OVERVIEW

1. First- and Third-Person Approaches
2. Saving the Phenomenological Appearances
3. Phenomenology and Theories of Consciousness
I. First- and Third-Person Approaches

- Until fewer than 150 years ago, theorists routinely tended to regard all mental states as conscious.
- Even Brentano, who acknowledged that “[a]n unconscious consciousness is no more a contradiction in terms than an unseen case of seeing” (1874/1973, 102), nonetheless insisted—though without any serious argument—that no mental state ever actually occurs without being conscious.

The recognition by Brentano and others that nonconscious mentality poses no conceptual conflict may have led many late 19th thinkers, Freud among them, to accept that mental states do actually occur without being conscious.

And this has pivotal implications for theorizing about consciousness.

For one thing, if mental states were all conscious, it would be tempting to see consciousness as a simple, unanalyzable property—as Moore regarded yellow.

That claim has less force if mental states aren’t all conscious, since then we must explain how the conscious states differ from mental states that aren’t conscious.
Countenancing mental states that aren’t conscious has another consequence. Some theorists approach the mind only in a first-person way (Nagel, Searle), others in just a third-person way (Quine, Dennett?). But any satisfactory understanding of mind must somehow accommodate both. Not only do we often know about our own mental states; we sometimes know pretty well about those of others. Also, when you say of yourself that you’re in pain or that you think that you mean the very same thing I mean when I say those things of you (P. F. Strawson 1958/59). The states we ascribe to ourselves are the same as those we ascribe to others.

The occurrence of mental states that aren’t conscious reinforces the need to accommodate our third-person access. As Freud recognized (e.g., 1915), if some mental states aren’t conscious, we must understand our own nonconscious states in the same third-person terms in which we understand the states of others. So there is nothing mysterious about mental states that aren’t conscious: We conceive of them just as we conceive of the conscious mental states of others. And their occurrence thereby underscores that consciousness and phenomenology cannot be our sole source of information about mental reality.
A first-person perspective seems to have pride of place because it seems to grasp mental reality *directly*—and so *perhaps more accurately*.

But *traditional claims of infallibility and incorrigibility are dubious*. We sometimes know *better than others* what mental states they are in—even with qualitative states others sometimes stand corrected.

Still, the *apparent directness* of our first-person access to mental states *seems* to make a third-person perspective hard to accommodate, since *directness would make little room—if any at all—for the causal ties that are central in a third-person understanding of mental states*.

On the other hand, a primarily third-person approach may seem unable to accommodate *the apparent directness* characteristic of our first-person access.

All this leads some to hold that we can’t accommodate both, and we’re left with an unavoidable choice between them.

But that’s not so. Though an exclusively first-person approach does not make room for the causal ties needed for any third-person approach, *the directness of our first-person access is arguably only apparent*—and so can likely be accommodated by a nuanced third-person approach (§III).
Might one’s access to others’ qualitative states always rely, possibly indirectly, on some first-person avowals?

Arguably not. We might, relying only on our third-person sense of when others are in pain, train a device that detects relevant neural states to determine when others are in pain. We could then use others’ first-person reports to confirm the device’s accuracy, but that accuracy would rely only on our third-person sense.

And in any case exclusively first-person access not susceptible to third-person test and correction would arguably be idle, since we can have no reason to rely on untestable pronouncements.

A higher-order theory can capture all this best, by representing the subjective directness of our first-person access as due to a higher-order awareness of our conscious states that is subjectively independent of any mediation and also independent of any observational input.

This lets us explain how third-person access can trump first-person access, and how we can have subjectively direct access to states that have the rich causal ties essential to our third-person access.

I’ll return in §111 to these higher-order theories. But let’s turn first to what it is for any theory to save the first-person, phenomenological appearances.
II. Saving the Phenomenological Appearances

► It's typically assumed—albeit tacitly—that *doing justice to our first-person, phenomenological intuitions means regarding those intuitions as true.*

► Why might that be? Our *pretheoretic intuitions about other things* are often incorrect—e.g., about how projectiles travel, whether the earth is flat, whether subjective color experiences resemble the visible properties of objects, and whether objects have absolute location.

► Even *pretheoretic intuitions about psychological functioning* are often wrong—e.g., about whether parafoveal vision is rich in detail and color.

► Sometimes false pretheoretic intuitions are close approximations to the truth; sometimes they're just wrong.

► Perhaps we expect more from first-person phenomenological impressions because *they seem unmediated*—and so perhaps (relatively) *immune from error.*

► *But again: Such lack of mediation is just a phenomenological impression,* which has little if any *independent, empirical substantiation*—indeed it's not obvious how it *could* have any.
Saving the appearances *can* amount to showing that the appearances are true. But that's not what it usually means. *We save the appearances by showing how the way things actually are gives rise to those appearances.*

So too with our *first-person impressions*: We needn't try to do full justice to them by showing that they're true; *rather, we can show why it is we have those phenomenological impressions—how they arise in us.*

Indeed, showing our impressions are correct *can't explain why we have those impressions*; things can be a particular way without seeming to be that way.

So saving the phenomenological appearances comes down to showing why our mental lives appear that way—*whether or not they really are that way.*

The role of phenomenology in building a theory of consciousness, then, is to *provide an accurate catalog of the appearances* we have of our mental lives, as data our theorizing must explain.

And as usual, *explaining the data must go beyond the data themselves.* Nor should our theorizing assume at the outset that those appearances accurately reflect the reality of our mental lives. Rather, our theorizing must initially just *explain why we have those appearances.*
Phenomenology is mental appearance: It is the way qualitative mental states appear to us—how we’re aware of them. So for phenomenology, appearance coincides with reality.

They might also coincide for the mental itself if mental states were all conscious. But since many mental states fail to be conscious, mental appearance does not exhaust the reality of the mental.

The phenomenological appearances reveal only how we are conscious, in a first-person way, of mental reality. To investigate that mental reality itself, we must go beyond phenomenology—beyond the first-person appearances.

We can best do that by relying on folk-psychological platitudes (Lewis 1966/71, 1972, 1999)—which are cast in third-as well as first-person terms.

Phenomenological appearances are real. And our explanation of mental reality must explain why it appears as it does. But it must explain more than just that. So we have a fruitful division of labor: We can explain some features of conscious states just as we explain their nonconscious mental counterparts. And we can also seek to isolate that additional factor which results in some mental states’ being conscious.
These observations have implications for theories of consciousness. I suggested that higher-order theories have the best shot at doing justice both to the first-person appearances that our conscious states present us with and to the rich causal properties that we know all our mental states have.

Is our higher-order awareness external to the states that are conscious? Or might it be intrinsic to those states? It might seem that intrinsicalism better saves the phenomenological appearances, since our awareness of our conscious states appears to be intrinsic to those states (Gennaro 1996, 2008; Natsoulas?).

But as with the appearance of directness, we must explain why we have this appearance, but needn’t see it as true. And we can best explain that simply by seeing our awareness of our conscious states as typically due to a distinct state of which we aren’t, in turn aware. Kriegel (2006) urges that we’re always aware—albeit usually peripherally—of our higher-order awareness. But that’s not so: Consider conscious visual perceiving that is relatively peripheral. And there are compelling independent reasons to deny that the higher-order awareness is intrinsic either to intentional or to qualitative conscious states.
III. Phenomenology and Theories of Consciousness

- These methodological points to one side, some may doubt that any theory not itself based on phenomenology can do justice to the first-person appearances.
- In closing, I’ll mention a few challenges of this sort and briefly sketch how a higher-order theory can deal with them.
- I’ll sometimes appeal to my own higher-order-thought (HOT) theory, but many points (though not quite all) will hold for any higher-order theory.

Can a higher-order theory save the idea that our first-person access is epistemically privileged?

- No. But we have only the appearance of privilege—and we can’t export from an apparent privilege to an actual privilege for the appearances.
- Also, we can capture that appearance in third-person terms. Suppose that a state is conscious if one is aware of it by having a HOT that one is in that state.
- That awareness will seem direct if it relies on no observation and on no conscious inference. And any such awareness will also appear to be privileged—relative to our ordinary observational access.
Such privilege aside, how can any third-person theory capture the way conscious qualitative states are “lighted up”—in contrast to nonconscious mental states? There is reason to think that HOTs do result in conscious qualitative character: We sometimes learn to taste wines or distinguish musical sounds by having words to classify distinct experiences. Having ‘tannin’ and ‘acid’ (or ‘clarinet’ or ‘oboe’) available to label mental qualities can, by itself, lead to mental qualities that are subjectively more fine grained.

Having words to pin the experiences on actually affects what it’s like for one to have those experiences.

How can learning words have that effect? It’s not credible that learning new words actually causes new mental qualities; no known mechanism could explain that. But words express thoughts. And since the words here describe one’s own mental qualities, the thoughts expressed are HOTs. Learning new words leads to becoming conscious of our mental qualities in more fine-grained ways; so the new subjective appearances must be due to new HOTs. The content of our HOTs does affect how our qualitative states are conscious.
This explanation may not resolve the **explanatory gap** (Levine 2001) that some see as separating conscious qualitative character from everything else.

But proponents of the explanatory gap insist on an explanation that will seem rational *independent of suitable theory*. And an explanation can do that only if it *explains the phenomenological data by appeal to the phenomenological itself*—within that closed circle of appearances.

Such an explanation will be *uninformative*: We can never informatively explain a phenomenon *in terms of itself*; we must in some way *go beyond the phenomenon we seek to explain*.

Nagel meant his famous (1974) rubric, ‘what it's like for one’, to capture not just the *lighted-up character* of conscious mental qualities, but also *their subjectivity*.

But we needn’t adopt Nagel’s question-begging construal of subjectivity as excluding the objective (Rosenthal 1983).

Rather, we should see *the subjective as being a special case of the objective*: It is that aspect of mentality that *depends to a large extent on individual perspective*—location, history, and other features of an individual’s points of view—and on *features of one’s perceptual apparatus* that affect mental qualities.
We can understand all these subjective features of an individual’s mental life in fully objective terms—indeed, in terms of the ways HOTs result in one’s being aware oneself as being in states with various mental properties.

Conflating mental appearance with mental reality adversely affects theorizing. If, e.g., mental reality were just its first-person appearance, then you couldn’t know whether your conscious sensation of red is the same as mine, or even if I have any conscious sensations at all.

But there’s more to mental reality than its first-person mental appearance.

So we can understand the mental qualities not just in first-person terms, but also in terms of their role in perceiving—by reference to location in a quality space that corresponds to that of the relevant perceptible properties.

That lets us explain why the actual quality spaces are asymmetric in ways that preclude such fanciful inter-personal quality inversion.

Such inversion seems conceivable only if we restrict ourselves to the first-person appearances our sensations present.

And since HOTs describe states in terms of their mental properties, inversion also won’t affect the first-person appearances.
What about our sense of the unity of consciousness?

Again, phenomenology, by itself, delivers only the appearance of unity. And the HOT model has the resources to explain these appearances: Each HOT represents its target as belonging to a self—and absent some reason to distinguish among the selves to which the various HOTs refer (as happens with dissociative identity disorder), a subjective impression of unity results.

It's altogether another question whether there is some underlying mental unity in addition to the subjective unity of the phenomenological appearances.

We also have a subjective impression of freedom—that our volitions are free and uncaused. That sense of freedom reflects our seldom being aware of our volitions as being caused by anything prior. We can explain that in turn as due to HOTs’ not representing their targets in respect of any such causal ties.

Indeed, we seldom have any phenomenological impression of mental states’ having any causal ties—which is arguably due to how HOTs represent those states. That doesn’t mean that our volitions are uncaused—just as there may be no real mental unity. But we do have a robust subjective impression of those things.
Summary

- Phenomenology provides some of the data that a theory of mind must explain: data about the first-person appearances of the underlying mental reality.
- Since mental states can occur without being conscious, appearance and reality cannot coincide for the mental itself—as they do for phenomenology.
- Phenomenology reveals only how we are conscious, in a first-person way, of the mental. To investigate mental reality itself we must go beyond phenomenology.

THANK YOU FOR YOUR ATTENTION