

To appear in *The Philosophy of Daniel Dennett*, ed. Bryce Huebner, Oxford University Press

## SEEMING TO SEEM

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

**Draft: Please do not quote**

**Abstract:** Dennett's account of consciousness starts from third-person considerations. I argue this is wise, since beginning with first-person access precludes accommodating the third-person access we have to others' mental states. But Dennett's first-person operationalism, which seeks to save the first person in third-person, operationalist terms, denies the occurrence of folk-psychological states that one doesn't believe oneself to be in, and so the occurrence of folk-psychological states that aren't conscious. This conflicts with Dennett's intentional-stance approach to the mental, on which we discern others' mental states independently of those states' being conscious. We can avoid this conflict with a higher-order theory of consciousness, which saves the spirit of Dennett's approach, but enables us to distinguish conscious folk-psychological states from nonconscious ones. The intentional stance by itself can't do this, since it can't discern a higher-order awareness of a psychological state. But we can supplement the intentional stance with the higher-order theoretical apparatus.

### I. First and Third Person

An apparent perplexity about the mental is that we know about mental states—thoughts, perceptions, sensations, and so forth—in two ways, and these two ways of knowing about mental states seem to have little if anything to do with one another. We know about our own mental states by first-person access, at least when those states occur consciously. Their being conscious, by itself and with no input from any other source, seems to tell us not only which states occur in one, but also what the mental nature of those states is.

The seemingly unmediated conscious access we have to our own mental states has led many to see the mental as resisting any sort of objective, scientific treatment, and perhaps on that account as being nonphysical. And because conscious first-person access to mental states seems subjectively to be the last word about their nature and occurrence, many have been tempted to see such access as also having a privileged epistemic status, as being incorrigible or even infallible.<sup>1</sup>

But we plainly also often know about the mental states of others. We can often tell when somebody is in pain, sees something red, wants a beer, or thinks it's going to rain. Such third-person access is plainly not incorrigible, much less infallible; but it's often very good. We sometimes see that somebody else is in a particular mental state even before that person is aware of being in that state, if indeed the person ever does.

Indeed, our knowledge about others' mental states is sometimes good enough to override another person's subjective sense of what mental state that person is in. You may sometimes know better than I what I want or even what I'm thinking or paying attention to. And we sometimes know better than another what emotional state that person is in. There is even reason to think one can be mistaken about whether one is in pain (e.g., Koyama et al 2005). First-person access to one's own mental states is not only not infallible; it is sometimes actually corrigible by appeal to others' observations.

Because first-person access to one's own conscious states and third-person access to others' mental states operate in such different ways, it can be difficult to do justice to both and, in particular, to know how the two ways of knowing about an individual's mental states fit with each another. How can the subjective grasp of one's own mental states square with the observation- and inference-based knowledge we have of the mental states of others?

One response might be to deny that first- and third-person access are about the same things. But this is unworkable. When you take me to be in some mental state, the state you take me to be in is typically of the same type as states that you are sometimes in yourself and have first-person access to. One indication that the states we ascribe in first- and third-person cases are of the same type is that if I take you to be in some mental state relying on third-person access and you, relying on your first-person access, deny being in that state, we are plainly contradicting one another. You are denying being in the very state that I assert you to be in. First- and third-person access are about states of the same type. We do not ascribe to others diminished or surrogate versions of the mental states we take ourselves to be in; the mental states we

---

<sup>1</sup> The two are not the same; first-person access might not be subject to being corrected even if it is not always correct.

take others to be in and ascribe to them are the genuine article, states of just the sort that sometimes consciously occur in us.<sup>2</sup>

A condition on any acceptable account of mind, accordingly, is to square these two types of access we have to mental states, to explain how the states we have first-person access to are states of the very same type as those we also often know others to be in by third-person observation and inference. An account of the nature of mental states must make clear how those states are subject to both kinds of access.<sup>3</sup>

The natural strategy is to begin with an account of mental states geared in the first instance to one kind of access and then show how that account can also accommodate access of the other kind. But that is easier said than done. First-person access presents itself subjectively as being wholly unmediated. Third-person access, by contrast, is plainly mediated by observation and inference, though the mediating observations and inferences are not always consciously explicit. So if we take this apparent lack of mediation at face value, any disparity between first- and third-person access should always be settled in favor of the first person. Mediation always allows for the introduction of error; though lack of mediation doesn't guarantee accuracy, it arguably wins in any contest with mediated access.

So if we take the apparent immediacy of first-person access seriously, we may see whatever inferences one might make about others' mental states as mere informal guesses as compared to first-person pronouncements. But we know that first-person access is on occasion less accurate than observation and inference by others. Giving pride of place to the first person unavoidably shortchanges the robust third-person access we do actually have to others' mental states.

It's also not clear what kind of state could be subject to wholly unmediated access. Some have been tempted by the thought that our access to abstract objects,

---

<sup>2</sup> Another strategy, recently popular, would be to say that we know in a first-person way about mental states, at least qualitative mental states, by applying special phenomenal concepts, whose modes of presentation are the mental qualities actually present to one (e.g., Alter *et al* 2006). But that mode of presentation would allow application of such concepts only to oneself, and no way has been given by proponents of this strategy to calibrate that application with concepts that apply across individuals. The problem about first and third person hasn't been solved, but just relocated from kinds of access to corresponding kinds of concept.

<sup>3</sup> Though this thought has been downplayed somewhat in the recent literature, it is hardly new. See, e.g., Strawson (1959, 104): "There would be no question of ascribing one's own states of consciousness, or experiences, to anything, unless one also ascribed, or were ready and able to ascribe, states of consciousness, or experiences, to other individual entities of the same logical type as that thing to which one ascribes one's own states of consciousness. The condition of reckoning oneself as a subject of such predicates is that one should also reckon others as subjects of such predicates."

such as numbers, is wholly unmediated; but how could we know about concrete states of ourselves in way that is absolutely unmediated? And how could states accessible in an absolutely unmediated way also be accessed in a third-person way? Third-person access would of course rely on causal ties the states have with observable occurrences, such as behavior. But first-person access seems to reveal no causal connections the states have with anything observable; indeed, such access, by itself, seems not even to represent mental states in respect of causal ties with one another.<sup>4</sup>

So taking first-person access as primary undermines not only our understanding of how mental states could be accessible in a reasonably reliable way from the third person, but also the importance to the nature of those states of their causal ties with behavior, sensory stimulation, other mental states. It is doubtless this that leads many who favor starting from the first person to downplay or sometimes even ignore altogether the need for an account to accommodate both types of access, and to deny that causal ties are essential to the nature of mental states.

It is partly for these reasons that Daniel Dennett's work on the mind has been so important and influential and has had such a deeply salutary effect. Dennett's appeal to the intentional stance holds that we ascribe to others the intentional states that make sense on an assumption of those individuals' rationality (1978, 1987, 1991b). It thereby paves the way for a full and rich account of mind that starts not from first-person access to mental states, but from the way we know about those states from the third person. And it helps undermine the now popular claim that a satisfactory account of mind must rest primarily, or possibly even exclusively, on first-person access. We can understand the mind by appeal to ascriptions of mental state based on behavior and context.<sup>5</sup> Dennett's endorsement of the primacy of a third-person approach echoes in this way arguments of W. V. Quine (1960, see esp. §45, 1985) and Wilfrid Sellars (1956, 1968).

It is important to note that the assumption of rationality that sustains the intentional stance does not require that one take all the intentional states one sees an individual as being in as rational, even given that individual's purposes and proclivities. The intentional stance applies to individuals some of whose intentional states depart to some extent from full rationality. But one cannot understand an individual as having thoughts and desires at all unless one takes those states to be rational on balance. The rationality assumption is in that way of a piece with a principle of charity that governs

---

<sup>4</sup> We do of course assume causal ties that mental states accessible from the first person have with behavior and with other mental states, but we infer such causal ties by applying folk-psychological generalizations, not because first-person access reveals them. I'm grateful to Pete Mandik for raising this concern, and for a helpful reading of an earlier draft.

<sup>5</sup> As Dennett stresses, use of the intentional stance is appropriate only when one can't explain the behavior of something by the less elaborate assumptions that figure in the design or physical stances.

how we understand one another's speech. On that principle, each of takes the speech acts of others to be true as much as possible, to embody valid inferences as much as possible, and to use words in standard ways as much as possible.<sup>6</sup>

Starting from first-person access makes it difficult to do justice to the way we have access to others' mental states from the third person and the way that access fits with our first-person access. This is in part because first-person access seems subjectively to be the last word about the nature of mental states and in part because such access seems to represent the nature of those states independent of the causal ties those states have with behavior, stimuli, and other mental states. So perhaps starting from the third person holds greater promise for an account that does justice to both kinds of access.

But those who favor the primacy of the first person will push back, urging that we will never be able to do justice to first-person access if we start instead with what it is about mental states that enables third-person access. So they will argue that we face a kind of standoff; according primacy to either type of access makes doing justice to the other and how the two fit together difficult if not impossible. Since we cannot do full justice to both, we must choose. And given that choice, they will conclude, first-person access plainly wins, since that is what is unique to and distinctive of the mental. We must accordingly give pride of place to the first person.

But that is not the choice we face. The apparent failure of first-person access to represent mental states in respect of at least many of their causal connections does likely doom any attempt to start with the first person and arrive at a satisfactory account of third-person access or, indeed, a full account of the nature of those states themselves. But the advocate of the primacy of the first person is arguably wrong that the opposite is also the case. We can build on a third-person account of the mental to arrive at an accurate and full explanation of the first-person aspect of mental states and of our conscious access to those states. We can usefully see Dennett's Consciousness Explained (1991a) and his extensive subsequent work on consciousness (e.g., 1996, 2006, 2008, 2015, and esp. 2005, ch. 2) in this light, as building a satisfactory account of first-person access and the first-person aspect of mental phenomena that rests on third-person considerations that fit with those that animate the intentional stance approach in (1987).

---

<sup>6</sup> Maximizing of truth, validity, and standard use of words must be by the lights of the individual doing the interpreting; there is no other way. And since it may not always be possible to maximize all three, one will sometimes have somehow to strike a balance. Because there are often alternative ways to balance the maximizing of the three, there will inevitably be conflicts of interpretation that cannot be settled in any independent, objective way. See Rosenthal (1989, §4; 2009, §6).

Evaluating this is of course a delicate matter. Can an account that fits with the third person deliver everything an advocate of a first-person approach would require? In particular, can an account that fits with an intentional-stance approach deliver those goods? Moreover, is everything that those who favor a first-person approach demand warranted? Might some demands that those who favor the first person simply be results of privileging the first person, and not independently motivated or warranted?

As we have seen, similar questions arise in evaluating whether a first-person approach can deliver a full and satisfactory account of the mind. Is first-person access subjectively privileged in the way it seems? Are the causal ties between mental states and observable occurrences and among the states really essential to the nature of those states?

Champions of the first person will urge that we accept the subjective privilege as true, and deny that causal ties are essential for a full account of mental states. It would then be reasonable to see third-person access as not that important to accommodate. The claim that causal ties are merely contingent to the mental states receives vivid expression with the intriguing idea that the qualitative states that occur in perceiving might be undetectably invertible from one individual to another. Perhaps, on this line, the mental quality that occurs when you see something red, for example, is the same type of mental quality that occurs when I see something green.

Such undetectable inversion would be possible—indeed conceivable—only if the causal ties those states have with observable stimuli were inessential to those states. And that would impair the kind of access we can have to others' mental states. One could only know that when somebody else sees something red the person is in the kind of state distinctive for that person when seeing red; one would not also know what kind of state it is, described in mental terms. Only first-person access could then reveal the nature of qualitative mental states; third-person access says nothing informative about the distinctively mental nature of others' states. And if causal ties are inessential to intentional states as well, similar considerations would apply there; one could know the nature of those only from the first-person.

By contrast, a view that accords primacy to third-person access rejects the possibility—indeed, even the conceivably—of such undetectable inversion. And our folk picture of mental states concurs. Causal connections with stimuli are not merely incidental to mental states, but an aspect of their mental nature; here the primacy of third-person access coincides with common sense. Our folk view holds that we ascribe to others states of the very same type we are aware of as occurring in ourselves; so that

view prevents us from even conceiving of the mental states others are in as undetectably inverted from our own.

Some will insist that even if undetectable inversion isn't possible, it is at least conceivable (e.g., Shoemaker, 2003, p. 336). But undetectably invertible states could not be mental states as our folk conception dictates. Compare conceiving of something observably exactly like water, though it isn't H<sub>2</sub>O (Rosenthal 2010, §§1, 3). We are in that case not conceiving of water, even if the imagined substance is superficially just like water.<sup>7</sup> By the same token, conceiving of undetectably invertible states is not conceiving of mental states. The pretense that we can imagine or conceive of undetectable inversion of mental states stems from the playful conceit that first-person access is all that matters—indeed, all that could conceivably matter—in determining the nature and occurrence of mental states. But we know on the slightest reflection that this isn't and couldn't be so.

All this has important consequences for evaluating the dueling strategies of starting with the first or third person in constructing an account of the mental on which both types of access fit together. If the champion of the first person argues that the conceivability of undetectable inversion is a pretheoretic datum that an account of mind must preserve, the natural reply is that it is not a datum at all, but at best an artifact of taking the first person as our starting point. Similarly with the alleged incorrigibility or infallibility of first-person access; these are not pretheoretic, commonsense data, but just the result from taking the first person by itself to reveal the nature of mental states. In evaluating an account that begins from the third person, we must exclude demands that rely only on taking first-person access to be our only genuine access to the mental.

The champion of the first person often appeals to undetectable invertibility, incorrigibility, and infallibility as pretheoretic intuitions about mental phenomena that any account must accommodate. But such alleged intuitions are typically just appealing ways to package substantive theoretical claims, passing them off as pretheoretic data that impose constraints on theorizing. Adapting Dennett's useful notion of an intuition pump,<sup>8</sup> we can think of these so-called intuitions as theory pumps: devices to get us to adopt a tacit but controversial theoretical approach.

---

<sup>7</sup> On cases like this see Kripke (1980, 130-140, esp. 130-134; cf. also 104); Kripke relies primarily not on the example of water, but on heat's being mean molecular kinetic energy.

<sup>8</sup> Dennett originally coined the term 'intuition pump' to describe "thought experiments ... [that involve] inviting the audience to imagine some specially contrived or stipulated state of affairs, and then—without properly checking to see if this feat of imagination has actually been accomplished—inviting the audience to 'notice' various consequences in the fantasy" (1991a, 282). Also: "an intuition pump—a story that cajoles you into declaring your gut intuition without giving you a good reason for it" (1991a, 397). See also Dennett (2013).

No such theory pumps would be needed, and so none would be invoked, were there independent, non-question-begging support for the theoretical approach in question. Here the alleged intuitions simply package the theoretically contentious and otherwise unsupported view that first-person access exhaustively reveals the nature of mental states.

But even if we took intuitions at face value, not as theory pumps but as genuinely pretheoretic convictions, there would still be two ways to accommodate them. We could, as the advocate of first-person primacy urges, take the convictions as imposing constraints on any satisfactory account. That would be to save the intuitive appearances by taking them to be accurate. But we need not accommodate intuitions in that way; we could instead simply seek to explain why we have the convictions in question and why they seem so compelling.

The second approach does not require assuming that the convictions are true; there are many compelling convictions that turn out not to be true. And even if they were true, we would need to explain why we have them. The traditional explanation smuggles in the controversial assumption that the mind is transparent to itself; the reason we have the intuitions is because such transparency leads automatically from the nature of mind to true intuitions about the mind. But assuming such transparency is patently question begging in the context of evaluating the intuitions. The very appeal to intuitions about the mind is simply a way to package the question-begging theoretical assumption that we must rely only on first-person access to learn about the nature of mind.

In what follows, I examine how well Dennett's treatment of consciousness based largely on third-person considerations does justice to first-person access and saves those first-person appearances that we have independent reason to respect. First-person access and the appearances it generates are a matter of mental states' being conscious; so evaluating how well an account saves legitimate first-person appearances hinges on how that account deals with consciousness. In section II, I briefly sketch Dennett's account of consciousness. In section III I raise some difficulties for that account. I conclude in section IV by showing that these difficulties point toward an alternative theory of consciousness, strongly similar in spirit to Dennett's but with some important differences.



## II. Dennett and Consciousness

Key to Dennett's views about consciousness is a type of problem case he raises about timing (1991a, chs. 5-6; Dennett and Kinsbourne 1992). It will sometimes happen that one has a conscious visual experience that is inaccurate due to the interference of a memory one has of a similar or associated object or scene. In Dennett's vivid example, you see a long-haired woman without eyeglasses running by, but the memory of having seen a different short-haired woman with eyeglasses intrudes, and your conscious experience is of seeing a short-haired woman with eyeglasses (1991a, 117-8). Perhaps the memory is of a friend and, though you don't know the passing woman, you consciously experience seeing her as your friend. We have all doubtless been subject to this kind of thing; memory does sometimes skew perception in this way.

The memory of your friend enters the causal stream relevant to your conscious experience and distorts the visual information sometime between retinal stimulation and stable conscious experience. But exactly when? One possibility is that it affects things before the perceptual input becomes conscious. But there is another possibility. Perhaps you first have an accurate conscious experience of the passing woman as having long hair and no eyeglasses, but the memory immediately intrudes, and your initial conscious experience is immediately replaced by an inaccurate conscious experience of a short-haired woman with eyeglasses. The initial conscious experience, moreover, is so brief that you have no memory of it; subjectively it's just as though you'd never had that initial conscious experience. The two types of occurrence, resulting from interference by a memory at different points, would be subjectively indistinguishable.

Folk psychology seems to offer nothing that favors one possibility over the other. Dennett colorfully labels the first possibility Stalinesque, since like Stalin's show trials consciousness presents us with false information; he calls the second Orwellian, because in that case consciousness would in effect rewrite history. Since neither subjective impressions nor folk psychology can help, Dennett sees no way to determine in any particular case when it is that memory contaminates the visual information, before the perceptual input comes to be conscious (Stalinesque) or after (Orwellian).

Dennett also surveys a variety of striking experimental results in which timing seems to pose a similar problem about conscious experience. In the phi phenomenon, two round patches of light appear briefly one after the other. If the spatial distance between them and the timing of their successive blinking are just right, observers don't experience two disks lighting up in succession, but rather a disk moving back and forth between the two locations. And if the stimuli differ in color, the moving disk appears to change color in midstream (Dennett 1991a, 114; Kolers et al 1976; Nelson Goodman

suggested testing for color change).

But the first blinking of the first light would not by itself result in the illusion of a moving disk. So what happens after that first blink stimulates the retina, sending visual information onto visual cortex, prior to the blinking of the second light? Does the visual input of the first light by itself simply never reach consciousness (Stalinesque)? Or does it become conscious only to be replaced so fast by a conscious experience of a moving disk that no memory remains of the conscious experience of the blinking of the first light by itself (Orwellian)? Again, neither subjective impressions nor folk psychology helps settle the question. Other experimental results pose similar quandaries.

Dennett argues that these phenomena have important consequences for understanding consciousness. If there is no nonarbitrary way to settle whether the initial perceptual input in these cases makes it to consciousness, it's to that degree indeterminate as to whether particular perceptual inputs occur consciously. And that, he argues, casts doubt about how determinate consciousness is in general. He concludes that "there is no reality of consciousness independent of the effects of various vehicles of content on subsequent action (and hence, of course, on memory)" (1991a, 132).

In the running-woman case, the visual input and the memory of one's friend compete with one another as to which will make it to and remain in consciousness. Similarly with the first blink and the illusion of motion generated by alternating blinks. So Dennett urges that we mustn't think that there are a number of well-defined psychological states, each of which may in the ordinary course of events make it to consciousness. Rather, there are a number of psychological drafts, often conflicting, and each competing with many others to make it to consciousness. This is Dennett's Multiple Drafts model of consciousness.

This conclusion is not all that surprising. It is natural to think that the mind involves many competing causal strands, and these play out in determining what makes it to consciousness. But Dennett also draws a more contentious conclusion from the puzzles about timing. If there is no way to tell whether a particular case occurred in a Stalinesque or Orwellian way, he concludes, then "there are no fixed facts about the stream of consciousness independent of particular probes" (1991a, 138; cf. 275). Such probes may include somebody's asking what one perceives or the demands of action or the like making it relevant what one consciously perceives. The facts of consciousness are fixed by the effects consciousness has on other things. It is this statement of his

view that Dennett labels “first-person operationalism” (henceforth “FPO”; 1991a, 132).<sup>9</sup>

Opposition to operationalism in general, Dennett notes, stems from the idea that there are facts not caught by whatever operationalist test one might propose. But operationalism about consciousness, he argues, is special; who would maintain that there are facts of consciousness not subject to first-person access? We must, he insists, “den[y] the possibility in principle of consciousness of a stimulus in the absence of the subject’s belief in that consciousness” (1991a, 132). In the cases described above, if the difference between a Stalinesque and an Orwellian occurrence would make no subjective difference to the individual in question and they don’t in any other way differ in outcome, there is simply no distinction to be drawn. The appearance of a distinction here is illusory. Operationalism in respect of an individual’s beliefs catches all the facts there are about consciousness.

It might not seem to matter all that much whether we can determine which of the Stalinesque and Orwellian mechanisms is operative in generating a particular case of conscious misperception, or even whether we can draw a tenable distinction between them. But FPO also rules illusory another distinction that figures more deeply in theorizing about consciousness. Dennett takes “‘writing it down’ in memory [to be] criterial for consciousness; that is what it is for the ‘given’ to be ‘taken’—to be taken one way rather than another” (1991a, 132). Perceptual input is conscious if, but only if, it’s taken in some way; that’s why there is no “consciousness of a stimulus in the absence of the subject’s belief in that consciousness.” The distinction between how consciousness is and how it seems is illusory. That’s the operationalism in first-person operationalism.

So on FPO, one cannot be wrong about the facts of consciousness in one’s own mental life; how consciousness seems to one is how it actually is. This may initially seem right, perhaps even obvious. Consciousness is itself just a matter of seeming; there is nothing to it except seeming. And this may seem to amount to endorsing the privilege that traditional theorists have accorded first-person access. If how it seems is how it is where consciousness is concerned, then one cannot be wrong about one’s own conscious states. Dennett lampoons the view he calls the Cartesian Theater, on which consciousness consists in the viewing by a metaphorical viewer of a stream of consciousness. But in adopting FPO, it appears that he may in effect be accepting a

---

<sup>9</sup> The idea that competition among many nonmental factors results in some folk-psychological conscious states may encourage another claim of Dennett’s, that the search for a neural correlate of consciousness “is probably a wild goose chase” because “the processes that elevate contents to consciousness are like the processes that elevate [evolutionary] lineage divergences into speciation events” (2009, 234). But the analogy with evolution is questionable; the competing factors may well result in states we can taxonomize along standard folk-psychological lines, allowing for a subsequent relatively specific process that results in those states’ becoming conscious.

version of the traditional Cartesian view that first-person access is infallible.<sup>10</sup>

The apparent distinction between Stalinesque and Orwellian mechanisms is an important opening wedge for Dennett. The reason there is, on his view, no fact of the matter of any sort about which mechanism is operative is that there is no subjective fact of the matter; which mechanism is operative would, by hypothesis, make no difference to how things seem subjectively to the individual. That's simply the way the alternative processes are described. The illusory character of a distinction between Stalinesque and Orwellian is of a piece with the view that how it seems for consciousness exhausts how it actually is. There being no way to determine which mechanism is operative reveals, Dennett urges, that there is nothing to consciousness beyond what the individual in question thinks there is.

And because how it seems for consciousness is how it is, there is no room for what Dennett stigmatizes as “the bizarre category of the objectively subjective—the way things actually, objectively seem to you even if they don't seem to seem that way to you!” (1991a, 132). Its seeming to one that it seems a particular way cannot amount to anything beyond its simply seeming that way to one. If one consciously sees a red square, it seems to one that there's a red square over there. What could it be, Dennett is asking, for it to seem to one that it seems to one that there's a red square over there? What could it be, that is, apart from its simply seeming to one that there's a red square over there?

If its seeming that it seems that there's a red square could differ from its simply seeming that there's a red square, we could distinguish the reality of seeming from how that seeming appears, how subjective seeming objectively is from how that seeming appears to one. So Dennett's denial that there are such levels of seeming, and hence a difference between how a case of seeming really is from how it appears, may strike one as straightforward common sense. There is nothing to how seeming appears, one may want to insist, beyond the seeming itself, no distinction here between appearance and reality. As Thomas Nagel puts it, when it comes to conscious experience “[t]he idea of moving from appearance to reality seems to make no sense” (1974, 444).

But Dennett's denial of levels of seeming and his insistence that there are no facts about consciousness apart from the states one believes one is in do not rest on a first-person approach to the mind such as that championed by Nagel. The reason we cannot distinguish seeming to seem from mere seeming, according to Dennett, is not that we have special access to our own mental states to provide us with infallible self-

---

<sup>10</sup> As Fred Dretske (2000, 138) wryly notes: “First-person operationalism sounds like warmed-over Cartesianism to me.”

knowledge. Rather, it is that there simply is nothing to our conscious experiences beyond what we believe there is. That is why we cannot wrong about them, and why the idea of seeming to seem can, according to Dennett, get no purchase. Dennett has in effect repackaged the traditional Cartesian idea of infallible access in an operationist approach to the mind, thereby removing its epistemic bite.

To the extent to which we can see Dennett's view as involving access to one's own mind at all, it is a feature of his operationalism about the mind. It does not rest on something special about how the mind works or on how we have access to our inner mental workings. It is simply that our beliefs about our conscious experiences are the last, and indeed the only, word about those conscious states. We aren't wrong about our conscious experiences because when it comes to conscious experiences there is nothing beyond our beliefs to be wrong about.

We will see reason to question whether such operationalized first-person privilege is tenable. But for now it's important to stress that the operationalist feature of Dennett's account allows him to construct a treatment of consciousness and first-person access that rests firmly on third-person considerations. Consciousness occurs, on Dennett's view, only if being conscious of something has an effect on memory, speech, or action, and the effect on memory is describable in speech. The multiple drafts of our psychological functioning get fixed as conscious states only once they can have discernible effects that are subject to interpretation.

One can adopt the intentional stance not only toward others, but also towards oneself. So it may seem inviting to see Dennett's view as holding that consciousness is a matter of self-interpretation, a matter of one's tacitly adopting the intentional stance towards oneself. This would fit comfortably with Dennett's denial of a metaphorical internal viewer of a stream of consciousness. As already noted, consciousness on Dennett's view is "what it is for the 'given' to be 'taken'—to be taken one way rather than another." And it is natural to see this taking as a kind of self-interpretation. It is self-interpretation on the intentional-stance assumption that one is on balance rational. And if Dennett's theory of consciousness rests in this way on the application to oneself of the intentional stance, it is an account of consciousness built solidly on a third-person approach to the mind.

As Dennett stresses, the intentional stance is thoroughly realist about the occurrence of mental states. There are patterns of behavior that one can discern only by adopting the intentional stance toward an individual; these patterns are real, and they would be missed without the intentional stance. The real patterns sustain the realist character of the intentional-stance account of mind (1987, ch. 2, postscript; 1991b).

So the intentional stance will be thoroughly realist when applied to oneself as well, in connection with consciousness. The states one interprets an individual to be in by using the intentional stance are states the individual is objectively in, though one will discern them only by adopting the intentional stance. The interpretivist character of the intentional stance does not undermine the independent objectivity of the patterns one discerns with that stance; Dennett's interpretivism readily accommodates realism about the mind.

But FPO cannot be simply the application to oneself of the intentional stance. For one thing, applying the intentional stance involves the discerning of real patterns of behavior that allow one to ascribe beliefs and desires given a background assumption of overall rationality. And one plainly does not ascribe mental states to oneself on the basis of observing patterns of behavior. The states one does self-ascribe are indeed of the sort one can also ascribe to others by noting relevant patterns in their behavior; but observing such patterns does not figure in one's self-ascription.

Rather, there is on Dennett's view a competition among various internal states as to which will dominate one's behavior, including the speech behavior by means of which one announces what conscious mental states one is in. And he argues that there is nothing beyond the results of that competition that could settle what one's conscious states are. There is nothing, for example, that could settle whether the visual input of the running woman with long hair and no eyeglasses makes it to consciousness and is then revised or is revised before making it to consciousness. FPO is operationalist in simply dismissing that question, since there is nothing operational to determine an answer.

And FPO embodies a third-person approach to the mind since it allows nothing except what a person says as a way to determine what that person's conscious states are. One will self-ascribe in response to a question's being put to one or some other probe, but "there are no fixed facts about the stream of consciousness independent of particular probes." One takes oneself to be in this or that state of consciousness, and that's the last and only word about it. So FPO is not simply a matter of applying the intentional stance to oneself, though it is in that spirit. Like the intentional stance we adopt toward others, FPO is interpretationist in that one's beliefs and reports of one's conscious states are the last word about what conscious states one is in. How one interprets oneself in respect of one's own conscious states settles the question of what one's conscious states are.

### §III. Seeming to Seem

How one takes one's conscious states to be is, according to FPO, the last word about the actual nature of those states. The operationalist character of FPO collapses any distinction one might seek to draw between what conscious states one is in and the way one takes one's conscious states to be, between one's conscious mental life and what one believes about one's conscious mental life. Dennett sees this as the right result. Consciousness is a matter of how things seem to us subjectively; so there can't be a tenable distinction between seeming and seeming to seem. If things seem to be a particular way, what could seeming to seem amount to other than simply seeming?

A satisfactory account of mind must do justice to both our first- and third-person takes about the mind, and it's likely that no account that takes first-person aspects of mind as primary can do the job. Starting from the first person is unlikely to succeed, in part because it deprives us of information needed to construct a satisfactory third-person account of mental functioning. Dennett's intentional-stance approach is grounded in the third person, and his FPO is designed to fit comfortably with that, taking one's own interpretation of one's conscious mental life as the first and last word about what one's conscious states are. In the absence of probes that result in one's taking oneself to be in particular conscious states, there is no fact of the matter about what conscious state one is in. So if FPO results in the collapse of any distinction between seeming and seeming to seem, we should, he would urge, accept that as the result of independently sound methodological considerations.

Those who favor an account of mind based on first-person access also tend to deny any distinction between seeming and seeming to seem. All seeming is intrinsically conscious on that approach. So consciousness cannot misrepresent our mental lives, and we are accordingly infallible about our mental goings on. And since we know of nothing else whose nature and occurrence we cannot be mistaken about, there is a difficulty, perhaps intractable, in giving an objective account of the kind of subjectivity that conscious seeming involves.

Since Dennett's version of infallibility and subjectivity doesn't rest on anything special about the nature of mental states and consciousness, typical advocates of a first-person picture of mind would likely deny that Dennett saves the phenomena those advocates see as crucial. But such a complaint about Dennett's version of subjectivity and infallible access is arguably without foundation. The interpretativism of FPO does result in a type of subjectivity that stands apart from the rest of objective reality, and has first-person beliefs about one's conscious states be the last word about them. This arguably saves the substance of what advocates of the first person insist on, though it

also dispels the intractable mystery inherent in an appeal to intrinsic consciousness.

But there is a different worry about Dennett's overall account, having to do with whether it enables first- and third-person access to fit comfortably together. Suppose that I believe that it's raining. Dennett's intentional stance should, at least on some occasions, enable you to tell that I have that belief. You adopt the intentional stance towards me, discern relevant patterns, and on that basis ascribe to me a belief that it's raining. That's all objective. You might make a mistake, but such a mistake, like any others about observable matters, can be corrected, and taking account of enough relevant patterns should do the trick.

But Dennett's FPO insists that "there are no fixed facts about the stream of consciousness independent of particular probes." How does that fit with your ability to tell that I believe that it's raining by discerning the relevant real patterns in my behavior? Does that mean that the intentional stance operates independent of "fixed facts about [an individual's] stream of consciousness"? Are the facts that can get fixed by particular probes independent of what the intentional stance enables one person to ascribe to another?

We can see how this difficulty arises by focusing on the claim of FPO that one's beliefs about one's own mental states are the last word about what they are. Suppose I believe that I believe it's sunny out and say that that's my belief. My pronouncements are the last word about what conscious beliefs and desires I have. But how can my belief and pronouncement be the last word if your adopting the intentional stance gives you third-person access to what beliefs and desires I have? Suppose the two deliver different verdicts. How does my having the last word square with your ascriptions grounded in third-person observation of my real patterns?

A natural answer to this quandary would be that the patterns you discern in my behavior enable you to ascribe beliefs and desires to me independent of whether those states are conscious, whereas my having the last word about what conscious states I'm in pertains only to my beliefs and desires only insofar as they are conscious. Your observations can reveal the states I'm objectively in, whereas my beliefs about what conscious states I'm in is the last word only about how my subjective stream of consciousness appears to me. How my stream of consciousness appears to me is itself objective; it appears to me in one way and not another, and my beliefs about that are the last word about how it appears to me. But that objective matter is different from the objective matter your observations of real patterns reveal, namely, beliefs and desires I have independently of how my mental life appears to me.



But that way of squaring things is not available to Dennett; indeed, he emphatically rejects it. For me to be in particular intentional states is for things to seem a particular way to me. So if we distinguish the intentional states I'm in from the intentional states I am aware of myself as being in, that would be to distinguish how things seem to me from how they seem to me to seem. And that's the very distinction Dennett decries. Countenancing seeming to seem as distinct from merely seeming "creates the bizarre category of the objectively subjective—the way things actually, objectively seem to you even if they don't seem to seem that way to you" (1991a, 132).

The point is worth stressing. When you, using the intentional stance, ascribe beliefs and desires to me, you ascribe states that pertain to how things seem to me. And when the facts of consciousness are fixed by my beliefs about what I am conscious of, I too ascribe to myself states that have to do with how things seem to me. The problem was how your ascriptions can be objective if my self-ascriptions are the last word. How can there be "no fixed facts about the stream of consciousness independent of" probes that determine what states I take myself to be in if your intentional-stance ascriptions are objective?

The inviting solution is that my beliefs about what states I'm in fix only what conscious states I'm in, whereas your application of the intentional stance to me objectively reveals what states I'm in independently of their being conscious. Indeed, it's likely that this is the only way to deal with the potential conflict. But the difficulty with that for Dennett is that the states ascribed using the intentional stance are themselves states of seeming. So if states of seeming independent of consciousness are distinct from conscious states of seeming, there is after all a distinction between seeming and seeming to seem, which Dennett is at pains to deny.

We cannot get Dennett's intentional stance and FPO to fit well together unless we accept a distinction between seeming and seeming to seem. But there is reason independent of that difficulty to endorse that distinction. Perceiving often occurs consciously, but it also occurs without being conscious, as in subliminal perception. In typical experimental work in masked priming (Breitmeyer *et al* 2006; Öğmen *et al* 2006), a stimulus is briefly presented, followed by another stimulus. If the target stimulus is presented on its own, subjects report seeing it, but when it is followed by a suitable mask they report seeing only the mask. Still, the target stimulus that subjects claim not to see influences subsequent psychological processing, and that is compelling evidence that the stimulus was seen, just not consciously.

Other cases involve stimuli presented to the blind field of a blindsight patient (Weiskrantz 1997). Though the individual sincerely denies seeing the stimulus, elicited

guesses are far above chance as to its shape and color and in some cases motion. This is so even with guesses about the emotional expression of faces presented to the blind field (de Gelder et al 1999). And there is nonconscious change detection; subjects report seeing no change, yet priming effects demonstrate that the change was seen, but not consciously (Fernandez-Duque and Thornton 2000; Laloyaux et al, 2003).

It would be arbitrary and groundless simply to deny that seeing occurs in these cases. The visual input here has the same downstream psychological effects that are characteristic of conscious seeing, effects for example that vary with the color and shape of the stimulus and that affect desires. The only difference is that the input does not result in visual states the individual is aware of. But the visual input does register psychologically in the way distinctive of seeing; it simply doesn't register consciously. We would need some reason to deny that subliminal states constitute genuine perceiving. One could just dig in one's heels and deny that perceiving can occur without being conscious. But what reason could there be for that apart from traditional Cartesian claims that the mental is necessarily conscious, claims that are themselves without independent support?

Perceiving, moreover, involves things' seeming a particular way to one; if one is presented with a red, square stimulus, it will seem that something is red and square. Since perceiving can occur without being conscious, so can seeming. Consciously perceiving something involves one's being consciously aware of that thing, whereas perceiving the thing nonconsciously involves being aware of it, but not consciously aware of it.

So there is after all something to "the way things actually, objectively seem to you even if they don't seem to seem that way to you." If one subliminally sees a red, square stimulus, it will not seem to one that one sees that stimulus; one will sincerely deny that one does. Still, we can experimentally establish that the relevant area of one's visual field seems to one to have a red, square stimulus in it; one will guess if pressed to do so that there is a red square in that place, guesses that are far above chance. And one will behave psychologically in other ways as though one has just seen something red and square. So if one sees the stimulus subliminally, there is a way in which it does seem to one that the stimulus is there even though one is unaware of its seeming that way to one.

But it will not seem to one that one sees a red, square stimulus, since one's seeing it isn't conscious. Things seem to one a particular way even though it does seem that they seem that way. There is after all room for a coherent "category of the objectively subjective—the way things actually, objectively seem to you even if they

don't seem to seem that way to you.”

Consciousness, according to Dennett, is “what it is for the ‘given’ to be ‘taken’—to be taken one way rather than another” (1991a, 132). Subliminally seeing a red square is taking there to be a red square, though one is unaware of that taking. Being unaware of that taking, one takes oneself not to take there to be a red square. It seems to one that there is a red square; it’s just that it doesn’t seem to one that it seems that way. Since consciousness is a matter of such takings, there are evidently two levels of taking we must consider: how we take things to be and our awareness of our taking things to be that way. These two levels must be factored into any complete account of perceiving.

Dennett has argued against the occurrence of “qualia as traditionally conceived” (2015, 3; cf. 1991, ch. 12, and 1998, ch. 8). The traditional conception he has in mind is qualia as “unanalyzable simples”; as he notes, consciousness seems to present qualitative mental states in that way (2015, 8). But if qualitative states occur not only consciously, but also without being conscious, the idea that their qualitative properties are unanalyzable simples loses its force, as does the denial of such properties occur.

We seldom talk in everyday situations about perceptual or other mental states that aren’t conscious. But that shouldn’t tempt us to deny to the subliminal cases the status of genuine perception. We rarely talk about our own mental states when they aren’t conscious because we are rarely aware of them; we would become aware of them only in some third-person way, say, by experimental results or by the application of the intentional stance to ourselves. We do sometimes note by observing others’ behavior that they are in psychological states they are unaware of, though social niceties typically inhibit our commenting on that. But our tendency not to talk about perceptual states that aren’t conscious doesn’t show that aren’t any.

Beliefs also involve things’ seeming to one to be a particular way. So even if one had qualms about genuinely perceptual subliminal states, it’s widely accepted that one can believe various things without being aware that one does. A belief’s not being conscious, moreover, does not prevent its being discernible from the intentional stance; so nonconscious beliefs are objective. Such nonconscious believing is another “way things actually, objectively seem to you even if they don't seem to seem that way to you.” We cannot deny that seeming to seem occurs that is distinct from mere seeming. And we accordingly cannot accept that what one takes one’s own states of seeming to be is the last word about those states of seeming.

We use the term ‘conscious’ in two ways, which it is important to distinguish. We

speak of seeing, believing, desiring, and other mental states as being conscious; being conscious is a property we ascribe to mental states. But we also speak of seeing, believing, and many other mental states as states of being conscious of things. Here being conscious of something is the same as being aware of it; it's not that the state is conscious, but that being in the state is a way of being conscious or aware of something.

Subliminally seeing a red, square stimulus is being subliminally conscious or aware of that stimulus; one is aware of the stimulus, just not consciously aware of it. It is a way of being aware of the red square, even though the state of seeing is not on that account a conscious state. Being conscious of something can dissociate from the relevant psychological state's being a conscious state.

If there were no seeming to seem distinct from merely seeming, we couldn't distinguish conscious from subliminal perceiving, nor conscious from nonconscious believing. Subliminal perception and other nonconscious mental states are cases of its seeming to one that things are a particular way without its also seeming to one that things seem to be that way. They are cases of being aware of something even though the psychological state in virtue of which one is aware of that thing is not a conscious state. Perceiving and believing couldn't fail to be conscious if there were no distinction between seeming and seeming to seem.<sup>11</sup>

Given Dennett's rejection of a distinction between seeming and seeming to seem, it's no surprise that he sometimes explicitly elides the distinction between one's being conscious of something and a mental state's being conscious. FPO, he tells us, "denies the possibility in principle of consciousness of a stimulus in the absence of the subject's belief in that consciousness" (1991a, 132). But adoption of the intentional stance will sometimes enable one to discern that the subject is conscious of something even if the subject sincerely denies being conscious of it. Dennett's denial that one can be conscious of something without realizing that one is not only rules out subliminal perception; it also encounters difficulty with possible conflicts between third-person ascriptions by the intentional stance of another person's mental states and the FPO insistence that that people's beliefs about their own mental states is the last word about those states.

Dennett of course recognizes that subliminal perception and unconscious

---

<sup>11</sup> Dretske's account of mental states' being conscious runs afoul of the same difficulty. On his view, a state is conscious if in virtue of being in that state one is conscious of something; "[e]xperiences and beliefs are conscious, not because you are conscious of them, but because, so to speak, you are conscious with them" (1995, 280-1). But nonconscious states make one aware of things no less than conscious states, just not consciously aware.

believing occur. But he holds it's a mistake to describe these things in ordinary folk-psychological terms. Rather, he insists, we should describe these occurrences in terms of what he calls "events of content fixation" (1991a, 365, 457f.). Thinking of these cases as genuine perceiving and believing, thereby applying the categories of folk psychology to them, fails to recognize their fleeting nature and the rough-and-tumble competition among them to precipitate into stable, well-ordered, conscious folk-psychological states.

Dennett's claim that nonconscious states are mere events of content fixation, and not genuine folk-psychological states, is of a piece with his treatment of the puzzles about timing. When the long-haired woman without eyeglasses runs by, events of content fixation that result from that visual input compete with events of content fixation that stem from the memory of one's friend with short hair and eyeglasses. The competition happens to result in a conscious perception of the short-haired friend with eyeglasses. And if there is no fact of the matter about whether the memory intrudes before or after the visual input makes it to consciousness, there are no folk-psychological states, properly so called, prior to one's conscious perception, only competing events of content fixation.

But this doesn't give the relevant nonconscious states their due. We taxonomize states in folk-psychological terms when they have mental content, even if they are fleeting and in competition for ultimate standing in our stream of consciousness. And the visual input of one woman and memory of another have the right kind of content, however fleeting and unstable the states may be. And folk-psychological states can fail to be conscious and still be long lasting and have a stable effect on our mental lives. There is no non-question-begging reason to deny full folk-psychological standing to contentful states one is unaware of being in.

If one recasts issues about mental states generally in terms of neural function, as Dennett sometimes does (e.g., 2015), then it may well seem inviting to dismiss states that aren't conscious as merely neural, not appropriately taxonomized in terms of folk-psychological categories. But the properties that matter to our folk-psychological taxonomizing are those that pertain to perceptible properties, such as color and sound, and intentional content; consciousness isn't needed for states to have those properties.

Our problem is to square the first- and third-person perspectives on the mental, in a way that does justice to both. Dennett's approach wisely starts from the third person, recognizing that first-person resources are too sparse to allow us to construct a satisfactory account of the third-person aspect of mind. Nonetheless, his approach arguably founders on a set of closely connected issues. On FPO one's beliefs about one's psychological states are the last word about what mental states one is in. And

Dennett rejects taxonomizing any nonconscious states as mental. So such states cannot occur without being conscious; states that aren't conscious cannot be genuine psychological states. States of seeming are folk-psychological states; so there are no states of seeming that aren't conscious. And since no states of seeming occur without being conscious, there is no room for a distinction between how things seem and how it seems to one that they seem.

The way FPO construes consciousness and the first-person aspect of mind accordingly makes for a difficulty in having them square with the third-person aspect of mind. If we accord a robust privilege to first-person beliefs about what one is conscious of, there are only two ways for our first- and third-person perspectives on psychological phenomena to fit together. One way, unavailable to Dennett, would be to downgrade our third-person access by denying that we often can tell what psychological states others are in. That is not only unrealistic; it also flies in the face of Dennett's sensible realism about the states we discern by the intentional stance.

The other way to square things is to downgrade the nature of the states discernible by the intentional stance when the individual in question is unaware of being in these states. That's Dennett's way; we mustn't taxonomize such states in folk-psychological terms, but should regard them instead as mere events of content fixation. But there is every reason to apply ordinary folk-psychological categories to such states. Given FPO, there seems no third way to have first and third person fit together. Dennett's account does not, after all, provide a successful path from the third-person aspect of mind to its first-person aspect.

#### **§IV. The Higher-Order Alternative**

The most promising way to get an account on which first- and third-person aspects mind fit comfortably together is to start from the third person. But it's important to bear in mind, as noted in section I, that we can do justice to intuitions that involve the first-person perspective without counting them all as true. Not all appearances in any area are accurate. When they aren't, however, it's seldom appropriate simply to dismiss them; we must explain why it is that they strike us as compelling. Even when we can't save the intuitive appearances because they aren't accurate, explaining why we have those appearances is a sound way to do justice to them.

The first-person appearance that has been causing trouble is the idea that first-person access as infallible, since that more than anything else blocks comfortable co-existence with any third-person approach to mental states. If the strategy of starting

with the third person and constructing a satisfactory first-person picture of mind must accommodate infallibility, it will almost certainly fail. By contrast, if first-person access is not infallible, it's open for one's subjective sense of what mental state one is in sometimes to be inaccurate; one may on occasion be aware of oneself as being in states of seeming that differ from the states of seeming one is actually in. It would then seem to one that one's states of seeming are different from what they actually are.

Dennett's FPO, by preserving what is in effect an operationalist version of infallibility, won't give us what we need. But we can get that with an account that is a close cousin of Dennett's. On that account, what it is for a mental state to be conscious consists in one's being aware of oneself, in the right way, as being in that state; a state is conscious if one has a suitable higher-order awareness that one is in that state. This is close to Dennett's view, since it accommodates his idea that the consciousness of mental states consists in what is given "to be taken one way rather than another."

But the alternative, higher-order account denies the operationalist way that FPO spells out what it is for the given to be taken. Being taken is not a matter of "'writing it down' in memory" or "of the effects of various vehicles of content on subsequent action (and hence, of course, on memory)" (1991a, 132). Rather, it is simply a matter of one's being aware of oneself as being in the mental state in question.<sup>12</sup> FPO insists that "there are no fixed facts about the stream of consciousness independent of particular probes" (1991a, 138; cf. 275). The alternative account takes probes as providing symptoms or expressions, verbal or otherwise, of the higher-order awareness in virtue of which mental states are conscious.

Such an account is compelling for independent reasons. Once we acknowledge that genuine folk-psychological states sometimes occur without being conscious, it's hard to avoid some type of higher-order theory about what it is for such states to be conscious. No mental state is conscious if one is wholly unaware of being in it. So if there is solid evidence that somebody thinks or sees something but the person sincerely denies doing so, the thinking or seeing simply isn't conscious. And that's equivalent to a state's being conscious only if one is aware of being in it.

Not any type of higher-order awareness will do; if one is aware of oneself as thinking or seeing something only because of applying a theory to oneself, experimental findings, observing one's own behavior, or taking somebody else's word for it, such thinking or seeing will not be conscious. One must be aware of oneself as being in the

---

<sup>12</sup> That higher-order awareness will typically, perhaps always, occur very slightly after the onset of the state it makes one aware of being in. But there is no reason to think that the process involves memory of any sort. I am grateful to Mandik for pressing this question.

state in some way that is independent of conscious observation and inference, that is, observation and inference of which one is aware. It must seem subjectively to one that one's awareness of the state does not rely on inference or observation.<sup>13</sup>

The traditional view is that first-person access is direct and unmediated. This goes well beyond what we have any reason to believe. Our first-person access, based as it is on the relevant state's being conscious, seems subjectively to involve no mediation by observation or inference. That doesn't show that no mediation occurs, only that if there is any it isn't conscious; we are unaware of any mediation that does occur. These considerations do justice to the first-person impression of actual lack of mediation not by crediting it with being true, but by explaining why it seems compelling.

Subjective impressions about lack of mediation aside, we can also explain why it's inviting to regard first-person access as infallible, again without taking that traditional doctrine to be true. There is no subjective check on the higher-order awareness we have of our conscious states. And because there is no subjective appeal beyond our higher-order awareness, we have a subjective impression that such awareness is the final word about our mental lives. But it is only the final subjective word about our mental states, the final word simply about how we are aware of the states, not about what those states actually are.

There is more to be said about the kind of higher-order awareness that figures in a mental state's being conscious. I have argued elsewhere (e.g., 1986, 2005) that we can best understand that higher-order awareness as consisting in one's having a thought to the effect that one is in the state in question. Having such higher-order thoughts (HOTs) provides a more satisfactory explanation than any other type of higher-order awareness.

One especially important advantage of construing one's awareness of each conscious state as a HOT to the effect that one is in that state is that it offers a ready explanation of why, barring special deficits or unusual circumstances, psychological states are verbally reportable just in case they're conscious. Thoughts are expressible in speech; if I have a thought that it's raining, I can express that thought by saying that it's raining. Similarly, if one has a thought that one is in a particular mental state, one can express that thought by saying that one is in that state. Expressing a HOT simply is reporting the state that the HOT is about. The coincidence of a state's being reportable and its being conscious barring special circumstances is best explained on the

---

<sup>13</sup> One's having consciously inferential or observational reason to be aware of oneself as being in a state does not prevent the state from being conscious so long as one is also aware of oneself as being in the state in a way that subjectively seems independent of inference and observation, contrary to an objection of Block's (2011, 446).



hypothesis that the way we're aware of a conscious state is by having a thought that one is in that state. No other type of higher-order awareness explains the ready reportability of states that are conscious.

HOTs need not be conscious, and indeed seldom are; they're conscious only if there's a third-order thought about the second-order thought. Conscious HOTs likely figure in introspective consciousness. When we introspect, we are not just aware of being in a mental state; we are also aware that we are thus aware. We attentively focus on our awareness of the introspected state; we are accordingly consciously aware of that state.<sup>14</sup>

Ordinary, nonconscious HOTs, by contrast, explain in virtue of what a conscious state differs from a mental state that isn't conscious. Since a mental state fails to be conscious only if one is wholly unaware of being in it, being aware of it in some way is necessary for the state to be conscious. Subjectively noninferential reportability points to HOTs as the way one is aware of one's conscious states. But as long as the HOT isn't itself conscious, one is aware of the state but not consciously aware of it; so there is no introspective awareness and no conscious judgment that one is in the state.<sup>15</sup>

Acknowledging genuinely psychological states that aren't conscious allows a good fit between first and third person. We have pretty reliable access to others' psychological states, essentially by the kind of tools Dennett describes in connection with the intentional stance. When we have solid third-person reason to ascribe a psychological to somebody that the person sincerely denies being in, we typically posit a psychological state that isn't conscious but is still taxonomized in terms of standard folk-psychological categories.

Such positing of nonconscious states will rely simply on folk-psychological considerations; novels and plays have been replete with examples long before the influence of Freud. There is no need to demote such nonconscious states to the status of mere events of content fixation. On a higher-order theory, such a state is genuinely folk-psychological despite its being unaccompanied by a suitable higher-order awareness.

FPO "brusquely denies the possibility in principle of consciousness of a stimulus

---

<sup>14</sup> Amy Kind (manuscript) has forcefully argued that such attentive focus is not present in all cases of what are generally regarded as introspectively conscious states. Still, many introspectively conscious states do involve some deliberate, attentive focus, and an account that appeals to third-order thoughts—i.e., to our being aware of being aware of our introspected states—arguably does apply to them.

<sup>15</sup> Contrary to Block's (1995, 235) apparent assimilation of consciousness that involves HOTs to monitoring or reflective consciousness, by which he evidently has in mind introspective consciousness. Such assimilation likely rests on a tacit assumption that HOTs are invariably conscious.

in the absence of the subject's belief in that consciousness" (1991a, 132). That denial concerns one's being conscious of observable objects and processes; it does not on that account bear on what it is for a psychological state to be a conscious state.

The consciousness of mental states is a matter of mental appearance; it is the way our mental lives appear to us. We are likely in a great many states of which we are wholly unaware but, because of their content, are appropriately described in folk-psychological terms. Mental states are conscious only if we appear subjectively to be in them. And because these states have content, they reflect the way things appear to us. So there is a second level of appearing when the states are conscious, the subjective appearance that one is in those states. That is what seeming to seem consists in.

But appearance can be illusory. Just as a first-order appearance of a red square can be illusory, so can the higher-order, subjective appearance that one is seeing a red square. If there is higher-order awareness, it must be that there can also be higher-order misrepresentation. This has seemed to many to be the undoing of any higher-order theory. But there is no reason to doubt that consciousness can misrepresent what mental states one is in (Weisberg 2008, 2011a, b; Rosenthal 2011, 2012).<sup>16</sup> Indeed, consciousness misrepresents whenever one is in a psychological state that isn't conscious, at least insofar as we typically assume if only tacitly that consciousness reveals all our current mental states. So consciousness in effect represents one as not being in any mental state that isn't conscious. And we are routinely aware of our color sensations in a relatively generic way, for example, simply as generic sensations of red, though the visual sensations doubtless reflect a far more specific shade. That too is a type of higher-order misrepresentation, albeit mild and innocuous.

Higher-order misrepresentation also occurs, however, in more dramatic ways. A striking example occurs in a particular type of change blindness in which the display changes during a saccade, when no significant retinal information reaches visual cortex. Using eye trackers to time the changes, John Grimes found a significant numbers of subjects failed to notice even highly salient changes, for example, a change of color from green to red in a parrot that occurs centrally in the display and occupies about 25% of it.<sup>17</sup>

After the parrot changes from green to red, retinal information of red presumably reaches visual cortex; one has a sensation of red. But if one notices no change, what it's like for one is that one is still consciously seeing green; one remains aware of

---

<sup>16</sup> And as Weisberg (2008, 2011a, b) forcefully shows, it would be no obstacle to misrepresentation by consciousness for the higher-order awareness to be intrinsic to the state it makes one aware of.

<sup>17</sup> Thanks to Dennett for having first alerted me many years ago to this work.

oneself as having a sensation of green. One's higher-order awareness misrepresents one's first-order qualitative mental state. Doubtless random saccades often result in this kind of effect in everyday life.

Michael A. Cohen and Dennett have recently argued against any view that dissociates consciousness from all cognitive function. They initially describe such cognitive function very inclusively, as including "attention, working memory, language, decision making, motivation etc." But they immediately go on to describe it more narrowly, simply as "verbal report, button pressing etc." (2011, 358). And they note in a follow up to their original article that the cognitive functions they focus on "are all the products of cognitive access" (2012, 140).

Cohen and Dennett stress that they are not identifying a mental state's being conscious with its having some such cognitive function, but rather simply denying that one can detect a conscious state in others independent of any such function (2012). And that appeal to cognitive function is in keeping with the importance of the third-person considerations urged in section I. Still, focusing on the broad list of cognitive functions Cohen and Dennett initially offer risks distracting from what it is in virtue of which mental states occur consciously.

Indeed, all the cognitive functions on the more inclusive list except sincere verbal report can arguably occur in connection with psychological states that aren't conscious (Rosenthal 2008; 2012, §5). Even attention, which Cohen and Dennett highlight (2011, 359-60), demonstrably occurs in the absence of conscious awareness (e.g., Norman *et al* 2013, 2015; van Boxtel *et al* 2010; see Montemayor and Haladjian, 2015, for a useful review). So not only must we avoid identifying those functions with consciousness; most can't reliably serve even as reliable indicators of a mental state's being conscious.

Cohen and Dennett's shorter list, "verbal reports, button presses," does reflect what it is for a mental state to be conscious, and singles out the pivotal indicator. A sincere denial of being in a mental state shows that if, despite that denial, the individual is actually in the state, that state is not conscious. Sincere denial is decisive because it reflects lack of any awareness that one is in the state. So a sincere report is similarly decisive that the reported state is conscious.<sup>18</sup> And since sincere reports express one's awareness of what one reports, such reports point to higher-order awareness of a state as constitutive of its being conscious.

---

<sup>18</sup> At least if the report does not rely on conscious inference or observation. The appeal to sincerity does not go beyond what is intersubjectively accessible; we can often tell, even if not perfectly, when others' speech acts are sincere.

Indeed, this is the upshot of the “perfect experiment” (2011, 361-2) Cohen and Dennett describe in support of their appeal to cognitive function. In that thought experiment we imagine the cortical area responsible for conscious color experiences having been severed surgically from every other cortical area.<sup>19</sup> Because the area is detached, the patient has no access to whatever occurs there, and so would sincerely deny having any color experiences. Cohen and Dennett conclude that so far as consciousness is concerned it doesn’t matter what happens in the detached area; sincere report and denial are the last word about consciousness. And though they don’t go on to say as much, the best explanation of why sincere report is decisive about mental states’ being conscious is that it expresses one’s higher-order awareness of those states, and that that’s constitutive of the states’ being conscious.

Cohen and Dennett’s main target is the view advanced by Ned Block (2005, 2007) that what Block calls phenomenal consciousness is wholly independent of and can occur in the absence of any form of cognitive access. That would also undermine any higher-order theory, since higher-order awareness is a type of cognitive access. Block proposes identifying phenomenal consciousness with particular neural correlates; we could then determine its presence neurally. But as Cohen and Dennett argue, one cannot establish neural correlates of any type of consciousness without an independent way to tell when such consciousness occurs.<sup>20</sup> Any such independent way will inevitably be psychological, and so detectable in others only by their behavior; hence the crucial role of the cognitive functions Cohen and Dennett appeal to.

Dennett would resist inferring from cognitive function to higher-order awareness; we should not, he would urge, construe the cognitive access he and Cohen insist on as a form of higher-order awareness. One way he might resist that construal would be by appeal to the intentional stance as the arbiter of objective psychological reality. The intentional stance underwrites ascribing first-order psychological states, but does not also sustain ascribing higher-order awareness of those states.

But the intentional stance is also neutral about whether the states ascribed to others are conscious; it tells us what psychological states others are in, but not whether they’re conscious. Ascription by way of the intentional stance rests on a background assumption of an individual’s rationality, and rationality is independent of whether the relevant psychological states in question are conscious. If believing that p and desiring

---

<sup>19</sup> Cohen and Dennett acknowledge that the thought experiment is highly unrealistic (2011, 361).

<sup>20</sup> Indeed, Block (e.g., 1995, 2005, 2007) arguably recognizes the need in practice to rely on signs of cognitive access to determine the presence of phenomenally conscious states in others, despite his insistence that phenomenal consciousness is independent of any such access (see esp. 2007, 487, noted by Cohen and Dennett, 2011, 360).

a makes it rational to do a particular thing, that's independent of whether the belief and desire are conscious (Rosenthal 2008, §2; 2012, §5). So the failure of the intentional stance to ascribe higher-order awareness can't tell against explaining a mental state's being conscious in terms of higher-order awareness, since the intentional stance is also silent about whether psychological states are ever conscious. Indeed, the intentional stance by itself doesn't even sustain ascribing cognitive access of the sort Cohen and Dennett acknowledge as crucial to the study of consciousness.

Even if the intentional stance is normally neutral about whether an ascribed state is conscious, one might urge that it need not always be.<sup>21</sup> If somebody says, 'I think it's raining', that speech act, taken literally, reports one's thought that it's raining, and it's natural to take that report to be independent of conscious inference or observation. So the speech act would express a HOT that one thinks it's raining, and we should accordingly ascribe to the speaker a conscious thought that it's raining.

But no appeal to explicit reports of one's mental states plays any role whatever in the intentional stance; rather, such an appeal would supplement the intentional stance. The intentional stance is concerned solely with which thoughts and desires we should posit, in a folk-psychological way, to make an individual's speech and other behavior rational. And the states that would make it rational to say 'It's raining' are exactly the same as those that would make it rational to say 'I think it's raining'; performing either speech act is wholly rational if one has a thought that it's raining, conscious or not.<sup>22</sup>

The exclusive reliance by the intentional stance on rationality results in its treating the two types of speech act as equivalent. But they clearly are not, since they differ in truth conditions; each can be true when the other is not. This echoes Wittgenstein's focus on use to the exclusion of truth conditions, which also leads to one's seeing as equivalent the speech acts of saying that p and saying that one thinks that p (Wittgenstein 1953, I, x).

It's inviting to speculate that this limitation on the intentional stance leads Dennett to his firm rejection of explaining consciousness by appeal to higher-order awareness (Dennett 1991, ch. 10, especially pp. 314-320). If the intentional stance is the last word about objective psychological reality but can't distinguish conscious from nonconscious psychological states, then if any psychological states are conscious, all of them are.

---

<sup>21</sup> I am grateful to David Pereplyotchik for raising this possibility.

<sup>22</sup> Verbally expressed thoughts are typically conscious, but their being conscious contributes nothing to the rationality of the speech acts that express them. Indeed, the only available account of why verbally expressed thoughts are conscious that doesn't simply stipulate an unexplained tie between consciousness and language makes no appeal to rationality. See Rosenthal 2005, ch. 10.

We must then demote any nonconscious states that function much as conscious psychological states do to the status of mere events of content fixation, and not countenance them as psychological states at all.

But though the intentional stance is blind to the difference between conscious and nonconscious psychological states, we can supplement that stance with the sincere reports and button presses that Cohen and Dennett appeal to, which serve as evidence of the cognitive access Dennett accepts as pivotal to the study of consciousness,. Doing so preserves the reliance on third-person considerations, but also provides the resources needed for an appeal to higher-order awareness in explaining consciousness, and with it a clear and robust distinction between seeming and seeming to seem.

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

## REFERENCES

- Alter, Torin, and Sven Walter (2007), eds., Phenomenal Concepts and Phenomenal Knowledge: New Essays on Consciousness and Physicalism, Oxford: Oxford University Press.
- Block, Ned (1995), "On a Confusion about a Function of Consciousness," Behavioral and Brain Sciences, 18, 2 (June): 227-247.
- Block, Ned (2005), "Two Neural correlates of Consciousness," Trends in Cognitive Sciences, 9, 2 (February): 46-52.
- Block, Ned (2007), "Consciousness, Accessibility, and the Mesh between Psychology and Neuroscience," Behavioral and Brain Sciences, 30, 5-6 (December): 481-499.
- Block, Ned (2011), "Response to Rosenthal and Weisberg," Analysis, 71, 3 (July): 443-448.
- Breitmeyer, Bruno G., and Haluk Ögmen (2006), Visual Masking: Time Slices through Conscious and Unconscious Vision, 2nd edn., New York: Oxford University Press.
- Cohen, Michael A., and Daniel C. Dennett (2011), "Consciousness Cannot Be Separated from Function," Trends in Cognitive Sciences, 15, 8 (August): 358-364.
- Cohen, Michael A., and Daniel C. Dennett (2012), "Response to Fahrenfort and Lamme: Defining Reportability, Accessibility and Sufficiency in Conscious Awareness," Trends in Cognitive Sciences, 16, 3 (March): 139-140.
- Dennett, Daniel C. (1978), Brainstorms, Cambridge, Massachusetts: MIT Press/Bradford Books.
- Dennett, Daniel C. (1987), The Intentional Stance, Cambridge, Massachusetts: MIT Press/Bradford Books.
- Dennett, Daniel C. (1991a), Consciousness Explained, Boston: Little, Brown and Company.
- Dennett, Daniel C. (1991b), "Real Patterns," The Journal of Philosophy, 88, 1 (January): 27-51; reprinted in Dennett (1998), pp. 95-120.
- Dennett, Daniel C. (1996), Kinds of Minds: Toward an Understanding of Consciousness, New York: Basic Books.

- Dennett, Daniel C. (1998), Brainchildren: Essays on Designing Minds, Cambridge, Massachusetts: MIT Press/Bradford Books.
- Dennett, Daniel C. (2005), Sweet Dreams: Philosophical Obstacles to a Science of Consciousness, Cambridge, Massachusetts: MIT Press/Bradford Books.
- Dennett, Daniel C. (2009), "The Part of Cognitive Science that Is Philosophy," Topics in Cognitive Science 1, 2 (April): 231-236.
- Dennett, Daniel C. (2013), Intuition Pumps and Other Tools for Thinking, New York and London: W. W. Norton and Company.
- Dennett, Daniel C. (2015), "Why and How Does Consciousness Seem the Way It Seems?", in Open MIND, ed. Thomas Metzinger and Jennifer M. Windt, Frankfurt am Main: MIND Group, 2015, pdf pp. 387-398, at <http://open-mind.net/collection.pdf>.
- Dennett, Daniel C., and Marcel Kinsbourne (1992), "Time and the Observer: The 'Where and When of Consciousness in the Brain,'" The Behavioral and Brain Sciences, 15, 2 (June): 183-201.
- de Gelder, Béatrice, Jean Vroomen, Gilles Pourtois, and Lawrence Weiskrantz (1999), "Non-conscious Recognition of Affect in the Absence of Striate Cortex," NeuroReport, 10, 18 (December 16): 3759-3763.
- Dretske, Fred (2000), Knowledge, Perception, and Belief: Selected Essays, Cambridge: Cambridge University Press.
- Fernandez-Duque, Diego (2000) and Ian M. Thornton, "Change Detection without Awareness: Do Explicit Reports Underestimate the Representation of Change in the Visual System?", Visual Cognition 7, 1-3 (January-March): 324-344.
- Grimes, John (1996), "On the Failure to Detect Changes in Scenes across Saccades," Perception, ed. Kathleen Akins, New York: Oxford University Press, pp. 89-110.
- Kind, Amy (manuscript), "Thin Introspection."
- Koyama, Tetsuo, John G. McHaffie, Paul J. Laurienti, and Robert C. Coghill (2005), "The Subjective Experience of Pain: Where Expectations Become Reality," Proceedings of the National Academy of Science, 102, 6 (September 6): 12950-12955.
- Kripke, Saul A. (1980), Naming and Necessity, Cambridge, Massachusetts: Harvard University Press.



- Laloyaux, Cédric, Arnaud Destrebecqz, and Axel Cleeremans (2003), "Implicit Change Identification: A Replication of Fernandez-Duque and Thornton (2003)," Journal of Experimental Psychology: Human Perception and Performance, 32, 6 (December): 1366-1379.
- Montemayor, Carlos, and Harry Haroutioun Haladjian (2015), Consciousness, Attention, and Conscious Attention, Cambridge, Massachusetts: MIT Press.
- Nagel, Thomas (1974), "What Is It Like to Be a Bat?," The Philosophical Review, 83, 4 (October): 435-50.
- Norman, Liam J., Charles A. Heywood and Robert W. Kentridge (2013), "Object-Based Attention without Awareness," Psychological Science, 24, 6 (June): 837-843.
- Norman, Liam J., Charles A. Heywood and Robert W. Kentridge (2015), "Exogenous Attention to Unseen Objects?," Consciousness and Cognition 35 (September): 319-329.
- Öğmen, Haluk, and Bruno G. Breitmeyer (2006), eds., The First Half Second: The Microgenesis and Temporal Dynamics of Unconscious and Conscious Visual Processes, Cambridge, Massachusetts: MIT Press.
- Quine, W. V. (1960), Word and Object, Cambridge, Massachusetts: MIT Press.
- Quine, W. V. (1985), "States of Mind," The Journal of Philosophy, 82, 1 (January): 5-8.
- Rosenthal, David (1986), "Two Concept of Consciousness," Philosophical Studies, 49, 3 (May): 329-359.
- Rosenthal, David (1989), "Philosophy and Its History," in The Institution of Philosophy, ed. Avner Cohen and Marcelo Dascal, Peru, Illinois: Open Court, 1989, pp. 141-176.
- Rosenthal, David (2005), Consciousness and Mind, Oxford: Clarendon Press.
- Rosenthal, David (2008), "Consciousness and Its Function," Neuropsychologia, 46, 3: 829-840.
- Rosenthal, David (2009), "Philosophy and Its Teaching," in A Teacher's Life: Essays for Steven Cahn, ed. Robert Talisse and Maureen Eckert, Lanham, Maryland: Lexington Books, pp. 67-83.
- Rosenthal David (2010), "How to Think about Mental Qualities," Philosophical Issues, 20: 368-393.
- Rosenthal, David (2011), "Exaggerated Reports: Reply to Block," Analysis, 71, 3 (July 2011): 431-437.

- Rosenthal, David (2012), "Higher-Order Awareness, Misrepresentation, and Function," Philosophical Transactions of the Royal Society B: Biological Sciences, 367, 1594 (May 19): 1424-1438.
- Sellars, Wilfrid (1956), "Empiricism and the Philosophy of Mind," in Minnesota Studies in the Philosophy of Science, I, ed. Herbert Feigl and Michael Scriven, Minneapolis: University of Minnesota Press, 253-329; reprinted in Sellars, Science, Perception and Reality, London: Routledge & Kegan Paul, 1963, and Atascadero, California: Ridgeview Publishing Company, 1991, pp. 127-196.
- Sellars, Wilfrid, Science and Metaphysics (1968), London: Routledge & Kegan Paul; reprinted by Atascadero, California: Ridgeview Publishing Company, 1992.
- Shoemaker, Sydney (2003), "The Inverted Spectrum," in Shoemaker, Identity, Cause, and Mind: Philosophical Essays, 2<sup>nd</sup> edn., Oxford: Clarendon Press, pp. 327-357.
- Strawson, P. F. (1959), Individuals: An Essay in Descriptive Metaphysics, London: Routledge.
- van Boxtel, Jeroen J. A., Naotsugu Tsuchiya, and Christof Koch (2010), "Consciousness and Attention: On Sufficiency and Necessity," Frontiers in Psychology, published 20 December 2010, doi: 10.3389/fpsyg.2010.00217, available at <http://journal.frontiersin.org/Journal/10.3389/fpsyg.2010.00217/full>
- Weiskrantz, Lawrence (1997), Consciousness Lost and Found: A Neuropsychological Exploration, Oxford: Oxford University Press.
- Weisberg, Josh (2008), "Same Old, Same Old: The Same-Order Representation Theory of Consciousness and the Division of Phenomenal Labor," Synthese, 160, 2 (January): 161-181.
- Weisberg, Josh (2011a), "Misrepresenting Consciousness," Philosophical Studies, 154, 3 (July): 409-433.
- Weisberg, Josh (2011b), "Abusing the Notion of What-it's-like-ness: A Response to Block," Analysis, 71, 3 (July): 438-443.
- Wittgenstein, Ludwig (1953), Philosophical Investigations, tr. G. E. M. Anscombe, Oxford: Basil Blackwell.