

To appear in Revue Roumaine de Philosophie,
67, 2 (July-December 2023), special issue ed.
Marius Augustin Draghici

Penultimate version

QUALITIES AND CONSCIOUSNESS

David Rosenthal

Philosophy and Cognitive Science

Graduate Center, City University of New York

Postal address: 425 Riverside Drive

New York, NY 10025-7731, USA

Email: davidrosenthal1@gmail.com

Word count: 9,695

Number of figures: 1

Abstract: Many today take mental qualitative character to be intrinsically conscious. But that view not only lacks serious support, but also results in significant undesirable consequences. That view can't be supported by claiming that we know about mental qualities only by way of consciousness, since we have reliable third-person knowledge about the qualitative states of others, and so about their mental qualities. Nor can so-called intuitions provide support, since they are arguably disguised theoretical claims cast as appealing one-liners. And taking mental qualities to be intrinsically conscious results in being unable to say anything informative about their nature, making it seem ineluctably mysterious. Happily, the view that consciousness is intrinsic to qualitative character is wholly optional. Compelling empirical findings demonstrate the occurrence of mental qualitative character that isn't conscious. So the way is open for an informative account of mental qualities based not on their being conscious, but on their role in perceptual discrimination. Such an account is richly informative about the nature of qualitative character and avoids any sense of mystery. And it fits well with an independent theory that explains both what it is for a state to be conscious and how conscious qualitative states differ from those that aren't conscious.

I. Qualities and Consciousness

Acknowledgment: This article is expanded from a talk presented at a conference in honor of Margaret Atherton and Robert Schwartz at the University of Wisconsin-Milwaukee, April 27, 2019. I am grateful for useful discussion on that occasion.

How is mental qualitative character related to consciousness? Today the most widely accepted view is that consciousness is simply built into qualitative character. Indeed, many now hold that no other view is even conceivable. On this popular view, qualitative character cannot occur at all without consciousness; consciousness is an essential aspect of every mental quality. Some even urge that we should think of mental qualities as types of consciousness, so that consciousness is simply a determinable of which the various qualities are determinates.

The view that consciousness is intrinsic to mental qualitative character sometimes arises, somewhat surprisingly, even with those who champion unconscious mental processing. Thus Freud, who held that "[t]he mental, whatever its nature may be, is in itself unconscious" (1966-74c, p. 283), nonetheless also flatly denied that feelings can occur unconsciously (1966-74a, pp. 177-8; 1966-74b, p.22).

Many who hold that consciousness must be built into mental qualitative character think so because it's also thought that we can learn about it only from first-person access. If we can't learn about mental qualities any other way, perhaps first-person access does fully reveal the nature of qualitative states. And the best explanation for that might be that mental qualities are indeed intrinsically conscious. Epistemological considerations are in this special case held to determine the metaphysics, something that most investigators strongly resist for phenomena of any other type.

But that line of thinking is unconvincing. We often do have third-person access to what qualitative states others are in: You can, for example, come to know from my behavior that I'm in pain. When you do, it will typically seem to you that you directly see that I'm in pain. And the state you come to know I'm in is a state of the same type as the state to which I have first-person access. It's undeniable that we do know in this way about many qualitative mental states of others.

In addition, if first-person access is conceived of as based on intrinsic consciousness, consciousness won't reveal any causal ties qualitative states have with behavior or stimuli. So first-person access so conceived does not reveal all the properties of those states. In particular, it does not reveal the very properties needed for third-person access. One might deny that those causal ties are part of the nature of qualitative states. But it's not clear what non-question-begging reason there could be for that denial. Those causal ties are by no means accidental, and they are typically highly reliable about others' states. And we often understand the nature of something by appeal to its causal proclivities. So we should count such ties as an aspect of those states' distinguishing natures.

The view that consciousness is built into qualitative character also makes a mystery of mental qualities, a mystery reflected in the idea that there's an explanatory gap (Levine 2001) or a hard problem (Chalmers 1995), and that mental qualities are undetectably invertible. That's because if being conscious is intrinsic to qualitative character, consciousness will be the first and last word about its nature; consciousness would then

dispositively overrule any conflicting information from any other source. And since consciousness, conceived as intrinsic, doesn't reveal connections that qualitative character has with anything else, we cannot then even give an informative description of what mental qualitative character consists in or of what the particular mental qualities are. It wouldn't help to say that some particular mental quality is what results from seeing, for example, a red tomato, since on undetectable inversion that quality could vary from one individual to the next.

This consequence of intrinsicism is conceded by many who hold that consciousness is intrinsic to mental qualities. Thus Ned Block, who holds that view in connection with his widely adopted notion of phenomenal consciousness, acknowledges that he "cannot define [phenomenal consciousness] in any remotely noncircular way" (1995, p. 230). As he puts it elsewhere, we can say nothing about what qualitative character is other than Louis Armstrong's quip about jazz: "If you gotta ask, you ain't never gonna get to know" (1978, p. 281). "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). Similarly, proponents of Frank Jackson's Mary (1986) are strikingly evasive about just what it is that Mary is supposed to learn from consciousness when she first sees red.

By contrast, if consciousness is not built into qualitative states, these mysteries do arise. There will then be qualitative states that aren't conscious at all, and there's no difficulty in explaining for unconscious states why it is that a particular brain state is associated with a particular mental quality, as opposed to a different mental quality or none at all. And undetectable quality inversion will seem possible, or indeed even conceivable, only if consciousness is built into mental qualities, so that we can know about them only by way of first-person access. Having third-person knowledge about mental qualities would readily enable the detection of any inversion. And there is no mystery about inversion on its own; it's mysterious only if it's undetectable. And if consciousness is not intrinsic to mental qualities, we can then also explain what it is for a state to be conscious by appeal to psychological factors that are not themselves conscious. More about that in section III.

The sense of mystery that some see as surrounding qualitative consciousness is of a piece with the difficulty in giving an informative description of what conscious qualitative character is. If one can't informatively describe a phenomenon, we won't be able to explain how it connects with anything else, and we won't be able to type instances of the phenomenon. And the difficulty in describing will by itself suggest a mysterious nature.

Despite all this, many currently see no alternative to the view mental qualities are intrinsically conscious, and so just accept the mysterious results. But that view is by no means obligatory. Nor does it even straightforwardly reflect traditional views about the mind. Traditional writers, from Aristotle through Descartes, Locke, Hume, and Kant, never spoke of mental states' being conscious, but only of our being conscious of our own mental states, conscious of them in a way that is subjectively unmediated.

Traditional writers did typically hold that we're conscious of all our mental states. And it's likely that speaking of mental state as conscious, instead simply of our being conscious of them, started in the late nineteenth century only because it became popular then to countenance mental states that aren't conscious, so that the one-place predicate, 'is conscious', as applied to mental states came to be especially useful.

And the contrast between 'is conscious' and 'is conscious of' matters. Even if being conscious of one's mental states were intrinsic to those states (e.g., Kriegel 2009), the property of being conscious of a state would be a distinct property from all that state's other mental properties. So there would be no reason to see the one-place property of being conscious as built into qualitative character. It would be more natural to think of it as a distinct, accompanying mental property.

In addition, if we cast things in terms of our being conscious of our mental states, there is also no reason to think that consciousness is in any way decisive about a conscious state's other mental properties. Consciousness would then compete with information that's available by way of third-person access. And if the property of being conscious of a state is not built into that state's other mental properties, there's no reason to think we're conscious of every qualitative state. Though traditional writers did for the most part held that we are, it's altogether unclear what serious, non-question-begging reason there could be for that claim if we're thinking in terms of being conscious of those states.

The view that consciousness is built into qualitative character, a view that's widely held only recently and is in any case wholly optional, results in intractable mysteries. Why then is that view now so widely held?

In addressing this question, many appeal simply to intuitions, which Saul Kripke has famously urged give us "more conclusive evidence ... about anything, ultimately speaking," than any other consideration can (1980, p. 42). But intuitions are not pretheoretic common sense. Knowing somebody's theoretical stance about a particular topic invariably enables one to predict reliably what intuitions that person will report having, and conversely.

The best explanation for this is that these intuitions are of a piece with those theoretical positions, and simply embody theoretical assumptions in appealing ways designed to disguise their theoretical nature. And theory can readily override intuitions, as Kripke himself maintains. If anybody were to have an intuition that heat isn't after all mean molecular kinetic energy, that person, Kripke argues, simply isn't thinking about heat at all (140-144). And we can override intuition by theory only if intuitions themselves are at bottom statements of a theoretical position.

Indeed, if intuitions didn't covertly embody tacit theories in this way, they would be merely stipulative, since those who appeal to intuitions happily describe them as not relying on any other considerations, and people also notoriously differ in the intuitions they claim to have. And there is no reason to take mere stipulations seriously. To reverse Daniel Dennett's nice trope (1991), intuitions are in effect theory pumps.

So it's a theoretical assumption, which requires theoretical evaluation, that consciousness is intrinsic to mental qualities, and not instead a distinct mental property that sometimes accompanies mental qualities. But if not intuitions, and if there are no compelling theoretical reasons, how can we explain why that intrinsicist view is now so widely held?

We describe nonqualitative conceptual states, such as thoughts and desires, in terms of their conceptual content, for example, a thought that it's raining or a desire that it rain. It's controversial how to account for such contents; but an appeal to consciousness does not seem helpful. Even phenomenal intentionality relies on the qualitative aspect, not on consciousness of the conceptual content itself. But it might seem, by contrast, that nothing except consciousness could explain what qualitative character is. And if so, perhaps we would be at a total loss to say anything at all about qualitative character unless we took consciousness to be built in.

But if we there is a good way to describe and explain the nature of qualitative character that's independent of consciousness, we would not need to rely on consciousness at all to describe its nature. And there would then be no reason to think consciousness is built into or an aspect of qualitative character.

The dialectic here is crucial. We must not take the claim that consciousness is built into or an aspect of mental qualities as a datum. Serious support is needed for that. And the only available support seems to be that there's allegedly no other way to know about them. Hence Thomas Nagel's notorious doubts about whether "it makes sense ... to ask what my experiences are really like, as opposed to how they appear to me" (1974, 448; Nagel's emphasis), so that "[t]he idea of moving from appearance to reality seems to make no sense" for conscious experiences (p. 444). These remarks of Nagel's about a distinction between appearance and reality in connection with consciousness in effect encapsulate the view that we can give no account of the nature of mental qualities that doesn't rest on consciousness. We'll come back to that in section IV.

But the availability of an account of the nature of mental qualities that's independent of consciousness would in any case undermine these doubts of Nagel's. And an account of mental qualitative character is available that makes no appeal to consciousness. We can give a revealing and full account, in psychologically relevant terms, independently of consciousness, of just what the mental qualities are.

II. Qualities without Consciousness

Many robust experimental findings point to mental qualitative character that occurs without conscious awareness. So we should be reluctant to accept any view on which that can't happen. Subjects guess with greater than 80 per cent accuracy about color stimuli that are masked, stimuli that they report not seeing. But since these masked

stimuli also prime for downstream psychological effects, plainly they were seen (Marcel 1983). So they must have been seen unconsciously. Similarly with blindsight patients for stimuli in their blind field, again with qualitative properties such as color (Weiskrantz 1986, 1997). And there are compelling reasons to see these findings as involving qualitative mental properties, not merely neural or subpersonal properties that are not mental at all.

For an especially dramatic and decisive empirical demonstration of this, consider the findings by Liam Norman and colleagues (2014) that a masked stimulus primed the mask if they match in surface color, but not in spectral reflectance (on this see also Kentridge 2017). Though the mask was consciously perceived, the masked stimulus was perceived only unconsciously. A conscious qualitative state was matched with another state that was totally unconscious. Since the matching occurred in respect of qualitative character, both states must have exhibited that qualitative character. So the unconscious state exhibited qualitative character that occurred unconsciously.

Since mental qualities do occur without being conscious, the way is clear to develop an account of qualitative character that makes no appeal to consciousness. Perceiving involves discriminating among the perceptible properties of stimuli: colors, shapes, sizes, sounds, odors, and so forth. Such discrimination occurs consciously, but also unconsciously. We discriminate stimulus properties in unconscious, subliminal perception no less than in perceiving consciously. And by manipulating stimuli, we can test for just noticeable differences (JNDs), differences between stimuli so small that were the stimuli physically any closer one would be unable to distinguish them at all.

We can then use JNDs for a particular range of stimuli to construct a quality space (QS) that captures all the stimuli in some range that an individual can discriminate. For color stimuli, the QS would reflect the three discriminable dimensions of hue, saturation, and brightness. And visual perception also involve JNDs between the spatial properties of size, shape, and location.

Discriminating stimuli requires one to be in states that differ in ways that reflect the way the stimuli differ for one. That's so independently of whether the discriminating is conscious or subliminal. And we know, for example, from Norman et al (2014), that unconscious perceptual states can be genuinely qualitative, as against subpersonal, nonmental states.

Still, let's first consider the conscious cases. Conscious perceptual discrimination plainly does operate by differences in mental quality. We consciously distinguish stimuli by being in conscious states that differ qualitatively in ways that correspond to perceptible differences among stimulus properties. So in the conscious case, the QS of discriminable stimuli will also map the mental qualities that enable one to discriminate those stimuli: That gives us a theory for what the mental qualities are in the conscious case: They are those properties of perceptual states, mapped by a QS of discriminable properties, in virtue of which an individual can make conscious JND discriminations.

On this QS theory, conscious mental qualities are fixed by relative location in a QS built on JNDs, at bottom, discriminative ability. If one thinks about mental qualities in terms of what it's like, this may seem unintuitive. Aren't conscious mental qualities fixed atomically one by one, independently of all the others, and so independently of any comparative considerations?

No. Conscious perceptual discrimination is a relational matter, a matter of distinguishing each perceptible property from others. And since conscious mental qualities figure in perceptual discrimination, they will also be fixed comparatively. Indeed, to fix conscious mental qualities in a fine-grained way, we often need to compare them introspectively. (More about that in section III.)

JNDs aren't transitive, or even symmetric (Goodman 1951). Moreover, subjects vary in discriminative ability and are conservative in judgments. But all that can be handled (see again Goodman 1951). What matters is whether the qualities that figure in conscious perceiving and are fixed by location in a QS are also responsible for unconscious perceptual discrimination.

And there's compelling a theoretical reason to think so. QS theory makes no appeal to consciousness to fix the conscious mental qualities. It appeals only to the discriminative ability that JNDs reveal. So whether a qualitative state is conscious shouldn't matter at all. In addition, in a number of experimental findings, though mental qualities occur consciously, they aren't sufficiently fine-grained consciously to reveal discriminable differences that subjects can perceive unconsciously. That points to an unconscious aspect of those conscious qualities that enables discrimination, for example, by forced-choice guessing or matching behavior.

An example of this occurs Diana Raffman's (2011) work, in which she presented subjects with adjacent patches alternating different and same, but when different always by less than conscious JNDs. Also, when the patches differed, it was always by way of increasing wavelengths. When subjects judged adjacent patches the same, a disk appeared with a hue randomly matching one of the patches, and subjects were instructed to adjust the disk to match the two patches they had judged identical.

The result was that "subjects' settings of the [central] disk progressed more or less systematically with the wavelengths of the patches, even though the members of the pairs in question had been judged 'same'" (118). Matching was unconsciously more fine-grained than conscious perceiving. Even when subjects consciously judge mental qualities for adjacent patches as same, the matching task revealed that they registered perceptually as different.

So there is a dissociation between how the properties that enable visual discrimination actually operate and how they are for consciousness, that is, how they are in respect of their subjective appearance. Qualitative states evidently can have an unconscious aspect that figures in perceiving, which isn't revealed in subjective appearance. So the property of being conscious cannot be built in.

An advocate of intrinsicism might contend that only the conscious cases are strictly speaking qualitative, and not the properties that underlie the discrimination in Raffman's matching task. But this is highly implausible. The striking aspect of Raffman's investigation is that the patches are all consciously seen. So it must be that some aspect of the qualitative character is responsible for her matching results. But because the matching results are not reflected in subjective awareness, they must be due to unconscious aspects of the conscious qualitative character.

And since conscious discrimination relies on mental qualities, there's no serious reason to think that unconscious discrimination doesn't also. Indeed, if discrimination in the unconscious cases were due only to some subpersonal or merely neural factor, those nonmental properties could also be responsible for discrimination even when the mental qualities are conscious. The conscious mental qualities would then be causally idle, far too high a price to pay to save a theory that is as best optional.

And there is more. Using very brief presentations (in the μsec range), Arnaud Beauny and colleagues (2020) determined thresholds at which subjects go from being able to subjectively detect stimuli to being able also to subjectively identify those stimuli. But even in the prior condition, when subjects can't subjectively identify stimuli, they could objectively identify them well above chance using forced choice. So there is unconscious identification of stimuli that are consciously detectable.

This again shows that a perceptual state can be conscious in respect of some properties of a stimulus, but not enough of those properties to enable conscious identification. And the identification of stimuli, which requires discrimination among stimulus types, is in these cases unconsciously more fine-grained than it is consciously. These results also show that discriminating that isn't conscious can be more fine-grained than when conscious, revealing unconscious JNDs (see also Scott et al 2010).

Might mental qualities figure in conscious discrimination, but not in unconscious discrimination? Might only subpersonal neural processes be operative there? No. As before, if subpersonal neural processes alone worked for unconscious discrimination, why wouldn't they work also for conscious discrimination, again making the conscious qualities idle. A double standard simply isn't workable.

And unconscious discrimination aside, our commonsense picture of conscious mental qualities, that is, of what it's like for one, fits well with a QS account. We typically describe what it's like to see a particular color by locating the quality among other color qualities.

Because QS theory explains mental qualities independently of subjective awareness, it explains qualitative character that occurs in subliminal perceiving and in blindsight. It also shows why undetectable inversion of mental qualities is inconceivable. If there were an axis around which a QS were symmetrical a creature would be unable to distinguish stimuli on one side of from those on the other. Any inversion around an asymmetric axis would be detectable. And there's no axis of symmetry for a QS of any

known modality. A workable concept of mental qualities conflicts with even imagining undetectable inversion.

QS theory provides a nice account of how mental qualities represent. Each mental quality is fixed by a relative location in a QS that corresponds to the location of some stimulus property. So we can take each mental quality to represent that corresponding stimulus property. Mental qualities represent, but they do so in a way quite different from the way concepts and conceptual contents do.

In perceiving, mental qualities represent an object's sensible properties, and conceptual content represents in some way what kind of thing the object is. Representation by mental qualities explains also the apparent appeal of nonconceptual content, and of phenomenal intentionality.

Jacob Berger (2018) has used the QS apparatus to defend representationalism, on which perception represents, but lacks any distinctively qualitative character. Berger's powerful and penetrating discussion is highly compelling, but I have some concern. By avoiding mental qualities, Berger's version of representationalism sidesteps the hard problem. But QS theory already sidesteps it, by casting mental qualities as independent of being conscious. In addition, Berger's representational properties will differ from one another in ways that reflect differences among the stimulus properties that give rise to them. So those representational properties will just be the mental qualities, fixed as on QS theory by relative location in a QS and by how they represent.

A nice byproduct of QS theory: We can extend it to provide an informative way to individuate the sensory modalities, which, unlike traditional proposals, begs no questions. Call a sequence of JND qualities that leads from one quality to another a JND bridge. Then a set of qualities belongs to a single modality if, but also only if, they're all connected by some JND bridge. This turns out to need fine tuning to deal with some odd empirical findings. But it avoids the difficulties of traditional ways of differentiating the modalities, such as sense organs, the physical nature of stimuli or media, and phenomenology (Rosenthal 2015; Young et al 2014).

One might seek to avoid unconscious mental qualities by denying that perceiving itself occurs without being conscious. But the occurrence of conscious perceiving is widely accepted, in both theoretical and empirical work. And it's unclear what theoretical motive there could be for denying that perceiving can occur unconsciously except the view, argued against above, that consciousness is intrinsic mental qualities. And there is in any case ample reason to disregard positive arguments against unconscious perception (Berger & Mylopoulos 2019; Kentridge 2017; Rosenthal forthcoming, §V).

We have competing theoretical stances. One approach relies on fixing mental qualities by perceptual role and the other approach on fixing them by what it's like. Both have some ties with common sense; so "intuition" won't decide between them. But both also make theoretical claims, and so are both subject to theoretical evaluation.

A perceptual-role approach fits better with empirical findings, and also underwrites a rich theoretical elaboration in QS theory, with testable predictions and explanations. One rarely if ever gets a deductive proof with competing theories. But the advantages of QS theory place it well ahead, including avoiding an explanatory gap, the hard problem, and undetectable inversion.

III. Conscious Qualities

Explaining mental qualities independently of consciousness avoids mysteries about qualitative character and consciousness, and fits well with both empirical findings and common sense. But if consciousness is not built into qualitative character, we need to explain why some qualitative states are conscious and others are not. And we would want the explanation to fit comfortably with QS theory. Since QS theory explains mental qualities independently of consciousness, we need to explain mental qualities by one theory and consciousness by a distinct theory. Can such a divide-and-conquer approach enable us to explain both phenomena successfully in terms appropriate to each?

If one is in some mental state but wholly unaware of being in that state, that state cannot be a conscious state. That's how we understand, in both folk and scientific psychological terms, what it is for a mental state not to be conscious. So it's a necessary condition for a state to be conscious that one is aware of being in that state in some way. And since it's reasonable to take being aware of something as equivalent to being conscious of that thing, that fits with the traditional approach to discussing consciousness mentioned earlier, on which we speak about being conscious of things, rather than of consciousness as a one-place property.

That gives us a necessary condition for a state to be conscious. But we can expand on that by specifying in what way one is aware of being in a mental state when that state is conscious. I've argued elsewhere that we're aware of being in a mental state that's conscious by having a thought to the effect that one is in that state, a thought one expresses a first-person report that one is in that state. This higher-order-thought hypothesis has a number of significant explanatory advantages (Rosenthal 2005, esp. chs. 2, 4, 9, and 10). But it's enough for present purposes that a state is conscious only if one has some higher-order awareness (HOA) of being in that state.

This HOA would rarely itself be conscious, since that would require a yet higher-order awareness about it, which would likely be rare. And indeed, we are very rarely aware of any such HOAs. And the HOA cannot rely on any inference or other mental mediation of which we are aware. That last provision explains the subjective sense of immediacy that characterizes the way we're subjectively aware of our conscious states.

Ernest Sosa (2003) and Block (2007) seek to dispel the higher-order character of these HOAs by construing them in a deflationary way: Being aware of conscious states, they

urge, is like smiling a smile. Just as there's only the smile, HOAs are not a distinct occurrence. But that won't do. To explain the contrast between mental states being conscious and not conscious, we need a contrast between there being an awareness or not. The HOA must be a distinct factor. No deflationary view can do justice to this contrast. Indeed, there is no other way to distinguish conscious from unconscious states in distinctively psychological terms. Sosa's and Block's proposal is in effect deflationary about that very distinction, reflecting the view that qualitative character never occurs without being conscious.

A qualitative state is conscious if one has a suitable HOA about it. And QS theory also tells us how that HOA represents the state. One is aware of the state in respect of its relative location in its QS. That fits with the point noted in section II that the comparative way mental qualities are described in QS theory matches the way we typically describe in folk terms what it's like for us to be in a conscious qualitative state. We describe what it's like for us by comparing the target state to other states that occur in response to currently present stimuli or to other states that would arise in response to objects we are familiar with.

Since perception involves the discriminating of stimuli and discrimination is itself a comparative matter, the mental qualities by which we perform such discriminations must be individuated comparatively. And that actually fits with the way we are subjectively aware of mental qualities. A vivid illustration of comparative subjective awareness is the way we're aware of qualities in less fine grain when they occur in succession than when we can compare several of them together. Stimuli with very close hues presented in succession may be consciously indistinguishable from one another, but when presented together the same stimuli are consciously distinguishable, indeed readily so. We're subjectively aware of mental qualities in comparison with one another, enabling us subjectively to assign to them a relative location in a QS. When such comparison is unavailable, it's significantly more difficult to discern relative location, resulting in far less fine-grained discrimination.

We can contrast this with other comparative effects that are due just to the mental qualities, independently of how we're subjectively aware of them. Though there's some controversy about how simultaneous color contrast works, it very likely occurs prior to subjective awareness, and indeed may even be largely retinal (e.g. Soranzo 2016). Thus we see the red squares on the upper and lower diagonals in figure 1 as differing in hue, though the vertical bar reveals that the red hues are identical. We see the upper and lower red squares as different because the different neighboring hues affect them. The apparent difference in the red-square hues is registered prior to and independently of subjective awareness, and subjectivity simply reflects that difference.

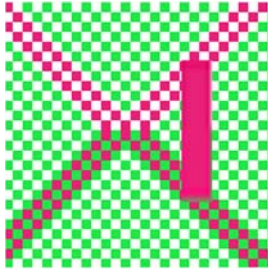


Figure 1

In this type of case, the comparative way our HOAs represent the qualities of our conscious qualitative states reflects the comparative nature of the mental qualities themselves, independently of consciousness. Mental qualities are fixed by relative location; and that is then how we're also subjectively aware of them.

As noted, if one conceives of subjectivity as built into mental qualities, subjectivity so conceived will say little anything informative about their nature, as reflected in Block's colorful appeal to Louis Armstrong's remark about jazz. By contrast, QS and HOA theories together explain what subjective awareness does actually tell us: For each qualitative state it says which other states it resembles and which not so much. Consciousness is informative in a QS way.

The independence of mental qualities from subjective awareness raises what many have seen as a problem. If the two are independent, why can't subjective awareness misrepresent the qualitative character of the state one is actually in? Such independence does allow this to happen, though contrary to some objectors it does not imply that it ever does. Still, we'll see in a moment that such misrepresentation by subjective aware plainly does actually occur.

First, however, let's look at the arguments that purport to show that it can't. Karen Neander (1998) and Joseph Levine (2001) imagine a case in which one has a mental quality of red but a HOA of having a green mental quality. Is what it's like then the having of a conscious red sensation? A conscious green sensation? Something else? Neander and Levine contend that there's no good answer, and so conclude that subjective awareness cannot, after all, be independent of qualitative character.

Levine in particular argues that if what it's like for one is green, the HOA is doing all the work, and the quality plays no role. Similarly, he maintains, if what it's like for one is red, the mental quality does all the work, and the HOA does nothing. But this assumes that all the work pertains solely to what it's like. And that's not so. Indeed, that assumption that it is simply channels the view that consciousness is intrinsic to qualitative character, so that there can only be a single mental factor.

But there are two mental factors, which do distinct things. The HOA does determine what it's like for one, and the mental quality independently determines perceptual processing, that is, how we react both psychologically and in behavior. So somebody

with a mental quality of red but a HOA of being in a green qualitative state would sincerely report seeing green, but would nonetheless be primed for red.

And we've already seen abundant evidence of priming, forced-choice guessing, and matching tests that reveal qualitative character independent of any subjective awareness of the relevant qualitative character. So there's no reason to doubt that such misrepresentation by a HOA could be revealed by suitable tests. No reason, that is, except the unsupported insistence that what it's like for one is all there is to qualitative character.

Still, one might not be satisfied. Could priming really reveal a red qualitative state when an individual is subjectively aware of seeing green? Indeed, how might we even induce such disparate states in a subject? Happily, there is striking change-blindness work by John Grimes (1996, currently being replicated Fallon et al 2020-2023) that speaks to the concern about inducing mental qualities that do differ from what it's like for that subject.

Grimes switched displays during saccades, when virtually no retinal input reaches V1. Many subjects missed the change; in arguably the most dramatic case, 18 per cent of subjects missing a salient change central to the display from green to red. Post-change, red retinal input must reach visual cortex. So the mental quality on its own will then be of red. But when subjects report no change, their subjective awareness presumably remained that of seeing green, a decisive disparity between subjective awareness and actual mental quality. We have an actual example of the very type of case that Neander and Levine imagine, and claim to be impossible.

One might object that subjective awareness on miss trials might not have continued to have content pertaining to the pre-change green color. Subjective awareness might have switched to red, but subjects were still unaware of there have been a change. This is implausible. But plausibility aside, the replication in Fallon et al (2020-2023) has demonstrated that subjective aware on miss trials often retains the pre-change content.

And in other, more traditional work on change blindness, it's been established that subjects are unconsciously aware of changes that they consciously miss (Fernandez-Duque & Thornton, 2000; replicated by Laloyaux et al 2006). Subjects unconsciously see the change even though subjective awareness continues to register the pre-change stimulus. Since visual cortex will often register the post-change stimulus, subjective awareness misrepresents the visual state the subject is in.

The change blindness in Grimes's experimental work is not an attentional effect, since changes occur with central, salient items, which likely attract attention. And in any case, subjective awareness and attention are largely independent. Attention occurs without awareness, even without object attention (Norman et al 2013). Attention simply isn't sufficient for subjective awareness. Familiar quotidian occurrences point to the same conclusion. Although parafoveal vision is poor, it is still conscious well into the periphery; but such visual states are typically unattended in any way, whether by

orientation or object attention. So attention cannot be a necessary condition for consciousness.

This last observation also raises a challenge for global-workspace theories (Dehaene & Naccache 2001), on which a state is conscious if its content is available for downstream processing. There are peripheral visual states that are plainly conscious but for which it is overwhelmingly unlikely that their content is thus available. An argument by Lionel Naccache (2018), may dispel this worry for global-workspace theory by in effect folding into global-workspace theory considerations proper to higher-order theory. But if so it is the higher-order factor that avoids the difficulty.

George Sperling (1960) famously presented subjects very briefly with a three by four matrix of letters. After the letters disappear, subjects report having consciously seen all twelve, but can identify only three or four of them. But if a subsequent tone directs subjects to one row, subjects can then identify three or four in that one row. The tone occurs only after the letters have disappeared. So subjects must register and retain information about most identities.

The question is how they do so. Block urges they do so consciously, since when informally asked that's what they say they do. But that's unconvincing. Subjects likely have no idea how they do it, and conscious imagery is likely the only hypothesis that would occur to them. In addition, if subjects take themselves to retain all the identities by way of conscious imagery, how would they explain the limit of three to four in the cued row? Subjects were not asked about that, making their informal, anecdotal suggestion that they retained identities by conscious imagery largely useless.

It's in any case highly unrealistic to hold that most identities are retained, or even registered, consciously. Perceptions are rarely if ever conscious in respect of all aspects of their qualitative character. Perceiving can be conscious and yet unconsciously register a lot of information. Indeed, that's typically the case. Thus Raffman's subjects unconsciously take in differences of hues consciously judged identical. Most identities are registered unconsciously. This is supported by subjects' perceiving all the items as alphanumeric even when a few of them aren't (Kouider et al 2010). And as for retention, there's compelling empirical evidence that the type of memory that's operative in Sperling's findings and related results is not conscious (Irvine 2014).

Incomplete representation by subjective awareness is a relatively trivial way in which consciousness often misrepresents. And consciousness also misrepresents whenever mental qualities occur without being conscious, since it then represents that those qualities don't occur. There being something it's like for one is its appearing to one that one is in a state with that qualitative character. There being something it's like for one is not due to the qualitative character itself. It is the mental appearance of being in a state with that qualitative character. A HOA makes one aware of oneself as being in a

qualitative state of a particular kind. That's what it is for a qualitative state to be conscious.

Indeed, that's what it is for qualitative states to be lighted up in the way the phrase 'what it's like' is meant to capture. A perceptual qualitative state is lighted up when one consciously sees or hears or otherwise perceives something. So a perceptual state's being lighted up simply is its subjectively seeming to one that one is perceiving that thing, that is, its appearing to one that one is in the relevant perceptual qualitative state. There is nothing more to being lighted up or to there being something it's like.

A standard objection to a higher-order theory of what it is for mental states to be conscious is that a HOA can't make a first-order state conscious unless the HOA is itself conscious. This is simply a misconception of higher-order theories. First-order states don't inherit the property of being conscious from a HOA. Rather, their being conscious simply is a matter of one's being aware of being those states.

And a HOA needn't be conscious to make one aware of their first-order targets, just as a subliminal perception needn't be conscious to make one aware of a subliminally perceived stimulus. In both cases one is aware of the stimulus or the first-order state, though one is not consciously aware of it. One is not, that is, aware of being aware of it.

IV. Methodological Considerations

Some deny that when it comes to consciousness, mental reality and mental appearance can differ, as appearance and reality always do in nonmental cases (e.g., Nagel, pp. 444, 448). But accommodating that distinction turns out to be methodologically pivotal for the study of consciousness. Consciousness is how a stream of mental occurrences subjectively appears to one. What first-person access tells us is how that stream of mental reality subjectively appears.

And we've just seen empirical findings that force that distinction between mental appearance and mental reality. To reject such a distinction is simply an oblique way of insisting on no independent basis that consciousness is intrinsic to mental qualitative character, so that consciousness and qualitative character cannot be distinct mental properties (Rosenthal 2018, 2022).

And as we have seen, relative location in a QS is not only the basis of an informative account of the nature of conscious mental qualities, but also reflects the way we describe, in ordinary folk-psychological terms, what it's like to have a particular qualitative experience. We do so comparatively, by appeal to other relevantly similar experiences, each described in terms of the stimuli that elicit those experiences.

Those, like Nagel, who reject a distinction between mental appearance and reality, also urge that we can't explain consciousness except in cognate terms, such as subjectivity, perspectives, and points of view. And they don't see this as a defect in the conception

of consciousness that they champion. Still, no account that simply redescribes the target phenomenon in essentially equivalent terms can be informative. Compare W. V. Quine's (1951) demonstration of how being confined to a closed family of terms undermines any notion of analyticity. To avoid an uninformative circularity, an account of consciousness must appeal to mental phenomena that are not themselves conscious. To achieve that, we must identify consciousness as the subjective appearance of a mental reality that is itself not conscious, so that mental appearance must be distinct from that mental reality.

Another methodological matter worth stressing: The JNDs that figure in QS theory are differences between stimuli, not between mental qualities, as they are sometimes construed, e.g., by Nelson Goodman (1951). Distinguishing mental qualities is subjective. And we can control stimuli: If they're any closer than JND, they're indistinguishable (on a suitable percentage of trials). Relying on JNDs also implies that similarity and difference aren't primary on QS theory. Those relations are constructed from discriminability, which is far more fine-grained and readily testable.

These considerations lead to another methodological implication, and also an apparent problem. The implication is that QS theory should not invoke the technique of multidimensional scaling (MDS), which relies on subjective similarity judgments. That's partly because of the subjectivity of the judgments, but also because relative similarity on QS theory is derived from discriminability, which can be controlled experimentally. MDS is highly useful as a practical shortcut in constructing a workable QS. But it cannot deliver the theoretically fine-grained and objectively testable results that JNDs can. So it cannot provide the theoretical basis for a scientific account of mental qualities.

The problem stems from applying the JND technique for QSs to Jackson's Mary. Mary, having been confined to seeing only achromatic color stimuli, is then presented with a red stimulus, producing a conscious experience altogether novel to her. The achromatic stimuli she had previously had access to, which are in shades from black to white, are totally unsaturated with respect to any hue. How can we locate that red in Mary's QS?

Casting things in terms of similarities and differences, which are relatively loose and impressionistic, obscures the problem, which emerges clearly when we switch to JNDs. Stimuli are JND when they would be indistinguishable if they were physically any closer. But assuming moderate saturation for the new red stimulus, it could be physically closer to any of the achromatic stimuli and still readily distinguishable from them. How can JNDs fix the location of a new red in Mary's QS?

The QS we would construct from Mary's JNDs would, according to the theory, fix the nature of any mental qualities she might have. But Mary's novel red experience is not JND from any she's ever had. How then can QS theory locate the new red relative to Mary's prior color qualities, all of which are achromatic? MDS could do it, but only in a

highly impressionistic way, and so without the reliable objectivity needed for a sound theory of mental qualities.

Growing up with visual sensations that are all achromatic might well make the new red sensation so unlike all of Mary's prior sensations that it would subjectively seem to her to lie outside her family of visual experiences altogether, somewhat like a distinct modality. Given the problem about JNDs, might that be what we should say?

That's too extreme. We should seek to count it as another visual experience, however hard it may be to locate among the others. And the extreme reaction is avoidable. Mary's new red differs from her earlier achromatic hues in being saturated. So one can construct a path of stimuli that differ very slightly in respect of saturation and which leads from Mary's achromatics to her new red. And neural processes that subserve Mary's vision would allow her to discriminate some, though not all, of the partly saturated stimuli in that path. So the theory can appeal to the potential JNDs that occur in that path to fix the location of the novel red that Mary is actually presented with. Though that actual red isn't close to the actual achromatic hues in respect of degree of saturation, the theory still has resources to fix its relative QS location.

A variant, due to Jacob Martin (in conversation): Mary* is presented with a green stimulus the same number of potential JNDs away from the achromatic stimuli as is Mary's red. It seems pretheoretically clear that what it's like for her would differ in the two cases. What it's like for one to see red and green would differ; so Mary*'s mental qualities should too. Can QS theory meet Martin's variant challenge?

Yes: Neural processes in the two subjects, or even in just one, would allow us to project potential JNDs between each new stimulus and other potential chromatic stimuli, and so to plot different bridges pathways of JNDs for red and green. Potential JNDs again let us fix the relative location of stimuli that we can't fix by appeal to actual JNDs.

There is a final concern about higher-order theories, which has methodological implications. If consciousness were intrinsic to qualitative character, consciousness would occur automatically with each qualitative state. But for consciousness to consist in distinct HOAs, we must explain how the HOAs themselves come to occur. And because mental qualities on their own determine perceptual functioning independently of consciousness, it can't be that HOAs occur because they're needed to do that.

But we can explain how HOA occur. They likely first arise in the earliest stages of infant perceptual development, in which qualitative inputs must constantly be calibrated across modalities. The infant must coordinate visual and tactile qualities pertaining to size and shape, visual and auditory qualities pertaining to location, and visual qualities with olfactory and gustatory qualities and with pleasurable and unpleasant qualities. In each case, the infant must determine whether a mental quality from one modality results from the same property or object as a mental quality from a different modality.

An advocate of a nativist answer to Molyneux's question would deny the need to coordinate qualities pertaining to size and shape, though perhaps not on that account spatial location as well. But no empirical support has been adduced for such nativism, and there is good provisional evidence against it (e.g., Held et al 2011). And the need to coordinate nonspatial qualities from different modalities is in any case indisputable.

In the course of the relatively constant need for cross-modal coordination, it will often initially be unclear whether mental qualities from two modalities result from the same object or property, leaving the infant puzzled. That will arrest smooth perceptual processing for a moment, leading the infant to wonder what's represented by one or another mental quality. And wondering that involves being aware of the relevant qualitative inputs, and those states of awareness are HOAs. HOAs pertaining to qualitative states initially occur due to early cross-modal calibration, and become entrenched with frequent occurrence. HOAs pertaining purely conceptual states, such as thoughts and desires, require a different, more complicated explanation (Rosenthal, 2005, ch. 10).

One might question whether infants do actually wonder about these things. But infants plainly are sometimes puzzled about perceptual inputs, and that puzzlement will often be accompanied by corresponding wondering about the perceptual situation. And how to coordinate inputs from distinct modalities will be a frequent source of such puzzlement, and consequent wondering.

Infants wouldn't conceptualize the mental qualities and what they represent as adult folk psychology conceptualizes them. But an infant can correlate inputs from distinct modalities only by psychologically representing perceptual inputs as resulting from some independent property or object. And each input is itself a qualitative state of the infant, and its resulting from something independent of the infant amounts to its representing that thing. Correlating inputs requires taking psychological account of what qualitative states represent, even if not in terms specific to folk psychology.

Because HOAs pertaining to qualitative states arise in this way, we can predict that qualitative states become conscious to the degree that successful perception requires significant cross-modal calibration. That applies both to humans and to nonhuman animals. So this explanation provides some leverage in coming to understand which perceptual states will be conscious in various types of creature.

Since cross-modal calibration has great utility, one might conclude that a state's being conscious itself confers some distinctive utility. But the utility here attaches just to the calibration, not also to the HOAs. It's unlikely that a state's being conscious does confer significant additional utility (Rosenthal 2008, 2012). And a methodological consequence of that is that we cannot learn about consciousness by investigating some utility it allegedly confers.

QS theory offers a sound way to explain the nature of qualitative character and to distinguish types of mental quality, all grounded in discriminative ability and richly

supported by robust empirical findings. Because the theory makes no appeal to consciousness, it avoids problems that arise when consciousness is construed as built in. Instead, a qualitative state is conscious when there's a HOA in virtue of which one is aware of being in that state in respect of a qualitative property with the relevant location in a QS. And since consciousness is independent of qualitative character, we get an informative explanation of consciousness that appeals only to psychological phenomena that are not themselves conscious. The independence benefits both accounts.

References

- Beauny, Arnaud, Adélaïde de Heering, Santiago Muñoz Moldes, Jean-Rémy Martin, Albert de Beir, and Axel Cleeremans (2020), "Unconscious Categorization of Submillisecond Complex Images," PLOS ONE, 15, 8, doi 10.1371/journal.pone.0236467 (August 12): 1-23
- Berger, Jacob (2018), "A Defense of Holistic Representationalism," Mind & Language, 33, 2 (April): 161-176.
- 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): 481-499.
- 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.
- Chalmers, David J. (1995), "Facing up to the Hard Problem of Consciousness," Journal of Consciousness Studies, 2, 3: 200-219.
- Dehaene, Stanislas, and Lionel Naccache (2001), "Towards a Cognitive Neuroscience of Consciousness: Basic Evidence and a Workspace Framework," Cognition 79, 1-2 (April): 1-37.
- Dennett, Daniel C. (1991), Consciousness Explained, Boston: Little, Brown and Company.

Fallon, Francis, Alan L. F. Lee, Brian Odegaard, Andrew Haun, and David Rosenthal (2020-2023), "Replication and Extension of Crucial John Grimes Experiment: Change Detection during Saccades," Templeton World Charity Foundation Project 0455.

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.

Freud, Sigmund, (1966-74a), "The Unconscious," The Complete Psychological Works of Sigmund Freud (Standard Edition), tr. and ed. James Strachey, London: The Hogarth Press, XIV, pp. 166-215.

Freud, Sigmund, (1966-74b), "The Ego and the Id," The Complete Psychological Works of Sigmund Freud (Standard Edition), tr. and ed. James Strachey, London: The Hogarth Press, XIX, pp. 3-66.

Freud, Sigmund (1966-74c), "Some Elementary Lessons in Psycho-Analysis," The Complete Psychological Works of Sigmund Freud (Standard Edition), tr. and ed. James Strachey, London: The Hogarth Press, XXIII, pp. 279-286.

Goodman, Nelson (1951), The Structure of Appearance, Cambridge, Massachusetts: Harvard University Press.

Grimes, John (1996), "On the Failure to Detect Changes in Scenes across Saccades," in Perception, ed. Kathleen Akins, New York: Oxford University Press, pp. 89–110.

Held, Richard, Yuri Ostrovsky, Beatrice de Gelder, Tapan Gandhi, Suma Ganesh, Umang Mathur, and Pawan Sinha (2011), "The Newly Sighted Fail to Match Seen with Felt," Nature Neuroscience, 14, 5 (May): 551-553.

Irvine, Elizabeth (2014), "Problems and Possibilities for Empirically Informed Philosophy of Mind," in New Waves in Philosophy of Mind, ed. Mark Sprevak and Jesper Kallestrup, London: Palgrave Macmillan, pp. 185-207.

Jackson, Frank (196), "What Mary Didn't Know," The Journal of Philosophy, 83, 5 (May): 291-295.

Kentridge, Robert W. (2017), "Sensation and Unconscious Perception," in "Does Unconscious Perception Really Exist? Continuing the ASSC20 Debate," Neuroscience of Consciousness, 3, 1: nix015, pp. 3-5.

Kouider, Sid, Vincent de Gardelle, Jérôme Sackur, and Emmanuel Dupoux (2010), "How Rich is Consciousness? The Partial Awareness Hypothesis," Trends in Cognitive Sciences, 14, 7 (July): 301-307.

Kriegel, Uriah (2009), Subjective Consciousness: A Self-Representational Theory, Oxford: Oxford University Press.

Kripke, Saul (1980), Naming and Necessity, Cambridge, Massachusetts: Harvard University Press.

Laloyaux, Cédric, Arnaud Destrebecqz., and Axel Cleeremans (2006), "Implicit Change Identification: A Replication of Fernandez-Duque and Thornton (2003)," Journal of Experimental Psychology: Human Perception and Performance, 32, 6 (December): 1366–1379.

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

Marcel, Anthony J. (1983), "Conscious and Unconscious Perception: An Approach to the Relations between Phenomenal Experience and Perceptual Processes," Cognitive Psychology, 15: 238–300.

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 (1974), "What Is It Like to Be a Bat?," The Philosophical Review, 83, 4 (October): 435-50.

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

Norman, Liam J., Charles A. Heywood, and Robert W. Kentridge (2013), "Object-Based Attention without Awareness," Psychological Science, 24, 6 (June): 836-843.

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.

Quine, W. V. (1951), "Two Dogmas of Empiricism," The Philosophical Review, 60, 1 (January): 20-43

Raffman, Diana (2011), "Vagueness and Observationality," in Vagueness: A Guide, ed. Giuseppina Ronzitti, Dordrecht: Springer, 107-122.

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

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

Rosenthal, David (2012), "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), "Misrepresentation and Mental Appearance," Trans/Form/Ação, 41, special issue ed. Alfredo Pereira: 49-74.

Rosenthal, David (2022), "Mental Appearance and Mental Reality," in Qualitative Consciousness: Themes from the Philosophy of David Rosenthal, ed. Josh Weisberg, Cambridge: Cambridge University Press, pp. 243-271.

Rosenthal, David (forthcoming), "Methodological Considerations for the Study of Mental Qualities," in Conscious and Unconscious States and Processes: Examining Their Nature, Similarities and Differences, ed. Juraj Hvorecký, Tomáš Marvan, and Michal Polák, London and New York: Routledge.

Scott, Ryan B., and Zoltán Dienes (2010), "Knowledge Applied to New Domains: The Unconscious Succeeds Where the Conscious Fails," Consciousness and Cognition, 19, 1 (March): 391-398.

Soranzo, Alessandro (2016), "Simultaneous Color Contrast," Encyclopedia of Color Science and Technology, ed. Ming Ronnier Luo, New York: Springer, pp. 1149-1152.

Sosa, Ernest (2003), "Privileged Access," in Consciousness: New Philosophical Perspectives, ed. Quentin Smith and Aleksandar Jokic, Oxford: Clarendon Press, pp. 273-292.

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

Weiskrantz, Lawrence (1997), Consciousness Lost and Found: A Neuropsychological Exploration, Oxford: Oxford University Press.

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

Young, Benjamin D., Andreas Keller, and David Rosenthal (2014), "Quality Space Theory in Olfaction," Frontiers in Psychology, 5, (January): Article 1.