

45. Uriah Kriegel (e.g., *Subjective Consciousness: A Self-Representational Theory*, Oxford: Oxford University Press, 2009) and Rocco Gennaro (e.g., *Consciousness and Self-Consciousness: A Defense of the Higher-Order Thought Theory of Consciousness*, John Benjamins Publishers, 1996) have advanced higher-order theories on which the higher-order content is intrinsic to the conscious states themselves. Such theories face difficulties, such as evidence that the onset of qualitative states occurs before they become conscious (e.g., Benjamin Libet, *Mind Time: The Temporal Factor in Consciousness*, Cambridge, Massachusetts: Harvard University Press, 2004, ch. 2, and “Neuronal vs. Subjective Timing, for a Conscious Sensory Experience,” in *Cerebral Correlates of Conscious Experience*, ed. Pierre A. Buser and Arlette Rougeul-Buser, Amsterdam and New York: North Holland, 1978, 69–82).

In any case, the issues discussed here arguably also arise for such intrinsicist higher-order theories.

46. "On the Failure to Detect Changes in Scenes across Saccades," *Perception*, ed. Kathleen Akins, New York: Oxford University Press, 1996, pp. 89–110.
47. The line separating disparity between HOA and mental quality and HOA without any relevant mental quality is somewhat arbitrary. If one has a sensation of red and is aware of oneself as having a sensation of green, is one aware inaccurately of that sensation of red? Or is one aware of oneself as having a sensation of green that one does not actually have? But that issue won't matter for present purposes.
48. Perceptual states that fail to be conscious often are also perceptually impaired, but not always (see n. 37). When perceptual impairment and failure to be conscious do co-occur, they likely aren't directly connected, but rather result from a single cause, such as low neural signal strength. So even when perceptual acuity is unimpaired, other factors may prevent formation of a HOA, and so block consciousness.

This raises the question whether there is any utility to a qualitative state's being conscious. For parallel doubts about the utility of intentional states' being conscious, see my "Consciousness and Its Function," *Neuropsychologia* 46, 3 (2008): 829–840.